VRIJE
UNIVERSITEIT
BRUSSEL

Master thesis

INVESTIGATING DOUBLE BUMP AIR SHOWERS WITH THE SKA

Vital De Henau

31/05/2024

promotor: prof Stijn Buitink
copromotor: Katarina Simkova and Mitja Desmet

Faculty of Sciences and Bioengineering Sciences

It's all a great mystery...Look up at the sky and you'll see how everything changes
–Antoine de Saint-Exupéry, The Little Prince

Abstract

Cosmic rays pose a fascinating scientific challenge. Their origin and composition remain unknown, the uncertainties are large in the transition region between Galactic and extragalactic sources (10^{16} - 10^{18} eV). These high-energy particles can collide with Earth's atmosphere causing a particle cascade called an extensive air shower (EAS).

Most EAS, exhibit a characteristic "universal" longitudinal profile, with a single peak. However, a fascinating and rare phenomenon known as the double bump shower deviates significantly from this typical profile. These double bump showers exhibit two distinct peaks, making them anomalous compared to the standard single-peaked EAS. This anomaly holds valuable clues about the primary cosmic ray particle and its interaction with the atmosphere. Between the universal and double bump profiles lies another category: stretched showers. Current cosmic ray experiments lack the resolution to distinguish between these different shower types. Simulations have shown that the upcoming SKA radio telescope will possess the sensitivity and resolution necessary to detect and reconstruct double bump showers. Therefore, we investigate the properties of double bump showers and the potential insights they offer into both the composition of cosmic rays and hadronic interaction models.

This thesis focuses on developing a reliable new method to identify double bump showers in simulations by analyzing their longitudinal profiles, ending up with almost twice as many double bump showers compared to the previous method. We will investigate the prevalence of these double bump showers across different cosmic ray primary particle masses and various hadronic interaction models used in the simulations. Then we will use these to identify shower parameters relevant for radio detection, R and L, as indicators for differentiating double bump showers from the more common universal cascades. The values found in this thesis are $R < 0.1$ or $R > 0.5$ and $L > 250$. The concept of "stretched showers" is introduced as an intermediate stage between the typical universal profile and the distinct double bump profile. By identifying a linear relationship between the rate of double bump showers and stretched showers we can be confident that these phenomena are related. Lastly, we will explore the potential of using these identified shower types in future SKA measurements to gain insights into hadronic interaction models and the composition of cosmic rays.

Acknowledgement

I am grateful to Professor Stijn Buitink for his guidance and support throughout my thesis journey. His expertise in the field of cosmic ray physics and his willingness to answer my questions, no matter how small, were instrumental in shaping this work.

My co-promotors, Katarina Simkova and Mitja Desmet, played a crucial role in the successful completion of this thesis. Their constant support, constructive feedback, and ability to keep me focused on the core research questions were invaluable. I am especially thankful for their willingness to answer even my seemingly "silly" questions and for ensuring I stayed on track without getting lost in unnecessary details, like making one to many needless plots.

I would also like to express my sincere appreciation to the colleagues I shared an office with and those who frequented it as a second social space. Their insightful feedback and, more importantly, their camaraderie made coming to the office a truly enjoyable experience, even with the occasional Nerf dart flying around.

Finally, I owe my deepest gratitude to my mother. Without her unwavering love and support, I would not have been able to complete high school, let alone this thesis. So thanks for making me the person I'm today.

Contents

1 Introduction

Cosmic rays will get a detailed section 2, high-energy particles originating from space, form a captivating and unique aspect of the universe. Their flux being predominantly composed of protons and atomic nuclei, these cosmic messengers exhibit energy levels surpassing those achievable in human-made particle accelerators. The question of their birthplace remains one of the most enduring mysteries in high-energy astrophysics. While supernova remnants (SNRs) and active galactic nuclei (AGNS) are prime suspects, definitive answers remain elusive. SNRs form the main galactic source but pulsar wind nebula also seem to be an option, while AGNs are thought to be the main extragalactic sources of cosmic rays, but other options are starburst galaxies or gamma-ray bursts. In between these is what is called the transition region.

Adding to the intrigue is the transition region, an energy regime between roughly 10^{16} and 10^{18} electronvolts (eV) where our understanding falters. Supernova remnants (SNRs), a popular candidate to be the main source within our galaxy, are believed to accelerate protons to a maximum energy of 2×10^{15}eV (the knee of the cosmic ray spectrum). Heavier nuclei can reach higher energy and the mass composition becomes progressively heavier towards the second knee. However, it is unlikely that SNRs can produce cosmic rays with energies exceeding the second knee. Cosmic rays from extragalactic components, on the other hand, are unlikely to reach us at energies much below the ankle because cosmic rays get trapped by (extra)galactic magnetic fields. Thus, there remains, a gaping hole in our knowledge in the $10^{16} - 10^{18}$eV energy range. Could these enigmatic particles obtain their energy in the strongly magnetized shock waves produced by supernovae involving Wolf-Rayet stars, or are they through a secondary acceleration cycle when they reach the galactic wind's termination shock?

These cosmic rays travel throughout the universe at near-light speed until they undergo a interaction. If they make it to the Earth they create extensive air showers (EAS), cascades of secondary particles triggered by cosmic rays interacting with Earth's atmosphere, which will be talked about more in section 3. These showers, though invisible to the naked eye, can be detected by various sensitive instruments, including scintillator detectors, fluorescence telescopes or radio arrays such as the Square Kilometre Array (SKA) radio telescope, providing crucial clues about the nature and origin of the primary cosmic ray. In section section 4 cosmic ray experiments will be treated in more detail.

Understanding the mass composition of cosmic rays is pivotal in unraveling their origin and the mechanisms behind their acceleration. These varying masses lead to unique interaction characteristics as they collide with air molecules in the atmosphere. By studying the compositional fingerprint of these interactions, we gain valuable clues about the astrophysical sources responsible for producing these cosmic messengers. One key difference lies in muon production. Different primary particle types, like protons or iron nuclei, produce varying numbers of muons, secondary particles that reach the ground and are detectable by ground-based observatories. This allows us to distinguish, to some extent, the nature of the primary cosmic ray based on the observed amount of muons.

However, an even more crucial aspect is the hadronic interaction cross-section. This essentially defines how deep the particle penetrates into the atmosphere (on average) before colliding collides with an atom in the atmosphere. Different particle types have varying cross-sections and initiate their cascade at different atmospheric depths. Atmospheric depth is a way to measure how much atmosphere a cosmic ray has traveled through. The different initial depths of the cascades, in turn, influence the depth at which the particle number is maximum (X_{max}) of the resulting extensive air shower, a measurable characteristic. The hadronic interactions governing the evolution of the EAS, are intricate and not fully understood. Currently, we cannot calculate cross-sections from first principles, relying instead on extrapolating from tuning models based

on particle collider data. As we push these models into the higher energy realms of cosmic rays, venturing beyond the familiar collider regime and insufficient measurements in the forward region, the models start to become increasingly uncertain. QGSJET, EPOS, and SIBYLL, widely used cutting-edge models, begin to diverge when the cosmic ray energies get out of reach of the collider experiments.

By refining our understanding of these interactions, we can improve the accuracy of hadronic interaction models, especially in the uncharted territories beyond collider energies. This, in turn, allows us to better interpret experimental data from extensive air shower observations, leading to a deeper understanding of cosmic ray composition, their origins, and helps refining the hadronic interactions models themselves.

Typical EAS exhibit a feature known as a universal longitudinal shower profile after they are transformed by moving X_{max} to a fixed position so that the variance of longitudinal development by species is reduced. One fascinating phenomenon observed in EAS is the double bump profile. Such an EAS does not conform to the universal profile and is considered anomalous. Unlike the typical single peak in particle count as the shower develops, these rare events exhibit two distinct peaks separated by a distance quantified by ΔX_{max}, offering clues about the primary cosmic ray and the hadronic interaction properties of the highest energy particles in the shower, created in the first interaction. In between the universal and double bump showers, stretched showers can be found, these are showers that exhibit a larger L value.

Imagine a cosmic ray, a high-energy particle, piercing through the atmosphere. In the initial collision, numerous secondary particles are born, some carrying a significant portion of the original energy. These high-energy particles can travel long distances before interacting again. One of two things can unfold. If these energetic secondaries travel further than expected, they delay the development of the main shower, resulting in an overall longer shower. This stretch in the longitudinal profile, quantified by the L parameter, extending beyond the typical depth of maximum development (X_{max}), provides valuable insights into the interaction properties of those energetic particles. Alternatively if the high-energy secondary travels even further a double bump shower occurs, creating a secondary peak in the particle count deeper in the atmosphere. This "echo" of the initial interaction offers unique information about the interaction cross sections. It will be visible in radio footprint measured by SKA, an example can be seen in figure 1.

By studying both long showers and double-bump events, we gain two crucial benefits: Constraining Hadronic Cross-sections: The behavior of these "wide" and "double-bumped" showers depends on the specific interactions between the energetic secondaries and air nuclei. Analyzing these events helps us refine and constrain the hadronic interaction models, which describe these high-energy collisions. This improved understanding is crucial for accurately interpreting extensive air shower data and constraining cosmic ray mass composition. Moreover, the fraction of "stretched" and "double-bumped" showers compared to universal showers is sensitive to the mass of the primary cosmic ray. Specifically, this can lead to huge improvements in proton-helium separation, which is crucial for the identification of cosmic-ray sources in the transition region.

By studying these rare events, we can refine our understanding of hadronic interactions, unveil the mass composition of the cosmic ray flux, and ultimately learn more about acceleration mechanisms and their sources residing in the transition region.

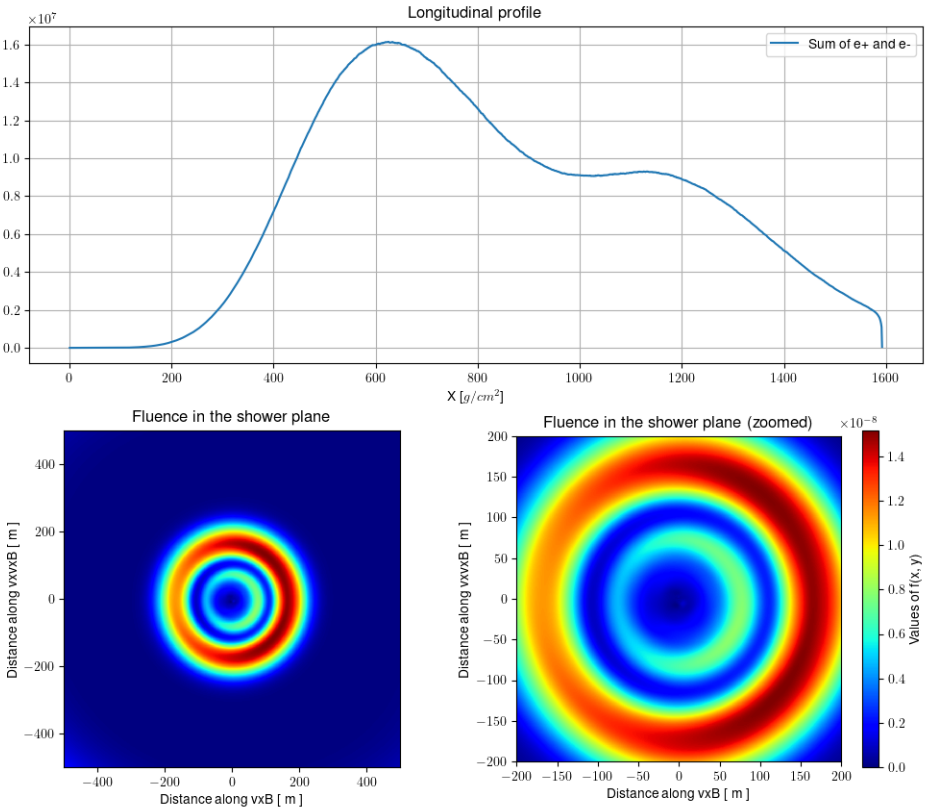


Figure 1: A simulation of the radio emission pattern on the ground from a double bump air shower induced by a proton primary. The data is presented in two key parts. The upper panel gives the longitudinal profile of the shower, in which two distinct peaks are visible. The two lower panels show the interpolated energy fluence contained in the radio emission, in the frequency range between $300 - 350$ MHz. In this frequency band, the two rings which reveal the double bump nature of the shower are most easily distinguished. This indicates that double bump showers will be recognisable by SKA. This shower was simulated using CORSIKA [1] with the CoREAS [2] plugin for the radio emission. A Fourier-based interpolation [3] as implemented in the NuRadioReco software [4] was used to construct the radio footprint. This plot was made by Mitja Desmet.

2 Cosmic rays

Have you ever looked up at the night sky, mesmerized by the vastness and wondering what lies beyond? Unbeknownst to most people, a constant rain of energetic particles, called cosmic rays, bombards Earth from the depths of space. But unlike the visible light we see, these cosmic rays remain largely invisible, shrouded in mystery.

Imagine particles with energies millions, of times greater than those achieved in our most powerful particle accelerators. These energetic travelers are cosmic rays, primarily protons and atomic nuclei, zipping through space at nearly the speed of light. The origin of these cosmic rays is one of the biggest unsolved puzzles in astrophysics.

While the definitive answer remains elusive, scientists suspect several celestial culprits. Supernovae, the explosive deaths of massive stars, are prime suspects. Black holes, with their immense gravitational pull, and other energetic astrophysical events are also potential accelerators. However, these are just educated guesses; the true sources are still hidden deep inside the universe.

2.1 History

It was Easter vacation in 1910, and physicist Theodor Wulf was not one to leave his scientific curiosity behind. He was known for developing portable electroscopes, and on his Parisian holiday, he brought a few along. Back then, prevailing theories suggested Earth itself was the source of mysterious penetrating radiation. Measurements taken over oceans, in salt mines, and even boxes filled with water supported this idea. Imagine then, Wulf's intrigue upon seeing the Eiffel Tower pierce the Parisian sky. It was the perfect opportunity to test this theory from a new vantage point. If the radiation originated primarily from the ground, its intensity should decrease exponentially with increasing height. While Wulf's experiment did not provide definitive answers, the idea of reaching higher altitudes to study cosmic phenomena would prove to be a fruitful one.

Hess, a German scientist and Jesuit priest, took the investigation a step further by conducting a series of balloon flights equipped with improved electroscopes. His observations defied the existing paradigm. Instead of a decrease in ionization with increasing altitude, he found a surprising rise beyond 1000 meters [5]. This revelation shattered the terrestrial origin theory and pointed toward a source beyond Earth's atmosphere.

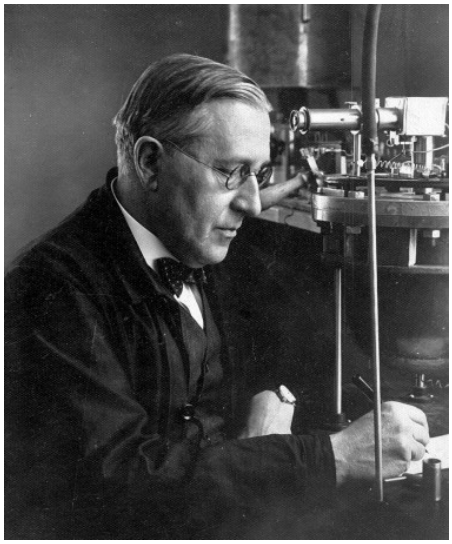
Hess's discovery received further support from other researchers like Pacini and Gockel, who conducted underwater [6] and balloon experiments, respectively. Their findings corroborated the idea of penetrating radiation originating from space. Despite these compelling results, the scientific community remained hesitant to fully abandon the terrestrial hypothesis.

The final confirmation came courtesy of another balloon experiment 2a by Hess himself [5]. Using a hermetically sealed electroscope, he conclusively demonstrated a steady increase in ionization with increasing altitude. Hess made earlier flights during a partial solar eclipse, meaning that the sun could be ruled out as a source, solidifying the extraterrestrial origin of cosmic rays. This groundbreaking discovery earned him the Nobel Prize in Physics in 1936.

The exploration of cosmic rays not only opened a new window to the universe but also played a crucial role in the discovery of fundamental particles. Early cosmic ray studies led to the identification of the positron, muon, and charged pions, paving the way for advancements in particle physics.

The term "cosmic ray" itself is attributed to Robert Millikan. Inspired by Hess's work, he devised an experiment involving measurements at different depths in high-altitude lakes. The observation that the ionization rate at the surface of a lower lake matched that of a shallower

(a) Victor F. Hess, center, departing from Vienna about 1911 [7].



(b) Victor F. Hess at his measuring apparatuses [8].

one at a higher altitude further cemented the cosmic origin of the radiation [9]. This realization solidified the term "cosmic ray" to describe this enigmatic phenomenon.

2.2 The Energy Spectrum of Cosmic Rays

Cosmic rays, come in a vast array of energies, spanning a mind-boggling range from a mere few GeV to hundreds of EeV. These cosmic messengers exhibit energy levels surpassing those achievable in human-made particle accelerators. The highest energy ever measured for a cosmic ray was $(3.2 \pm 0.9) \times 10^{20}$ eV [10], equivalent to a baseball traveling at about 100 km/h.

This spectrum and the energies compared to particle accelerators are depicted in Figure 3, this is not a smooth curve but rather a series of drops and changes in slope, revealing clues about the origin and acceleration mechanisms of these enigmatic particles. In general, the flux follows a power law:

$$\frac{dF}{dE} \propto E^{\alpha} \tag{1}$$

The flux decreases steeply as a function of energy, from about 1000 particles per second and m^2 at energies around the 10^9 eV, to less than one particle per century and km^2 above 10^{18} eV.

The hallmark features of the cosmic ray spectrum are the aptly named "knee" (from which point $\alpha \simeq -3.1$), "ankle" (from which point $\alpha \simeq -2.7$), and a possible "2nd knee". The names were chosen in analogy to a human leg that can change its slope at its knee and ankle. These distinct bends in the curve, signifying changes in the power law governing the particle flux, hint at different cosmic ray acceleration processes at play. The knee, located around 3×10^{15} eV, separates a region dominated by lower-energy cosmic rays, likely originating from within our galaxy, from a realm of ultra-high-energy cosmic rays (UHECR) with a steeper decline. The ankle, at around 3×10^{18} eV, marks the threshold for a milder decrease in particle flux.

One possibility is that the ankle signifies the limit of cosmic ray acceleration within our galaxy. Beyond this point, UHECRs might originate from distant extragalactic sources. Then cosmic rays reach the theoretical barrier known as the GZK cutoff (5×10^{19} eV). This limit

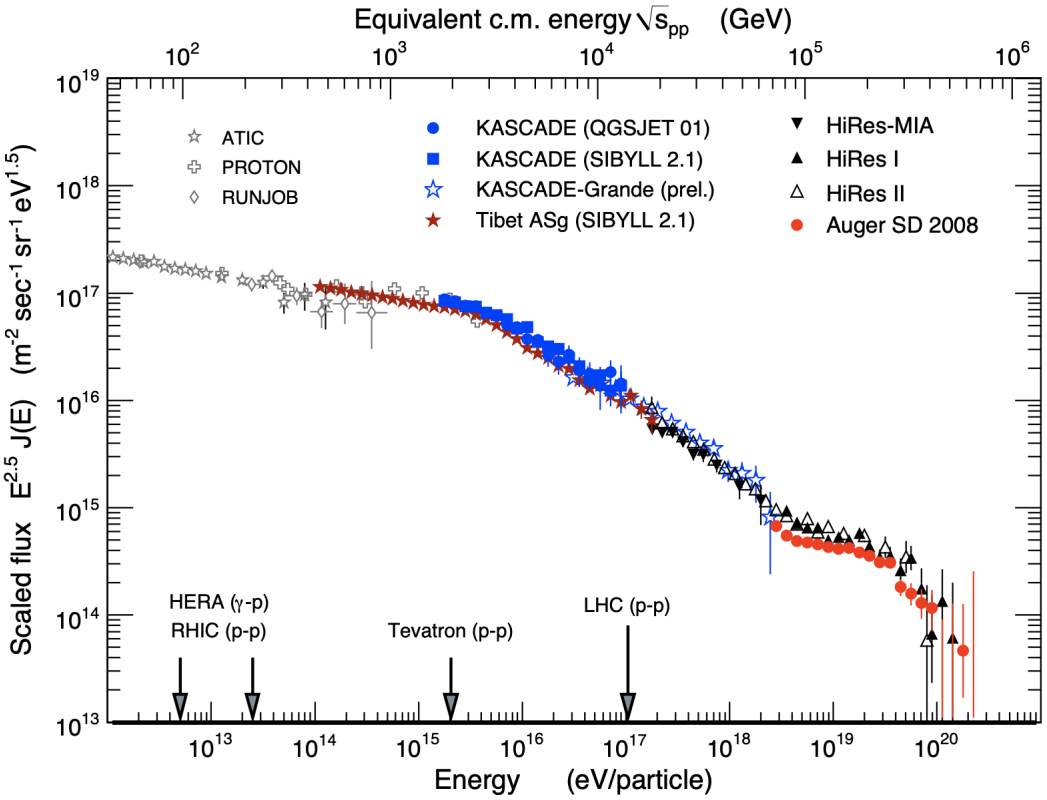


Figure 3: Cosmic-ray energy spectrum from various experiments and comparable energies of particle collision experiments compiled by[11].

arises from interactions with the cosmic microwave background radiation. However, it is likely not the only reason the spectrum ends. There is also evidence that extragalactic sources have an intrinsic limit to the maximum particle energies they can produce.

The composition of cosmic rays, and the types of particles they consist of, also play a role in understanding the spectrum.

2.3 Composition of Cosmic Rays

Cosmic rays, unlike the well-defined composition of our solar system, are a mysterious mix of elementary particles and atomic nuclei. This fascinating diversity in their makeup is a consequence of several factors. The abundance of a specific element in cosmic rays depends on its availability at the source, the processes that produce them, the mechanisms that accelerate them to such high energies, and the effects they experience as they propagate through the galaxy. By studying this cosmic ray fingerprint, we can piece together the origin story of these energetic particles and their journey across the vast of space.

Figure 4 showcases a comparison: the relative abundances of elements in cosmic rays versus those found in our solar system. Interestingly, both exhibit the "odd-even effect," where elements with even atomic numbers (Z) and stronger binding energies are more abundant.

However, there are two key distinctions between these compositions. The first difference lies

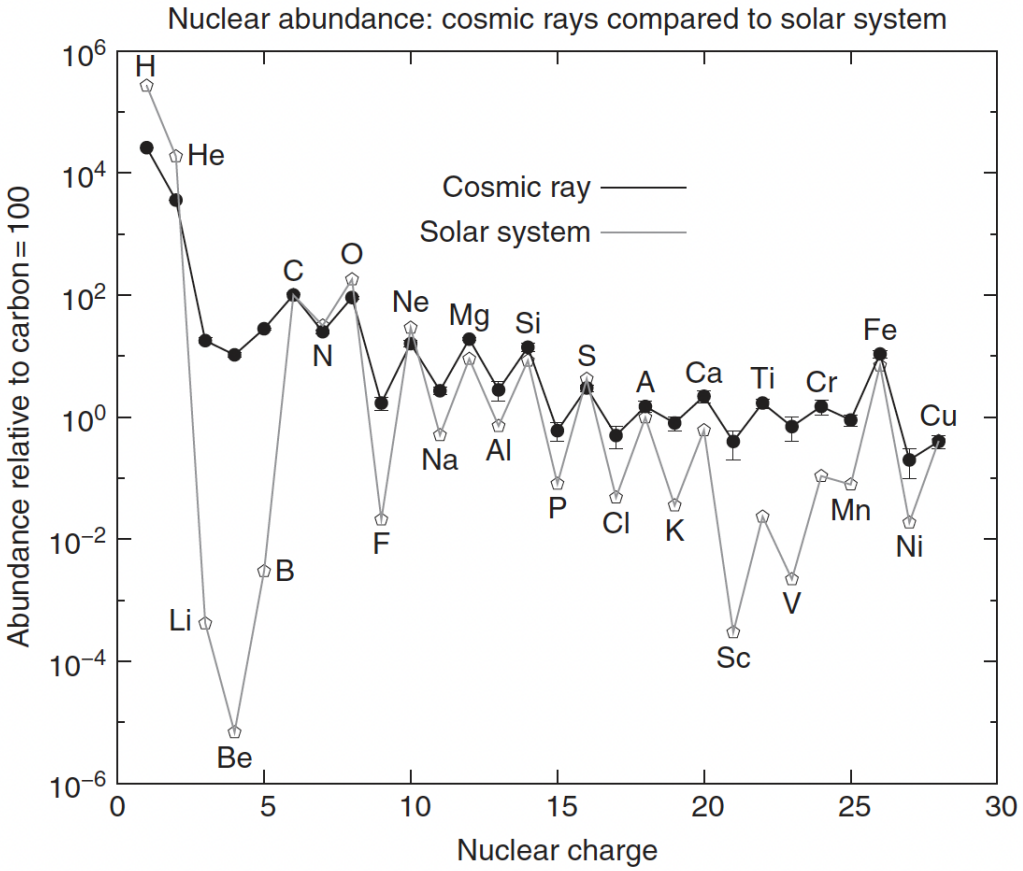


Figure 4: The cosmic ray elemental abundances measured on Earth (filled symbols connected by solid lines) compared to the solar system abundances (open symbols), all relative to carbon = 100 [12].

in the abundance of nuclei with $Z \sim 1$. Cosmic rays show a significantly lower abundance of these elements, compared to protons [12], than what is observed in our solar system. The exact reason behind this disparity remains unclear. It could be linked to the difficulty of ionizing hydrogen for injection into the acceleration process, or it might reflect a truly different elemental makeup at the source of cosmic rays.

The second major difference is well-understood and provides valuable insights into the propagation and confinement of cosmic rays within our galaxy. Two specific groups of elements, Lithium, Beryllium, and Boron (Li, Be, B), and Scandium, Titanium, Vanadium, Chromium, and Manganese (Sc, Ti, V, Cr, Mn), are surprisingly abundant in cosmic rays compared to their scarcity in the solar system. These elements are not typically produced during the final stages of stellar nucleosynthesis.

Their presence in cosmic rays can be attributed to a process called "spallation." Collisions with the interstellar medium (ISM), the dusty and gaseous material between stars in our galaxy, fragment heavier nuclei into lighter ones. By studying the spallation cross-sections (probability of a collision causing fragmentation), scientists can estimate the amount of matter a cosmic ray

traverses before reaching Earth.

On average, cosmic rays travel through a column density of roughly 5 g/cm^2 of interstellar material. Considering the average density of protons (ρ_N) in the galactic disk (around one proton per cubic centimeter), this translates to a path length of:

$$l = \frac{X}{\rho_N \ m_p} \approx 1000 \ kpc \tag{2}$$

However, this is likely a lower limit. Cosmic rays might spend time in the galactic halo, a more diffuse region surrounding the disk. Therefore, the actual distance traveled is likely much larger than the 0.1 kpc, which represents the half-thickness of the galactic disk. This significant path length implies that cosmic ray confinement within the galaxy is a diffusive process. These energetic particles essentially wander through the galaxy for extended periods before escaping into intergalactic space.

2.4 Sources and Acceleration of Cosmic Rays

Cosmic rays, come in a vast array of energies. This spectrum is not a smooth curve, hinting at different origins and acceleration mechanisms for these particles. The underlying mechanism for efficiently accelerating microscopic particles like cosmic rays to such tremendous energies (up to 10^{20} eV) remains an area of active research.

"Bottom-up" scenarios propose that these particles are accelerated from lower-energy precursors near their source. The primary contenders for this role are diffusive shock acceleration, a variant of the Fermi mechanism operating in SNRs, and acceleration in very intense electric fields [13].

Our exploration of cosmic ray powerhouses begins with the most familiar star – our Sun. Surprisingly, the Sun itself contributes to the lowest energy cosmic rays, those in the range of $0.1 - 10$ GeV [14]. These energetic particles erupt from the Sun during violent solar events like flares and coronal mass ejections. However, the Sun's contribution does not end there. At the higher end of the solar cosmic ray spectrum, we encounter a fascinating group called Anomalous Cosmic Rays. Unlike their solar counterparts, these particles have a very different elemental makeup compared to the typical galactic cosmic rays. The origin story of Anomalous Cosmic Rays unfolds in the heliosphere, the bubble carved out by the Sun's wind in the interstellar medium. Here, neutral gas from interstellar space interacts with the heliosphere. Solar ultraviolet radiation strips away electrons from these neutral atoms, transforming them into charged ions. Once charged, the solar wind's magnetic field takes over, transporting these ions towards the outer regions of the heliosphere, see 5 for diagram.

This journey becomes an acceleration opportunity. A process known as diffusive shock acceleration comes into play [16]. The termination shock is a boundary marking the thinning of the solar wind. Ions repeatedly collide with this shock wave, gaining energy with each bounce. This diffusive shock acceleration is the driving force behind the energetic Anomalous Cosmic Rays.

As cosmic ray energies climb beyond the realm influenced by our Sun, we venture into the heart of our Milky Way galaxy to explore the primary accelerators – supernova remnants (SNRs). These colossal structures, the expanding shells left behind by the explosive deaths of massive stars, are prime suspects for propelling cosmic rays to GeV energies and beyond.

The vast size and long lifespan of SNRs make them ideal cosmic ray factories. A typical galaxy like ours experiences roughly three supernovae per century. If just a small fraction ($5 - 10\%$) of the immense kinetic energy unleashed in such an explosion gets transferred to accelerating particles, it would be sufficient to account for the observed energy density of galactic cosmic rays

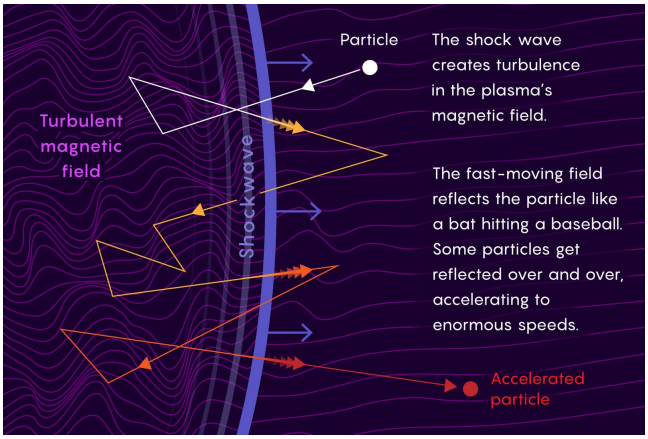


Figure 5: Schematic of cosmic ray acceleration by magnetic fields [15].

[17]. For example, for 10 $M_\odot$ ejected from a type II supernova with a velocity $v \sim 5 \times 10^8$ cm/s every 30 years:

$$L_{SN} \sim 1.872 \times 10^{45} \text{GeV/s} \tag{3}$$

The acceleration mechanism at play within SNRs resembles the diffusive shock acceleration seen in the heliosphere. Particles, like protons and atomic nuclei, get swept back and forth across the shock front, a boundary where the supersonic stellar wind slams into the surrounding interstellar medium. With each crossing, the particle gains a significant energy boost. However, this acceleration process is not a guaranteed one-way trip. There's a chance the particle might escape the shock region, ending its acceleration journey.

The maximum energy (E_{max}) a particle can achieve within an SNR is determined by its charge (Z_e), the shock velocity ($V_s = c\beta_s$), the intensity of the magnetic field (B), and the time spent in the acceleration zone (T).

$$E_{max} \simeq Z_e \beta_s B T V_s \tag{4}$$

Calculations suggest that for typical conditions in a Type II supernova and the interstellar medium, this upper limit is around 10^{14} eV for protons [18].

The galactic cosmic ray spectrum exhibits a fascinating feature known as the "knee" around 10^{15} eV. This point marks a shift in the slope of the spectrum, indicating a change in the dominant processes governing cosmic ray acceleration. Below the knee, the spectrum follows a power law with an index of $2.70 - 2.75$. However, above the knee, the spectrum steepens, suggesting a decline in the number of cosmic rays at these higher energies.

The origin of the knee is a topic of ongoing research. One explanation suggests it arises from interactions between cosmic rays and background particles during their journey through the galaxy. These interactions could lead to the fragmentation of heavier nuclei into lighter ones, potentially shaping the cosmic ray spectrum between 10^{16} eV and 10^{17} eV.

However, data suggests a trend towards heavier elements in this energy range, contradicting the fragmentation hypothesis. Another explanation focuses on limitations within the acceleration process itself. Models suggest that SNRs might have an inherent maximum energy achievable for particles based on their charge and the strength of the magnetic field at the source.

Additionally, models describing the diffusion of cosmic rays through the galaxy propose a "rigidity limit." Above a certain rigidity (ratio of a particle's momentum to its charge), nuclei

become less susceptible to the magnetic fields that confine them within the galaxy, leading to leakage and a drop in their abundance at higher energies [19].

Both these explanations, limitations within SNR acceleration and rigidity-dependent leakage, predict a knee-like structure in the spectrum of heavier cosmic rays around the observed energy range of 4×10^{16} eV to 1.2×10^{17} eV. Recent observations by the KASCADE-Grande experiment support this rigidity-dependent scenario [20].

The knee in the cosmic ray spectrum likely arises from a combination of these factors. The maximum achievable energy within the source (SNR) and the leakage of particles from the galaxy due to diffusion likely work together to shape this significant feature in the cosmic ray energy distribution [21].

The cosmic ray spectrum takes a fascinating turn beyond the "ankle," a point around 10^{18} eV where the dominance of galactic origins wanes [22]. Here, the challenge of unraveling the sources and acceleration mechanisms for these Ultra-High-Energy Cosmic Rays (UHECRs) intensifies.

The search for UHECR origins extends beyond the Milky Way, towards a diverse range of potential extragalactic sources. Here, the focus shifts to environments with colossal energy output and powerful magnetic fields. There is a limit to how energetic particles can get in an accelerator. The maximum energy (E_{max}) depends on the particle's charge (Z_e), the strength of the accelerator's magnetic field (B), the shock velocity in units of c (β), and the size of the accelerator itself (R).

$$E_{max} \simeq Z_e \beta (\frac{R}{kpc})(\frac{B}{\mu G}) \tag{5}$$

Only a few astrophysical sites are capable of accelerating cosmic rays to energy as high as 10^{20} eV, see figure 6.

Active Galactic Nuclei (AGNs), the luminous cores of galaxies harboring supermassive black holes, are prime suspects. Their immense size and formidable magnetic fields offer the potential to propel particles to the required energies. Particularly promising are AGN jets, collimated outflows capable of accelerating protons to 10^{20} eV. However, heavy nuclei might face challenges escaping the AGN's central region due to interactions with photons, suggesting a predominantly protonic composition for UHECRs originating from AGNs.

Radio galaxies, a class of active galaxies including AGNs, also emerge as candidates. Their vast dimensions and magnetic fields allow for particle acceleration to even higher energies, reaching 10^{21} eV. Notably, Centaurus A, the closest radio galaxy to Earth, holds particular interest for future observations.

Gamma-ray bursts (GRBs), the brightest electromagnetic events in the universe, are another possibility. Their short-lived, ultra-relativistic jets provide the necessary conditions for Fermi acceleration, potentially propelling particles to 10^{20} eV.

Starburst galaxies are galaxies undergoing intense bursts of star formation, churning out stars at a much faster rate than normal. The rapid production of stars means a frequent occurrence of supernovae. Starburst galaxies are also rich in gas clouds. These dense clouds serve as places where the energetic protons from supernovae can interact and produce neutral pions, which decay into gamma rays. This process contributes significantly to the overall gamma-ray emission from these galaxies.

The transition from galactic to extragalactic cosmic rays likely occurs between the knee (around 10^{15} eV) and the ankle (around 10^{18} eV). Two main approaches attempt to explain the observed changes in the spectrum at this energy range.

One approach proposes a clear distinction between the sources responsible for cosmic rays below and above the ankle. Here, the knee marks the limit of galactic acceleration, with extragalactic sources taking over and shaping the spectrum beyond the ankle. However, smoothly

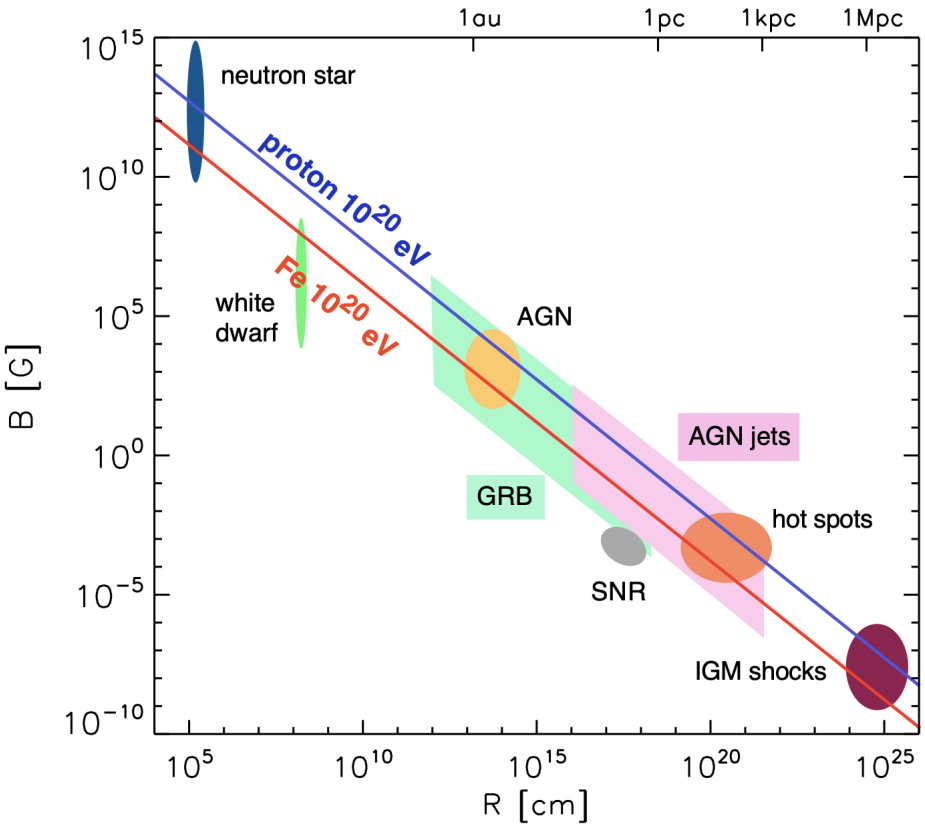


Figure 6: Astrophysical objects which can be the source of UHE cosmic rays are sorted according to their magnetic field strength and size. Objects below the diagonal line are not suitable sources [23].

merging the spectra from these two distinct origins remains a challenge.

Another approach suggests that the same class of objects might be responsible for acceleration up to the ankle, with the knee-to-ankle transition arising from the propagation of cosmic rays within the galaxy. As cosmic rays travel through interstellar space, they undergo various processes that alter their composition and energy. Interactions with the interstellar medium can lead to fragmentation (splitting of nuclei) and energy loss. Additionally, the galactic magnetic field deflects their trajectories, complicating the task of pinpointing their origin.

The quest to solve the UHECR mystery continues, with a diverse range of potential sources and acceleration mechanisms under scrutiny. By combining advancements in theoretical modeling, observations from ground-based detectors, and potentially future space-based missions, we may one day unravel the origins of these enigmatic cosmic visitors.

At energies above 10^{18} eV, a UHECR proton can interact with a CMB photon to create an electron-positron pair.

$$\gamma + \gamma_{CMB} \rightarrow e^- + e^+ \tag{6}$$

While this process only results in a small energy loss (around 0.1%), repeated interactions can significantly deplete the cosmic ray's energy over long distances.

The journey of ultra-high-energy cosmic rays (UHECRs) across vast distances in the cosmos is not without its challenges. As their energy climbs, they encounter a formidable obstacle – the cosmic microwave background (CMB), a faint echo of the Big Bang permeating the universe. This seemingly innocuous radiation field presents a significant hurdle for UHECRs exceeding a certain energy threshold, introducing a concept known as the GZK cutoff, predicted by Greisen [24], and Zatsepin and Kuzmin [25].

The culprit behind this energy barrier lies in the interaction between UHECR protons and CMB photons. When a high-energy cosmic ray proton collides with a CMB photon, it can undergo two key processes:

Photo-pion production, for even higher energies, exceeding $10^{19.6}$ eV, a more dramatic interaction becomes possible. Here, the UHECR proton interacts with a CMB photon to produce a delta resonance ($\Delta+$), a short-lived subatomic particle. The $\Delta+$ quickly decays, carrying away a significant portion of the cosmic ray's energy (around 20%).

$$\begin{aligned} p + \gamma_{CMB} \rightarrow \Delta^+ &\rightarrow \pi^0 + p \\ &\rightarrow \pi^+ + n \\ &\rightarrow p + e^- + e^+ \end{aligned} \tag{7}$$

The GZK cutoff essentially demarcates the energy limit beyond which these interactions become so frequent that cosmic ray protons effectively lose their ability to travel across vast cosmological distances.

The GZK effect applies not only to protons but also to heavier nuclei, albeit with some variations. Heavier cosmic ray nuclei can interact with the CMB and infrared background radiation (IRB) through photodisintegration, where a photon knocks out one or more nucleons (N) from the nucleus.

$$\begin{aligned} A + \gamma_{CMB,IRB} &\rightarrow (A-1) + N \\ &\rightarrow (A-2) + 2N \\ &\rightarrow A + e^- + e^+ \end{aligned} \tag{8}$$

The main channels for photodisintegration involve knocking out a single nucleon, either a neutron (γ, n) or a proton (γ, p). Double-nucleon emission processes, like $(\gamma, 2n)$, (γ, np), or $(\gamma, 2p)$, are about ten times less likely than single-nucleon emission. Similar to protons, pair production remains a relevant energy loss mechanism for heavy nuclei. However, the specific energy thresholds required for photodisintegration and pair production, as well as the overall shape of the flux suppression at high energies, differ between protons and heavier nuclei.

Observations from various cosmic ray experiments, including HiRes [26], Pierre Auger Observatory [27], and Telescope Array [28], have confirmed the presence of a flux suppression at high energies, consistent with the GZK cutoff. This finding strongly suggests that UHECRs with energies exceeding the predicted limit have likely not interacted with the background radiation field during their long journeys.

However, a lingering question remains. Could this suppression simply reflect the limitations of the cosmic ray sources themselves, perhaps an inability to accelerate particles beyond a certain energy? While the GZK effect provides a compelling explanation for the observed decline in the UHECR flux at high energies, the possibility of source-imposed limitations cannot be entirely ruled out [29]. Mass composition plays a crucial role in understanding the origin of UHECRs. If the sources could accelerate particles beyond the GZK cutoff energy, the highest energy particles would be those with the largest interaction lengths: protons and iron nuclei. Particles with intermediate masses would disintegrate more readily during their journey to Earth. However, if the observed cosmic ray spectrum reflects the maximum energy achievable by the sources

themselves, we would expect a gradual transition from lighter to heavier nuclei with increasing energy. This scenario resembles the rigidity-dependent "knee" observed in the spectra of cosmic rays from supernova remnants, where a change in composition marks the transition from lighter to heavier nuclei. The Pierre Auger Observatory boasts the most precise measurements of mass composition for UHECRs to date. These measurements favor the latter scenario, indicating a slow enrichment of heavier nuclei like iron at the highest energies. Unfortunately, Auger's limited fluorescence detector duty cycle restricts the collection of reliable depth of shower maximum (X_{max}) data for the highest energy events. X_{max} is a key parameter sensitive to mass composition. To address this limitation, Auger is undergoing a major upgrade (AugerPrime) that will significantly improve the capabilities of its particle detectors, particularly in separating different mass components. This upgrade will provide valuable insights into the origin and composition of UHECRs.

3 Extensive Air Showers

Cosmic rays, energetic particles bombarding Earth from outer space, put on a spectacular show when they enter our atmosphere. These high-energy collisions trigger a cascade of secondary particles, predominantly electrons, positrons, photons, and pions, in a phenomenon known as an Extensive Air Shower (EAS).

The initial interaction with an air molecule sets off a chain reaction, with secondary particles interacting further, creating even more particles. This cascade continues until the individual particles no longer have enough energy to keep the party going, see figure 7 and 8 illustrates the decomposition of an EAS into particle groups.

EAS are fascinating for scientists because directly measuring cosmic rays above a certain energy threshold (around 10^{15} eV) is impractical because of the very low flux at these energies. Instead, we use the atmosphere as a giant calorimeter, studying the EAS itself to understand the properties of the original cosmic ray particle. Scientists employ a variety of detective techniques, from large ground detectors to telescopes that capture the faint glow of the shower, to piece together the story of these cosmic visitors.

There are two main types of cascades within an EAS: the electromagnetic (EM) cascade and the hadronic cascade. The EM cascade, is initiated and dominated by electrons, positrons, and photons. These particles readily interact with the air, creating more particles with each interaction. The hadronic cascade, driven by pions, is more complex. Pions can either interact with air nuclei to create additional particles or decay into muons, which are able to travel to ground level without interacting. While hadronic interactions dominate at higher energies, eventually, the energy gets transferred to the EM component, leading to a combined shower.

By studying the properties of EAS, such as the depth in the atmosphere where the shower reaches its peak size and the number of muons produced, scientists can learn about the nature and energy of the original cosmic ray particle.

3.1 Heitler's model for EM showers

In an EAS, a cosmic ray slams into an air molecule, triggering a cascade of secondary particles electrons, positrons, and photons. Heitler's model [31] provides a simplified yet powerful tool to understand the key features of this EM shower.

Heitler's model depicts the EM shower as a perfect binary tree, see figure 9. High-energy photons, electrons, and positrons travel a fixed distance (determined by the radiation length of the medium) before interacting. At each interaction point, two things happen: Electrons and positrons radiate a single photon via a process called bremsstrahlung. Photons split into electron-positron pairs. Crucially, each interaction halves the energy of the parent particle, with the daughter particles sharing this energy equally. This branching process continues until the individual particles no longer have enough energy (critical energy E_c) to keep the shower going (around $80 - 85$ MeV in air).

Heitler's model successfully captures two important aspects of EM showers that can be observable. Particle Number and Primary Energy, the number of particles at the shower's peak (N_{max}) is proportional to the initial energy of the cosmic ray (E_0). The shower reaches its maximum size when all particles fall below the critical energy, halting particle production.

$$N_{max} = \frac{E_0}{E_c} \tag{9}$$

In the realm of cosmic rays interacting with our atmosphere, a crucial concept is atmospheric depth. This depth is not measured in meters or kilometers, but rather in units of gram per

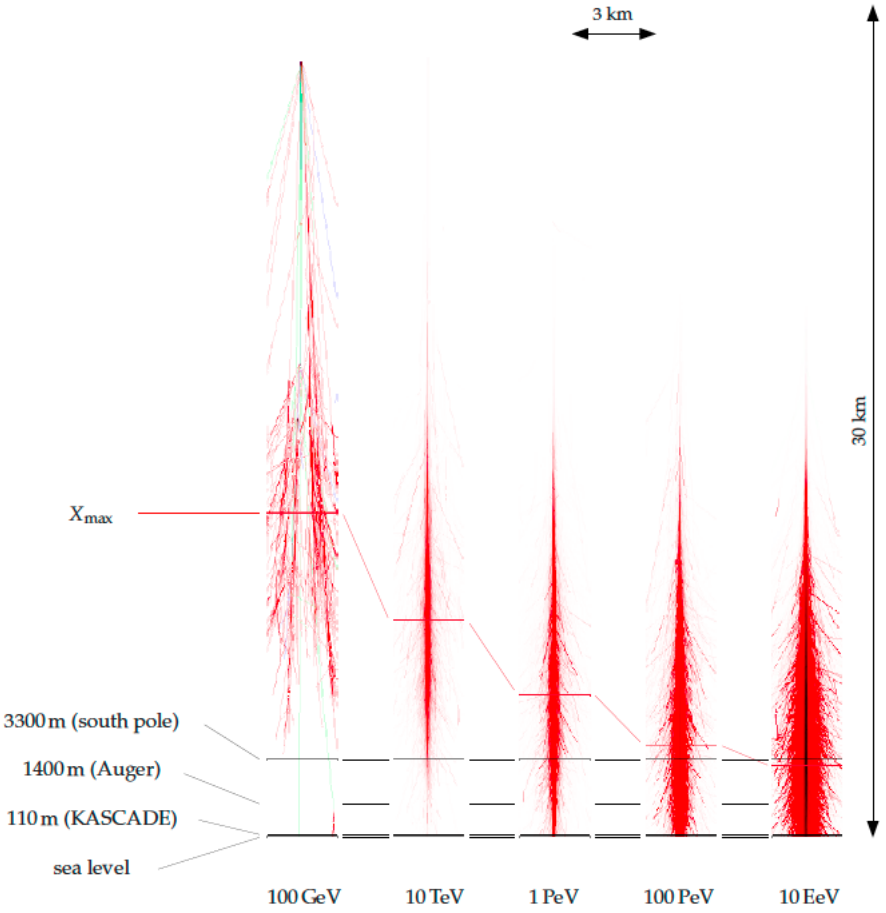


Figure 7: Simulated proton induced air showers at different primary energies E with CORSIKA. Due to the enormous difference in the total number of particles ($N \sim E$) the used intensity scale is different at each energy. The impression of the shower starting point moving deeper into the atmosphere with increasing energy is just an artifact of the changing intensity scale. In fact, higher energy showers tend to start higher in the atmosphere. The line marked by "X_{max}" identifies the shifting depth of the shower maximum, which increases approximately proportional to $\log(E/eV)$. Typical observation levels in altitude above sea level are indicated as horizontal black lines. Low energy air showers do not reach any ground based detectors, so they can only be studied using balloons or satellites. Starting from $\sim$ 1 TeV air showers are penetrating the atmosphere, which makes observations possible at high altitude detectors. With increasing energy X_{max} moves closer to the observation level, reaching it at some point [30].

square centimeter ($g\ cm^{-2}$). It represents the amount of material (air) a cosmic ray traverses on its path through the atmosphere. Higher atmospheric depth signifies a greater amount of air traversed, and lower depth indicates a lesser amount.

While atmospheric depth does not directly correspond to physical height, it provides a more relevant measure for studying cosmic ray interactions. Imagine two cosmic rays entering the atmosphere at different angles. One might travel a longer horizontal distance but encounter

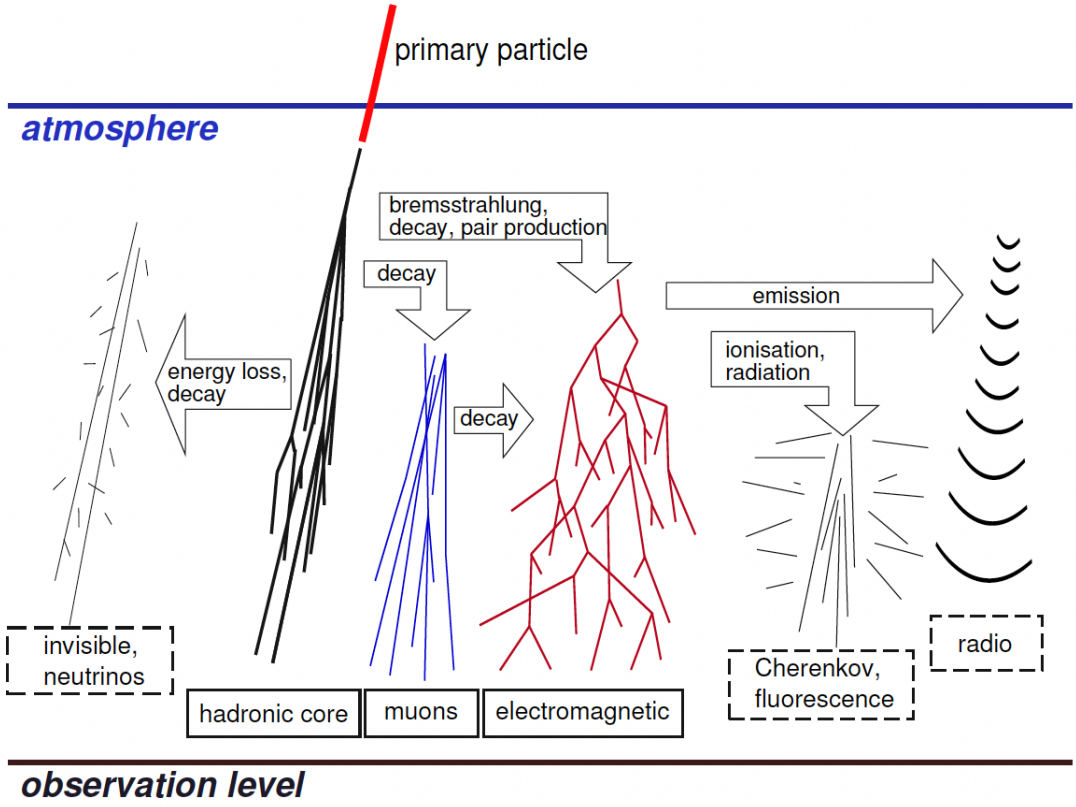


Figure 8: A primary cosmic ray nucleus hitting the atmosphere initiates a cascade of secondary hadrons, forming the inner core of an air shower. Muons are produced mostly due to the decay of charged pions. Electromagnetic particles are permanently produced mainly by decaying neutral pions and also decaying muons. In fact, after a few hadronic interaction lengths most of the shower energy is transferred to the electromagnetic shower component [30].

less air (lower depth) compared to another that travels a shorter horizontal distance but plows through a thicker layer of air (higher depth). The depth of maximum shower development, the depth in the atmosphere where the shower reaches its peak size (X_{max}) is logarithmically proportional to the initial cosmic ray energy.

$$X_{max} = X_0 + \lambda_r \ \ln\left(\frac{E_0}{E_c}\right) \tag{10}$$

With X_0 the depth of the first interaction and λ_r the radiation length in air $\lambda_r \approx 37$ g/cm^2.

While elegant, Heitler's model has its limitations. It overestimates the number of particles and the electron-to-photon ratio. The model's limitations include neglecting absorption of particles above the critical energy, multiple photon emission during bremsstrahlung, and the faster energy loss of electrons compared to photons.

Despite these limitations, Heitler's model provides a valuable foundation for understanding the core principles of EM showers within EAS. It offers a starting point for more complex simulations that incorporates these additional details.

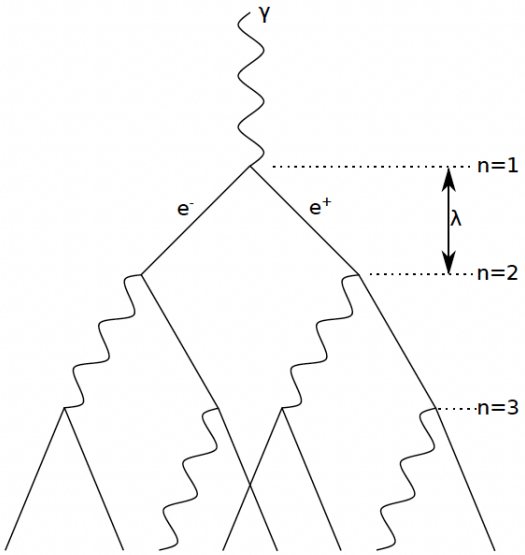


Figure 9: Schematic view of an electromagnetic cascade [12].

3.2 Hadronic showers

While Heitler's model provides a powerful tool for understanding EM showers, it does not account for all the particles produced in a cosmic ray collision so Matthews extended it [32]. This is where the hadronic component comes in. Hadronic showers are initiated by two particles that are made up of quarks, like cosmic rays made up of protons and iron nuclei, with air molecules, schematic 10.

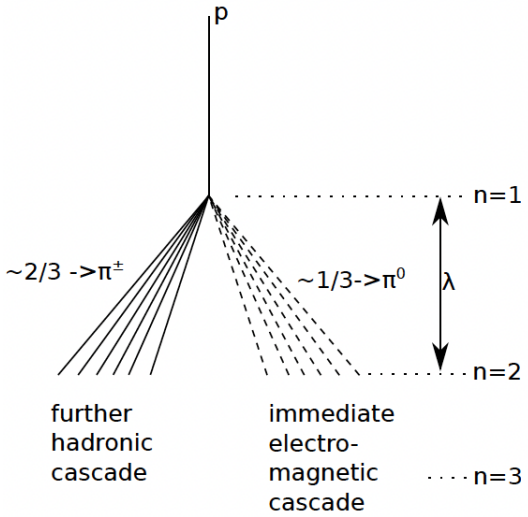


Figure 10: Schematic view of a hadronic shower [12].

Hadronic showers are dominated by pions, the lightest mesons. These come in three varieties: neutral pions (π^0), positively charged pions (π^+), and negatively charged pions (π^-). Neutral pions have an extremely short lifespan (8.4×10^{-17} s)and almost instantly decay into photons, initiating their electromagnetic sub-showers within the EAS.

Hadronic interactions follow a cascading process similar to the electromagnetic component.

A high-energy hadron collides with an air molecule, producing multiple pions (N_{ch}). The energy of the primary hadron gets divided equally among the produced pions. So the energy per charged pion in the nth interaction layer is given by:

$$E_\pi = \frac{E_0}{(3/2N_{ch})^{n_c}} \tag{11}$$

After a certain number n_c of generations, E_π falls below E_c^π. The number of interactions needed to reach E_c^π is calculated as:

$$n_c = \frac{ln(E_0/E_c^\pi)}{ln(3/2N_{ch})} \tag{12}$$

Charged pions can interact further or decay into muons and muon neutrinos via $\pi^\pm \longrightarrow \mu^\pm + \nu_\mu$, depending on their energy. As energy decreases, the probability of decay increases. The interaction length, the average distance a hadron travels before its next collision, is a crucial parameter. It's roughly constant for pion energies relevant to EAS, between 10 and 1000 GeV.

The critical energy (E_{crit}) is the energy at which a charged pion's decay length (l_{decay}) becomes smaller than the distance to the next interaction point. Below this energy, pions decay rather than interact further. The decay length is defined as:

$$l_{decay} = \gamma c\tau; \tag{13}$$

Where c is the speed of light, γ is the Lorentz factor, and τ is the lifetime of the particle.

Muons are a defining feature of hadronic showers. Unlike charged pions, muons have a much longer lifespan, allowing them to travel significant distances through the atmosphere before decaying. This is why muons are observed even deep within an EAS. To obtain the number of muons in the shower one assumes that all charged pions decay to muons when they fall below their critical energy:

$$N_\mu = (N_{ch})^{n_c} \tag{14}$$

From which we can get the following:

$$N_\mu = \left(\frac{E_0}{E_\pi^c}\right)^\beta \tag{15}$$

Interestingly, the number of muons does not increase linearly with energy like the number of electrons. Additionally, the value of β is influenced by both the average number of pions produced in each interaction (pion multiplicity) and the inefficiency of hadronic interactions (inelasticity).

The Matthews model offers a simplified yet insightful look at hadronic showers. It assumed:

- A constant number of charged pions are produced at each interaction.
- The energy is shared equally among charged and neutral pions.
- Neutral pions decay instantly into photons.
- Charged pions either interact or decay at the critical energy.

This model successfully predicts the number of muons in a shower, which does not increase linearly with energy like the number of electrons. It also helps estimate the depth in the atmosphere where the shower reaches its peak size (X_{max}).

The hadronic shower is more intricate than this simplified model suggests. Factors like the production of heavier mesons and variations in interaction lengths can influence the actual shower development. However, the core principles outlined here provide a solid foundation for understanding this crucial component of Extensive Air Showers.

3.3 Nuclear primaries

EAS are triggered by cosmic rays colliding with air molecules. But what if the cosmic ray itself is not a single particle, but a nucleus made of many protons and neutrons? This is where the concept of nuclear primaries comes in.

Imagine a nucleus with an atomic mass number A. We use a clever trick called superposition to understand how it interacts in an EAS. They treat the nucleus as if it were A individual particles, each carrying $1/A$th of the total energy. So, a shower initiated by a nucleus is essentially a superposition of A proton showers happening simultaneously.

This approach reveals some key differences between showers from nuclear primaries and those from single protons.

Because a nucleus breaks into multiple particles, each with a lower energy, there are more chances for charged pions to decay into muons, leading to a higher overall number of muons in the shower.

$$N_{\mu}^{A}(E_0) = \sum_{i}^{A} N_{\mu}^{p}(E_0/a) = N_{\mu}^{p}(E_0)A^{1-\beta} \tag{16}$$

Nuclear primaries reach their peak particle number (X_{max}) shallower in the atmosphere compared to protons. This is because the lower energy of each nucleon in the nucleus means they reach the threshold energy for particle production sooner, causing the shower to develop higher up.

$$X_{max}^{A}(E_0) = X_{max}^{p}(E_0/A) = X_{max}^{p}(E_0) - \lambda_r A \tag{17}$$

This equation shows that as the number of nucleons (A) increases, the second term gets larger, pushing X_{max} shallower.

The superposition model is a simplification, and real-world showers exhibit fluctuations. However, on average, showers from heavier nuclei like iron will have:

Lower X_{max} compared to showers from protons with the same energy. A higher number of muons. Less shower-to-shower variation in properties like X_{max} compared to proton showers, as the sum of many sub-showers averages out fluctuations. By studying these characteristics, scientists can use EAS observations to infer the nature of the primary cosmic ray particle, including whether it's a single proton or a heavier nucleus.

3.4 Radio emission mechanisms

Radio emission is a valuable method for cosmic ray detection because it allows for a larger effective detection volume compared to particle detectors, for a similar cost. The key players in radio emission are electrons and positrons, the abundant charged particles in air showers. These charged particles interact with the surrounding environment in interesting ways, leading to the emission of radio waves. Thankfully, we can use classical field theory, a well-established framework, to calculate this emission.

When do these charged particles (or the medium around them) start emitting radio waves? Several mechanisms have been identified several emission mechanisms, some experimentally confirmed to be significant contributors to the overall radio signal.

Imagine a charged particle moving through a magnetic field, like electrons in an air shower experiencing Earth's magnetic field, schematic left 11. This interaction, described by the Lorentz force, deflects the particle's path. This interaction is the key player behind the geomagnetic effect, a dominant mechanism for radio emission from air showers.

Earlier theories likened the geomagnetic effect to synchrotron radiation, where charged particles moving in a circular path emit radio waves. However, this analogy falls short. In air showers, electrons and positrons do not travel in smooth circles. Instead, they constantly collide with air molecules, similar to electrons moving through a conductor.

This interplay between the magnetic field's acceleration and collisions with air molecules leads to a fascinating phenomenon. Electrons and positrons develop a net drift in opposite directions, perpendicular to the shower axis, due to the Lorentz force (F_L).

$$\vec{F_L} = q\vec{v} \times \vec{B} \tag{18}$$

Where q denotes the particle charge, v is its velocity vector and B is the magnetic field vector. These drifts create a current, aptly named a transverse current.

As the air shower unfolds, the number of secondary particles initially increases, reaches a peak, and then declines. This dynamic variation in the transverse current leads to electromagnetic radiation. The relativistic speeds of these particles (near the speed of light) lead to a unique radio signature. The emission is compressed into short, powerful pulses that are beamed forward, along the shower axis. Additionally, these pulses span a broad range of frequencies. The observer's location plays a role too. As you move away from the shower axis, the radio pulses broaden, and the detectable frequencies become lower, for geometrical reasons.

While the macroscopic picture involves the collective motion of charged particles, a microscopic view reveals that individual particle accelerations contribute to the radio emission. This acceleration, induced by the changing magnetic field, is what ultimately produces the radio waves.

The radio waves emitted due to the geomagnetic effect exhibit linear polarization. The electric field vector, responsible for this polarization, aligns with the Lorentz force. This alignment translates to a polarization direction along the product of the particle's velocity (approximated by the shower axis) and the magnetic field.

It's important to note that while any charged particle can undergo this process, electrons and positrons dominate the radio signal. Their high charge-to-mass ratio makes them the most efficient radio wave emitters in this scenario. Muons, with their significantly larger mass, play a negligible role.

As an air shower develops, it disrupts the peaceful coexistence of positive and negative charges in the air, schematic right 11. This happens in two ways. Ionization: Energetic particles from the shower knock electrons off air molecules, creating positive ions and free electrons. Annihilation: Some positrons in the shower annihilate electrons, further disrupting the charge balance. The result is a temporary excess of negative charge, which is not uniformly distributed but rather concentrated around the shower axis, because of the relativistic speeds of the particles. This shifting charge imbalance, creates variations in the electric field. These variations in the electric field, not unlike Cherenkov radiation (light emitted by particles exceeding the speed of light in a medium), lead to the emission of radio waves. These radio waves, like those from the geomagnetic effect, exhibit linear polarization. However, the Askaryan effect's polarization holds a key difference. The electric field vector, responsible for polarization, points radially outward from the shower axis. This means the polarization direction depends on the observer's location relative

to the shower axis – unlike the geomagnetic effect where it's aligned with the shower axis itself. While the geomagnetic effect reigns supreme for most air shower radio emission, the Askaryan effect takes center stage for particle showers in dense media like ice or lunar regolith. Due to the shorter length scales of these showers, the Askaryan effect dominates the radio emissions at much higher frequencies (around GHz). Additionally, the geomagnetic contribution fades away entirely for air showers that arrive parallel to Earth's magnetic field. This leads to the Askaryan effect becoming the dominant emission mechanism for showers with a small angle relative to the magnetic field.

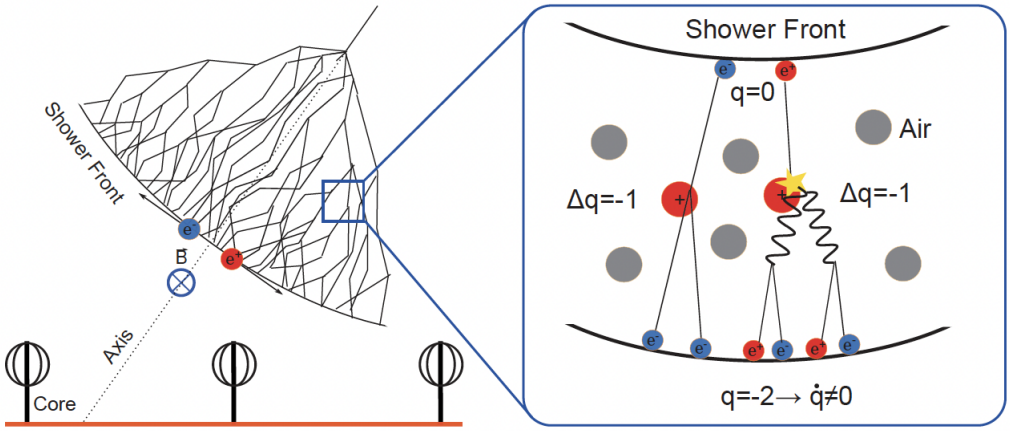


Figure 11: Schematic representation of the main contributions to the radio emission from air-showers [33]. Left: Geomagnetic effect. A current produced by deflection of electrons and positrons emits a radio signal polarized perpendicularly to the magnetic field and shower axis. Right: Askaryan effect. Variation of the net current leads to a changing electrical potential, which produces radially polarized radio emission.

When the electric field vectors produced by these mechanisms superpose, a fascinating interplay occurs. Depending on the observer's location relative to the shower axis and the magnetic field, these contributions can add up either constructively or destructively.

This interplay leads to complex asymmetries in the radio signal, particularly along the east-west direction for vertical air showers (the direction perpendicular to both the shower axis and the magnetic field). Figure 12 depicts this phenomenon through a visualization of the "radio footprint" (two-dimensional radio lateral distribution function). The degree of asymmetry is influenced by the relative strengths of the geomagnetic and charge excess contributions. These strengths, in turn, depend on the angle between the shower axis and the magnetic field (geomagnetic angle) and the local magnetic field strength.

3.5 Typical Development of Extensive Air Showers

The EAS begins with the first interaction. Here, a single cosmic ray particle interacts, creating a burst of secondary particles. These particles then interact further, leading to an exponential growth in their number. This rapid growth does not last forever. Soon, the shower reaches a depth in the atmosphere known as X_{max}, where the number of secondary particles reaches its maximum (N_{max}). After reaching X_{max}, the number of secondary particles starts to decrease

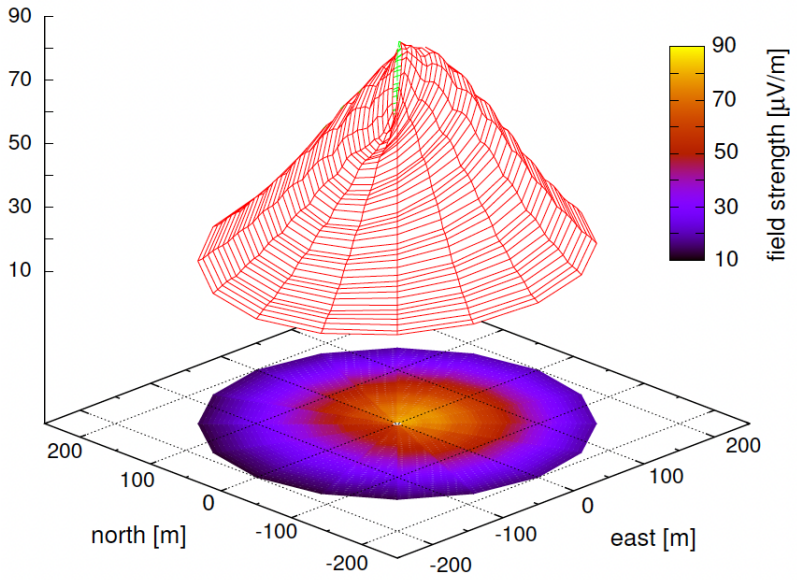


Figure 12: Simulation of the total electric field amplitude in the $40-80$ MHz band for a vertical cosmic ray air shower at the site of the LOPES experiment. The asymmetry arises from the superposition of the geomagnetic and charge-excess emission contributions [34].

with each interaction length. However, this decrease is slower compared to the initial explosive growth. The Gaisser-Hillas function [35] is used to mathematically describe this evolution of particle numbers in an air shower:

$$N(X) = N_{Max}\left(\frac{X - X_0}{X_{Max} - X_0}\right)^{\frac{X_{Max}-X_0}{\lambda}} e^{\frac{X_{Max}-X}{\lambda}} \tag{19}$$

This function considers 4 key parameters the maximum number of particles (N_{max}), the atmospheric depth at which N_{max} is reached X_{max}, λ and X_0 are fitted parameters that are correlated with the shower starting depth and attenuation after maximum which can vary depending on the energy and type of the cosmic ray particle.

The growth and decline of an air shower is not always perfectly smooth. Several factors can cause fluctuations in the shower profile. The first Interaction, the depth at which the first cosmic ray interaction occurs is a major source of variation. A deeper first interaction leads to a slightly different shower profile compared to a shallower one. Cosmic ray species, the type of cosmic ray particle initiating the shower also plays a role. Showers from heavier cosmic rays, like iron, fluctuate less because their energy is spread amongst their constituent nucleons, which can be seen in figure 13.

3.6 Universality of the Shape of Longitudinal Profiles

We saw earlier that the longitudinal profile of an air shower can vary slightly from event to event. Scientists have developed a clever technique to overcome these variations and reveal a

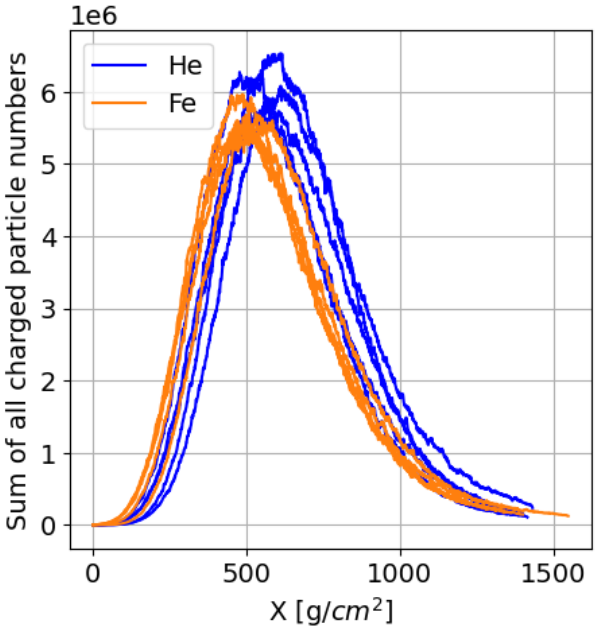


Figure 13: Longitudinal shower profile for 5 iron and helium simulations between $10^{16} - 1.1 \times 10^{16}$ eV.

more universal picture of shower development. This method focuses on the depth (X_{max}) where the shower reaches its peak number of particles. We virtually shift all showers so their X_{max} values align at a fixed depth. Remarkably, this process also eliminates the variations caused by the first interaction depth. This transformation utilizes a modified version of the Gaisser-Hillas function with specific substitutions [36]:

$$\begin{aligned} X' &= X - X_{max}, \\ R &= \sqrt{\lambda / |X_0'|}, \\ L &= \sqrt{|X_0'| \lambda}. \end{aligned} \tag{20}$$

The resulting equation 21 essentially creates a new profile where all showers, regardless of the initial cosmic ray or the first interaction depth, appear similar:

$$N(X) = N_{max} \left(1 + R\frac{X'}{L}\right)^{R^{-2}} e^{-\frac{X'}{R\,L}}. \tag{21}$$

In this distribution, L is related to the variance, and R to the skewness. Imagine revisiting Figure 13, which depicted separate profiles for proton and iron showers. When we apply this transformation and re-plot the same data, the showers become virtually indistinguishable. Figure 14 showcases this beautifully, with all the proton and iron shower profiles collapsing on top of each other.

This transformation highlights the underlying universality of the typical air shower profile. While the overall shape becomes remarkably similar, the transformed profiles still exhibit some minor dependencies. A slight influence of the original cosmic ray particle type remains in the

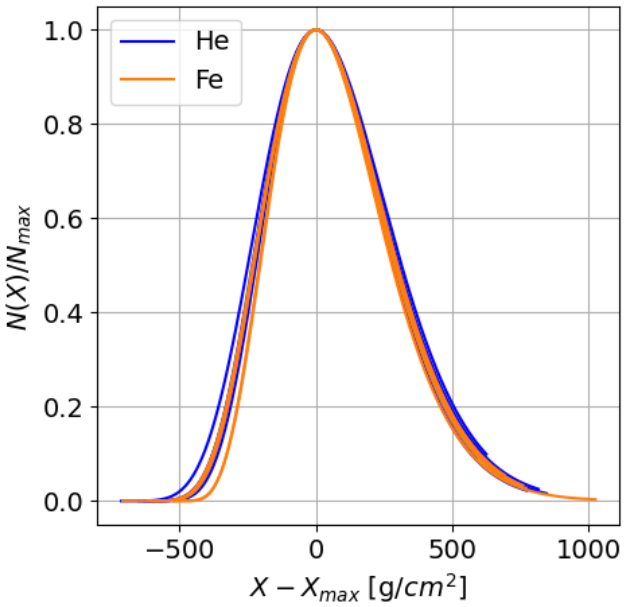


Figure 14: The same 5 iron and helium simulations from 13 plotted under the reduced Gaisser-Hillas transformation.

transformed profiles, reflected by parameters R and L in the reduced Gaisser-Hillas function. The width (L) of the transformed profile shows a weak increase with higher energies as can be seen in table 1. The specific model used to simulate hadronic interactions can slightly affect the R and L values.

Table 1: R and L values for each systematic uncertainties of the measured energy bins, along with the statistical and systematic uncertainties [37].

Energy [eV]	$< log_{10}[E/eV] >$	N	R			$L[g/cm^2]$		
			$< value >$	stat.	syst.	$< value >$	stat.	syst.
$10^{17.8} - 10^{18.0}$	17.90	7829	0.260	0.006	$^{+0.039}_{-0.040}$	226.2	0.4	$^{+5.7}_{-4.9}$
$10^{18.0} - 10^{18.2}$	18.09	5648	0.244	0.007	$^{+0.037}_{-0.039}$	227.6	0.4	$^{+5.6}_{-4.5}$
$10^{18.2} - 10^{18.5}$	18.33	4780	0.252	0.007	$^{+0.035}_{-0.037}$	229.1	0.5	$^{+5.6}_{-4.3}$
$10^{18.5} - 10^{18.8}$	18.63	1907	0.267	0.009	$^{+0.034}_{-0.035}$	231.4	0.7	$^{+6.2}_{-4.1}$
$10^{18.8} - 10^{19.2}$	18.97	1026	0.264	0.010	$^{+0.033}_{-0.034}$	233.3	0.8	$^{+7.0}_{-4.0}$
$< 10^{19.2}$	19.38	342	0.264	0.012	$^{+0.023}_{-0.035}$	238.3	0.9	$^{+7.3}_{-4.0}$

To truly visualize the power of this transformation, we simulated a large number of proton and iron showers across various energies and zenith angles (angle of arrival). When they averaged these transformed profiles (Figure 15), the resulting curves for both types of cosmic rays (even with a large difference in mass) almost completely overlapped.

This demonstrates that despite variations in the initial particle and shower details, the average transformed profile exhibits remarkable universality. Even the slight discrepancies are mainly due to the heavier iron nuclei.

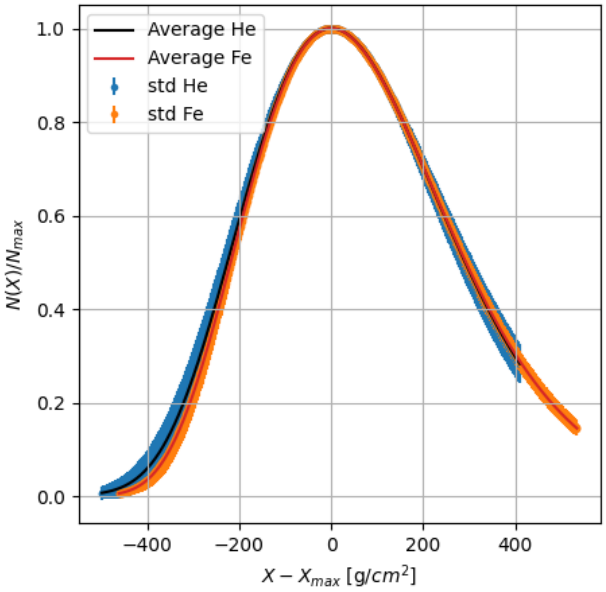


Figure 15: 1000 different iron and helium simulations with energies between $10^{16} \times 10^{18}$ eV plotted under the reduced Gaisser-Hillas transformation.

3.7 Anomalous Extensive Air Showers

Extensive Air Showers (EAS) are typically well-behaved, with a predictable rise and fall in the number of secondary particles as they travel through the atmosphere. But what if, on rare occasions, things get a little strange? Anomalous air showers are exactly showers that deviate from the expected pattern (shower universality), see figure 16.

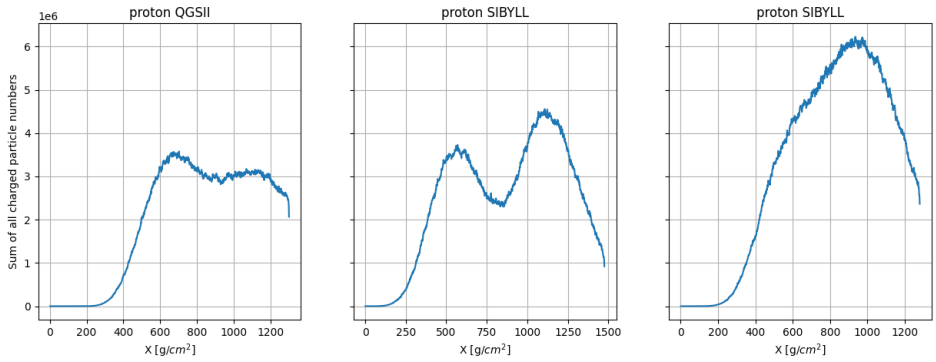


Figure 16: Examples of anomalous shower profiles simulated with Corsika.

Imagine a secondary particle born in an air shower. Normally, it interacts with air molecules fairly quickly, transferring its energy and creating more particles. However, in anomalous showers,

a secondary particle can defy the odds and travel much deeper into the atmosphere than usual before interacting. The chance of a particle penetrating some distance greater than ΔX is given by:

$$P(\Delta X) = e^{-\frac{\Delta X}{\lambda}}. \tag{22}$$

With λ is the hadronic interaction length in air. The chance of this happening depends on a few factors. Firstly the interaction length this is the average distance a particle travels before interacting:

$$\lambda = \frac{M_{molar}}{\sigma N_A}. \tag{23}$$

With M_{molar} as the molar mass of the target, N_A is Avogadro's number, and σ is the cross-sectional interaction length of the beam particle. The average molar mass of the elements composing dry air remains relatively constant (~ 28.97 g/mol [38]) up to an altitude of 90 km. This consistency is largely attributed to the dominance of nitrogen in the air's composition. To illustrate the key concepts, we'll restrict our attention to the inelastic component of the proton-proton cross-section (σ_{pp}^{inel}). This component is modeled by the following equation:

$$\sigma_{pp}^{inel} = 65\left(1 + 0.237\ln\left(\frac{E}{200GeV}\right) + 0.01\ln^2\left(\frac{E}{200GeV}\right)\right). \tag{24}$$

The depth to which a cosmic ray particle penetrates the atmosphere depends on two key factors: its energy and its type. Higher energy particles and those with lower mass are generally more likely to penetrate deeper before interacting with air molecules.

The Heitler model provides a simplified view of how a heavier cosmic ray nucleus interacts with air. It essentially treats the nucleus as a collection of individual protons. For these heavier nuclei (with a larger number of nucleons), the total energy is distributed amongst all its components. This distribution leads to a lower probability of any single nucleon within the nucleus having enough energy to penetrate deeply. As a general rule, the interaction length increases with the particle's energy. However, the chance of a deeply penetrating nucleon (from either a light or heavy nucleus) actually decreases with increasing energy. This might seem counter intuitive, but it's a consequence of the specific details of particle interactions. The Heitler model helps visualize why heavier nuclei have a lower probability of producing deeply penetrating particles. Because the model treats a nucleus as multiple individual proton showers, it becomes clear how the number of potential deep-penetrating particles diminishes as the total number of nucleons in the nucleus increases. Figure 17 illustrates the probability of traveling a distance greater than ΔX before interacting for cosmic ray protons at different energies. The figure supports the notion that lower energy cosmic ray particles, especially those with a lower atomic mass, are more likely to penetrate deeper than their higher-energy counterparts. This deeper penetration can lead to the formation of anomalous shower events. While the simplified model suggests a $\sim 5\%$ chance for a 10^{19} eV proton to penetrate beyond 300 g/cm^2, more complex simulations incorporating all relevant physical interactions suggest a lower probability at the same energy. These simulations provide a more accurate picture of the actual penetration depths for cosmic rays.

[39] investigated the possibility of leading particles causing the double-bump profile. They analyzed the correlation between the fitted difference in peak depths (ΔX_{max}) and the difference in slant depths (ΔX) between the two most significant interaction points (Figure 18). Here, X_{first} represents the interaction depth of the primary particle, and X_n represents the interaction depth of the nth leading particle. Only interactions with a minimum inelasticity of $k > 0.15$ were considered. The results show that for most showers, the distance between the shower maxima (ΔX_{max}) is a good estimate of the actual difference in slant depths (ΔX). However, a cluster

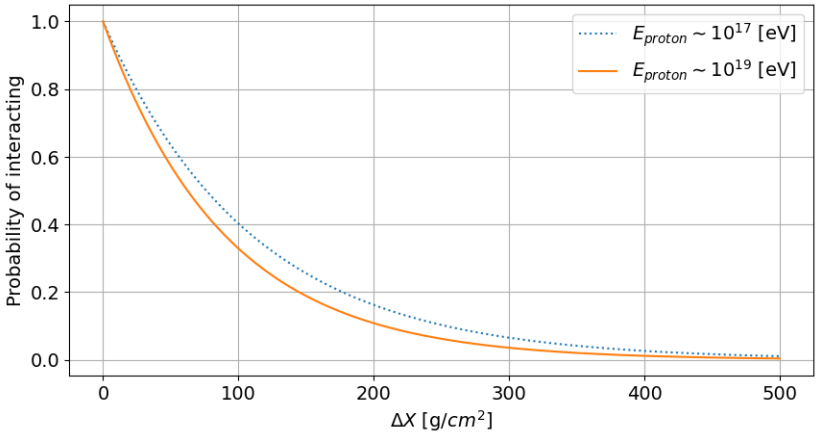


Figure 17: The probability of proton cosmic rays interacting with air for two energy levels.

of events deviates from the diagonal trend, exhibiting large ΔX_{max} values despite small ΔX values. These events likely originate from deeply penetrating sub-showers not initiated by leading particles. Consequently, ΔX_n and ΔX are not appropriate indicators for these specific events.

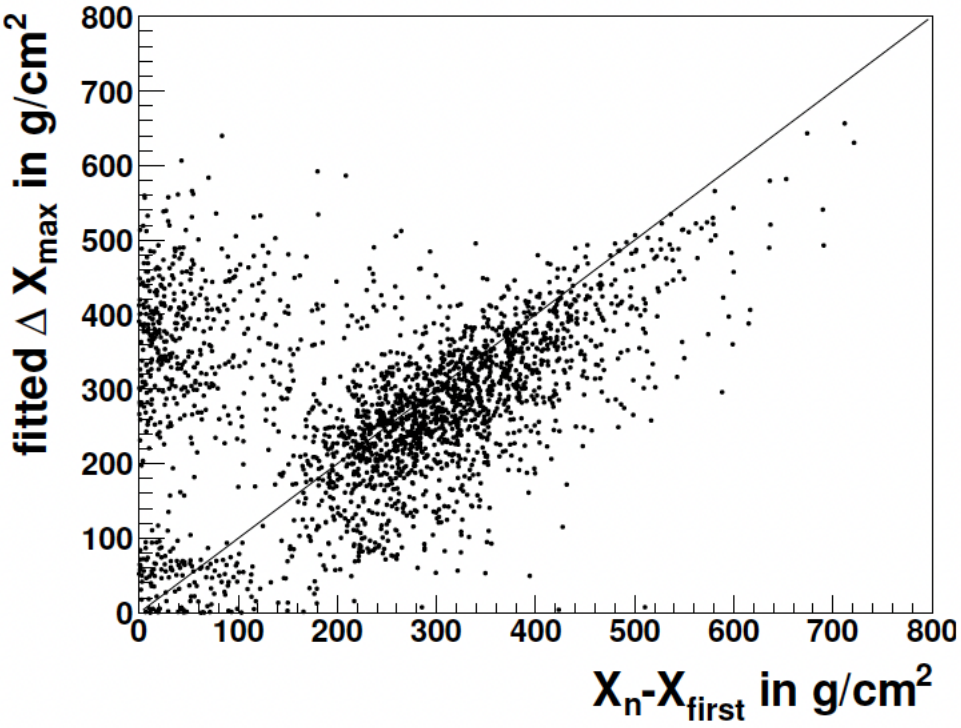


Figure 18: Correlation between ΔX and ΔX_{max} for protons simulated with SIBYLL [39]

4 Cosmic rays experiments

Cosmic rays bombard Earth with incredible energies. But how do we actually study these elusive particles? Scientists have devised a range of ingenious experiments to capture the secrets these cosmic rays hold. One daring approach involves sending instruments aloft, either on balloons or even satellites. These high-altitude detectors aim to directly measure the cosmic ray particles themselves before they get absorbed by the atmosphere. This allows scientists to study the properties of the particles, like their energy and type. Another strategy involves setting up giant detectors on the ground, forming vast surface arrays to gather data out of the low flux. These arrays do not measure the cosmic rays themselves, but rather the secondary particles created when cosmic rays slam into the atmosphere, triggering EAS. By studying the characteristics of these showers, scientists can piece together information about the original cosmic rays that caused them. Cosmic rays also unleash radio waves as they hurtle through the atmosphere during an EAS. By deploying radio detectors, scientists can pick up these faint signals. This technique provides valuable insights into the properties of the cosmic ray and the development of the air shower itself. Imagine a giant, three-dimensional X-ray machine for cosmic rays! This is essentially what the Square Kilometre Array (SKA), a future radio telescope under construction, aims to achieve. With its dense network of antennas, SKA will use a technique called radio tomography to reconstruct air showers in exquisite detail, offering unprecedented views of cosmic rays. These diverse detection methods, from high-altitude chases to radio eavesdropping, are all crucial tools in the cosmic ray detective kit. Each approach sheds light on different aspects of these enigmatic particles, helping us unravel the mysteries of the cosmos and the exotic phenomena that give rise to cosmic rays. Key features of the primary cosmic ray spectrum, along with their suspected origins and prevalent detection techniques, are summarized visually in Figure 19.

4.1 Air-shower detection techniques

The study of cosmic rays, high-energy particles originating from various astrophysical sources, involves sophisticated air-shower detection techniques. These methods aim to unravel the intricacies of cascades produced when cosmic rays interact with Earth's atmosphere. Pioneering surface particle detectors, which distinguish muonic and electromagnetic components reaching the surface, set the foundation. Imaging techniques, employing fluorescence and air-Cherenkov telescopes, excel in precision by focusing on narrow-angle imaging of the sky. Non-imaging techniques, including air-Cherenkov and radio emission measurements, offer wide-angle coverage but necessitate indirect air-shower reconstruction. Satellites, such as TUS and JEM-EUSO, explore monitoring from above. Notably, radio detection, tested decades ago and revitalized recently, emerges as a powerful non-imaging technique, providing insights into air-showers with primary energies exceeding 10^{16} eV and higher. Each technique contributes uniquely to cosmic-ray detection, offering varied perspectives on these high-energy phenomena.

4.2 Particle Detectors

Surface particle detectors hold the distinction of being the longest-established technique for studying EAS. These arrays have served as the cornerstone of air-shower detection since their inception by Pierre Auger in 1938 [41], marking the same year as the debut of the iconic Superman comic book.

These detectors focus on the two key components of an air shower that reach the Earth's surface: muons and the electromagnetic component. There are two main types of surface particle detectors.

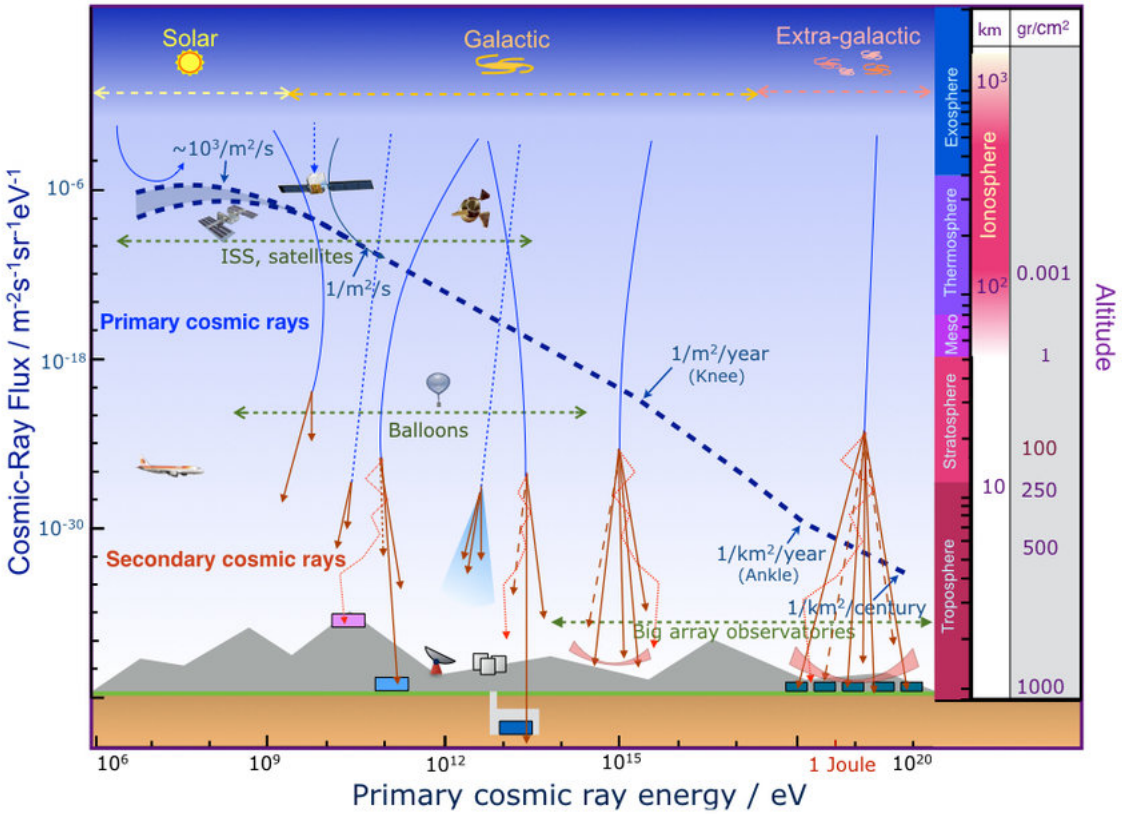


Figure 19: Primary cosmic ray spectrum as a function of the energy. The blue dashed line represents the cosmic rays flux using the scale units shown on the left vertical axis. Below energies $\sim 10^{14}$ eV primary cosmic rays are measured directly with detectors placed in satellites and balloons. At higher energies indirect measurements done by ground-based arrays of detectors are used. The right vertical axis shows the atmosphere layers, their heights and the corresponding air pressure [40].

Water-Cherenkov Detectors: These detectors utilize a large volume of water. When an energetic particle races through the water at a speed higher than the speed of light in water, it emits a faint glow called Cherenkov light. By capturing these light signals in multiple water tanks, scientists can infer the properties of EAS such as energy and arrival direction. A notable example being the water-Cherenkov detectors of the Pierre Auger Observatory.

Scintillators: These are more compact detectors that use materials that scintillate (emit a flash of light) when struck by energetic particles. The key advantages of surface particle detectors are their simplicity, robust construction, and ability to operate continuously, day and night.

Despite their strengths, surface detectors have limitations. They only capture the particles reaching the ground, which represents just a slice of the entire air shower. This restricts their ability to pinpoint the location of the shower maximum. Furthermore, interpreting the data heavily relies on hadronic models that simulate how particles interact in the atmosphere.

One way to improve the sensitivity to the type of cosmic ray that initiated the shower is to separate electrons and muons in the detector, as demonstrated by the KASCADE-Grande

experiment [42]. This technique, known as electron/muon separation, provides valuable insights into the chemical composition of the primary cosmic ray.

4.3 Imaging Techniques

Among the most powerful tools for studying air showers are imaging telescopes, which come in two main flavors: fluorescence and Cherenkov. Unlike surface detectors that only measure particles reaching the ground, these telescopes capture the light emitted by the air shower itself, providing a much more detailed picture of the shower's development.

Fluorescence telescopes act like giant cosmic-ray cameras. They use mirrors to focus the faint light emitted by nitrogen molecules in the atmosphere as they relax after being excited by energetic particles in the air shower. This light provides valuable information about the shower's development, essentially allowing scientists to see the shower unfold.

The other imaging technique, the air Cherenkov telescope, focuses on a different kind of light. When charged particles in the shower, specifically electrons and positrons, move faster than light through the air, they emit a faint glow called Cherenkov light. Air Cherenkov telescopes use mirrors to capture this light, revealing the path of these speedy particles within the shower.

This technique offers several advantages. Air Cherenkov telescopes can detect air showers with lower energies compared to fluorescence telescopes. They can differentiate between air showers initiated by electromagnetic particles (like gamma rays) and those initiated by hadronic particles (like protons and neutrons). This is crucial for studies of gamma-ray astronomy at TeV energies.

Imaging telescopes offer exceptional precision in studying air showers. However, this precision comes at a cost. To achieve such accuracy, the telescopes require meticulous calibration to ensure they are measuring the light signal correctly. These telescopes can only operate during clear, moonless nights, significantly limiting their observing time (duty cycle). Their maintenance is challenging since they are technically much more complicated compared to particle detectors.

While most air-shower detection happens on the ground, there is another approach – space-based observatories. These satellites would aim to directly observe the faint fluorescence light emitted by nitrogen molecules in the atmosphere during ultra-high energy cosmic ray ($> 10^{19}$ eV) air showers. Satellites offer a unique advantage an unobstructed view of the entire sky, free from limitations like weather or night/day cycles that ground-based detectors face. This opens doors to studying the highest-energy cosmic rays, which are much rarer but hold immense potential for unlocking mysteries about the universe's most extreme environments.

Several exciting satellite missions have been proposed to explore air showers from space. TUS (Tailored Ultraviolet Spectrometer) on Lomonosov satellite. This mission aimed to utilize a satellite specifically designed for cosmic ray detection. KLYPVE (Korean Large Yield Photon Experiment) on ISS, this experiment planned to leverage the International Space Station (ISS) as a platform. JEM-EUSO (Extreme Universe Space Observatory on Japanese Experiment Module), another ISS-based experiment designed for high-energy cosmic ray detection.

4.4 Non-Imaging Techniques

While imaging telescopes offer a detailed view of air showers, there's another powerful approach – non-imaging techniques. These techniques prioritize capturing the big picture, sacrificing some detail for wider sky coverage. Unlike imaging methods that focus on a specific part of the sky, non-imaging techniques employ wide-angle detectors that can monitor a much larger area.

Non-imaging detectors achieve this by capturing the faint electromagnetic emissions produced during an air shower. Showers initiated by high-energy cosmic rays interacting high in the atmo-

sphere are then identified by searching for time coincidences of signals in neighboring detector stations. The arrival direction of the cosmic ray can be determined from the time delay of the shower front reaching the different detectors, the schematic of the detection principle can be seen in figure . Additionally, the size of the signals in the detectors allows for an estimation of the primary cosmic ray's energy.

Essentially, air shower detectors operate by effectively using the atmosphere as a giant calorimeter to estimate the total energy of the incoming particle. Since the primary particle itself is not directly detected, the relationship between the ground measurements and the primary particle's properties (including the type of atomic nucleus) has some inherent uncertainties. This highlights the importance of cross-checking these indirect measurements with direct methods whenever possible.

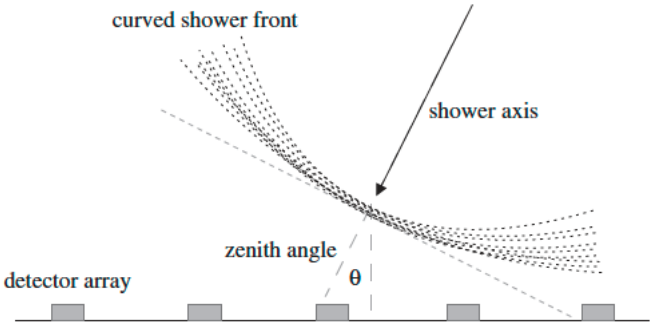


Figure 20: Detection principle and geometry reconstruction of air showers with surface detector arrays.

Radio waves are also emitted during air shower development, and non-imaging detectors can pick up these signals. The key difference from imaging techniques is that non-imaging detectors don't capture the specific details of the shower's development. They only measure the total electromagnetic footprint of the shower on the ground. Since they lack detailed spatial information, scientists need to use sophisticated reconstruction techniques to infer the properties of the original cosmic ray that caused the air shower. Despite this indirect approach, non-imaging techniques can achieve impressive precision, rivaling that of imaging methods in some cases.

4.5 Radio Detection

While the history of air-shower detection stretches back decades, the radio technique offers a unique and relatively new approach. Pioneered over 50 years ago, radio detection has seen a resurgence in interest over the past decade. Unlike imaging techniques that rely on light, radio detection focuses on the faint radio waves emitted during air-shower development. This non-imaging technique is particularly sensitive to showers with primary cosmic ray energies exceeding 10^{17} eV. Radio waves offer several advantages for studying air showers. Radio signals are particularly sensitive to the depth in the atmosphere where the shower reaches its peak particle count (shower maximum). Radio detection offers a calorimetric measurement of the energy deposited by the electromagnetic component of the shower. In simpler terms, it provides a way to estimate the total energy of the shower. Unlike ground-based detectors that can be affected by weather or night/day cycles, radio detectors can operate almost 24/7, offering a high-duty cycle. These advantages make radio detection a valuable complement to other air-shower detection techniques. By combining radio data with information from ground-based detectors

and telescopes, scientists can gain a more comprehensive understanding of these fascinating cosmic phenomena.

Scheduled for full operation at the end of the 2020s, the Square Kilometre Array (SKA) represents a revolutionary leap forward in radio astronomy. This meticulously planned array, characterized by its dense electrical configuration (see Fig. 21), will be the largest and most powerful low-frequency radio observatory ever built. SKA boasts thousands of specialized radio antennas operating across a broad frequency band, capable of observing a vast array of astronomical phenomena.

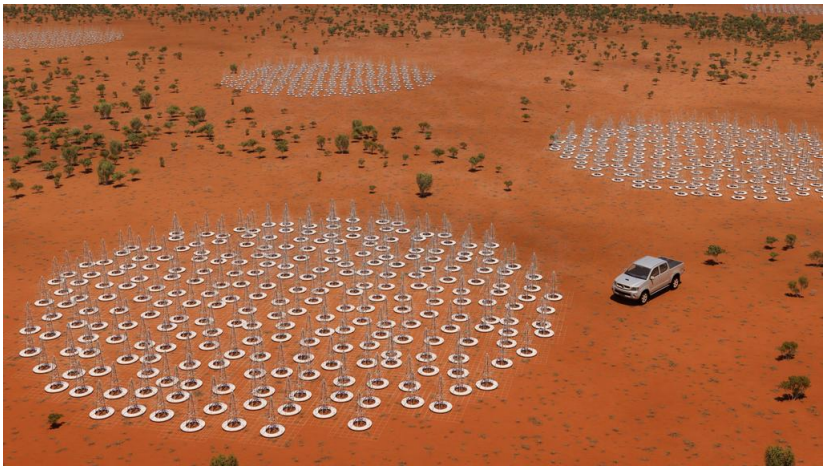


Figure 21: An artists impression of the aerial view of the SKA's low frequency aperture arrays in Western Australia [43].

One of SKA's strengths lies in its ability to revolutionize radio detection of cosmic rays, particularly in the energy range around 10^{17} eV [44]. With its extreme antenna density (over 60.000 antennas within a square kilometer, hence the name), SKA offers superior capabilities in reconstructing crucial parameters of cosmic rays, such as arrival direction, energy, and the depth of the shower maximum (X_{max}) with a resolution of less than 10 g/cm^2 [44]. This detailed information, visualized through radio emission footprint simulations (see Fig. 22), is invaluable for studying cosmic ray interactions and air showers.

SKA's versatility extends beyond its primary astronomical objectives. The dense array and high-resolution measurements enable researchers to perform "radio tomography" of air showers. This advanced technique involves calculations to effectively deconvolve the detailed structure of air showers, providing valuable insights into cosmic ray interactions with the atmosphere.

The size of SKA's detector array influences the detectable cosmic ray energy range. While the upper limit is around 10^{18} eV, SKA's sensitivity allows for detection down to roughly 10^{16} eV using beamforming techniques employed with the dense antenna layout. This lower energy range, between the "knee" and "ankle" in the cosmic ray spectrum, is particularly intriguing as it likely holds the key to understanding the transition from Galactic to extragalactic cosmic ray sources. Additionally, it may harbor a secondary Galactic component arising from cosmic ray reacceleration at the Galactic termination shock or within the powerful shocks associated with Wolf-Rayet supernovae [45].

Determining the composition of cosmic rays is crucial for understanding the astrophysics governing the "knee" to "ankle" region. High-accuracy measurements are essential, as some models predict subtle transitions between elements with similar masses. For instance, scenarios involving

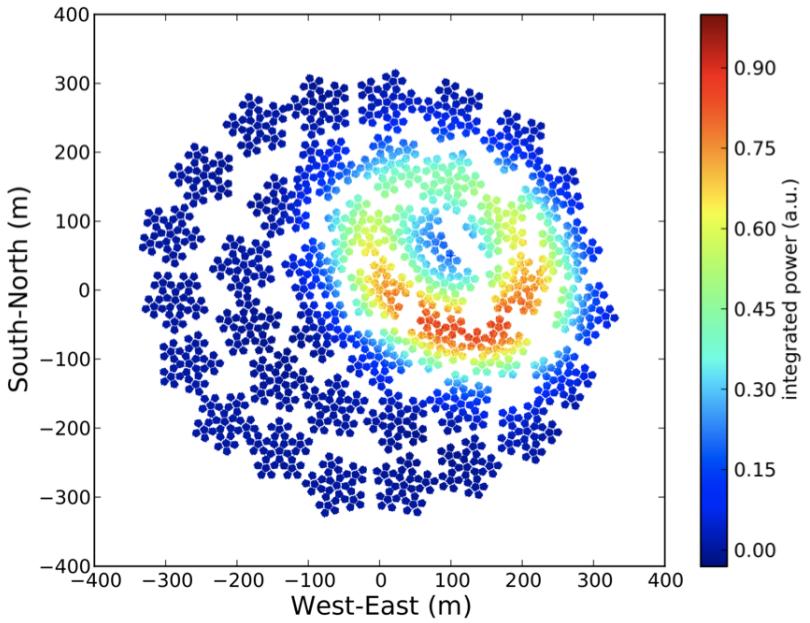


Figure 22: CoREAS-based simulation of how the air-shower radio emission would be sampled SKA (this simulation is based on an old antenna layout). The SKA will measure the radio-emission footprint with an extremely dense and homogeneous array of antennas, yielding superior reconstruction quality on an event-to-event basis [44].

helium-rich Wolf-Rayet stars suggest the transition from Galactic to extragalactic cosmic rays might be marked by a shift from a helium-dominated to a proton-dominated flux. Uncertainties in hadronic interaction models, which describe how cosmic ray particles interact with matter, further complicate this observational challenge. Different models predict varying relationships between the primary cosmic ray mass and the average X_{max} of the air showers they generate.

SKA's exceptional antenna density allows for the reconstruction of air showers with unprecedented detail. Monte Carlo simulations indicate that traditional radio reconstruction methods will achieve a precision of $6 - 8$ g/cm^2 for X_{max} and 3% for primary energy [46]. However, to fully unlock SKA's potential for cosmic ray science, researchers are actively developing novel analysis techniques to extract even more information from air shower data. One promising approach involves reconstructing the shower length (L), which offers a powerful way to determine the proportion of protons within the cosmic ray flux.

By combining its exquisite sensitivity with advanced analysis techniques, SKA promises to usher in a new era of discovery in both radio astronomy and cosmic ray physics.

5 Simulating anomalous air showers

5.1 Corsika explanation

CORSIKA (COsmic Ray SImulations for KAscade) is a comprehensive Monte Carlo program designed to investigate the evolution of extensive air showers (EAS) initiated by photons, protons, nuclei, or any other particle. Originally developed to support simulations for the KASCADE experiment at Karlsruhe, CORSIKA has undergone continuous refinement and evolved into a widely utilized tool across research groups. Its applications span from Cherenkov telescope experiments at $E_0 \approx 10^{12}$ eV to the highest observed energies ($E_0 \approx 10^{20}$ eV).

Since its inception in 1989, CORSIKA has undergone numerous extensions and improvements. To address the challenge of extrapolating hadronic interactions to higher energies not covered by experimental data, five different hadronic interaction models are available in CORSIKA: VENUS, QGSJET, DPMJET, SIBYLL, and the phenomenological HDPM generator. The program also incorporates GHEISHA routines for handling low-energy hadronic interactions.

CORSIKA has been extensively employed by the KASCADE group for calculations involving p, α, O, Fe, and γ primaries in an energy range of 10^{11} eV $\leq E_0 \leq 10^{16}$ eV. Laboratories worldwide use CORSIKA to interpret and understand cosmic ray experiments, evaluating particle numbers, lateral and energy distributions, arrival times, and other features, comparing them with available experimental data.

The CORSIKA program utilizes a right-handed Cartesian coordinate system specifically designed to account for Earth's magnetic field, which can be seen in figure 23.

- Positive x-axis points towards magnetic north.
- Positive y-axis points west.
- Positive z-axis points upwards.

The coordinate system's origin is set at sea level. This specific orientation becomes crucial when considering the Earth's magnetic field, which significantly impacts charged particle trajectories within air showers. CORSIKA Defaults to Karlsruhe, Germany.

The zenith angle (θ) it represents the angle between the particle's direction and a perfectly vertical line pointing straight down. The azimuthal Angle (ϕ) is measured in the xy-plane. It signifies the angle between the positive x-axis (magnetic north) and the projection of the particle's momentum vector onto the xy-plane. Imagine looking down on the xy-plane, with positive x towards magnetic north. The azimuthal angle is measured counter-clockwise from this positive x-axis to the particle's momentum projection.

At the core of the Monte Carlo method lies the concept of random numbers, allowing us to model the inherent uncertainties within complex systems. For accurate simulations in CORSIKA, a robust random number generator is essential, especially for today's demanding calculations.

CORSIKA leverages the state-of-the-art random number generator RANMAR. This generator, implemented in the CERN program library, is widely recognized for its quality in computational physics. RANMAR is a pseudorandom number generator, meaning it produces a deterministic sequence of numbers that appear random. These numbers are uniformly distributed, meaning all possible values within a specific range have an equal chance of appearing.

EAS simulations using Monte Carlo techniques often face a significant challenge, the computational time required scales roughly with the primary cosmic ray energy. This becomes particularly problematic for ultra-high energy particles with energies exceeding 10^{16} eV.

CORSIKA offers a solution to this predicament through a technique known as *thinning* [47]. This method focuses on secondary particles with energies below a user-defined threshold, expressed as a fraction ($\epsilon_{th} = E/E_0$) of the primary particle's energy.

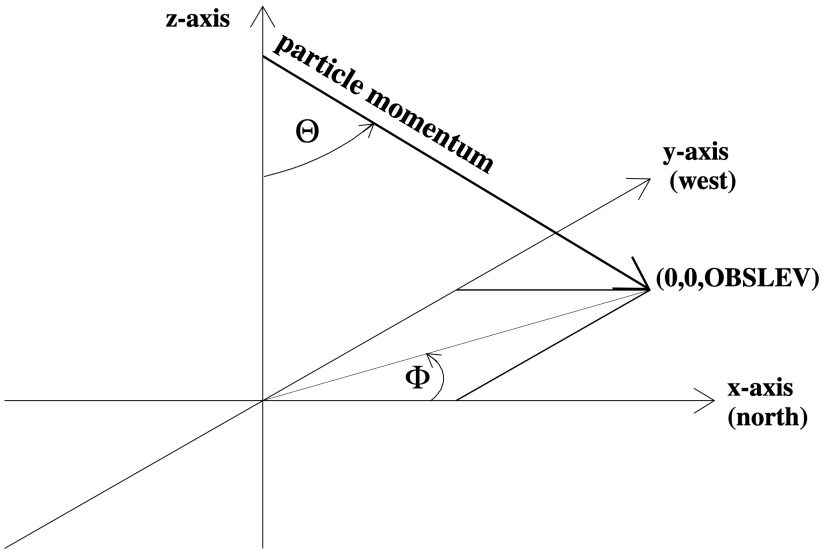


Figure 23: Coordinate system in CORSIKA.

All secondary particles generated during the simulation are evaluated against the thinning threshold. If the combined energy of all secondary particles falls below the threshold, only one particle is chosen to be tracked further:

$$\epsilon_{th} E_0 > \sum_j E_j. \tag{25}$$

This selection is done probabilistically, with the probability being proportional to the particle's individual energy,(E_i):

$$p_i = E_i / \sum_j E_j. \tag{26}$$

All other low-energy particles are discarded from the simulation. To ensure energy conservation, the surviving particle's initial weight is multiplied by a factor of $w_i = 1/p_i$. This effectively compensates for discarding the lower-energy particles. If only a portion of the secondary particles' combined energy falls below the threshold, those particles have a chance to survive based on the same probability calculation:

$$p_i = E_i / \epsilon_{th} E_0. \tag{27}$$

If they do survive, they inherit a weight factor of $w_i = 1/p_i$. When the total secondary particle energy surpasses the threshold, thinning allows for multiple particles to be retained. Again, the selection is probabilistic, and surviving particles receive the corresponding weight factor. The key benefit of thinning lies in significantly reducing the number of low-energy particles that need to be tracked during the later stages of the simulation. This is particularly impactful since the number of secondary particles typically grows exponentially with shower development. Thinning is implemented in both the hadronic interaction routines and the EGS4 code used by CORSIKA [47]. Table 2 demonstrates how the thinning threshold significantly reduces computation time for simulating 10^{15} eV proton showers with vertical incidence, employing the QGSJET hadronic interaction model and the default EGS4 settings. Another method to speed up the simulations is

Table 2: Computing times for various thinning levels [1].

ϵ_{th}	none	10^{-6}	10^{-5}	10^{-4}	10^{-3}
Time (min)	98	51	7.2	1.2	0.16

called CONEX. Unlike traditional Monte Carlo methods that rely on weighted particle tracking, CONEX utilizes a different approach for simulations below a specific energy threshold. It leverages parametrizations. This allows CONEX to solve the cascade equations describing shower development analytically, bypassing the need for individual particle tracking. This analytical approach offers a significant speed advantage, particularly for simulating showers initiated by primary particles with energies exceeding 10^{17} eV.

5.2 Hadronic interaction models

CORSIKA simulates hadronic interactions, which are crucial for understanding the development of extensive air showers (EAS), using a combination of models depending on the particle's energy. We will give a short overview of the models employed.

QGSJET [48], SIBYLL [49], or EPOS [50]. These well-established models handle interactions at high energies ($E_{cm} > 12$ GeV or $E_{lab} > 80$ GeV). They are described in detail within the referenced literature. Each model has its own strengths and limitations, and researchers often compare results from multiple models to gain a more comprehensive understanding.

The choice of hadronic interaction model can significantly impact the results of EAS simulations. For accurate simulations, it is important to select models appropriate for the energy range of interest. The transition energy between high and low-energy models can be adjusted within CORSIKA, but the default value ($E_{cm} = 12$ GeV or $E_{lab} = 80$ GeV) is a common starting point. Multiple high-energy models (e.g., SIBYLL, QGSJETII, and EPOS) are available, each with its own theoretical underpinnings and predictions for particle interactions, see 3 for an overview.

Table 3: Basic features of the interaction models used [1].

	QGSJET	SIBYLL	EPOS
Gribov-Regge	+		+
Minijets	+	+	
Sec. Interactions			+
N-N Interactions	+		+
Superposition		+	

Quark-Gluon String Model with Jets (QGSJET-II) is an advanced iteration of the QGSJET model that incorporates non-linear interactions, providing a more comprehensive picture of high-energy hadron-hadron collisions. QGSJET-II leverages the concept of pomerons, which represent parton (quark and gluon) cascades that bridge the gap between the projectile and target particles during an interaction. These pomerons can be classified as soft (low momentum transfer) or semi-hard (higher momentum transfer), allowing the model to handle a wider range of interaction scenarios. For hard diffractive processes, QGSJET-II introduces the concept of minijet production. These minijets are collimated sprays of energetic partons that contribute to the overall particle production in the interaction. When simulating collisions between atomic nuclei, QGSJET-II employs Glauber theory. This theory provides a probabilistic framework for calculating the impact parameter (distance of closest approach) between the colliding nuclei, which influences the nature of the interaction.

EPOS stands for "Extensive Particle Origin Simulation." It adopts a quantum mechanical framework that ensures energy conservation during particle interactions. Additionally, it employs a multiple scattering approach based on parton ladders. EPOS envisions hadron-hadron interactions as exchanges of multiple pomerons, visualized as ladders on Feynman diagrams. These pomerons are further understood as two fragmented color strings, which ultimately fragment into the final-state particles. The colliding particles (projectile and target) are treated as off-shell remnants, meaning they can be left in an excited state after the interaction. Notably, the theoretical underpinnings of EPOS are particularly adept at explaining the data observed at the Relativistic Heavy Ion Collider (RHIC), a landmark experiment that studied collisions of heavy nuclei.

SIBYLL is specifically designed for simulating extensive air showers, making it a computationally efficient choice. It treats hadron-hadron, nucleus-hadron, and nucleus-nucleus collisions with distinct approaches. SIBYLL relies on a fragmented string model for all collision types. In hadron-hadron interactions, the nucleons (protons and neutrons) are envisioned as splitting into quark-diquark systems, which then fragment into the final-state particles. When simulating collisions between atomic nuclei, SIBYLL employs a superposition model. This simplifies the calculation by assuming that only a portion of the nucleons from each nucleus actively participate in the interaction, with the remaining nucleons acting as spectators. Glauber theory is again used to determine the number of participating and spectator nucleons. Spectator nucleons can still decay into lighter nuclei after the collision.

One of the most striking differences between these models lies in their predictions for the proton-air cross-section, see figure 24. SIBYLL generally predicts the highest cross-section across all energy ranges, while QGSJET-II predicts the lowest. EPOS falls between the two at lower energies, but its predictions diverge significantly from SIBYLL at ultra-high energies ($> 10^{18}$ eV).

The choice of a specific model depends on the researcher's simulation goals and the desired level of computation speed. While the right model for high-energy simulations remains an open question, faster models can be advantageous for preliminary studies or tasks requiring rapid computation. However, relying solely on a faster model might compromise the validity of results for in-depth analysis. Utilizing all three models allows researchers to assess the sensitivity of their findings to the choice of the simulation model, providing a more comprehensive understanding of the potential model dependence in their results. To investigate the influence of the hadronic interaction model on the identification of anomalous air showers, this thesis utilizes simulations with QGSJET-II, EPOS, and SIBYLL.

5.3 Creation of EAS databases

First I created my own small data set, consisting of only 1000 proton showers and 1000 helium showers to test the following analysis. While having a small enough dataset to manually check the results afterwards. For the results in this thesis I used a databases of showers available within the LOFAR collaboration.

To construct a comprehensive database of EAS, A two-pronged approach within the Corsika software is used. Steering files defined the transition point between detailed Monte Carlo simulations and more efficient cascade equations (CONEX), the starting depth of the shower within the atmosphere, and the observer's altitude. Additionally, Corsika commands provided control over the shower energy range (10^{16} to 10^{18} eV), energy spectrum, zenith angle (15° to 45°), and the type of primary cosmic ray particle (protons, helium, carbon, silicon, or iron).

To account for systematic uncertainties in modeling the interactions between cosmic rays and air molecules, three different hadronic interaction models were employed: SIBYLL (version:

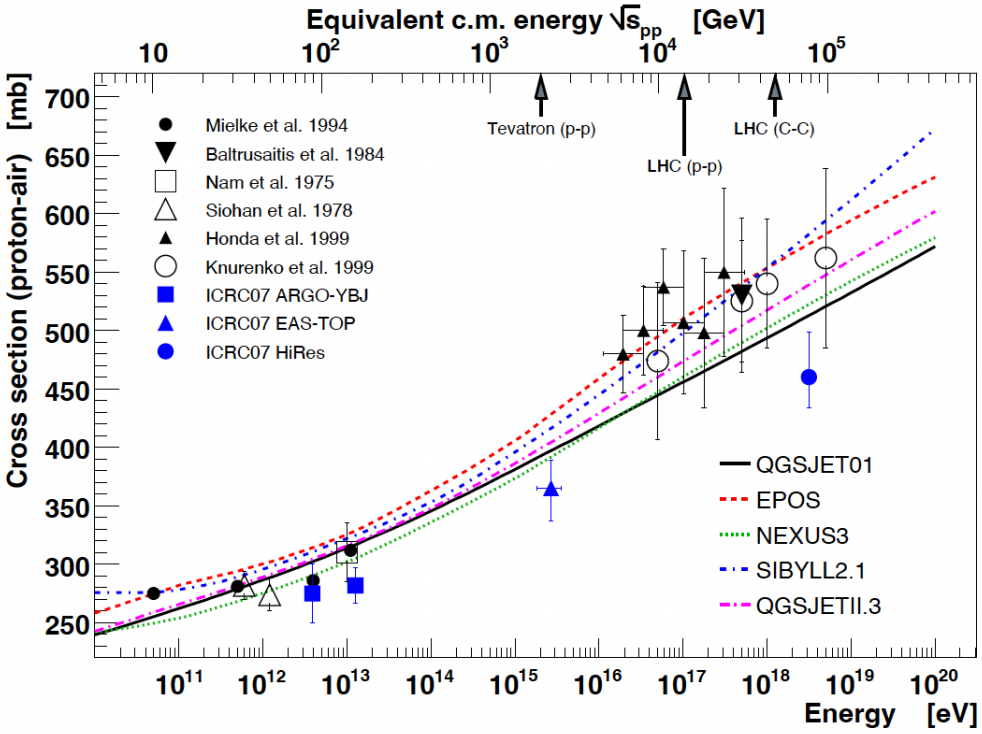


Figure 24: A comparison of the proton-air cross-section between different high energy models [30].

SIBYLL 2.3d.), QGSJETII (version: QGSJETII-04), and Epos (version: EPOS-LHC). For each unique combination of primary cosmic ray particle type and hadronic interaction model, a total of 20.000 simulations were executed. This comprehensive approach, combining controlled simulations, captured data, and multiple hadronic interaction models, has resulted in a robust database of EAS events.

5.4 Fitting procedure

Our analysis employs two fitting strategies to identify double bump profiles in the longitudinal profiles of the simulated air showers. The first approach, based on [39],involves fitting the profile with both a standard Gaisser-Hillas function (Eq. 19) and a reduced Gaisser-Hillas function (Eq. 21). We then repeat the fitting using the sum of two standard Gaisser-Hillas functions and the sum of two reduced Gaisser-Hillas functions.

To ensure a meaningful chi-square (χ^2) value during the fitting process, a careful choice is required for the weight factor (V_i) assigned to each data point. While CORSIKA software provides "exact" shower sizes without statistical fluctuations, a simple Poissonian approach ($V_i = N_i$) would lead to an unwanted energy dependence in the sensitivity of the χ^2 value. To achieve a constant sensitivity across different energies, we adopt the weight factor:

$$V_i = k\frac{N_i}{E}. \tag{28}$$

Where k is a constant chosen such that the uncertainties are reasonable,

$$\frac{\sqrt{\sum V_i}}{\sum N_i} = 0.006. \tag{29}$$

This value will depend on the exact settings of the CORSIKA simulation, it strongly depends on the vertical step width for sampling of the longitudinal development (in g/cm^2) the closer these steps are together the lower the value should be set. This normalization ensures equal χ^2 probabilities for all energies, allowing us to focus on the intrinsic evolution of the fraction of anomalous profiles with respect to energy. Additionally, to minimize the influence of the long tail of the deeply penetrating muonic component (which is not well described by the four-parameter Gaisser-Hillas function), any profile point where the ratio of the square root of its weight and its shower size is greater than 0.3 is excluded from the fitting process,

$$\frac{\sqrt{V_i}}{N_i} > 0.3. \tag{30}$$

Finally, to introduce realistic statistical fluctuations during the analysis, each profile point is adjusted by a random value drawn from a Gaussian distribution with a standard deviation equal to the square root of its corresponding weight factor ($\sqrt{V_i}$), this is done in order to obtain a meaningful χ^2.

When fitting for double bumps, which involve eight free parameters, the optimization process can easily become trapped in local minima. To mitigate this issue and guide the fitting towards a more optimal solution, we employ an initial guess for the fitting parameters. This initial guess is based on the shower profile itself. The position of the first peak (X_{max}) is identified as the data point with the maximum shower size. The X_{max} for the potential second peak is estimated by searching for a local minimum within a window of $\pm 300 \mathrm{g/cm^2}$ around the first peak's X_{max}. The remaining initial guess values are chosen based on a combination of physical expectations and experience with fitting air shower profiles. This approach helps the fitting algorithm converge on physically meaningful solutions for the double bump profiles.

Following the fitting process, to ensure that the identified peaks are easier to identify. In a double bump scenario, the largest sub-shower will always be N_{max_1}. The smaller sub-shower will thus always be N_{max_2}.

5.5 Identifying anomalous air showers

Our quest to identify anomalous air showers within the simulated data utilizes two distinct methods.

The first method builds upon the approach presented in [39], which we will refer to as the "old method". It leverages a double Gaisser-Hillas function fitting strategy to identify profiles that deviate significantly from the expected single peak behavior. Here's a breakdown of the key criteria used in this method:

The fitting process with two Gaisser-Hillas functions should offer a demonstrably better fit compared to a single Gaisser-Hillas (GH) function. All the fitting is done in python with scipy curve fit. This improvement is quantified by a ratio of chi-squared values exceeding a threshold:

$$\frac{\Delta\chi^2}{\chi^2_{\text{single GH}}} > 0.75. \tag{31}$$

Where $\Delta\chi^2$ is the difference between $\chi^2_{\text{single GH}}$ and $\chi^2_{\text{double GH}}$. Essentially, the double function significantly reduces the overall fitting error compared to the single function. This differs slightly

from the method in [39], in which they only looked at $\Delta\chi^2$. The choice to divide by $\chi^2_{\text{double GH}}$ was made such that the errors cancel out.

The two peaks (X_{max_1} and X_{max_2}) identified by the double Gaisser-Hillas function fitting should be well-separated in depth within the atmosphere. A minimum separation of is required:

$$\Delta X_{max} > 300 \text{ g/cm}^2. \tag{32}$$

This separation suggests the presence of two distinct sub-showers within the overall air shower profile.

Both fitted X_{max} values (X_{max_1} and X_{max_2}) representing the peak locations of the sub-showers must lie within a reasonable physical range. We require them to fall between 200 g/cm^2 and 2000 g/cm^2 of atmospheric depth. This ensures the identified peaks are physically plausible within the context of air shower development.

Each of the sub-showers identified by the double Gaisser-Hillas function fitting should carry a significant fraction of the initial cosmic ray energy.

$$N_{max_2} \geq 20\% N_{max_1}. \tag{33}$$

This criterion helps eliminate cases where the double peak might be an artifact of the fitting process rather than a genuine signature of an anomaly.

The second method we employ for anomaly detection is based on the Akaike Information Criterion (AIC) [51]. The AIC is a valuable tool in statistical modeling, particularly for model selection. It provides an estimate of the relative quality of different models in representing a given dataset.

The AIC takes into account a model's ability to fit the data (goodness-of-fit) while penalizing its complexity (number of parameters). This effectively avoids both overfitting (a model that captures too much noise in the data) and underfitting (a model that fails to capture the underlying trends). The AIC calculation is represented by:

$$\text{AIC} = 2k - 2\ln(L). \tag{34}$$

Where k is the number of parameters in the model and L is the maximum likelihood of the model fitting the data, in our chase:

$$\text{AIC} = 2k + \chi^2. \tag{35}$$

A lower AIC value indicates a better model relative to others being compared. Additionally, the likelihood ratio between two models can be estimated using the formula:

$$\exp((\text{AIC}_{min} - \text{AIC}_i)/2). \tag{36}$$

Where AIC_{min} is the minimum AIC value (best model) and AIC_i is the AIC value of the model under consideration. A higher likelihood ratio signifies a stronger preference for the model with the lower AIC. In the context of our analysis, we will compare the AIC values obtained from single and double Gaisser-Hillas function fits for each shower profile. A significantly lower AIC for the double Gaisser-Hillas function fit would be indicative of an anomalous air shower with a double peak structure being a more statistically plausible explanation for the observed data compared to a single peak.

Figure 25 exemplifies how this analysis works. It presents the findings for a dataset of 20.000 simulated proton showers generated using the QGSJET-II hadronic interaction model. The analysis investigates two methods, normal and reduced GH fits, for identifying double bumps within these showers. The x and y axes of each plot represent the unique parameters associated

with the normal and reduced GH functions for single bumps, respectively. The background in each panel is a 2D histogram, representing the overall distribution of showers classified as single bumps by the corresponding method (normal or reduced). Areas with a denser concentration of points in the histogram favor a single-bump model for those parameter values. Scattered on top of the histograms are points depicting showers identified as double bumps using the respective method, plotted at the location that corresponds to the parameters found by the single GH fit. The color of these points indicates a measure of likelihood, the lighter the color the stronger the preference for a double bump fit over a single bump fit. Points with a black ring around them represent double bumps identified by both the "old method" and the new AIC method. Points without a black ring are double bumps identified solely by the current (normal or reduced) method based on the AIC criterion. This distinction allows for a visual comparison of how well the new methods agree with the existing approach in identifying double bumps.

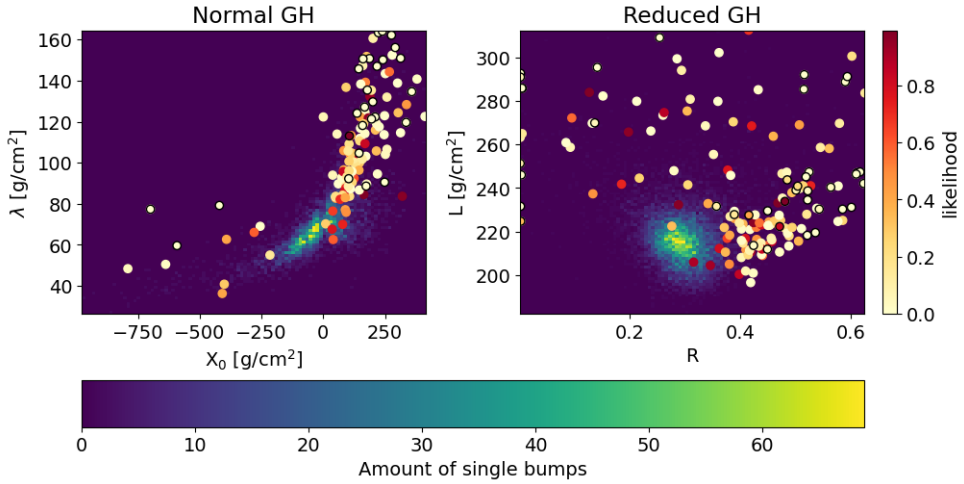


Figure 25: Comparison of normal and reduced GH fits for identifying double bumps in simulated proton showers. The analysis investigates 20.000 showers simulated with the QGSJET-II hadronic interaction model.

This plot intentionally excludes some double bump showers for better clarity. The x and y axes are restricted to the range of values occupied by single bumps. Figure 26 shows the plot without this limitation, allowing visualization of all double bumps.

Figure 25 reveals a clear distinction between the locations of single and double bumps in the parameter spaces defined by the normal and reduced GH functions. For the normal GH functions, single bumps tend to cluster in the bottom left portion of the plot, following an upward trend with increasing X_0 values. In contrast, most double bumps identified by the normal method reside in the upper right corner, suggesting a different region in the parameter space. Interestingly, the "old method" primarily captured double bumps positioned along an extension of the single bump curve. The new (AIC-based) method, however, exhibits a wider range of double bumps with more spread-out locations. Turning to the reduced GH method, single bumps are concentrated around $R \sim 0.3$ and $L \sim 220$. Double bumps identified by this method fall into three distinct regions. The first region, with the high density of double bumps, is located at $R > 0.35$ and $200 < L < 240$. The second region encompasses highly asymmetric showers with very low R values. The last region lies between the first two and includes showers with higher L values, indicating a larger variance in their profiles. Bellow you can find some

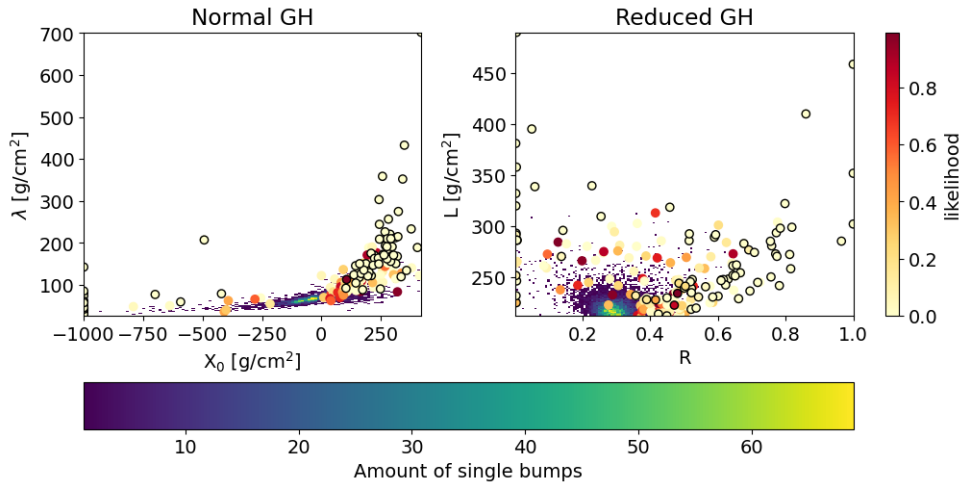


Figure 26: Distribution of double bumps identified by the normal and reduced GH methods (including those outside the single bump region). This figure complements Figure 25, which focuses on the core distribution of single bumps for improved clarity.

values comparing the old and AIC method and the normal and reduced GH:

- Amount of double bumps found with AIC methode: 316.
- Amount of extra double bumps found with AIC methode: 241.
- Amount of double bumps found by both normal and reduced GH: 301.
- Amount of double bumps found by normal GH and not reduced GH: 22.
- Amount of double bumps found by reduced GH and not normal GH: 7.

This analysis reveals a clear distinction between the new (AIC-based) method and the old method for identifying double bumps. The new method detects a significantly larger number of double bumps compared to the old approach because the old method put some very harsh criteria on the double bump showers. This was done to select only showers that would be candidates for detection with fluorescence at Auger. Interestingly, the normal and reduced GH methods exhibit a high degree of agreement in their identification of showers. The few discrepancies observed are likely due to inherent randomness in the fitting process and tend to involve showers with higher likelihood values indicating less certainty in the classification. To ensure a more robust selection of double bumps, we will establish a likelihood 36 threshold of 0.1. This value is chosen by examining the distribution of likelihoods in Figure 27. It is clear that the majority (over 50%) of the double bumps identified by the AIC method fall below this threshold. This means we will only consider showers identified as double bumps by the AIC method if their associated likelihood value is smaller than 0.1. This threshold helps us focus on showers with a higher degree of certainty in their double bump classification, while potentially excluding those with greater fitting ambiguity, for instance the top right plot in figure 27 has the lowest likelihood value because it is unquestionably a double bump shower.

Looking at an heavier cosmic ray (iron), figure 28, we can see that the single bumps have a smaller variance around $R \sim 0.35$ and $L \sim 220$ which is a slight increase in R while L stays the same. The type of double bumps is different compared to the proton cosmic ray case, unlike

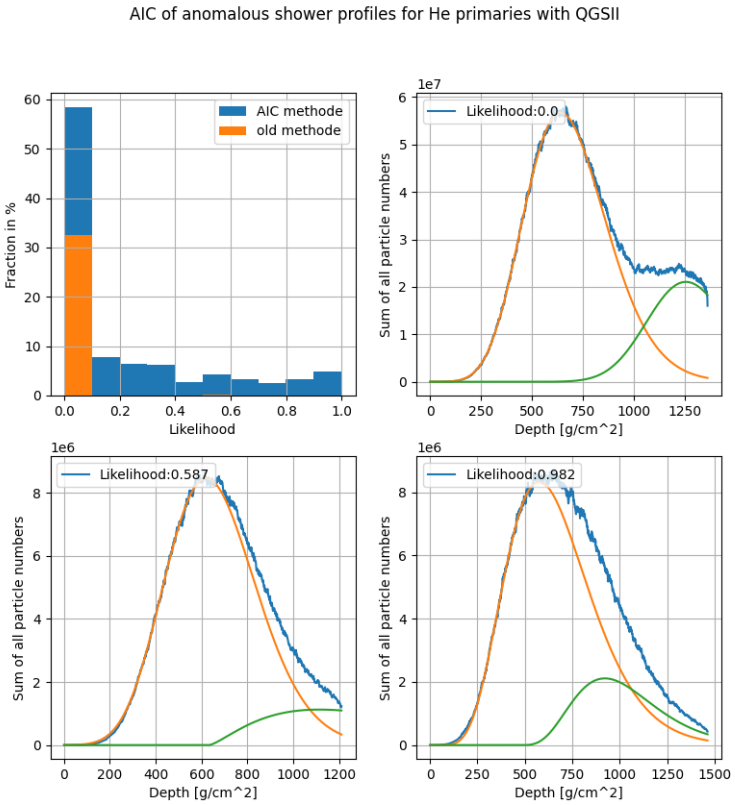


Figure 27: Comparison of Shower Profile Likelihoods using AIC and Existing Method for Helium Primaries with QGSJET-II Hadronic Interaction. Three example showers and the fitted sub showers are shown, each corresponding to a different bin (low, medium, and high) in the likelihood distribution.

before there is only 1 region. The double bumps can only be found around larger X_0 and λ or $R > 0.35$ and $200 < L < 240$.

All other plots showing the distribution of double bump showers compared to single bump showers can be found in the appendix A.1.

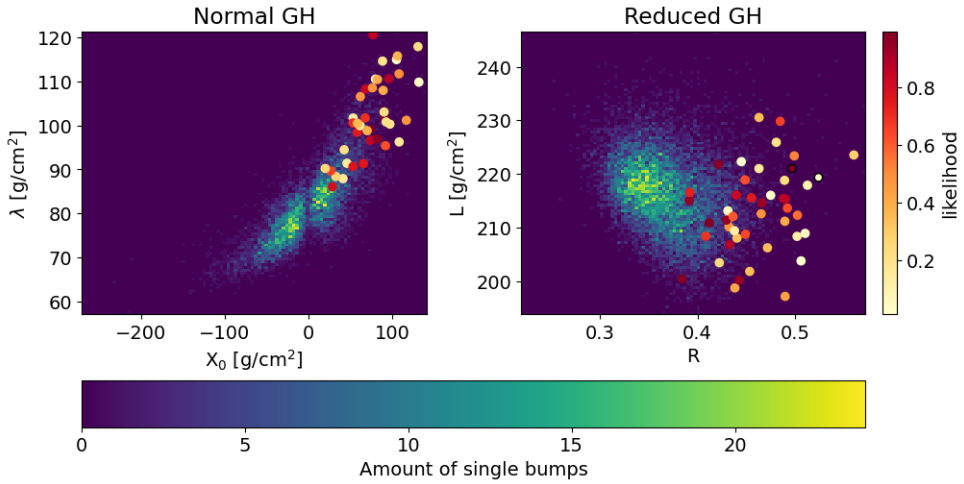


Figure 28: Comparison of normal and reduced GH fits for identifying double bumps in simulated iron showers. The analysis investigates 20.000 showers simulated with the QGSJET-II hadronic interaction model.

5.6 Frequency of double bumps

Following the shower profile selection criteria from the previous section (AIC method and the old method), we can determine the fraction of anomalous showers within the simulated sample. The frequency of double bumps for each primary cosmic ray particle and all hadronic interaction models is given by table 4.

Table 4: The frequency of double bumps for both the old and the AIC method (with likelihood < 0.1) for all the primary cosmic ray particles and hadronic interaction models.

	AIC method	old method		AIC method	old method
QGSJET-II	0.59%	0.27%	P	0.91%	0.46%
			He	1.11%	0.61%
			C	0.66%	0.22%
			Si	0.23%	0.03%
			Fe	0.05%	0.01%
EPOS	0.53%	0.24%	P	0.96%	0.50%
			He	0.92%	0.49%
			C	0.66%	0.21%
			Si	0.14%	0.03%
			Fe	0.03%	0%
SIBYLL	0.63%	0.29%	P	1.23%	0.73%
			He	1.09%	0.52%
			C	0.59%	0.17%
			Si	0.24%	0.03%
			Fe	0.03%	0%

Figure 29 illustrates the relationship between the primary cosmic ray particle type and the frequency of double bumps detected by the AIC method. As expected, the frequency of double bumps generally decreases with increasing particle mass. However, the QGSJET-II hadronic interaction model exhibits an anomaly compared to the other hardonic interaction models, the frequency of double bumps for proton primaries is lower than for helium.

As shown in Figure 30, nearly all simulations predict a negligible number of anomalous showers for iron primaries. This aligns with expectations as spectator nucleons in iron can only carry a small fraction (less than 2%, or 1/56) of the primary energy. Heavier fragment nuclei, while containing more energy, have a large interaction cross-section leading to a short mean free path in the atmosphere, further limiting the production of anomalous profiles. The AIC and old method produce nearly identical curves, with a constant offset arising from the stricter criteria applied in the new method. This consistency strengthens the validity of the results.

Interestingly, anomalous shower profiles are observed with similar frequency for helium and proton primaries. The overall fraction of these profiles decreases with increasing energy, as expected due to the rising interaction cross-sections implemented in the simulations.

L is a interesting parameter in air shower physics. It is obtained by fitting a single reduced Gaiser-Hillas function to the measured profile of a shower. This function helps quantify the width of the longitudinal evolution. Ideally, the measured values of L should be close to the reference values provided in Table 1. However, a significantly higher measured L value compared to the expected range might indicate a double bump shower. In such cases, the high L could be the function's attempt to account for the presence of two bumps in the shower profile, characteristic of double bump events. Figure 31 shows a lower prevalence of double bumps for heavier elements (carbon, silicon, and iron) compared to lighter elements, particularly for L values below 240. This

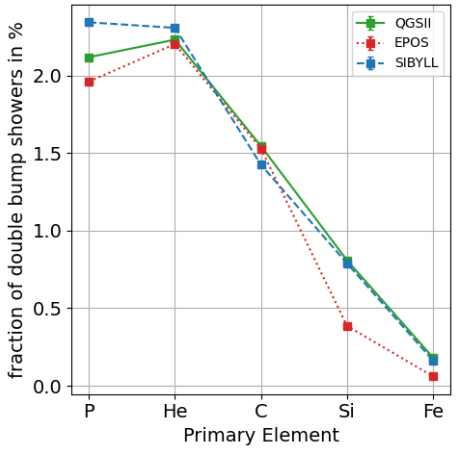
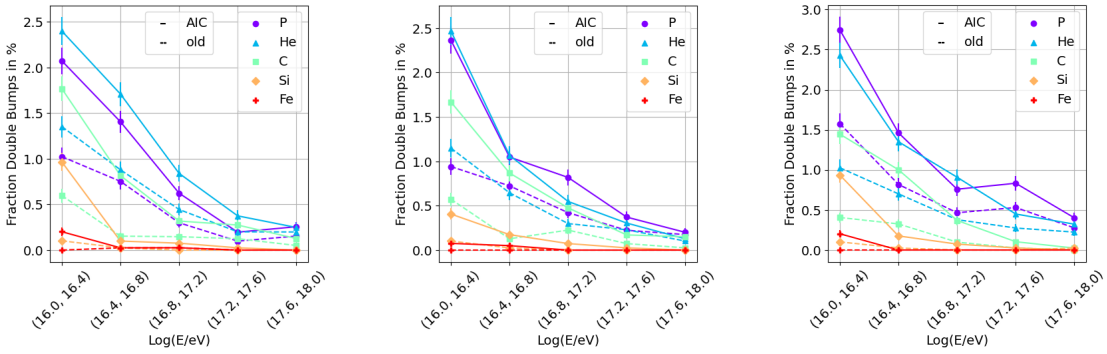


Figure 29: The frequency of double bump showers (between 10^{16} and $10^{16.5}$ eV) identified by the AIC method for all the primary cosmic ray particles (proton, helium, carbon, silicon and iron) and hadronic interaction models (QGSJET-II, EPOS and SIBYLL).

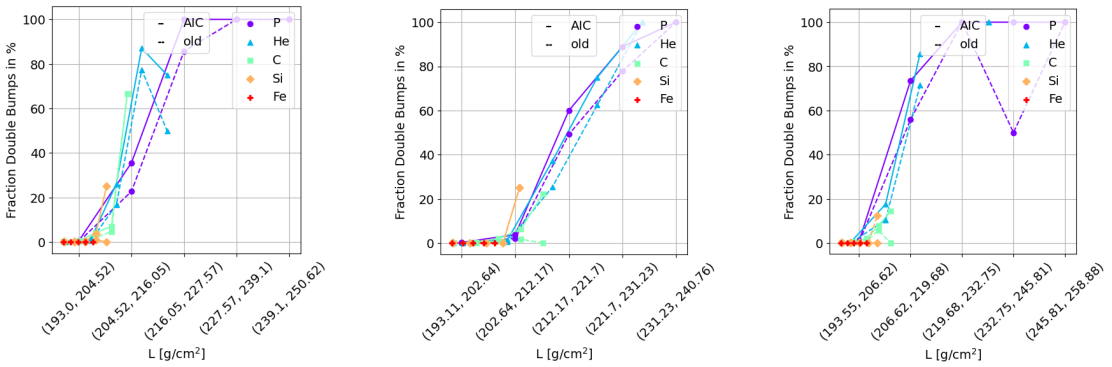


(a) QGSJET-II (b) EPOS (c) SIBYLL

Figure 30: Fraction of anomalous shower profiles for different primaries for different hadronic interaction models, plotted against $\log(E/eV)$. Each data point represents the percentage of anomalous shower profiles with double bumps detected for different primary particles (Proton, Helium, Carbon, Silicon, Iron).

can be attributed to the superposition model used to simulate these elements, which inherently leads to lower variance in the shower profile and consequently. The fraction of double bumps increases sharply for L values exceeding 250. This suggests that showers modeled with a single reduced Gaiser-Hillas function at these L values might be better represented by a double-bump model hence refitting using a double Gaiser-Hillas and comparing the results is warranted.

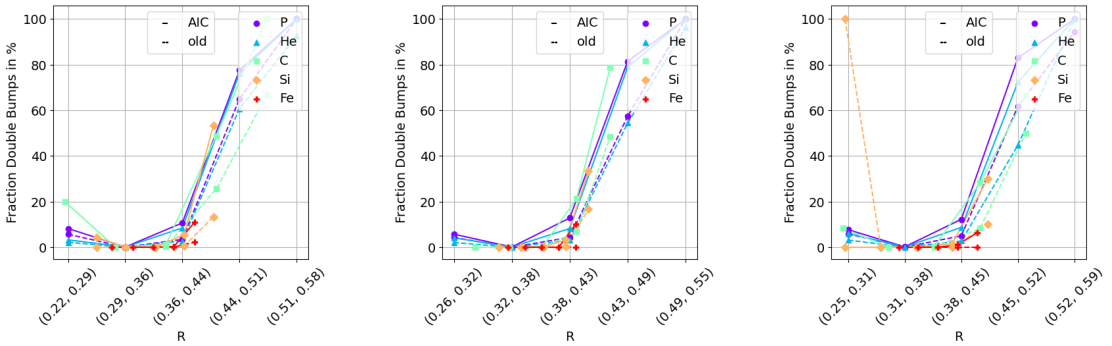
Figure 32 reveals a significant rise in the fraction of double bumps for low R values. The specific reasons behind this will be explored later, but it is noteworthy that the magnitude of this jump varies depending on the hadronic interaction model used. This observation might hold potential for differentiating between such models. As expected, there are near-zero double bumps observed for R values around 0.3. However, the fraction of double bumps increases for showers with R values approaching either extreme ($R < 0.1$ or $R > 0.5$), indicating a need for



(a) QGSJET-II (b) EPOS (c) SIBYLL

Figure 31: Fraction of anomalous shower profiles for different primaries with QGSJET-II Hadronic interaction model, plotted against L. Each data point represents the percentage of anomalous shower profiles with double bumps detected for different primary particles (Proton, Helium, Carbon, Silicon, Iron).

re-evaluating showers modeled with these R values and seeing if they might be better represented by a double-bump model.



(a) QGSJET-II (b) EPOS (c) SIBYLL

Figure 32: Fraction of anomalous shower profiles for different primaries with different hadronic interaction model, plotted against R. Each data point represents the percentage of anomalous shower profiles with double bumps detected for different primary particles (Proton, Helium, Carbon, Silicon, Iron).

Figure 33 shows some example showers of proton double bump showers simulated with the QGSJET-II hadronic interaction model which have interesting values for the parameters of the reduced Gaiser Hillas function (L and/or R). The left most shower is one with a very low R value, the middle shower has a large L and R value, the right most shower has a large R value but a lower L value.

Figure 34 shows the same showers but with fitted sub-showers shown and the total shower which is the sum of these sub-showers. This shower is a double bump which has a smaller sub-shower before the main sub-shower, the bulk of the main shower development is cut of due to happening at a low depth and thus reaching the surface of the earth. The right shower has a

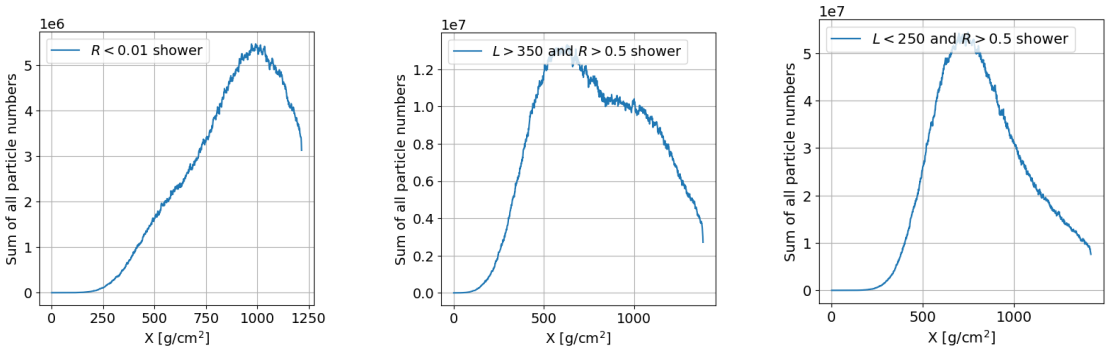


Figure 33: Example proton double bump showers simulated with QGSJET-II hadronic interaction model which have extreme parameters for L and/or R from the single reduced Gaiser Hillas fit.

low R value because the smaller sub-shower precedes the main sub-shower. The middle shower has a higher R value and also a high L value, in this case the main sub-shower develops before the smaller sub-shower. This explains why the R value is larger compared to the left plot. The large L value is due to the single bump Gaiser Hillas needing to compensate for both peaks. The right most plot looks the least like a double bump shower by eye, and shows why we need an automated method to identify them. The lower L value is due to the fact that the second sub-shower is not large enough to influence it. The larger R value indicated that the smaller sub-shower happens after the main sub-shower.

Figure 34 shows the same showers but with fitted sub-showers and the total shower profile, which is the sum of the individual sub-showers. The leftmost panel depicts a smaller sub-shower precedes the dominant sub-shower. However, the main shower development occurs at a depth exceeding the detectable depth, leading to a truncated profile at the Earth's surface. The middle panel depicts a shower with a higher R value. This is attributed to the preferential development of the main sub-shower before the smaller one. The single Gaiser-Hillas fit attempts to compensate for both peaks, resulting in a larger L value compared to the left panel. The rightmost panel exemplifies the limitations of relying solely on visual identification for double bumps. This shower exhibits a subtler double bump profile, highlighting the necessity of an automated method to capture such features effectively. The lower L value observed here can be explained by the smaller size of the second sub-shower, which has minimal impact on the overall shower profile. Conversely, the larger R value suggests that the second sub-shower develops after the peak of the main sub-shower.

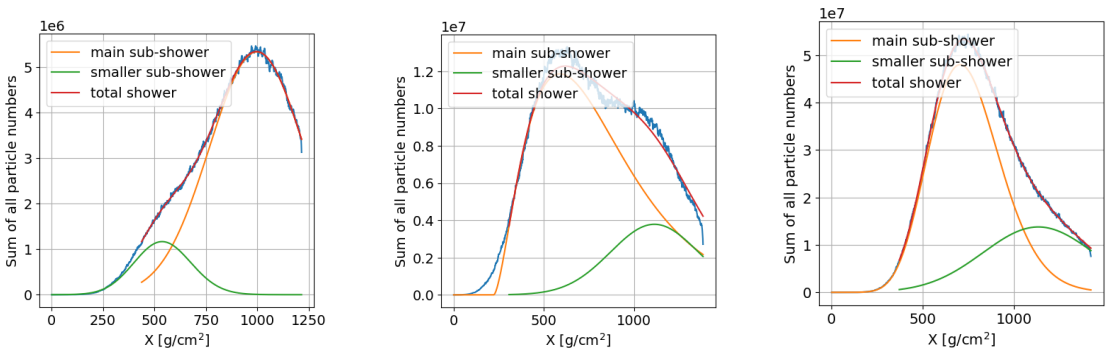


Figure 34: The same showers as 33 with the sub-showers which make up the bumps shown.

5.7 Classifying types of double bumps

By examining the distributions of primary energy (E_0) and the distance between the double bumps (ΔX_{max}) as shown in figure 35, we observe that the overall number of double bumps identified decreases as the primary energy increases. Interestingly, the AIC method consistently finds a higher number of double bumps across all energy ranges compared to the existing method. The number of double bumps increases with the absolute value of ΔX_{max}, creating a "hill and valley" structure centered around $\Delta X_{max} \sim 0$. Notably, the AIC method primarily identifies additional double bumps within regions where the old method does not find any. This is because the old method employs an arbitrary limit of $\Delta X_{max} > 300$ g/cm^2 for double bump identification, a restriction not present in the AIC method. The fact that the AIC method finds more double bump showers compared to the old method for $\Delta X_{max} > 300$ g/cm^2 comes from double bumps with a $N_{max_2} < 20\% N_{max_2}$ and/or from the double bumps that failed the $\frac{\Delta\chi^2}{\chi^2_{\text{single GH}}} > 0.75$ cut off.

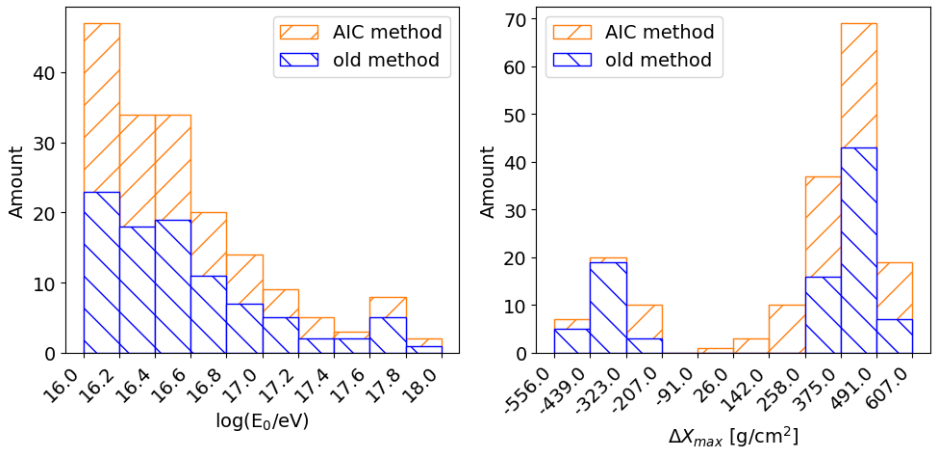


Figure 35: Primary energy (E_0) and distance between the double bumps (ΔX_{max}) for different shower classifications based on the old method and the AIC method for different shower classifications based on the old method and the AIC method for 20.000 proton primaries showers simulated with the QGSJET-II hadronic interaction model.

Comparing the results from helium cosmic rays to those from iron the differences are clear, see figure 36.

Consistent with expectations for heavier cosmic rays, iron primaries only exhibit double bumps at lower energies ($\log(E_0/eV) < 16.4$. The AIC method identifies slightly more double bumps compared to the existing method within this energy range. The distance between the double bumps in depends on the mass of the primary particle, the heavier the particle the smaller $|\Delta X_{max}|$, this is because the probability of a heavier nucleus remaining intact and traveling a large distance before its first interaction is very low. This does not hold for proton primary because they do not have this fragmentation of the nucleus.

Notably, no double bumps with negative ΔX_{max} values are observed for iron primaries. Negative ΔX_{max} indicates a smaller sub-shower preceding the larger one (occurring at a higher depth in the atmosphere). This finding is particularly interesting because the frequency of these "early" double bumps depends heavily on the chosen hadronic interaction model. This study shows that SIBYLL predicts a significantly higher occurrence of early double bumps in both

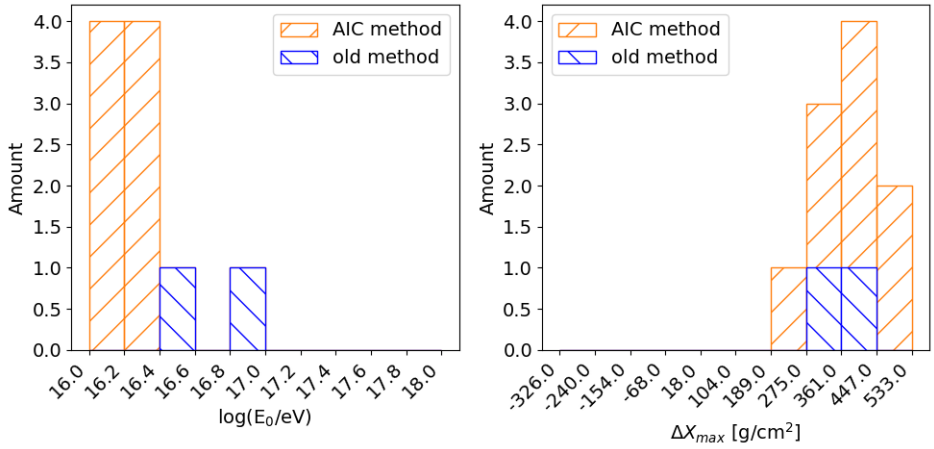


Figure 36: Primary energy (E_0) and distance between the double bumps (ΔX_{max}) for different shower classifications based on the old method and the AIC method for 20.000 iron primaries showers simulated with the QGSJET-II hadronic interaction model.

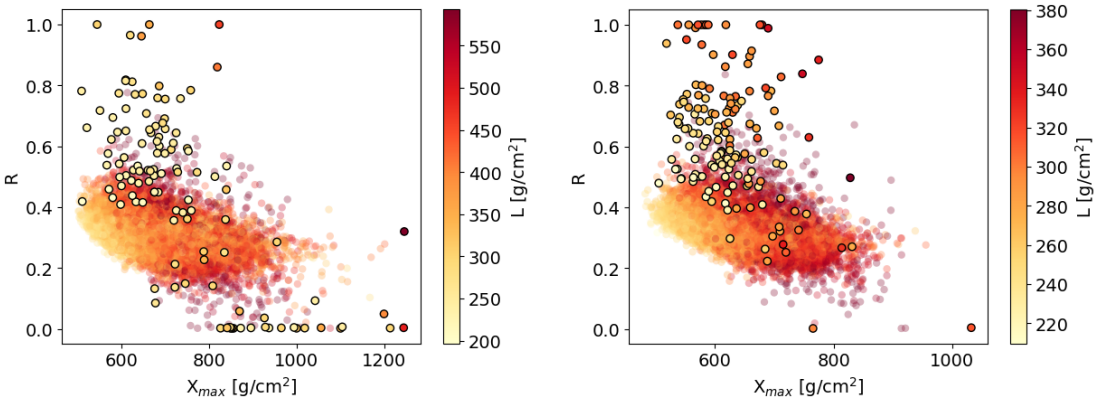
proton and helium showers compared to QGSJET-II and EPOS. The following table 5 details the frequency of early double bumps for different models and all the plots can be found in A.2. We observe a decrease in the percentage of early double bumps for heavier cosmic rays (carbon, silicon, and iron) compared to lighter ones. This can be understood by considering the superposition model used to simulate these interactions. For early double bumps to occur, a significant portion of the nucleus needs to travel a longer distance in the atmosphere before interacting, causing a separation in the shower development. However, the probability of a heavier nucleus remaining intact and traveling a large distance before its first interaction is very low. This is because heavier nuclei are more likely to break apart (fragment) during their initial collision with air molecules. As a result, the shower development for heavier nuclei tends to be more compact, leading to a lower prevalence of early double bumps, particularly for elements like carbon, silicon and iron.

	P	He	C	Si	Fe
QGSJET-II	38%	8%	1%	0%	0%
EPOS	22%	4%	1%	0%	0%
SIBYLL	52%	19%	4%	0%	0%

Table 5: The percentile distribution of early double bump showers (the smaller of the two sub-showers reaches its peak depth (X_{max}) at a greater depth compared to the main shower). The table indicates that a higher percentile of double bumps exhibit this behavior at deeper depths.

In figure 37 the parameters of a single bump fit are shown, X_{max} and R values are plotted against each other with the color indicating the L value for the proton and helium showers simulated with the QGSJET-II hadronic interaction model. The general shapes for the single bump showers are the same but the differences show themself by looking at the double bump showers, which are indicated by dots with black circles around them. In the case of proton showers, there are two groups of double bump showers, the higher R and low X_{max} double bumps and the low R and high X_{max} double bump showers. As stated before these low R

showers are "early" double bumps while the high R showers are the "late" double bumps. This shows a major way in which double bumps differ between proton and helium.



(a) Proton (b) Helium

Figure 37: Fitted X_{max} and R (and L indicated by colour) for showers generated by proton and helium primaries with the QGSJET-II hadronic interaction model.

The reason for these two distinct R values can be found by looking at figure 38, which shows all the proton showers for each hadronic interaction model combined. There is a clear relation between the early double bumps ($\Delta X_{max} < 0$) and low R values which are obtained by fitting the double bump shower with a single bump reduced GH.

In the case of a large distance and $\Delta X_{max} > 0$ larger R values are found but it seems to depend on both the ratio of the size of the 2 sub-showers and the distance between them, higher N_{max_2}/N_{max_1} and higher $X_{max_2} - X_{max_1}$ showers have a higher R value if $\Delta X_{max} > 0$.

Looking at left plots of figure 38 the relation between L and $X_{max_2} - X_{max_1}$ or $|\Delta X_{max}|$ and N_{max_2}/N_{max_1} becomes harder to tell. We do not expect the L values to depend on the sign of $X_{max_2} - X_{max_1}$ because it measures the broadness of the shower. There should be a correlation between $|\Delta X_{max}|$ and L, this can be seen by thinking about the fact that the L parameter needs to cover the profile of both bumps. The same logic can be used to argue that a correlation between N_{max_2}/N_{max_1} and L is expected. This can be tested by calculating the Pearson correlation coefficients, this is only testing for linear correlations, and the p-values for all the parameters, the results are given in the table 6.

		X_{max_2}-X_{max_1}		$\lvert\Delta X_{max}\rvert$	N_{max_2}/N_{max_1}	
R	double bump type	early	late	both	early	late
	r	0.12	0.49	0.31	0.33	0.36
	p-value	2.06×10^{-1}	2.33×10^{-29}	2.80×10^{-14}	2.81×10^{-4}	1.20×10^{-15}
L	double bump type	early	late	both	early	late
	r	−0.28	0.078	0.1	0.24	0.54
	p-value	1.93×10^{-3}	9.29×10^{-2}	1.52×10^{-2}	8.32×10^{-3}	1.63×10^{-37}

Table 6: The Pearson correlation coefficient (r) and the p-values for all the combinations X_{max_2}-X_{max_1}, $|\Delta X_{max}|$ and N_{max_2}/N_{max_1} correlated with R and L for both early and late double bumps for all the proton simulations (QGSJET-II, EPOS and SIBYLL) combined.

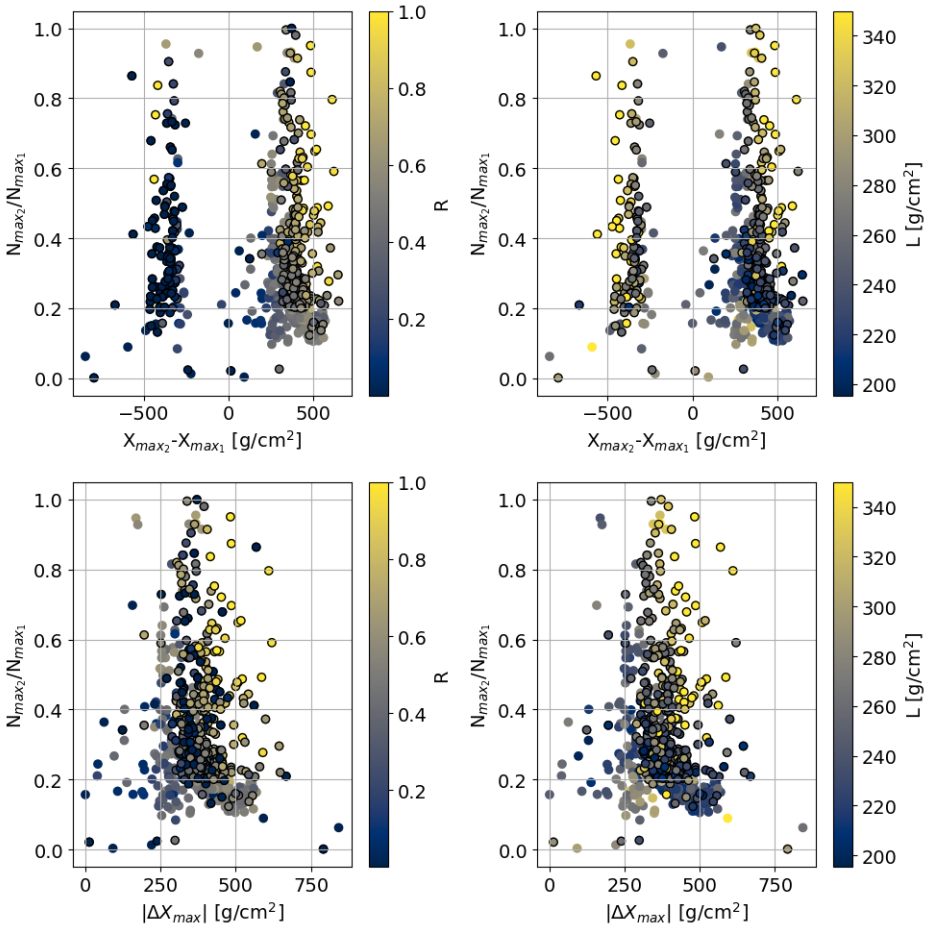
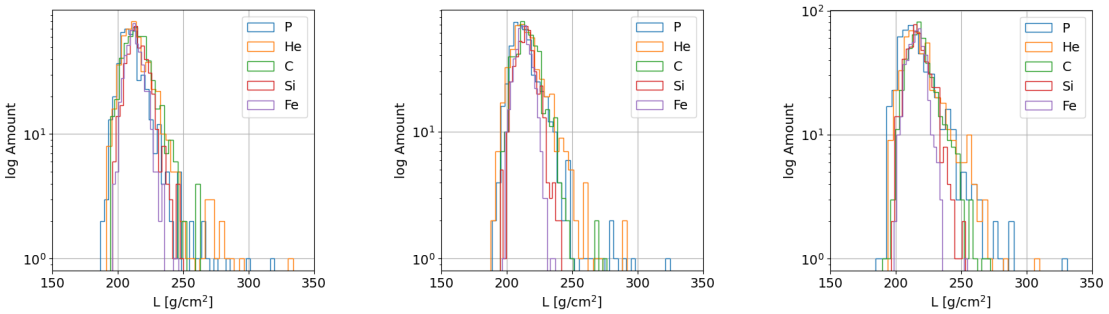


Figure 38: Peak density ratio (N_{max_2}/N_{max_1}) on the y-axis versus the difference in peak depths (X_{max_2}-X_{max_1}) or the absolute value ($|\Delta X_{max}|$) on the x-axis. With the color bar indicating the R or L value of the double bump shower if they were to be fitted as a single bump shower with a reduced GH fit. For different shower classifications based on the AIC method and the old method (circles with a black outline) for all the proton primaries showers simulated (QGSJET-II, EPOS and SIBYLL) combined.

In table 6 we can see that as expected the correlation between X_{max_2}-X_{max_1} and R for early double bump showers is weak while it is strong for late double bump showers. Both early and late double bump showers show a positive correlation between N_{max_2}/N_{max_1} and R. For late double bumps the correlation is strong between N_{max_2}/N_{max_1} and L meanwhile for early double bumps for both X_{max_2}-X_{max_1} and N_{max_2}/N_{max_1} there seems to be a moderate correlation between them and L. The successful detection of early double bump showers (those with the leading sub-shower reaching its peak near the detectors) remains an open question. These showers produce strong radio emission concentrated in a small region around the shower core due to the proximity of the main sub-shower's peak to the detectors. This concentrated emission might pose detection challenges for SKA. If this proves to be the case, it will be crucial to factor in this potential bias when analyzing SKA data to estimate the frequency of double bump showers.

5.8 Relation to stretched showers

Figure 39 visualize the distribution L for showers between 10^{16} and 10^{17} eV, categorized by the primary particle type. Notably, all three hadronic interaction models exhibit good agreement. While the peak values across elements are similar, the high-L tails differ significantly. Consequently, focusing on the width of the distribution or the tail shape becomes a more reliable observable, as it is less susceptible to larger systematic uncertainties.



(a) QGSJET-II (b) EPOS (c) SIBYLL

Figure 39: Histograms of fitted L for all elements and interaction models for showers between 10^{16} and 10^{17} eV.

Stretched showers lay between the universal and double bump showers, these are showers that exhibit a larger L value, and thus they are the showers that lay in the high-L tails of 39. The first idea as to how to define stretched showers is as showers with L values higher than the $L_{cut} = 2\sigma$ values which is calculated by fitting a Gaussian to all single bump showers with the same primary particle, simulated with the same hadronic interaction models and within the same energy range. An example of how this works is shown in figure 40.

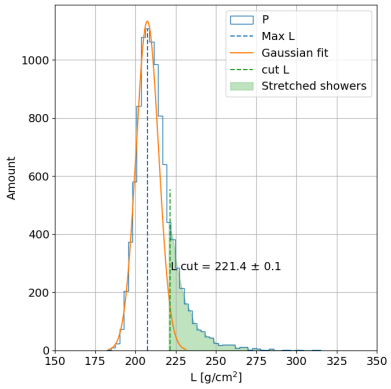
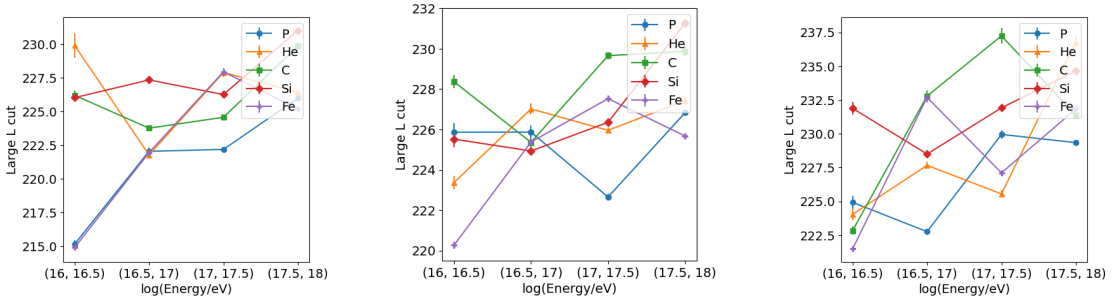


Figure 40: Histograms of fitted L for single bump proton showers simulated with QGSJET-II in the energy range $10^{16} - 10^{17}$ eV. With a Gaussian fitted to the rising part of the distribution, L_{cut} is places at 2σ. Stretched showers are defined as showers with L values higher than the L_{cut}.

This method results in different L_{cut} depending on the energy, this is shown in figure 41. This method is not stable and thus this method can not be used to determine stretched showers.

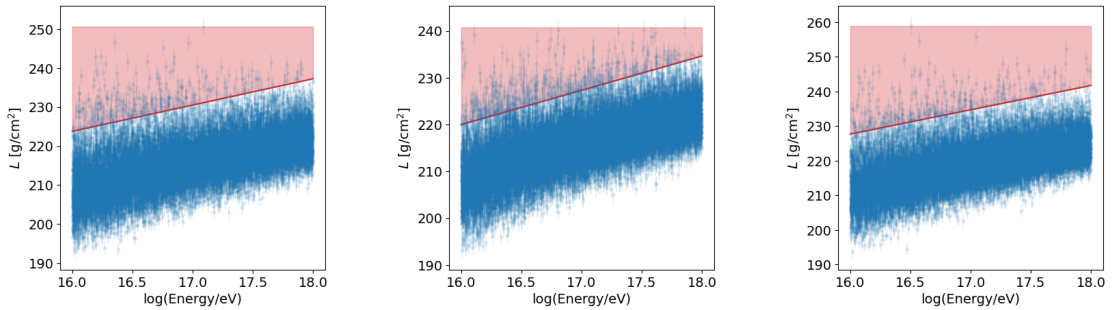
A problem in this method comes from the fact that we define $L_{cut} = 2\sigma$ which means even for a perfect Gaussian $\sim 2.5\%$ of the showers would still be counted as stretched showers.



(a) QGSJET-II (b) EPOS (c) SIBYLL

Figure 41: The cut of values for L (L_{Cut}) for when a shower gets identified as a stretched shower in terms of energy range for the different hadronic interaction models.

Instead we will use the difference in the L distribution between primary cosmic ray nuclei. Iron nuclei, with a larger interaction length, tend to exhibit a more symmetrical L distribution compared to lighter nuclei like protons. This property can be exploited to define a threshold for stretched showers. A linear fit $(ax + b)$ is applied to the relationship between shower energy and the corresponding L values. Showers with L values exceeding a specific threshold above this fit are classified as stretched showers. This threshold is is set at a number of standard deviations above the fitted line, for this thesis 10σ. Figure 42 illustrates this concept visually, and shows why the 10σ was chosen. It is important to note that the specific parameters of the linear fit will vary depending on the chosen hadronic interaction model. The corresponding fit coefficients for different models are presented in Table 7.



(a) QGSJET-II (b) EPOS (c) SIBYLL

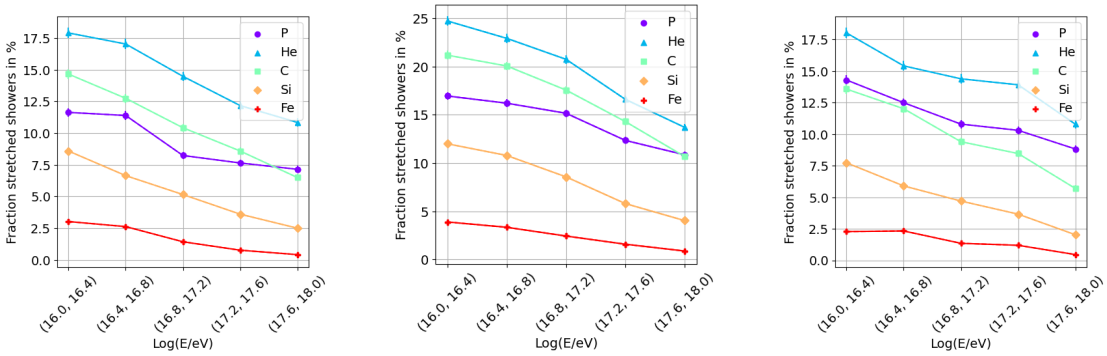
Figure 42: L values plotted against energy for iron primary cosmic ray showers. A linear fit is made and the L cut-off is defined as the showers which lie over 10σ.

Figure 43 depicts the variation in the percentile of stretched showers across different primary cosmic ray nuclei and hadronic interaction models, as a function of energy. Unlike the results observed in Figure 41, a clear downward trend emerges for the fraction of stretched showers at higher energies. This trend aligns with the previously discussed observation for double bump showers, where their prevalence decreases with increasing energy.

Table 7: Fitted $(ax + b)$ parameter values from fitting energy against L for iron primary cosmic rays for the different hadronic interaction models and the average of the three.

	QGSJET-II	EPOS	SIBYLL	average
a	6.73	7.32	7.01	7.02
b	116.12	102.88	115.53	111.4

The fraction of stretched proton showers consistently falls below both helium and carbon showers. This differs from the behavior seen with double bump showers, where the dominant fraction (proton or helium) depended on the specific hadronic interaction model. Notably, the QGSJET-II model suggests that the stretched carbon shower fraction dips below the stretched proton fraction at higher energies. Additionally, the EPOS model exhibits a significantly higher overall fraction of stretched showers compared to other models.



(a) QGSJET-II (b) EPOS (c) SIBYLL

Figure 43: Fraction of stretched shower profiles for different primaries and different Hadronic interaction model, plotted against energy. Each data point represents the percentage of stretched shower profiles for different primary particles (Proton, Helium, Carbon, Silicon, Iron).

The exact fraction of stretched showers for each primary cosmic ray particle and hadronic interaction model can be found in table 8.

Figure 44 shows the correlation between the fractions of stretched showers and double bump showers, identified using the AIC method. The plot depicts this relationship for various primary cosmic ray nuclei and different hadronic interaction models. A clear linear trend emerges between the fractions of stretched showers and double bumps, suggesting a potential underlying connection between these two shower types. This probably comes from the fact that these two shower types originate from the same idea, a leading secondary particle traveling a significant distance more in the atmosphere compared to the others. However, an interesting exception is observed for proton primaries. Across all models, proton showers consistently appear as outliers in this relationship. Furthermore, the figure reveals that the fraction of double bumps for protons and helium falls within one standard deviation for all models. This contrasts with the stretched shower fraction, where the difference between protons and helium is consistently above 3% for every hadronic interaction model.

Figure 45 shows the correlation between stretched shower and double bump fractions, identified using the AIC method, for various primary cosmic rays and hadronic interaction models. This analysis is further divided by different energy ranges. The key observation from splitting

QGSJET-II	8.2%	P	9.2%
		He	14.5%
		C	10.6%
		Si	5.3%
		Fe	1.6%
EPOS	12.2%	P	14.2%
		He	19.7%
		C	16.7%
		Si	8.2%
		Fe	2.4%
SIBYLL	8.4%	P	11.3%
		He	14.5%
		C	9.8%
		Si	4.8%
		Fe	1.5%

Table 8: The frequency of stretched showersfor all the primary cosmic ray particles and hadronic interaction models.

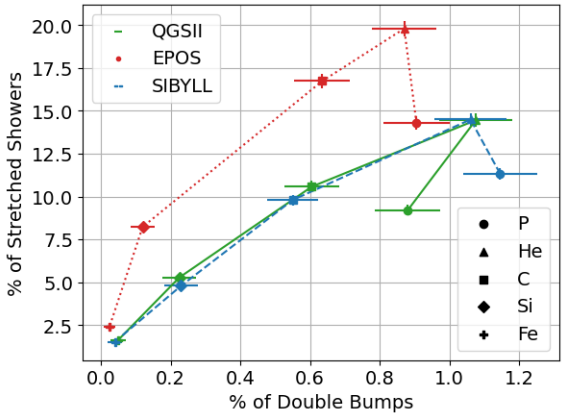
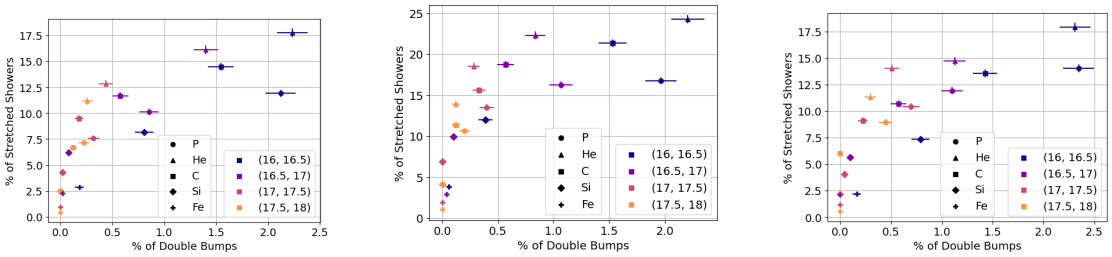


Figure 44: The percentile of stretched showers plotted against the percentile of double bumps identified using the AIC method for different primary cosmic rays and different hardonic interaction models.

the data by energy is the dependence of the linear relationship's slope on energy. Higher energy ranges exhibit a steeper slope, because the fraction of double bump showers goes down rapidly at higher energies. Notably, protons continue to deviate from this trend across all energy ranges. The figure also reveals a more pronounced increase in the double bump fraction compared to the stretched shower fraction when transitioning to lower energies. This effect becomes evident when focusing on a single particle type and connecting the fractions across energy ranges. This pattern, even observed for protons, suggests a logarithmic-like trend. The slower decrease in the stretched shower fraction compared to double bumps at higher energies provides a potential explanation for this observation.

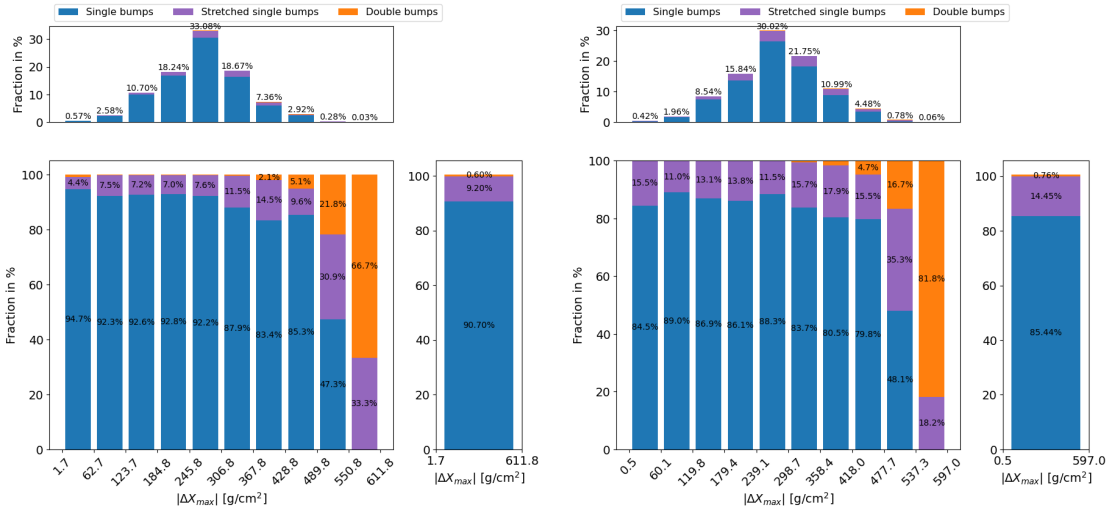
Another way to see the connection between single bump showers, stretched showers and double bump showers is by looking at the fraction of each type in function of a shower parameter.



(a) QGSJET-II (b) EPOS (c) SIBYLL

Figure 45: The percentile of stretched showers plotted against the percentile of double bumps identified using the AIC method for different primary cosmic rays and different hardonic interaction models split for different energy ranges.

Figure 46 shows a distribution of the absolute value of the fitted $|\Delta X_{max}|$ obtained by fitting a double Gaiser-Hillas function with QGSJET-II proton and helium shower profiles to all showers. Each bar in the histogram is further subdivided based on the identified shower type (single bump, stretched, double bump). This visualization effectively highlights the transition between stretched and double bump showers as the absolute value of $|\Delta X_{max}|$ increases. Consistent with previous observations, the figure reveals a lower fraction of stretched showers for proton primaries compared to helium primaries. This trend also holds true for the double bump fraction.

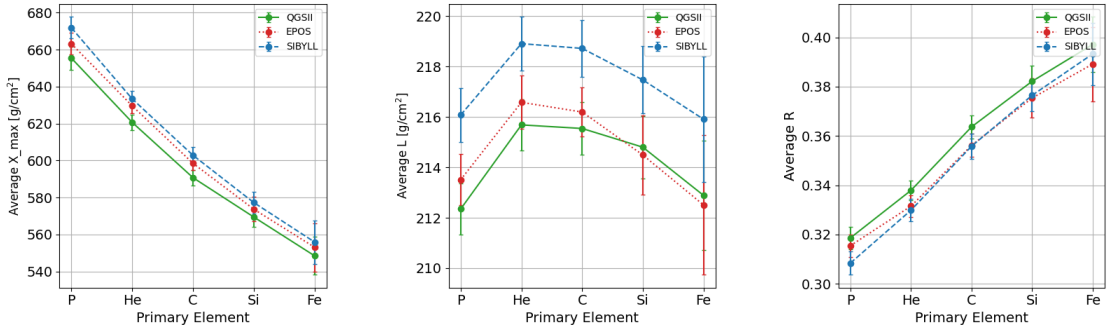


(a) Proton (b) Helium

Figure 46: Fraction of anomalous shower profiles for proton and helium primaries with QGSJET-II . The bottom left subplot shows the fraction of showers identified as having a single bump for different bins of $|\Delta X_{max}|$. The bars are colored differently to represent single bumps, stretched showers and double bumps. The top subplot shows the same data as the bottom left subplot the the total bar heights are now representative of the actual fraction they represent in the data. The right subplot shows the fraction of showers for each category for the entire data set.

5.9 Effect on SKA measurements

Figure 47 presents the dependence of various shower parameters on the primary element for different hadronic interaction models (QGSJETII, SIBYLL, and EPOS). The plot displays the average values of X_{max}, R, and L for both single as a function of the primary element. Most shower parameters exhibit a consistent trend across primary elements, changing monotonically. This reflects the influence of increasing mass on shower development. An interesting exception is observed for L. L shows a sharp rise from protons to helium, followed by a steady decrease for heavier elements. This unique behavior of L for helium can be exploited for its identification. By combining a measurement of L with any of the other shower parameters (X_{max} or R), it is possible to effectively isolate protons from all other elements in the primary cosmic ray flux.



(a) Average X_{max} (b) Average L (c) Average R

Figure 47: Results from simulations. Average X_{max}, L and R as a function of primary element for different hadronic interaction models (QGSJETII, SIBYLL, and EPOS).

Figure 48 visualizes the average X_{max} and L across all primary cosmic ray particles which generated a single bump, primary energies ranges (given in the orange box), and hadronic interaction models (QGSJETII, SIBYLL, and EPOS). Additionally, the figure displays the fraction of double bumps detected for each specific element, energy range, and interaction model. For each combination of energy and interaction model, the elements form a triangle on the X_{max} against L plot. Protons consistently occupy one corner of the triangle, while all other elements lie on the opposite side. Notably, for some elements, this opposite side exhibits a curved shape, and protons never share the same curve as the other elements. The average L predicted by the models diverges significantly at higher energies. In particular, SIBYLL consistently predicts a higher average L compared to QGSJETII and EPOS. The actual cosmic ray flux is likely a mixture of various elements. As a result, the average X_{max} and L for cosmic rays would fall within the triangle defined by the different elements in the simulation. This allows observations of X_{max} and L to be used as a tool to validate hadronic interaction models, even with limited knowledge of the actual mass composition of cosmic rays. Measurements of X_{max} and L by the Square Kilometre Array (SKA) telescope can provide a unique data set to probe the details of hadronic processes occurring during cosmic ray showers.

Figure 48 will be affected by whether or not double bump showers are recognized and removed from the sample. Looking at figure 49 the average L, R and X_{max} are shown for both single and double bump showers. The double bump showers will cause the average R and L values to go up compared to single bumps. Interesting is that the average R for double bumps drops off for proton primary double bump showers and this because of the high fraction of early double bump showers compared to helium. The biggest shift will be in average L as those double bump

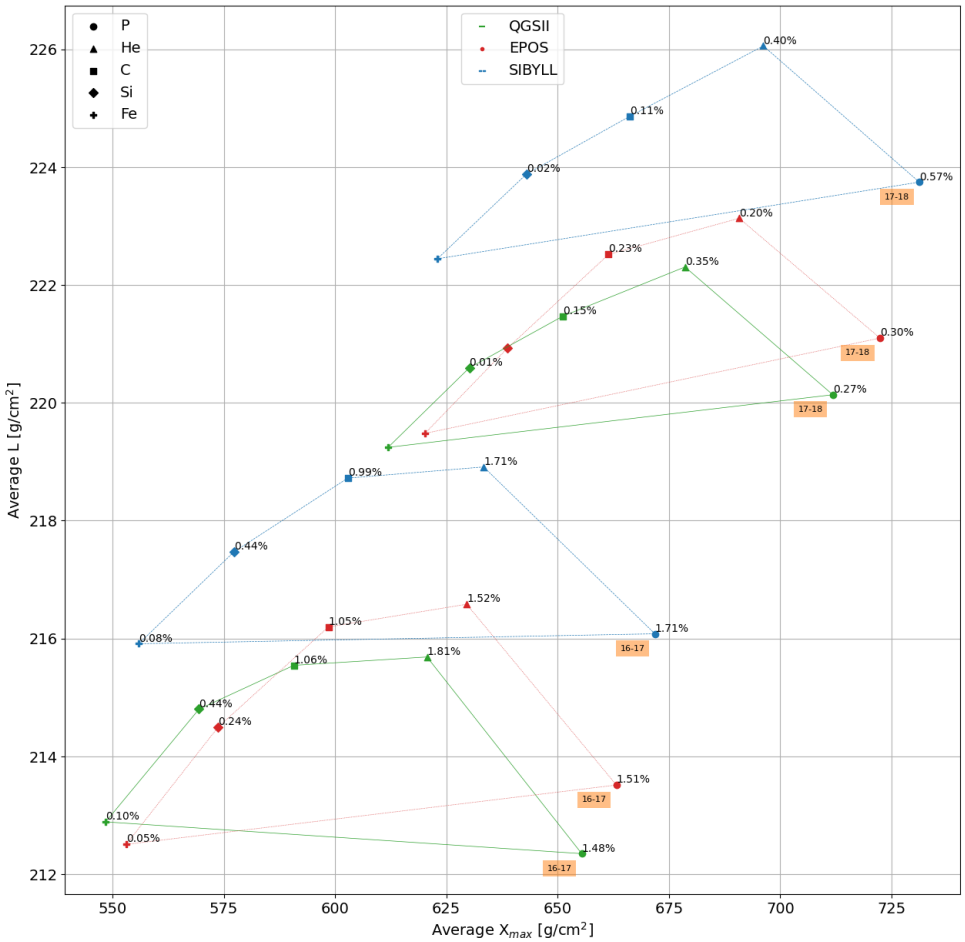
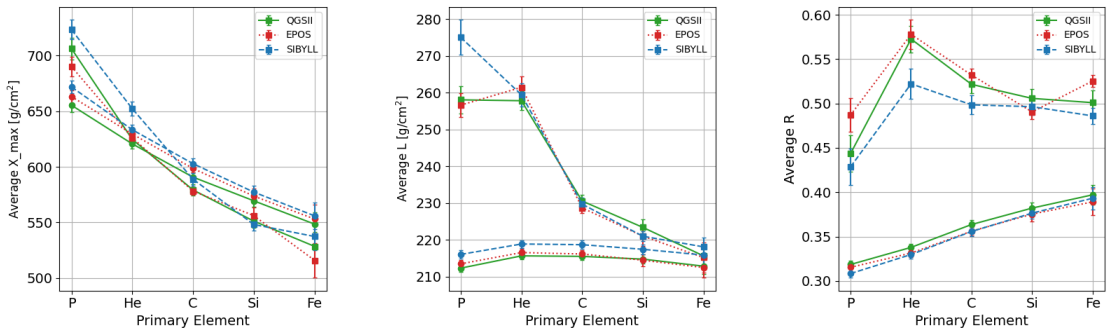


Figure 48: Average L versus average X_{max} for different energies and hadronic interaction models. In each case, the elements form a triangle with protons on one corner and all other elements on the opposite side. Next to each point is the percentile of double bumps detected for that given value.

showers generated by proton and helium primary cosmic rays have very high L values compared to single bump showers. The average X_{max} values for double bump showers follow the same trend as single bump showers but average X_{max} goes down steeper. This makes the difference between proton and helium primary showers bigger, going from a difference around 33 g/cm^2 for single bump showers to at least double for double bump showers.

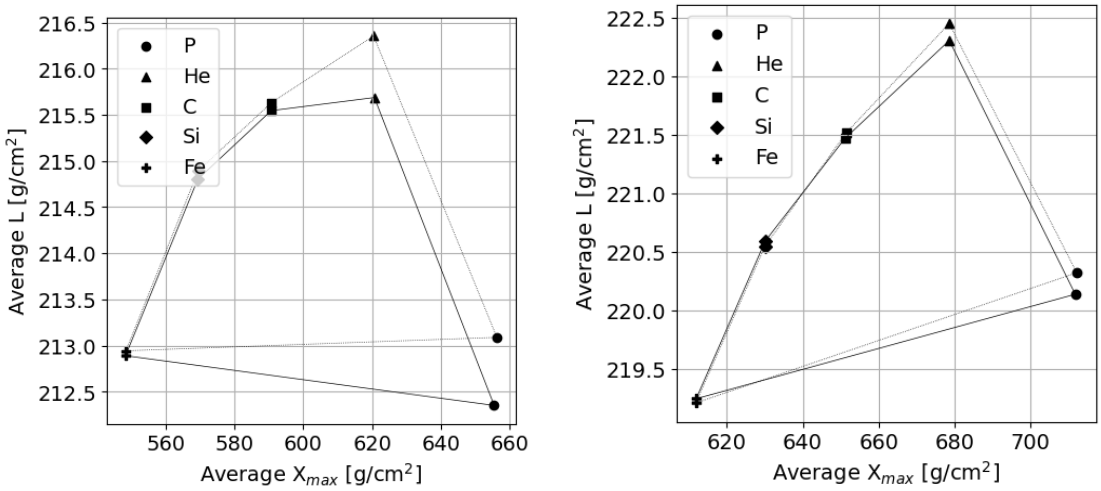
The shift due to double bump showers is shown in 50 for two different energy ranges ($10^{16} - 10^{17}$ eV and $10^{17} - 10^{18}$ eV), for the QGSJET-II hadronic interaction models. The dotted line is before filtering the double bump showers out and the solid line is after filtering the double bumps out. As expected the shift at the higher energies is smaller. The shift in average L is the largest, especially for proton and helium. While the average values of the shower parameter L exhibit slight variations (order of a few g/cm^2) across different primary cosmic ray elements, a robust prediction for the systematic uncertainty associated with L reconstruction using SKA is still lacking. This highlights the significant challenge in achieving the required accuracy.



(a) Average X_{max} (b) Average L (c) Average R

Figure 49: Results from simulations. Average X_{max}, R, L for both single (circles) and double bumps (squares) as a function of primary cosmic ray particle for different hadronic interaction models. The solid lines represent the values for single bump showers while the dashed lines represent the values for double bump showers.

Consequently, filtering out double bump showers, particularly at lower energies, might become warranted. This plants the idea of our head of using the fraction of stretched showers instead of average L since it is a more robust measurement, which we will come back to shortly.



(a) Energy range: $10^{16} - 10^{17}$ eV. (b) Energy range: $10^{17} - 10^{18}$ eV.

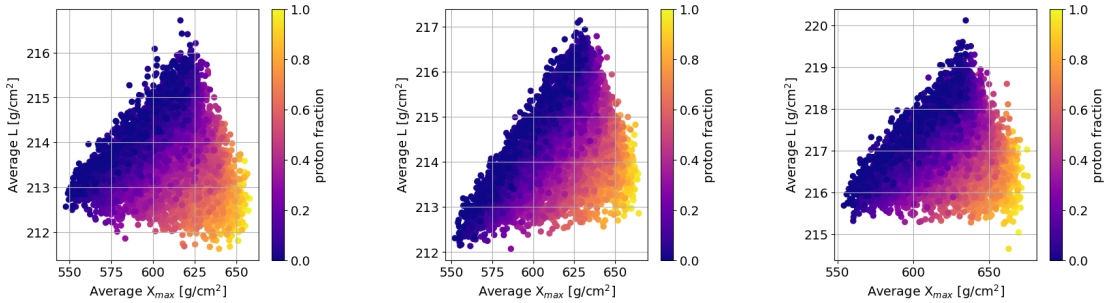
Figure 50: Average L versus average X_{max} in energy ranges for the QGSJET-II hadronic interaction model. Showing the shift in the triangle caused by the filtering of double bump showers. The dotted line is before filtering and the other is after.

To test the impact of reconstructing L on the mass composition analysis, we create random astrophysical models containing varying mixtures of three elements: proton, helium, and iron. For each model, 1000 simulated cosmic ray showers are generated. The specific mixture of primary cosmic rays used in each simulated shower event is determined by a parameter called the "percent jump." The process starts with a model containing 100% proton showers. Subsequent

models are created by decreasing the proton fraction by the specified percent jump, with the remaining fraction being between helium and iron nuclei such that all possible combinations of helium and iron fractions within the remaining portion are modeled. The number of models created depends on the chosen percent jump. The formula for this number is:

$$\left(\frac{100}{\text{percent jump}} + 1\right)\left(\frac{100}{\text{percent jump}} + 2\right)\frac{1}{2}. \tag{37}$$

This calculation reflects the increasing number of helium and iron combinations as the proton fraction decreases. The model creation process continues until the proton fraction reaches 0%. Each model represents a single data point in the final plot. We then analyze the average X_{max} and average L of showers for each model. The results are visualized in Figure 51, with a percent jump chosen at 1%, where data points are color-coded based on the proton fraction in the corresponding model. Again the triangular-shaped distribution from figure 48 appears, now filled in with points representing mixed-composition models.



(a) QGSJET-II (b) EPOS (c) SIBYLL

Figure 51: Each dot represents a set of 1000 showers of a random mix of elements (proton, helium and iron). The color indicates the proton fraction. The dot is plotted at the position corresponding to the average X_{max} of the sample and average L for the different hadronic interaction models.

Using the fraction of stretched showers will cause the differences between the different primary particles to increase compared to using average L. This makes the fraction of stretched showers a more robust observable that can tolerate higher systematic uncertainties compared to the average L value. Important to note is that the definition of when a shower is a stretched shower does not depend on the hadronic interaction model in this plot, the linear fit uses the average values in table 7. From an observational point of view needs to choose a single definition, and then see how the different models behave under that definition. The results of this change are shown in figure 52. The first thing to notice is that like before now SIBYLL is the hadronic interaction model that stands out the most. This is expected because the fraction of stretched showers for the EPOS model was consistently higher than the other hadronic interaction models.

The shift due to double bump showers is shown in 53 for two different energy ranges ($10^{16} - 10^{17}$ eV and $10^{17} - 10^{18}$ eV). The dotted line is before filtering the double bump showers out and the other line is after filtering the double bumps out. As expected the shift in fraction of stretched showers will have less of an impact compared to the shift in figure 50.

While stretched showers are significantly more frequent, particularly at higher energies, than double-bump showers, they are observationally more challenging to identify. However, preliminary studies show promise for the reconstruction of L [52]. One study demonstrated that the

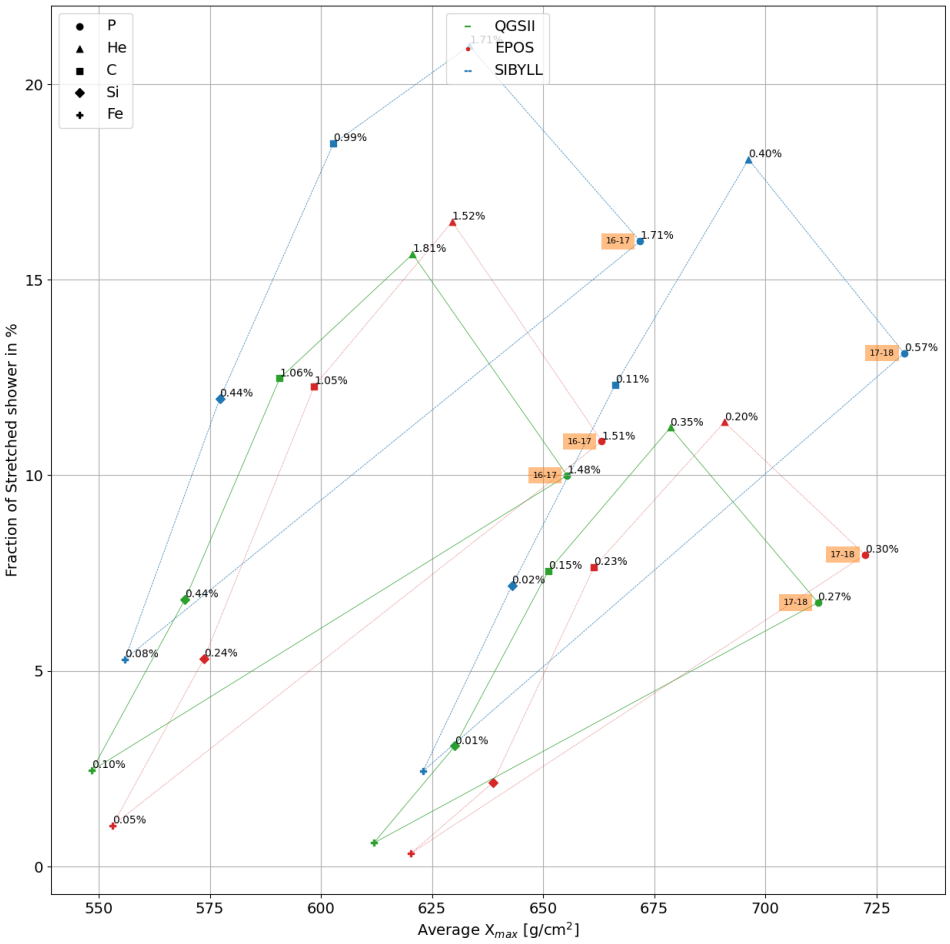
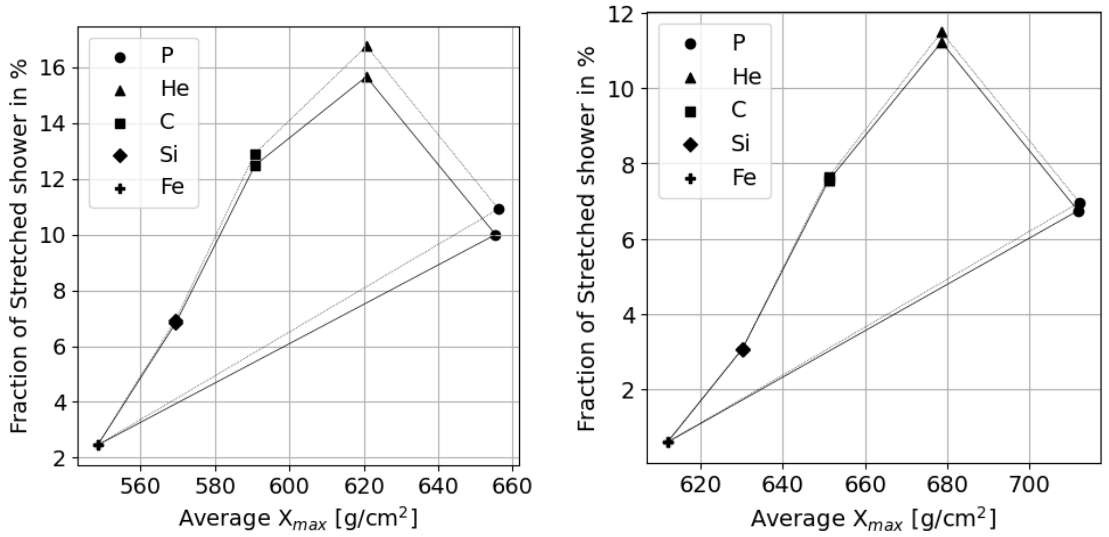


Figure 52: Fraction of stretched showers (definition used is average values in table 7) versus average X_{max} for different energies and hadronic interaction models. In each case, the elements form a triangle with protons on one corner and all other elements on the opposite side. Next to each point is the percentile of double bumps detected for that given value.

LOFAR radio telescope exhibits sensitivity to L, as evidenced by its ability to reconstruct X_{max} based on the observed shower profile. However, a degeneracy exists between L and X_{max} reconstruction using LOFAR. Small shifts in X_{max} can lead to similar reduced χ^2 values, making it difficult to simultaneously reconstruct both parameters with high accuracy. Looking towards the future, the possibility of L reconstruction with the Square Kilometre Array (SKA). A promising method involves using a linear combination of R and L, while fixing X_{max}. This approach has been successfully applied in simulations, yielding a reconstructed L value $L = 203.8$ g/cm^2 that closely matches the Monte Carlo truth value ($L = 204.6$ g/cm^2). In principle, this technique could be further developed to enable the simultaneous reconstruction of both L and X_{max}.

To test the improvement of using the fraction of stretched showers compared to average L on the impact of reconstructing the mass composition analysis, we again employ the same random astrophysical models containing varying mixtures of three elements: proton, helium, and iron. For each model, 1000 simulated cosmic ray showers are generated. We then analyze

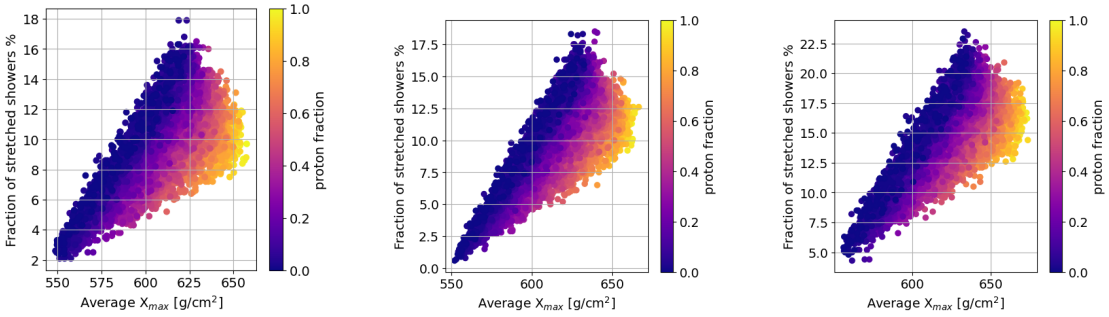


(a) Energy range: $10^{16} - 10^{17}$ eV.

(b) Energy range: $10^{17} - 10^{18}$ eV.

Figure 53: Fraction of stretched showers versus average X_{max} in energy ranges for the QGSJET-II hadronic interaction models. Showing the shift in the triangle caused by the filtering of double bump showers. The dotted line is before filtering and the other is after.

average X_{max} and the fraction of stretched showers for each model. The results are visualized in Figure 54, where data points are color-coded based on the proton fraction in the corresponding model. Again the triangular-shaped distribution from figure 52 appears. Models with 0% protons consistently fall on the left side of this triangle, while the specific position within the curve is determined by the unique mixing ratios of helium and iron. As the proton fraction in the model increases, the entire curve shifts towards the right, resulting in a clear color gradient in the figure. This demonstrates the direct influence of the mass composition on both X_{max} and the fraction of stretched showers.



(a) QGSJET-II

(b) EPOS

(c) SIBYLL

Figure 54: Each dot represents a set of 1000 showers between $10^{16} - 10^{17}$ eV of a random mix of elements (proton, helium and iron). The color indicates the proton fraction. The dot is plotted at the position corresponding to the average X_{max} of the sample and the percentage of stretched showers for the different hadronic interaction models.

6 Conclusion

Cosmic rays, pose a fascinating scientific challenge. Their origin and composition remain unknown, the uncertainties are large in the transition region between 10^{16} and 10^{18} eV. These energetic particles interact with Earth's atmosphere, triggering a cascade of secondary particles known as an extensive air shower. This thesis proposes novel techniques and refinements of existing methods to unlock the secrets of this enigmatic region. By leveraging the unprecedented sensitivity and resolution of the upcoming Square Kilometre Array (SKA) radio telescope, this work aims to develop a robust analysis method for identifying and characterizing two rare shower types: double-bump and stretched showers. Double-bump showers deviate significantly from the standard "universal" profile by exhibiting two distinct peaks in particle density during their development. These anomalous showers hold valuable clues about the nature of the primary cosmic ray particle and its interaction with the atmosphere. They can also shed light on the hadronic interaction cross-sections, which govern the behavior of these high-energy collisions. Stretched showers, on the other hand, represent an intermediate stage between the universal profile and the distinct double-bump signature.

This thesis has yielded notable advancements in our understanding of anomalous shower profiles, specifically double bump showers and stretched showers.

1. We established the AIC method as a superior technique for identifying double bump showers compared to the old method. This approach eliminates arbitrary limitations and reveals a larger population of double bump showers (almost twice as many), including those previously excluded by these arbitrary limitations and a higher number within the previously searched region.

2. We compared the prevalence of these double bump showers across different cosmic ray primary particle masses and various hadronic interaction models used in the simulations.
By applying the AIC method, we were able to compare the frequency of double bump showers across different hadronic interaction models (QGSJET-II, EPOS, and SIBYLL). This comparison provides valuable constraints for refining these models. EPOS generates the lowest amount of double bump showers (0.53%) and SIBYLL generates the largest amount (0.63%) . Anomalous shower profiles are observed with similar frequency for helium and proton primaries across different energy ranges. The overall fraction of these profiles decreases with increasing energy. Nearly all simulations predict a negligible number of anomalous showers for iron primaries.

3. We were also able to find some limits for when a single bump fit is likely to be a double bump shower, $R < 0.1$ or $R > 0.5$ and $L > 250$.

Moreover we identified two distinct types of double bump showers: early double bumps (smaller sub-shower preceding the larger one) and late double bumps (larger sub-shower preceding the smaller one). The amount of early bumps found depend on the chosen hadronic interaction model, with SIBYLL exhibiting a much higher prevalence of early double bumps (52%). Notably, the frequency of double bumps for protons and helium is comparable, but early double bumps are more than twice as frequent for protons compared to helium. The successful detection of early double bump showers remains an open question. If this turns out to be the case that they are not detectable, it will be crucial to factor in this potential bias when analyzing SKA data to estimate the frequency of double bump showers.

4. We developed a novel approach for identifying stretched showers, which bridge the gap between single and double bump showers. The fraction of stretched showers serves as an indica-

tor of the mass composition of the primary cosmic ray. Interestingly, the frequency of stretched showers for protons is even lower than that for carbon which is lower than that of helium, offering a means to differentiate between proton and helium primary particles. The observed linear correlation between the fractions of stretched and double bump showers suggests a connection between these shower types, possibly explained by the leading particle concept.

5. We demonstrated the importance of incorporating double bump showers into the triangle plot, which are used to distinguish between hadronic models and infer mass composition. Replacing the average L parameter with the fraction of stretched showers in this plot leads to more pronounced differences between models, particularly highlighting the SIBYLL hadronic interaction model. Enhanced p/He separation holds promise for pinpointing the onset of the extragalactic component within the cosmic-ray flux. Additionally, it could aid in the identification of a potential helium-rich secondary galactic component arising from supernova remnants of Wolf-Rayet stars [45].

7 Outlook

The study of cosmic ray showers and their intriguing features such as stretched and double bump showers is far from over. Several promising avenues for future research could be explored:

A comprehensive understanding of the secondary particles generated in the first few interactions within cosmic ray showers is crucial for elucidating the mechanisms behind double bump and stretched showers. Future investigations could focus on:

Quantifying the contribution of specific secondary particle types. This would involve identifying the dominant particle types contributing to each shower feature and analyzing their individual roles. Studying how the energy spectrum of secondary particles influences the development of double bumps and stretched showers could provide valuable insights. Imagine successfully observing and characterizing double bump showers with SKA. Key measurements include the ΔX_{max} and N_{max_1}/N_{max_2}. These measurements hold valuable information about the properties of the hadronic cross sections of the particles from the first interaction: ΔX_{max} reflects the distance traveled by the particle that initiates the second sub-shower. This, can be used to estimate the particle's energy before interacting with the atmosphere. While a single shower cannot definitively identify the initiating particle, simulations can help us understand the expected N_{max_1}/N_{max_2} ratio's for different primary particles (e.g., protons vs. pions) responsible for the second sub-shower. By comparing the measured ratio to these simulated values, we can infer the most likely type of particle that caused the second bump.

Current definitions of stretched showers often rely on specific hadronic interaction models. Future research could explore alternative definitions that are independent of the chosen model and directly linked to the actual distance traveled by secondary particles within the shower. A robust definition based on the actual distance traveled by secondary particles would allow for more consistent and reliable identification of stretched showers across different models and experimental datasets. Connecting stretched showers directly to the travel distance of secondary particles could offer direct insights into the underlying mechanisms responsible for their formation.

Radio detection offers a unique perspective on cosmic ray showers, capturing the electromagnetic radiation emitted by charged particles during their interactions. Future research could investigate the potential of radio footprints to directly identify and classify double bump showers. Analyzing the specific radio emission patterns associated with double bump showers could lead to the development of methods for their direct detection and characterization. This would allow us to test if SKA will be able to measure these early double bump showers and if not, to make appropriate calibrations to the expected measurements.

A Appendix

A.1 Normal and reduced Gaisser Hillas plots of single and double bump showers

Comparison of normal and reduced GH fits for identifying double bumps in simulated different primaries. The analysis investigates 20.000 showers simulated with the different hadronic interaction models.

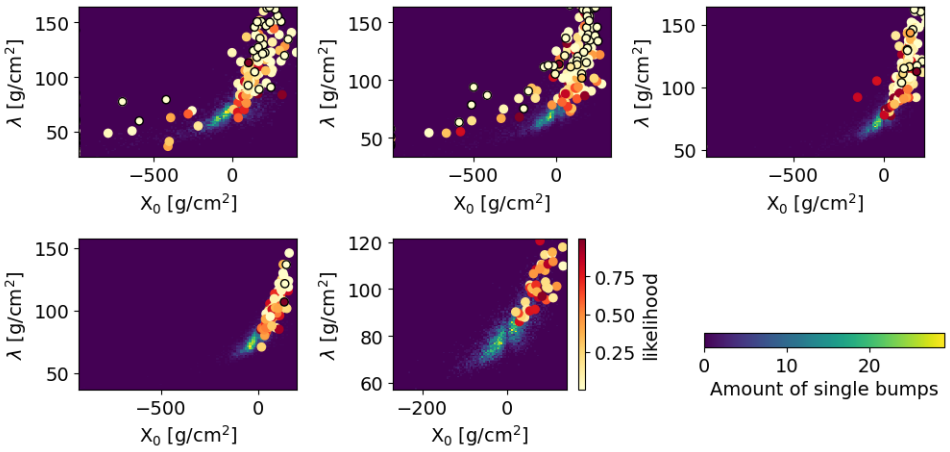


Figure 55: Showers generated with QGSJET-II, with proton, helium, carbon on the first row in that order and on the second row silicium and iron in that order.

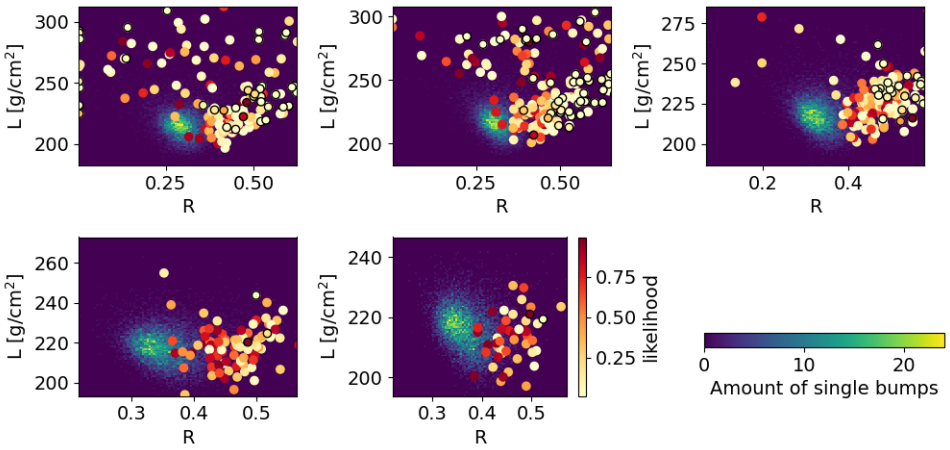


Figure 56: Showers generated with QGSJET-II, with proton, helium, carbon on the first row in that order and on the second row silicium and iron in that order.

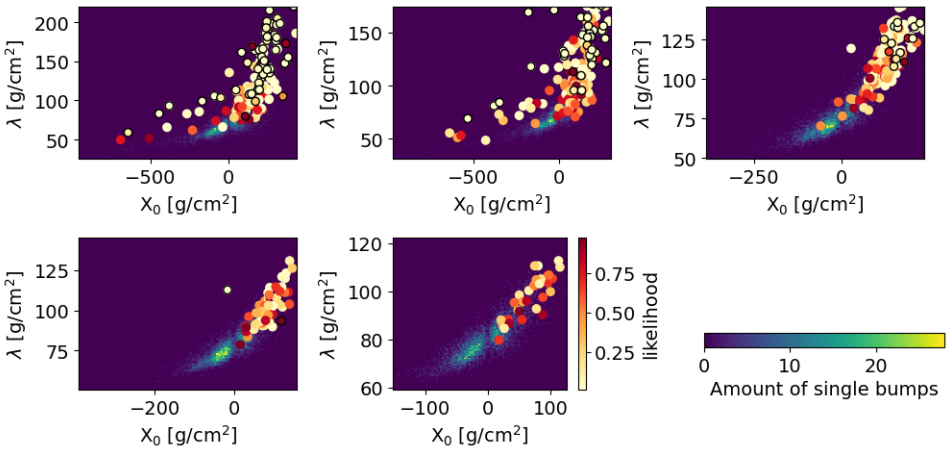


Figure 57: Showers generated with EPOS, with proton, helium, carbon on the first row in that order and on the second row silicium and iron in that order.

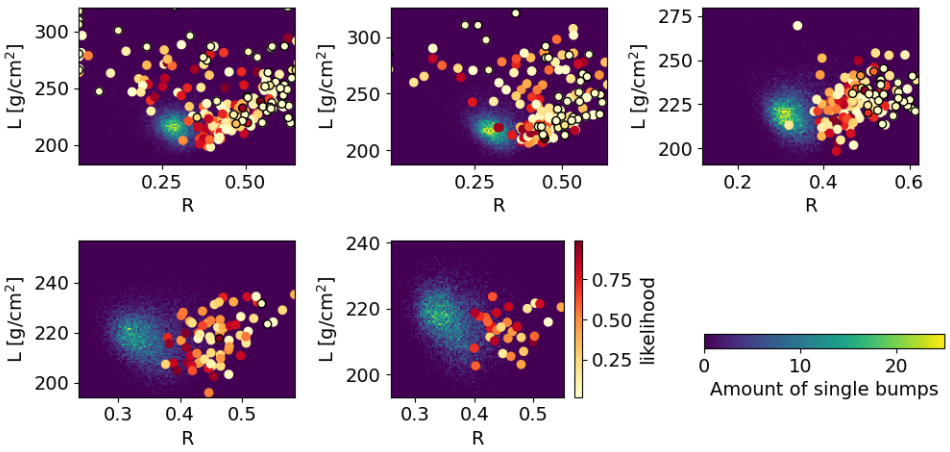


Figure 58: Showers generated with EPOS, with proton, helium, carbon on the first row in that order and on the second row silicium and iron in that order.

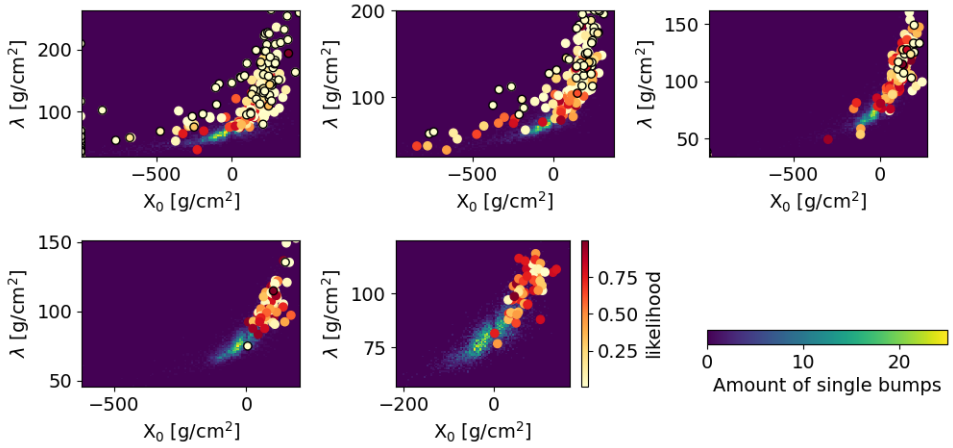


Figure 59: Showers generated with SIBYLL, with proton, helium, carbon on the first row in that order and on the second row silicium and iron in that order.

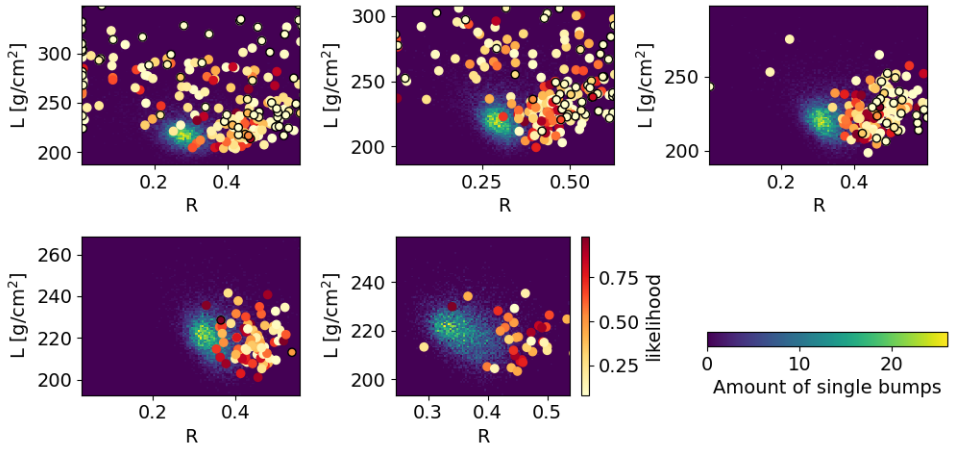


Figure 60: Showers generated with SIBYLL, with proton, helium, carbon on the first row in that order and on the second row silicium and iron in that order.

A.2 Distributions of energy and distance between double bumps

These figure compares the distributions of primary energy E_0 and distance between the double bumps ΔX for different shower classifications based on the old method and the AIC method for 20.000 different primaries showers simulated with different hydronic model.

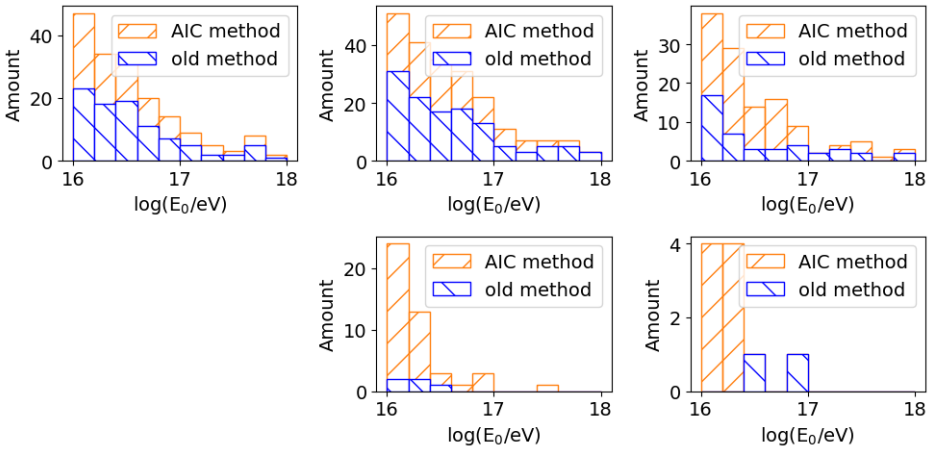


Figure 61: Showers generated with QGSJET-II, with proton, helium, carbon on the first row in that order and on the second row silicium and iron in that order.

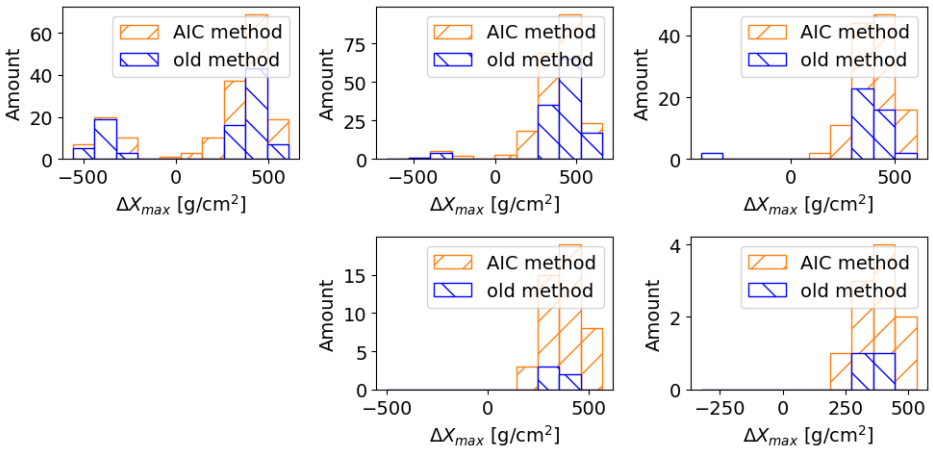


Figure 62: Showers generated with QGSJET-II, with proton, helium, carbon on the first row in that order and on the second row silicium and iron in that order.

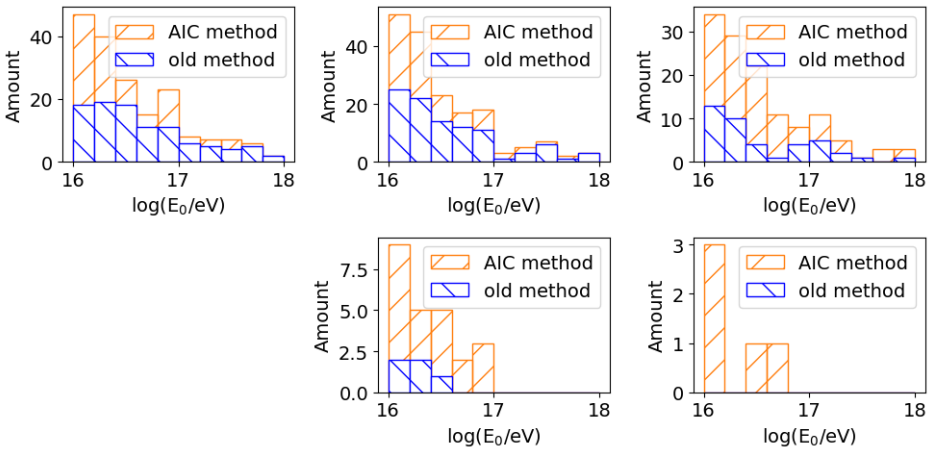


Figure 63: Showers generated with EPOS, with proton, helium, carbon on the first row in that order and on the second row silicium and iron in that order.

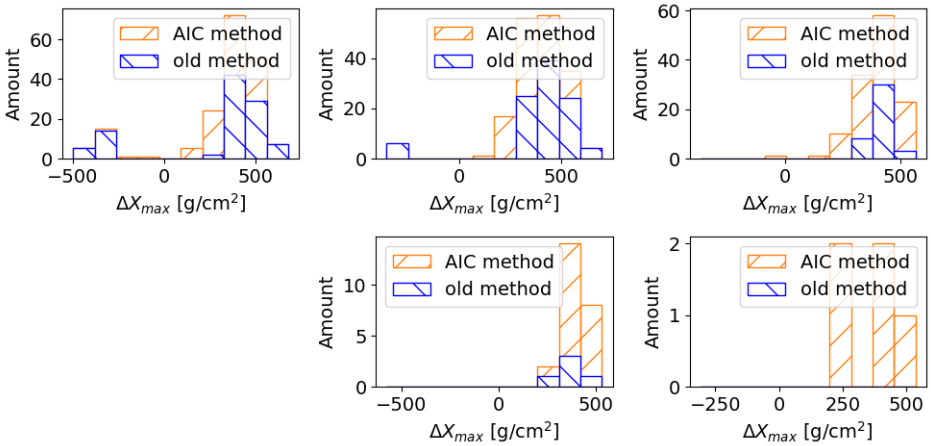


Figure 64: Showers generated with EPOS, with proton, helium, carbon on the first row in that order and on the second row silicium and iron in that order.

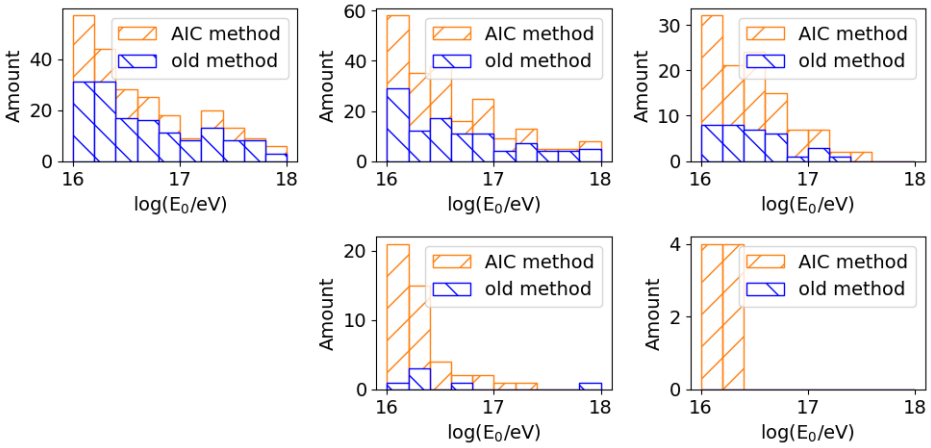


Figure 65: Showers generated with SIBYLL, with proton, helium, carbon on the first row in that order and on the second row silicium and iron in that order.

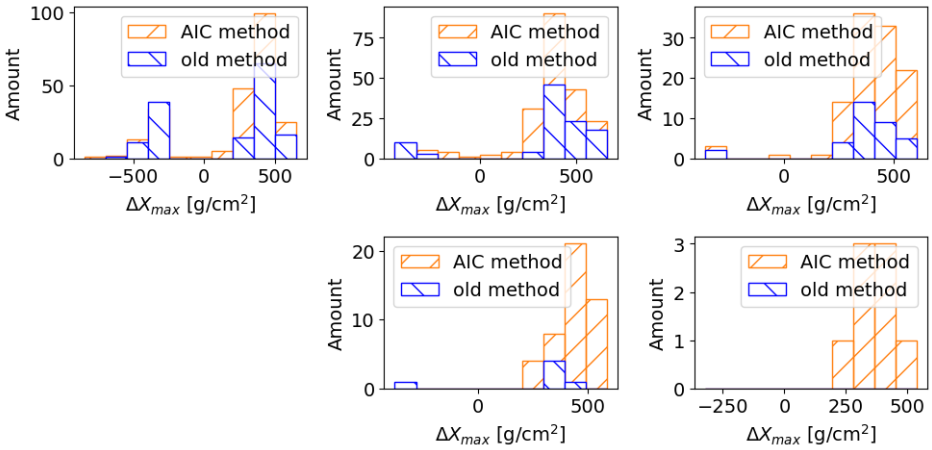


Figure 66: Showers generated with SIBYLL, with proton, helium, carbon on the first row in that order and on the second row silicium and iron in that order.

References

[1] D. Heck, J. Knapp, J.N. Capdevielle, G. Schatz and T. Thouw, *CORSIKA: a Monte Carlo code to simulate extensive air showers.* (1998).

[2] T. Huege, M. Ludwig and C.W. James, *Simulating radio emission from air showers with coreas*, in *AIP Conference Proceedings*, AIP, 2013, DOI.

[3] A. Corstanje, S. Buitink, M. Desmet, H. Falcke, B. Hare, J. Hörandel et al., *A high-precision interpolation method for pulsed radio signals from cosmic-ray air showers*, *Journal of Instrumentation* **18** (2023) P09005.

[4] C. Glaser, A. Nelles, I. Plaisier, C. Welling, S.W. Barwick, D. García-Fernández et al., *Nuradioreco: a reconstruction framework for radio neutrino detectors*, *The European Physical Journal C* **79** (2019) .

[5] V.F. Hess, *Über Beobachtungen der durchdringenden Strahlung bei sieben Freiballonfahrten*, *Phys. Z.* **13** (1912) 1084.

[6] D. Pacini, *Penetrating radiation at the surface of and in water*, 2017.

[7] "On Its Centenary, Celebrating a Ride That Advanced Physics the new york times." https://www.nytimes.com/2012/08/07/science/space/when-victor-hess-discovered-cosmic-rays-in-a-hydrogen-balloon.html.

[8] P.M. Schuster, *The scientific life of victor franz (francis) hess (june 24, 1883–december 17, 1964)*, *Astroparticle Physics* **53** (2014) 33.

[9] R.A. Millikan and G.H. Cameron, *The origin of the cosmic rays*, *Phys. Rev.* **32** (1928) 533.

[10] D.J. Bird, S.C. Corbato, H.Y. Dai, J.W. Elbert, K.D. Green, M.A. Huang et al., *Detection of a cosmic ray with measured energy well beyond the expected spectral cutoff due to cosmic microwave radiation*, *The Astrophysical Journal* **441** (1995) 144.

[11] J. Blümer, R. Engel and J.R. Hörandel, *Cosmic rays from the knee to the highest energies*, *Progress in Particle and Nuclear Physics* **63** (2009) 293–338.

[12] T.K. Gaisser, R. Engel and E. Resconi, *Cosmic Rays and Particle Physics*, Cambridge University Press, 2 ed. (2016).

[13] W.I. Axford, *The Origins of High-Energy Cosmic Rays*, **90** (1994) 937.

[14] R. Battiston, P. Blasi, M.T. Brunetti, D. De Marco, P. Lipari, B. Sacco et al., *Cosmic ray physics from low to extreme energies: Status and perspectives*, *Adv. Space Res.* **37** (2006) 1834.

[15] "High-energy cosmic ray sources get mapped out for the first time." https://www.wired.com/story/high-energy-cosmic-ray-sources-get-mapped-out-for-the-first-time/.

[16] B. Klecker, *Anomalous cosmic rays: Our present understanding and open questions*, *Advances in Space Research* **23** (1999) 521.

[17] V. Ginzburg and Syrovattskiĭ, *The Origin of Cosmic Rays*, Errata slip inserted, Pergamon Press; [distributed in the Western Hemisphere by Macmillan, New York] (1964).

[18] R. Diesing, *The maximum energy of shock-accelerated cosmic rays*, *The Astrophysical Journal* **958** (2023) 3.

[19] J.R. Hörandel, *Models of the knee in the energy spectrum of cosmic rays*, *Astroparticle Physics* **21** (2004) 241–265.

[20] W.D. Apel, J.C. Arteaga-Velázquez, K. Bekk, M. Bertaina, J. Blümer, H. Bozdog et al., *Kneelike structure in the spectrum of the heavy component of cosmic rays observed with kascade-grande*, *Physical Review Letters* **107** (2011) .

[21] J.R. Hörandel, N.N. Kalmykov and A.V. Timokhin, *Propagation of super-high-energy cosmic rays in the galaxy*, *Astroparticle Physics* **27** (2007) 119–126.

[22] A. Letessier-Selvon and T. Stanev, *Ultrahigh energy cosmic rays*, *Reviews of Modern Physics* **83** (2011) 907–942.

[23] K. Kotera and A.V. Olinto, *The astrophysics of ultrahigh-energy cosmic rays*, *Annual Review of Astronomy and Astrophysics* **49** (2011) 119–153.

[24] K. Greisen, *End to the cosmic ray spectrum?*, *Phys. Rev. Lett.* **16** (1966) 748.

[25] G.T. Zatsepin and V.A. Kuzmin, *Upper limit of the spectrum of cosmic rays*, *JETP Lett.* **4** (1966) 78.

[26] R. Abbasi, T. Abu-Zayyad, M. Al-Seady, M. Allen, J. Amann, G. Archbold et al., *Measurement of the flux of ultra high energy cosmic rays by the stereo technique*, *Astroparticle Physics* **32** (2009) 53–60.

[27] J. Abraham, P. Abreu, M. Aglietta, C. Aguirre, D. Allard, I. Allekotte et al., *Observation of the suppression of the flux of cosmic rays above* $4x10^{19}$ *ev*, *Physical Review Letters* **101** (2008) .

[28] T. Abu-Zayyad, R. Aida, M. Allen, R. Anderson, R. Azuma, E. Barcikowski et al., *The cosmic-ray energy spectrum observed with the surface detector of the telescope array experiment*, *The Astrophysical Journal* **768** (2013) L1.

[29] A.V. Olinto, *Cosmic rays at the highest energies*, in *Journal of Physics Conference Series*, vol. 375 of *Journal of Physics Conference Series*, p. 052001, July, 2012, DOI [1201.4519].

[30] R.M. Ulrich, *Measurement of the proton-air cross section using hybrid data of the Pierre Auger Observatory*, Ph.D. thesis, 2008. 10.5445/IR/200070733.

[31] W. Heitler, *The Quantum Theory of Radiation*, Dover Books on Physics Series, Dover Publications (1984).

[32] J. Matthews, *A Heitler model of extensive air showers*, *Astropart. Phys.* **22** (2005) 387.

[33] D. Kostunin, *Reconstruction of air-shower parameters with a sparse radio array*, Ph.D. thesis, 01, 2015. 10.5445/IR/1000053435.

[34] T. Huege, *Radio detection of cosmic ray air showers in the digital era*, *Physics Reports* **620** (2016) 1–52.

[35] T.K. Gaisser and A.M. Hillas, *Reliability of the Method of Constant Intensity Cuts for Reconstructing the Average Development of Vertical Showers*, in *International Cosmic Ray Conference*, vol. 8 of *International Cosmic Ray Conference*, p. 353, Jan., 1977.

[36] S. Andringa, R. Conceição and M. Pimenta, *Mass composition and cross-section from the shape of cosmic ray shower longitudinal profiles*, *Astroparticle Physics* **34** (2011) 360.

[37] A. Aab, P. Abreu, M. Aglietta, I. Albuquerque, J. Albury, I. Allekotte et al., *Measurement of the average shape of longitudinal profiles of cosmic-ray air showers at the pierre auger observatory*, *Journal of Cosmology and Astroparticle Physics* **2019** (2019) 018.

[38] L.E. Miller, *"molecular weight" of air at high altitudes*, *Journal of Geophysical Research (1896-1977)* **62** (1957) 351 [https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/JZ062i003p00351].

[39] C. Baus, R. Engel, T. Pierog, R. Ulrich and M. Unger, *Anomalous Longitudinal Shower Profiles and Hadronic Interactions*, in *32nd International Cosmic Ray Conference*, vol. 2, pp. 206–209, 11, 2011, DOI [1111.0504].

[40] J. Garzon, J. Collazo, J. Cuenca-García, D. Castro, J. Otero, M. Yermo et al., *Tragaldabas. first results on cosmic ray studies and their relation with the solar activity, the earth magnetic field and the atmospheric properties*, .

[41] P. Auger, P. Ehrenfest, R. Maze, J. Daudin and R.A. Fréon, *Extensive cosmic-ray showers*, *Rev. Mod. Phys.* **11** (1939) 288.

[42] W. Apel, J. Arteaga-Velázquez, K. Bekk, M. Bertaina, J. Blümer, H. Bozdog et al., *The kascade-grande energy spectrum of cosmic rays and the role of hadronic interaction models*, *Advances in Space Research* **53** (2014) 1456.

[43] "Construction of square kilometre array deep space telescope to begin next year." https://www.midwesttimes.com.au/news/midwest-times/construction-of-square-kilometre-array-deep-space-telescope-to-begin-next-year-ng-b881573623z.

[44] T. Huege, J.D. Bray, S. Buitink, D. Butler, R. Dallier, R.D. Ekers et al., *Ultimate precision in cosmic-ray radio detection — the ska*, *EPJ Web of Conferences* **135** (2017) 02003.

[45] S. Thoudam, J.P. Rachen, A. van Vliet, A. Achterberg, S. Buitink, H. Falcke et al., *Cosmic-ray energy spectrum and composition up to the ankle: the case for a second galactic component*, *Astronomy amp; Astrophysics* **595** (2016) A33.

[46] A. Corstanje, S. Buitink, J. Bhavani, M. Desmet, H. Falcke, B. Hare et al., *Prospects for measuring the longitudinal particle distribution of cosmic-ray air showers with ska*, in *9th International Workshop on Acoustic and Radio EeV Neutrino Detection Activities,*, Proceedings of Science, Proceedings of Science, Oct., 2023, DOI.

[47] W.R. Nelson, H. Hirayama and D.W. Rogers, *Egs4 code system*, .

[48] S. Ostapchenko, *Qgsjet-ii: towards reliable description of very high energy hadronic interactions*, *Nuclear Physics B - Proceedings Supplements* **151** (2006) 143–146.

[49] R.S. Fletcher, T.K. Gaisser, P. Lipari and T. Stanev, *sibyll: An event generator for simulation of high energy cosmic ray cascades*, *Phys. Rev. D* **50** (1994) 5710.

[50] K. Werner, *The hadronic interaction model epos*, *Nuclear Physics B - Proceedings Supplements* **175-176** (2008) 81.

[51] P. Stoica and Y. Selen, *Model-order selection: a review of information criterion rules*, *IEEE Signal Processing Magazine* **21** (2004) 36.

[52] A. Corstanje, S. Buitink, J. Bhavani, M. Desmet, H. Falcke, B.M. Hare et al., *Prospects for measuring the longitudinal particle distribution of cosmic-ray air showers with SKA*, *arXiv e-prints* (2023) arXiv:2303.09249 [2303.09249].