Mass Composition of Ultra-High Energy Cosmic Rays:

A Study Based on Elongation Rate of Depth of Shower Maximum Measured at the Pierre Auger Observatory

DISSERTATION

zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften

vorgelegt von
Livingstone O. Ochilo, M. Sc.

eingereicht bei der
Naturwissenschaftlich-Technischen Fakultät
der Universität Siegen

Mai 2016

Gutachter der Dissertation: Prof. Dr. Markus Risse
Prof. Dr. Claus Grupen
Prof. Dr. Ivor Fleck
Prof. Dr. Otfried Gühne

Datum der Disputation: 17. Juni 2016
Internetpublikation der Universitätsbibliothek Siegen:
urn:nbn:de:hbz:467-10184

Abstract

An analysis of the mass composition of UHECRs based on the $X_{\max}$ elongation rate (ER) within the energy range $10^{17.8} - 10^{20.0}$ eV is performed, using $X_{\max}$ data collected by the Pierre Auger Observatory from December 2004 to December 2012. The development of ER with energy is investigated by estimating the ER and the rate of change of ER with energy in sliding windows. In order to make it easier to observe the effect of a change in the proton fraction on the overall primary cosmic ray composition, the data has been truncated into 'light' and 'heavy' components based on a cut on the $X_{\max}$ value above which 50% of helium showers produced with EPOS-LHC survive. The 'light' component should be dominated by proton and helium nuclei. The behavior of the ER of the 'light' part potentially allows one to distinguish changes in the mass composition from changes in the properties of hadronic interactions. The rate of change of ER with energy is observed to be negative and differs from zero for $\lg(E/\text{eV}) \approx 18.2 - 18.5$ by $\sim 3\sigma$: primary mass is increasing and the rate of change of primary mass might not be constant. The increase in primary mass might be mostly due to a reduction of the fraction of protons. The 'heavy' subset seems to be consistent with a constant primary mass. An interpretation of the data as being due to a pure proton composition with an unexpectedly large increase in interaction cross-section might be in tension with the behavior of ER of the 'light' component.

Zusammenfassung

Es wird eine Analyse der Massenzusammensetzung der ultrahochenergetischen kosmischen Strahlung vorgestellt, die auf der X_{max}-Elongationsrate (ER) im Energiebereich $10^{17.8} - 10^{20.0}$ eV basiert und Daten des Pierre-Auger-Observatoriums verwendet, die zwischen Dezember 2004 und Dezember 2012 aufgenommen wurden. Die Entwicklung der ER mit der Energie wird untersucht, indem die ER und die Änderungsrate der ER mit der Energie in kleinen Energiebins bestimmt wird. Um die Beobachtung des Effekts einer Änderung im Anteil von Protonen an der gesamten primären Massenzusammensetzung zu erleichtern, wurden die Daten in eine "leichte" und eine "schwere" Komponente eingeteilt mit Hilfe eines Schnitts auf X_{max}, der so definiert ist, dass 50% der mit EPOS-LHC simulierten Helium-Luftschauer überleben. Die "leichte" Komponente sollte durch Protonen und Helium-Kerne dominiert sein. Die Untersuchung der ER der "leichten" Komponente ermöglicht es, Änderungen der Massenzusammensetzung und Änderungen in den Eigenschaften der hadronischen Wechselwirkungen zu unterscheiden. Die berechnete Änderungsrate der ER mit der Energie ist negativ und verschieden von Null für $\lg(E/\mathrm{eV}) = 18.2 - 18.5$ mit einer Signifikanz von $\sim 3\sigma$: Die primäre Masse steigt und es gibt Hinweise auf eine nicht konstante Änderungsrate der primären Masse. Der Anstieg der primären Masse kann vorwiegend aus einer Verringerung des Protonanteils resultieren. Die "schwere" Komponente scheint verträglich mit einer konstanten primären Masse zu sein. Eine Interpretation der Daten als reine Proton-Zusammensetzung mit einem unerwartet starken Anstieg des Wechselwirkungsquerschnitts scheint im Widerspruch zu stehen zum beobachtenen Verhalten der "leichten" Komponente.

Dedication

To my wife Sarah and children Marvel, Joe and Kevin

Acknowledgements

I gratefully acknowledge the scholarship grant from the German Academic Exchange Service (DAAD), which funded this study. I would like to thank Prof. Markus Risse for allowing me to join his group and for accepting to supervise my work. His advice and support have been very invaluable for the success of this study. I would also like to thank Dr. Alexey Yushkov not only for his invaluable advice in the course of this work, but also for being a source of inspiration and encouragement. Much appreciation to Dr. Mariangela Settimo for inducting me into the Auger community. I would like to thank the Pierre Auger Collaboration for allowing me to be a part of the team handling its data and for the invaluable help obtained from various members of the collaboration which enabled this study to succeed. In particular, I acknowledge the help and support of the members of the Auger group in the University of Siegen.

Contents

1 Introduction

The discovery of cosmic rays can be traced back to 1912, following experiments conducted by Victor Hess. His observation that the rate of discharge of an electroscope increased with altitude led him to conclude that the "rays" originated from a source outside the atmosphere of the Earth. Indeed, cosmic rays are now known to be charged nuclei that originate from outside the solar system. As the study of these rays progressed, a major development occurred when a cosmic ray event with energy exceeding 10^{20} eV was detected at the Volcano Ranch experiment in 1962. The term Ultra High Energy Cosmic Rays (UHECR) has come into use to refer to cosmic rays with energy exceeding 10^{18} eV.

Since the first detection of an UHECR event, it has remained an interest of physicists to understand the nature of such very energetic cosmic rays. Among the questions which have not been definitively answered concerning UHECRs are their sources, the acceleration mechanisms that lead to such enormous energies, and their chemical composition.

A major challenge in collecting UHECR data is the low flux involved: at an energy of 10^{20} eV, for example, the expected flux is less than one particle per square kilometer per century. As a result of the very low flux of UHECRs, it is not possible to detect them directly by the use of techniques such as satellite-borne experiments. It is in this context that the Pierre Auger Observatory has been designed and installed to collect data over a large area of approximately 3000 km^2. The objective of the Observatory is to measure the energy spectrum, arrival directions and nuclear identities of the UHECRs with high statistics [1]. The Auger Observatory has been collecting reliable data since 2004, although its construction was fully completed in 2008.

The only method of studying UHECRs is by reconstructing their properties from the extensive air showers initiated when primary particles interact with atmospheric nuclei. In this procedure, one can infer the mass of the primary particle by comparing experimental observables with detailed simulations of air showers [2]. However, an uncertainty is introduced due to the fact that the characteristics of hadronic interactions assumed in interaction models are extrapolated from collider experiments, which are orders of magnitude below the energy of UHECRs. Hadronic interaction models make different physical assumptions in performing the extrapolations. Un-

certainties in hadronic interaction models get propagated during the determination of composition, and in addition, there is a shower-to-shower fluctuation in the primary composition. Hence it is not possible to identify the composition of a shower event by event.

One of the extensive air shower observables which is commonly used in the study of the composition of primary cosmic rays is the depth of shower maximum, X_{max}. This refers to the slant depth at which the energy deposited by a shower per unit mass of atmosphere traversed reaches its maximum, such that the number of particles in the shower also reaches a maximum. From the evolution of X_{max} with the logarithm of energy, the slope, known as elongation rate (ER), can be obtained. In this thesis, a description of the use of ER analysis to study the mass composition of cosmic rays in the energy range $10^{17.8}$ eV to 10^{20} eV as collected by the Pierre Auger Observatory in the period December 2004 to December 2012 is presented. A study based on the ER is more robust than that based on the mean of the distributions of X_{max}, represented as $\langle X_{\mathrm{max}} \rangle$.

Although previous analyses of the Pierre Auger X_{max} data which were based on a subset of the current data obtained ERs from good fits to $\langle X_{\mathrm{max}} \rangle$ values, the abrupt change in ER reported might be non-physical. This is because a look at the measured cosmic ray flux shows that it is a smooth function of energy. It is therefore natural to expect that the ER should also change gradually with energy. In this work, the possible gradual change of ER is investigated by calculating its value over short overlapping intervals of energy (sliding windows) in the energy range covered by the Pierre Auger Observatory.

Similarly, previous analyses of Auger X_{max} data have consistently concluded that the CR primary composition gets lighter up to $\lg(E/\mathrm{eV}) = 18.2$, before beginning to get heavier after this point. However, it has not been confirmed whether the increase in the mean primary CR mass is caused by the disappearance of the lighter nuclei or an increase in the fraction of heavier nuclei. There is also the possibility that if the primary flux is dominated by protons, then the observed increase in primary mass may simply be due to an accelerated increase in the proton interaction cross-section. Given that in most astrophysical scenarios the lighter components of the primary cosmic ray flux are expected to disappear earlier as the energy increases, this thesis separates the Auger data into 'light' and 'heavy' subsets, so that the behaviour of ER in the 'light' subset can be scrutinized more easily. If indeed the increase in the primary mass observed in the full data set is mainly due to a reduction in the proton fraction, then the lighter part should display more lightening as compared to the combined data. In the case of an increasing proton interaction cross-section, a comparison of data with such a simulated scenario can paint a picture of whether this is likely or not.

This thesis is structured as follows. In chapter 2, ultra high energy cosmic rays are described in detail. This includes their sources and how they are propagated. Mention is made of a few astrophysical scenarios under which UHECRs can be produced. The determination of the mass composition of UHECR by the use of depth of shower maximum, as well as the current experimental results are addressed in chapter 3. A detailed description of the Pierre Auger Observatory is given in chapter 4, including the procedures for event reconstruction and the different components of

the Observatory. The procedure that was used in selecting the data that has been analyzed in this thesis is described in chapter 5. Since the data is for events detected simultaneously by the fluorescence and surface detectors, the procedures for both of them are given. The analysis of the data is presented in chapter 6 and finally the results are summarized in chapter 7.

2 Ultra-High Energy Cosmic Rays

The initial discovery of ionizing radiation from outside the atmosphere of the Earth which were subsequently named cosmic rays was followed by studies in which the measurable energy of the cosmic rays increased more and more. The result of studies of the effect of geomagnetic field on cosmic rays predicted that their spectrum should extend to at least 10 GeV. Later, in 1938, the order of magnitude of the energy increased by about 5, thanks to the discovery of air showers (see e.g. [3]). The current level of knowledge puts the extreme end of the spectrum at over 10^{20} eV.

The discovery of UHECRs was greatly aided by the studies already done earlier that laid the groundwork for the necessary techniques and principles. The discovery of extensive air showers towards the end of the 1930s and the coincidence technique developed by Bothe and Rossi [4] were later to prove useful in the design of UHECR detectors. The UHECR event with an energy of 1.0×10^{20} eV, recorded in 1962 by the Volcano Ranch experiment [5], remains one of the highest energy events ever. The energy was later revised to 1.4×10^{20} eV [6]. This amount of energy is remarkable because it exceeds the rest mass of a cosmic ray particle by far. It is a macroscopic energy, yet is observed in this case in a microscopic particle. Astrophysical sources that are capable of producing such large amounts of energy are still not known. The most energetic cosmic ray event so far reported had an energy of 3×10^{20} eV [7].

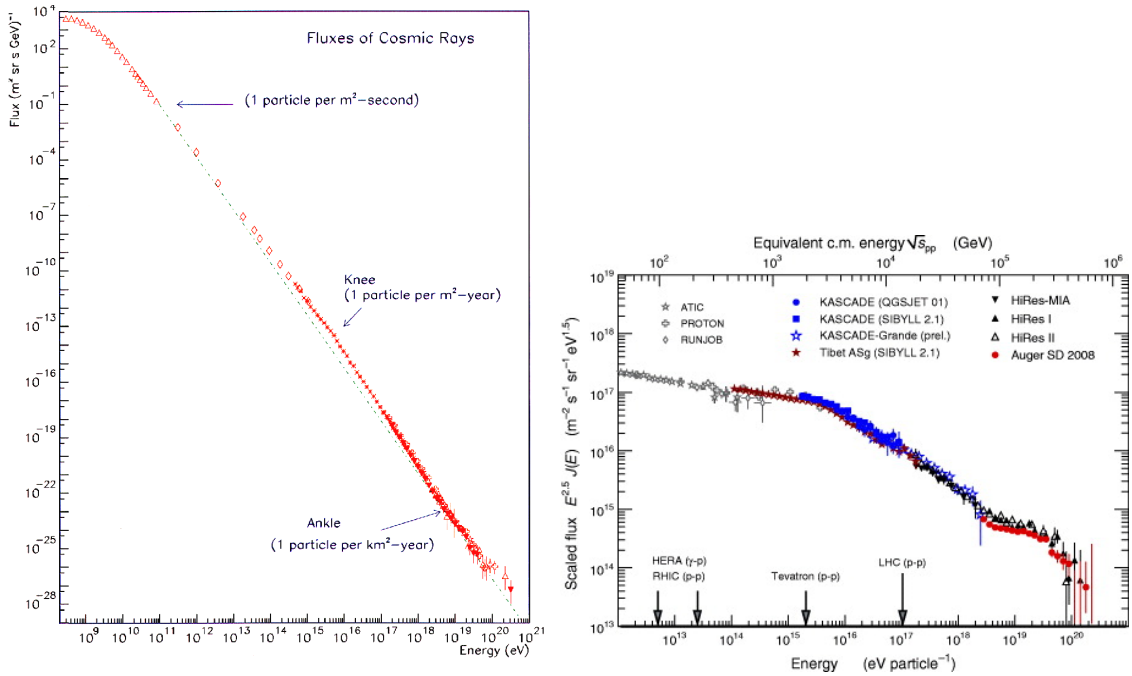
2.1 Properties of UHECR

2.1.1 The Cosmic Ray Spectrum

Operating within different ranges of energies, including the UHECR range, several experiments have measured the flux of cosmic rays i.e. the number of particles per (m^2 sr s eV). The resulting combined spectrum of cosmic rays as a function of energy is essentially a power law of the form

$$\frac{dN}{dE} \sim E^{-\gamma} \tag{2.1}$$

as represented in Fig. 2.1. The integrated flux in different parts of the spectrum are



(a) CR spectrum (b) Scaled CR spectrum

Figure 2.1: *(a)The spectrum of cosmic rays as recorded by different experiments. (b) The all-particle CR spectrum scaled by $E^{2.5}$ in order to reveal more clearly the knee and the ankle ([8] and [9] respectively).*

as indicated. The spectrum covers 11 decades of energy, beginning at $\sim 10^9$ eV, and $\sim$ 30 orders of magnitude of flux. The features in the spectrum are more clearly visible when the flux is scaled by a convenient factor such as in Fig. 2.1(b). The spectral index, γ, in Eqn. 2.1, steepens from 2.7 to 3.1 at around 10^{15} eV, a feature referred to as the "knee" and flattens to 2.6 at $\sim 10^{18.5}$ eV, a feature commonly known as the "ankle". A slight change of the spectral index at energy $\sim 10^{17.5}$ eV from 3.1 to 3.3 is called the second knee. These features of the spectrum can be related to transitions from cosmic rays arriving from one class of sources to those arriving from another. They could also be related to acceleration and propagation processes. These are discussed in more details in section 2.2.

At low energies, as a result of the large cosmic ray flux, detectors on board balloons or satellites can be used for the direct observation of cosmic rays. However, at energies in the UHECR range, the low expected flux ranging from 1 particle per (km^2 sr year) to 1 particle per (km^2 sr century) makes it impossible to observe the cosmic rays directly. The Pierre Auger Observatory was initially designed to study cosmic rays from energy $10^{18.0}$ eV and above i.e. UHECRs. Subsequent upgrades have lowered the minimum energy to $10^{17.0}$ eV, hence the whole region around the ankle and second knee is covered. The large surface area of 3000 km^2 covered by the surface detector of the observatory is necessitated by the low flux anticipated in this energy range.

2.1.2 Chemical Composition

At energies above 10^{14} eV, it is not possible to measure directly the abundance of individual elements in the cosmic ray spectrum. Instead, the mean mass of the cosmic rays at a given energy can be estimated by analyzing for example the mean atmospheric depth at which the air shower initiated by the primary cosmic rays reaches its maximum development, following the suggestion in [10]. This is discussed in more details in sections 3.1 and 3.2. The mass of the primary particle is inferred from detailed comparisons of experimentally observed $X_{\max}$ with the results obtained from air shower simulations. Due to uncertainties in hadronic interactions at the highest energies resulting from extrapolation of interaction parameters from data obtained at much lower center of mass energies at the LHC, systematic uncertainties are inevitably introduced.

A recent detailed review of the possible approaches that can be used in inferring mass composition at the highest cosmic ray energies is presented in [2]. Some of the approaches discussed therein are mentioned here briefly. In each case, an extensive air shower observable that is sensitive to the primary mass is selected and used to probe the primary mass.

The discovery of high energy photons and neutrinos that are produced via the process that was proposed by Greisen [11] and independently by Zatsepin and Kuz'min [12], known as the GZK effect (see Eqn. 2.2), would prove unambiguously that cosmic ray composition at ultra-high energies is light.

$$\begin{aligned} p + \gamma_{\mathrm{CMB}} &\rightarrow p + \pi^0 \\ \pi^0 &\rightarrow \gamma\gamma \end{aligned} \tag{2.2a}$$

$$\begin{aligned} p + \gamma_{\mathrm{CMB}} &\rightarrow n + \pi^+ \\ \pi^+ &\rightarrow \mu^+ + \nu_\mu \\ \mu^+ &\rightarrow e^+ + \nu_e + \bar{\nu}_\mu \end{aligned} \tag{2.2b}$$

In the case of heavy primaries, much lower neutrino fluxes are expected in the EeV range compared to protons. In addition to the study of $X_{\max}$ of cosmic rays, the measurement of particle densities and arrival times of air showers arriving at the ground can be used to estimate the mass of the primary cosmic rays which initiated the shower [13–16]. The most important particles whose properties are normally used in this regard are muons at the ground level. Data from surface detectors can also be used to determine the distance to the shower maximum [13–16].

However, it should be noted that conclusions of experimental results on the composition of UHECRs based on the analysis of the depth of shower maximum have so far not been unanimous. This is discussed further in section 3.4. The interpretation of mass composition from the depth of shower maximum is the subject of this study.

2.1.3 Anisotropy of UHECR

The observation of a cutoff in the spectrum of UHECRs, when interpreted in the light of the predictions of [11, 12], implies that UHECRs are propagated through

relatively short lengths of up to $\approx$ 250 Mpc [17]. At such scales of length, the distribution of matter in the universe is inhomogeneous: it takes the form of clusters and superclusters [17]. Neglecting the deflection caused by the intervening magnetic flux, the flux of the highest energy cosmic rays should therefore be anisotropic. This should particularly be the case if the cosmic rays possess energy above a few tens of EeV and have only a small charge. A study of the anisotropy of UHECRs should reveal whether they are galactic or extra-galactic in origin.

A natural assumption in most astrophysical scenarios is that the number of sources is large enough such that their distribution represents the distribution of matter. Taking this to be the case, anisotropy at the Earth is therefore expected to depend only on the nature and size of UHECR deflections. Hence a fit to the measured anisotropy yields parameters that influence the deflections. Such parameters include charge composition and cosmic magnetic fields. Experimental measurements of anisotropy in the arrival directions of UHECRs have been presented e.g. [18–20].

In the study whose results are reported in [18], the initial data collected by the Pierre Auger Observatory from 1 January 2004 to 31 August 2007 whose integrated exposure for event selection amounted to 9000 km^2 sr yr, was analyzed for anisotropy. For cosmic rays with energies exceeding approximately 6×10^{19} eV, an anisotropy was reported at a confidence level greater than 99%. The analysis was based on the correlation within an angle of separation of 3.1° between the arrival directions of UHECRs and the locations of nearby active galaxies in the Véron-Cetty and Véron (VCV) catalogue. However, subsequent studies using progressively higher amounts of data have obtained levels of anisotropy that are less and less important statistically, and also lower correlating fraction [21, 22]. In the latest study carried out by the Auger experiment on anisotropy using a data set that was three times larger than the previous one [20], none of the tests for anisotropy carried out produced any statistically significant result. One cannot therefore tell whether the primary composition is dominated by protons as assumed in the dip model (see Sec. 2.2.5), nor can one make any deduction as to the the maximum distance of the sources from the Earth.

Some other results from the Pierre Auger Observatory [19, 23, 24] seem to challenge the models which treat cosmic rays up to the ankle to be galactic in origin. The dipolar anisotropy of the direction of arrival of particles has been found in these studies to be surprisingly small compared to the expectations of the models.

2.2 Astrophysical scenarios

2.2.1 Sources of UHECR

One of the most intriguing questions in astroparticle physics as yet is the source of UHECRs. Cosmic rays are known to be composed of protons and other nuclei, with the heaviest one with significant abundance being iron. Considering that UHECR travel cosmological distances, their propagation is expected to be influenced by the cosmic microwave background (CMB), cosmic infrared background (CIB) and photo disintegration among other factors. A consequence of extragalactic nuclei interaction

with the CMB is the GZK effect.

When considering possible candidate sources of UHECRs, one must take into account the processes through which the high energy they possess can be attained. The first acceleration process that was proposed, the Fermi type [25], was assumed to be stochastic. This was later replaced by acceleration at astrophysical shocks, which is a more efficient process. However, estimates and calculations suggest that diffusive shock acceleration can only account for cosmic rays of energies up to a few GeV e.g. protons of up to $\approx 10^5$ GeV [26, 27].

An alternative that one can turn to that is capable of acceleration to energies up to 10^{20} eV are extra-galactic objects. Based on the principles of Fermi's first order acceleration, the accelerated nuclei must be confined to the region where acceleration takes place i.e. in the magnetic field of the extra-galactic object. It is thus possible to calculate the maximum theoretical acceleration energy to be [28, 29]:

$$E_{max} \sim \beta c Z e B L \,, \tag{2.3}$$

where βc is the characteristic velocity of the magnetic scattering centers, Ze represents the charge of the nucleus, B is the strength of the magnetic field in μG and L is the linear dimension of the accelerating object in parsec. This is usually called the Hillas' criterion. Based on this criterion, a plot of B against L, which includes the extra-galactic objects fitting in given regions (the Hillas' diagram) can be made, as shown in Fig. 2.2. Only very few objects such as highly magnetized neutron stars, active galactic nuclei, and lobes of giant radiogalaxies can reach the indicated minimum energy, even before accounting for efficiency. It is thus apparent that acceleration of nuclei from such sources is unlikely. A model where the nuclei are injected at a low energy, and later acquire a high energy due to acceleration, is known as a *Bottom–Up* model [30]. As a result of the limitations of the Bottom–Up models in explaining the observed high energy cosmic rays, a different category of models based on exotic particle physics scenarios were floated as alternative explanations for the sources of UHECRs. This second category of models are collectively known as *Top–down* models.

2.2.2 Top–down models

In a top-down model, it is supposed that there is no acceleration, but instead, a very massive particle decays, and its subsequent disintegration produces the observed UHECRs. Such a particle, referred to as an X particle, is supposedly trapped in the galactic halo and has a mass that is greater than 10^{12} GeV and a lifetime greater than the age of the universe [31]. Possible candidates include superheavy dark matter (SHDM), topological defects (TD) such as monopoles or strings [6] and Z-bursts (ZB) [32].

The SHDM is explained as having been produced in the first place during reheating caused by inflation, and is assumed to be contained in cold dark matter [31]. Likewise, the topological defects are considered to be left over from phase transitions in the early universe caused by spontaneous breaking of symmetries [33, 34]. The Z-bursts, on the other hand, are assumed to occur when ultra-high energy neutrinos

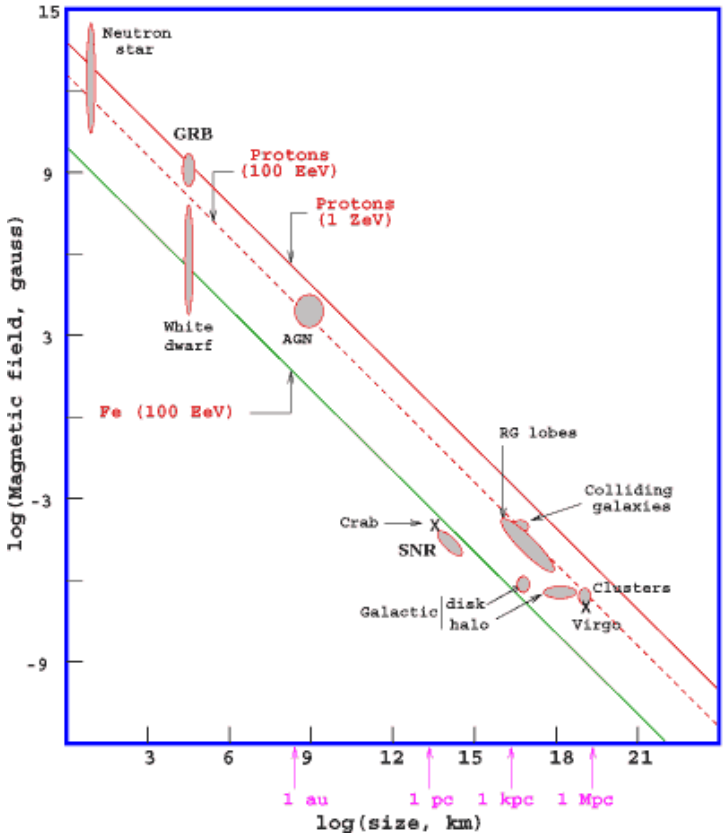


Figure 2.2: *The Hillas' diagram showing the limits of magnetic field and size of candidate objects for the acceleration of CRs (adapted from [28]). An object below a given line cannot accelerate proton or iron to energies beyond the one indicated for the line.*

originating from remote sources annihilate with relic background neutrinos at the Z-resonance [35, 36].

All the top-down models predict that at the highest energies, the photon flux attributed to them should be much higher than the expected flux of photons from the interaction of extragalactic nucleons with the CMB [32]. By estimating the upper limit of the photon flux of the top-down models, assuming the data spectrum, one can test the validity of these models. The flux predictions should hold at least at the highest energies, if it is to be expected to be true at lower energies.

The results for the tests for the validity of the Top-down models have generally not been positive. An examination in 1999 of the then available world data set concluded that less than 10% of UHECRs originate from relic particles in the halo [37]. The most recent Pierre Auger results [38] strongly disfavour the production of photons as a result of the decay of heavy primordial particles predicted by top-down models. This is because the models cannot account for a significant part of the UHECR flux observed by Auger. The results are summarized in Fig. 2.3.

The availability of data on both $X_{\max}$ and the the all-particle spectrum has encouraged many authors to come up with a variety of scenarios of UHECR bottom-

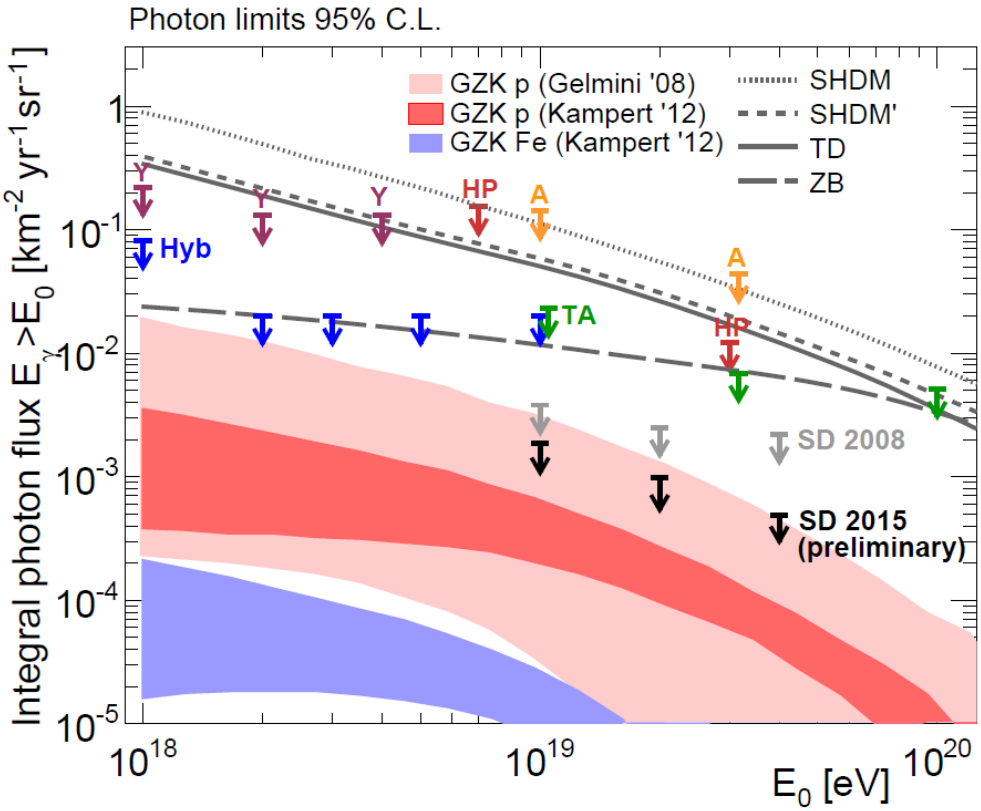


Figure 2.3: *Upper limits at 95% C.L. to the diffuse flux of UHE photons from different observatories: Pierre Auger hybrid (Hyb) and SD, Telescope Array (TA), Yakutsk (Y), Havera Park (HP), AGASA (A). Also included are predictions of the top-down models superheavy dark matter (SHDM), topological defects (TA) and Z-burst (ZB) as well as cosmogenic models [38].*

up models. The models generally assume that the nuclei are injected at the sources following a spectrum of the form $E^{-\alpha}$ at maximum energies in excess of 10^{20} eV. The nuclei then get transformed by interaction with the CMB and CIB through the processes of pair production and photo-disintegration such that by the time they arrive and are detected on the Earth, their arrival directions, energy and even chemical composition are different from the ones at the sources. A comparison of the calculated composition of UHECR at the Earth for a given scenario with the experimentally measured result is an indicator of how well the scenario describes the sources and propagation of UHECR. The evolution of $X_{\max}$ with energy, which is directly influenced by the nature of UHECR spectrum, is an indicator of how the average mass of the measured UHECR varies with energy. Other data that may give a clue as to the suitability of a model are anisotropy and energy spectrum.

Each scenario differs fundamentally from the others in its interpretation of the experimental data. Some three of the most common models are discussed in this section.

2.2.3 The standard model of galactic cosmic rays

The essential assumption in this model is that the particles injected at the sources of UHECR are accelerated so as to attain a maximum energy $E_{\max}$, which is proportional to the charge Z of the injected particle [39, 40]. Beyond energy $E_{\max}$, the source spectra are exponentially attenuated. This model treats the cosmic rays with energies below the knee as originating from, and being accelerated by, supernova remnants (SNRs). The acceleration of particles occurs at the supernova shock through first order Fermi acceleration [41]. A necessary condition for this is that the magnetic field in the shock region has to be amplified by a factor of $100 - 1000$ compared to the interstellar field. The particle acceleration is regulated by the resulting streaming instability.

The net effect of acceleration by the supernova shock is that particles in a narrow energy range around $E_{\max}(t)$ escape from the upstream region, while lower energy particles remain within the shock and undergo continuous acceleration [41]. The latter particles escape at much later times. The total flux injected by SNRs is then given by an integration over time of the superposition of particles escaping from upstream, and the flux of particles accumulated behind the SNR shock, in all supernova events. In modern nonlinear theories of particle acceleration at shocks, particles that are confined in the shock region make a significant contribution to the cosmic ray spectrum (see e.g. [42]). The standard model is characterized by $E_{\max}(t)$ that is dependent on rigidity via

$$E_{\max} \approx 2Z \times 10^{15}\ \text{eV}, \tag{2.4}$$

where Z is the charge of the nucleus. At $E \gtrsim E_{\max}$, a sharp cutoff is predicted for the spectra of all nuclei. The diffusion of cosmic rays in the interstellar medium has to be accounted for before the final predictions of the standard model can be expected to compare well with measured values. However, calculations of flux done taking into account diffusion coefficient, $D(E) \propto E^{\beta}$, have not been conclusive [43].

From Eqn. 2.4, given that the observed proton knee is $E^{p}_{\text{kn}} \approx (2-3) \times 10^{15}$ eV, the standard model predicts that the galactic cosmic ray spectrum should coincide with the iron knee, $E^{\text{Fe}}_{\text{kn}} \approx (5-8) \times 10^{16}$ eV. The limitation of the standard model, as revealed in this fundamental conclusion to which it leads, is that it fails to account for the cosmic rays observed between $E^{\text{Fe}}_{\text{kn}}$ and the ankle at $E_a \sim 1 \times 10^{19}$ eV, which are supposed to be galactic.

2.2.4 The ankle model

A marked flattening of the cosmic ray spectrum at $\sim 3 \times 10^{18}$ was first observed by Havera Park [44], Akeno [45] and Yakutsk [46] experiments. Recent results from the Pierre Auger Observatory have revised the position of this feature to be at $\sim 5 \times 10^{18}$ eV [47]. The feature has traditionally been explained by the ankle model, named after the shape it gives to the spectrum around this energy. This model essentially considers the ankle to be the result of a transition from galactic to extragalactic cosmic rays [27]. In this model, the extragalactic spectrum, which is relatively flat, intersects the steep galactic spectrum, resulting in the ankle at

an energy just above the point of intersection. It is expected that at around the energy of the ankle, the galactic magnetic field should lose its efficiency to confine charged particles since the gyroradius of a charged particle becomes comparable to the thickness of the galactic disc. For this reason, any cosmic ray detected beyond the ankle is presumed to originate from outside the galaxy. The generation spectrum of the extragalactic component of ~ 2 according to this model seems to fit well with that predicted for Fermi acceleration at non-relativistic shocks ($\gamma = 2 - 2.5$) and at ultra-relativistic shocks ($\gamma_g = 2.2 - 2.3$) [41]. A dip at $E \sim 5 \times 10^{18}$ eV can be explained using the ankle model.

Since the galactic and extragalactic fluxes are equal at the ankle, the galactic component at this point needs to have undergone a large acceleration in order to attain the high energy. The required acceleration is higher than the maximum allowed by the standard model by a factor of $30 - 40$ [48]. Furthermore, most of the ankle models e.g. [27] predict a large fraction of galactic cosmic rays to be present in the measured flux at energy 10^{19} eV and even beyond. There is hence an energy gap that is not reasonably accounted for by the ankle model.

A stringent constraint on the ankle model is given by the chemical composition it predicts. In the energy range $(10^{18} - 10^{18.6})$ eV, the model predicts a heavy galactic cosmic ray component. However, deductions made concerning composition based on the depth of shower maximum measured by HiRes and TA experiments show a light composition throughout this range [48]. The Pierre Auger results show a composition that initially gets lighter up to $10^{18.3}$ eV, before beginning to get heavier beyond this point [49]. The ankle model is therefore strongly disfavoured.

2.2.5 The dip model

As noted above in the case of the standard model of galactic cosmic rays, the ankle model leaves a gap in energy between the iron knee and the ankle that is not accounted for. An alternative interpretation that takes care of this gap is provided by the pair-production dip model, in which the transition from Galactic to extragalactic components occurs at $\sim 10^{17}$ eV. In this model, an assumption is made that the sources of extragalactic cosmic rays accelerate only protons, hence all the other nuclei are absent in the spectrum by the time it is measured at the Earth. It was first proposed by Berezinsky and Grigorieva [50] and later studied further e.g. in [51]. A flux distribution calculated using the dip model is shown in Fig. 2.4.

The dip is viewed as coming about as a result of e^+e^- pair production caused by the interaction of extragalactic protons with CMB photons, leading to energy losses. The consequence is a feature in the propagated spectrum often referred to as the pair production dip. The ankle is described by the pair production dip since it forms a flat intrinsic part of the dip. Flux suppression in the dip model can be described in a scenario whereby the sources inject protons according to the power law $E^{-\beta}$, which is attenuated exponentially by a scale parameter $E_{\rm cut}$ of the order of 10^{20} eV. An assumption made is that of a cosmological evolution of source luminosity parametrized as $(1 + z)^m$ and a continuous source distribution.

The dip was observed with good statistical significance by the Fly's Eye, Yakutsk, Akeno-AGASA and HiRes detectors at $1 \times 10^{18} \leq E \leq 4 \times 10^{19}$ eV [52]. The

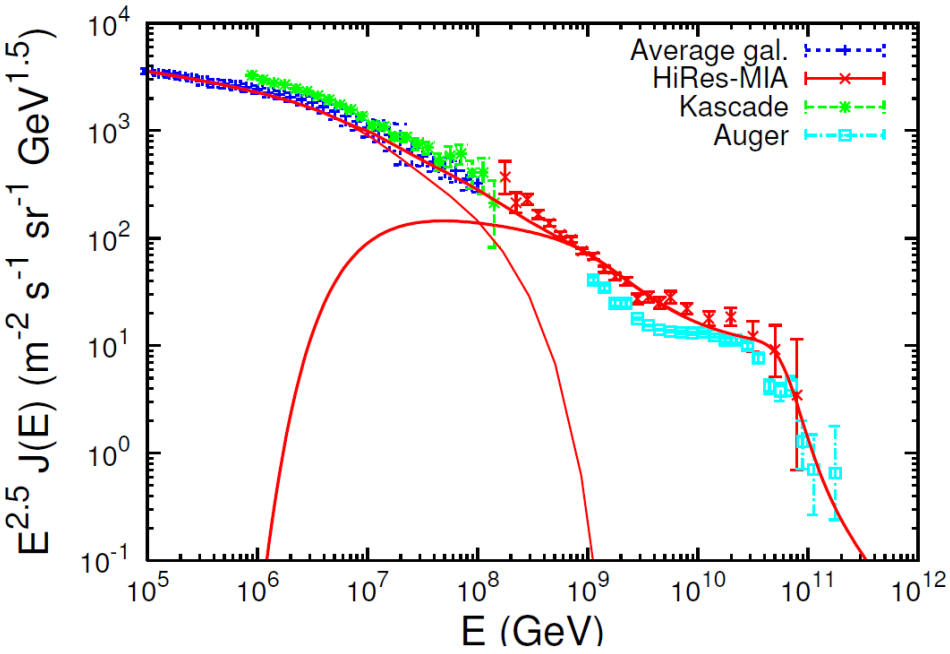


Figure 2.4: *The dip model explanation of the transition from galactic to extragalactic CR. The galactic flux begins at* 10^5 *GeV and intersects the extragalactic flux at* $\sim 10^8$ *GeV. The calculated combined spectrum is fitted to data from the experiments indicated [39].*

results of the Pierre Auger data on the one hand show a feature similar to the dip, but on the other hand the chemical composition at the dip contradicts the basic assumption of the dip model: that the primary composition is pure proton. Using the data collected by the Auger Observatory from 1 January 2004 to 31 December 2012 [53], a reasonably good description of the spectrum, assuming the dip model, was obtained by choosing $m = 5$ and $\beta = 2.35$ as shown in Fig. 2.5. However an analysis of mass composition for the same data that was done based on the variation of the depth of shower maximum with energy concluded firmly that the composition at the ankle region was heavier than proton. Similar conclusions have been made in subsequent analyses [49].

The extragalactic spectrum of protons predicted by the dip model becomes flat at energies below $E_{\rm cr} \approx 1 \times 10^{18}$ eV, while the galactic spectrum gets very steep ($\propto E^{-3.1}$). Hence the two spectra have to intersect at some point below $E_{\rm cr}$, in a similar manner to the case of the ankle model. The dip model agrees very well with the standard model [41].

The generally accepted idea that the ankle marks a transition from galactic to extragalactic cosmic rays and explained by different models, can only be checked by using experimental data. The models can be tested using energy spectra, anisotropy or chemical/mass composition data. In the case of the dip model, the ER can provide a signature manifested in a sharp transition from an iron-dominated to a proton-dominated composition at energy $E \approx 1 \times 10^{18}$ eV [41]. The sharp change predicted

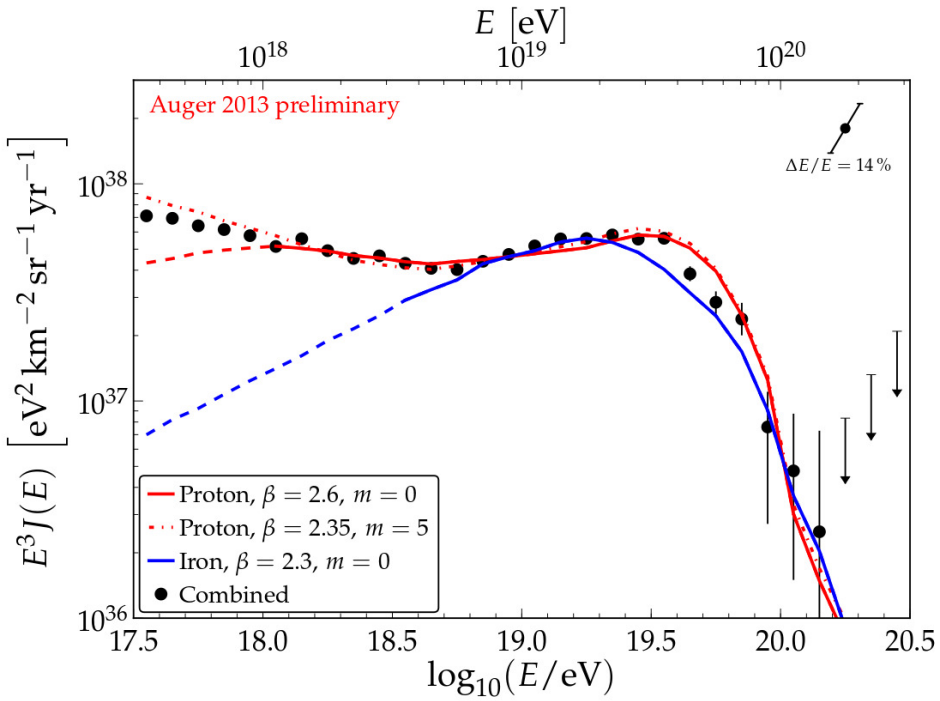


Figure 2.5: *The combined Auger spectrum as at 2013 compared to distributions obtained from spectra generated assuming different parameters of the dip model [53]. For more details, see text.*

by the dip model from iron-dominated galactic to proton-dominated extragalactic cosmic rays below the ankle is expected to cause a steep dependence of $\langle X_{\mathrm{max}} \rangle$ as a function of energy in the interval 10^{17} to 10^{18} eV. Iron nuclei should contribute more to the $\langle X_{\mathrm{max}} \rangle$ observed below 3×10^{17} eV, while above 10^{18} eV, protons are expected to dominate. A small admixture of nuclei in the extragalactic flux is allowed.

A comparison of calculated X_{max} as a function of energy, based on the dip model, has been found to compare well with the experimental measurements of Fly's Eye, HiRes-Mia, HiRes and TA experiments [41, 48]. However, the Pierre Auger results show a composition that is heavy [49], which is a contradiction of the dip model.

2.3 Extensive Air Showers

Due to the transverse momentum acquired by the secondary particles at their production and their interaction with air particles, the cascade of shower particles spreads out. The phenomenon whereby the particles of an air shower resulting from the same primary cosmic ray are detected within a negligible time interval over a large area is called an extensive air shower. The surface area covered by particles in an extensive shower range from approximately 10^4 m^2 at 10^{15} eV to 10^7 m^2 at 10^{20} eV. The number of particles contained are approximately 10^6 and 10^{11} respectively [54].

2.3.1 Interactions of UHECR

The upper atmosphere of the Earth is continually bombarded by elementary particles, protons and atomic nuclei originating from outside the solar system, collectively referred to as cosmic rays. The interaction of cosmic rays with air particles in the atmosphere results in the production of charged ($\pi^{\pm}$) and neutral (π^0) pions. Since the lifetime of π^0 is relatively short, it decays quickly into two gamma rays. The γ-rays then produce e^+e^- pair, which in turn produce more γ-rays, and the process is repeated. The charged pions survive for longer due to their longer lifetime, and hence may collide with air particles before decaying, mainly into a muon and a neutrino. In this manner, a cascade or shower of particles moves towards the ground. A sketch of the development of a shower is shown in Fig. 2.6. As the shower particles continually lose energy at every interaction, the production of new particles stops at some point, at which the shower reaches its maximum. The depth at which a shower reaches its maximum depends on the energy of the primary particle and its type.

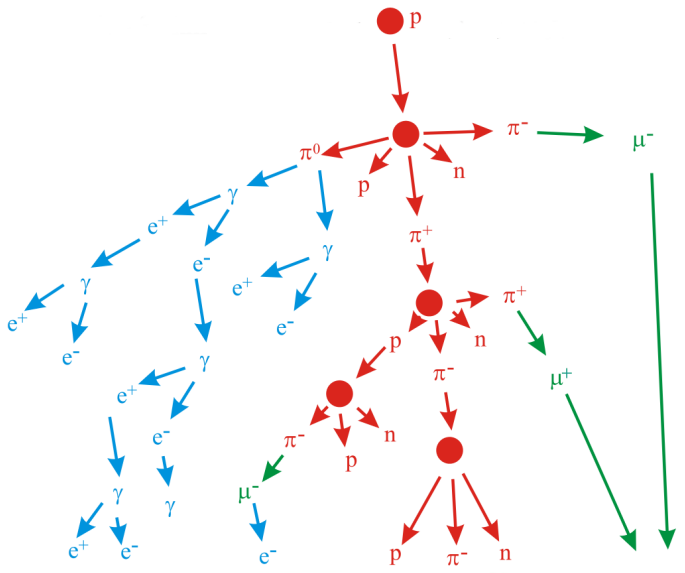


Figure 2.6: *A sketch of the development of an air shower in the atmosphere showing the electromagnetic (blue), hadronic (red) and muonic (green) components (adapted from [55]). The neutrino component is not shown.*

An air shower may be viewed as made up of several components. The electromagnetic component is propagated by the production of photons through the decay of π^0. The interactions involving the charged pions constitute the hadronic component, while the muons resulting from the final decay of $\pi^{\pm}$ form the muon component. The neutrinos produced alongside the muons may be considered as an additional component, but they are not detectable by the Pierre Auger Observatory.

The electromagnetic and hadronic cascades not only develop longitudinally in the atmosphere but also spread out laterally. The extent of lateral spreading is determined by the multiple scattering of electrons and positrons in the electromagnetic

cascades and in hadronic cascades by the transverse momenta of the secondary particles at production. In the case of an UHECR, the lateral spread can cover several square kilometers.

2.3.2 Heitler model

A simple model was proposed by Heitler [56] to represent the development of extensive air showers. In its simplest form, this conception of extensive air shower cascades has the advantage that it predicts accurately the most important features of electromagnetic showers [57]. It provides a powerful illustration of the relationship between shower observables and interaction physics at a microscopic level, although it is limited in describing real showers. The description of the model given here is based on the discussion of the Heitler model in [57]. A more realistic treatment of the depth of shower maximum is given in Secs. 3.1 and 3.2.

2.3.2.1 Basic features of Heitler model

The model considers electrons, positrons and photons to undergo two-body interactions repeatedly resulting in the production of either a photon of bremsstrahlung or e^+, e^- pair. The distance between two successive splits is fixed and is related to the radiation length. When the individual e^+, e^- energies drop below a critical value, E_c^e, multiplication ceases at once. A schematic view of such a cascade is shown in Fig. 2.7(a).

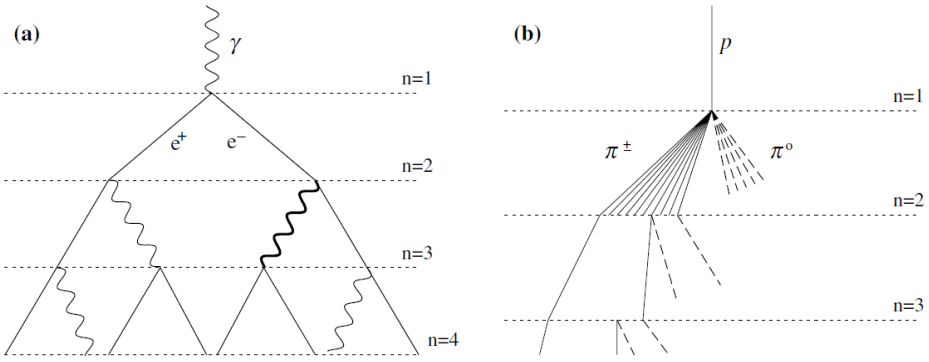


Figure 2.7: *A sketch of the development of (a) an electromagnetic shower and (b) a hadronic shower as expected by the Heitler model (adapted from [57]). The distance between two successive levels, λ_e, is related to the radiation length, λ; $d = \lambda \ln 2$. Decaying neutral pions are represented in the hadron shower using dashed lines. Only a few pion lines are shown after the $n = 2$ level.*

Hadronic interactions are approximated in a similar manner, as depicted in Fig. 2.7(b). A proton interacts with an air nucleus, causing the release of pions. The neutral pions decay to photons instantaneously, resulting in electromagnetic sub-showers. Thus the energy of the initiating particle is divided into the hadronic and electromagnetic channels, with the electromagnetic channel accounting for a

third of it. As in the case of the electromagnetic shower, the $\pi^{\pm}$ travel fixed distances and then interact. This is repeated until a critical energy E_c^{π} is reached, beyond which only decays occur, resulting in the production of muons. The critical energies are approximately 150 GeV and 10^9 GeV for charged pions and neutral pions respectively. A more realistic approach has to take into account the energy carried away by the leading particle.

2.3.2.2 Electromagnetic cascades

The total number of particles in the shower after k generations is 2^k, where two successive generations are separated by a length, λ_e, given by $\lambda_e = \lambda \ln 2$ and λ is the radiation length in the medium. Assuming that energy is always divided equally between the daughter particles, each particle in the k^{th} generation possesses energy $E_0/2^k$. Hence,

$$E_c^e = \frac{E_0}{2^{k_{\max}}} = \frac{E_0}{N_{\max}} \tag{2.5}$$

and

$$N_{\max} = \frac{E_0}{E_c^e} \, . \tag{2.6}$$

The number of generations required before the shower maximum is reached is

$$k_{\max} = \frac{\ln N_{\max}}{\ln 2} = \frac{\ln(E_0/E_c^e)}{\ln 2} \, . \tag{2.7}$$

The depth of the shower maximum is then

$$X_{\max} = k_{\max}\lambda_e = \frac{\ln(E_0/E_c^e)}{\ln 2}\lambda_e = \frac{\lambda_e}{\ln 2}\ln E_0 - \frac{\lambda_e \ln E_c^e}{\ln 2} \, . \tag{2.8}$$

The Eqn. 2.8 may be written in the form

$$X_{\max} = D \ln E_0 - C \tag{2.9}$$

where $D = \lambda_e/\ln 2 = 2.3\lambda$ is the elongation rate. It can be seen from Eqn. 2.6 that the maximum number of particles in a shower is proportional to the energy of the primary particle initiating the shower. Similarly from Eqn. 2.8, one can conclude that the depth of the shower maximum is proportional to the logarithm of the energy of the primary particle which initiated the shower. These two fundamental predictions of the Heitler model have been confirmed through the results of detailed simulations and by experiments.

2.3.2.3 Proton-initiated air shower

A model can be derived for air showers that are initiated by hadrons, based on the Heitler model for electromagnetic showers. Some assumptions have to be made in order to simplify the calculations. A cosmic ray particle is considered to interact with an air nucleus after travelling a fixed length $\lambda_p \ln 2$, where λ_p is the interaction length of a proton. After every interaction, the ratio of charged pions to neutral

pions produced is 2:1, and energy is equally distributed among them. These are the only particles produced in an interaction. The neutral pion decays instantly into two gamma rays, producing electromagnetic showers, while the charged pions continue on and interact in the next layer of air. The interactions stop when the pion energy falls below a critical value E_c^π, at which point the pions decay instantly into muons and neutrinos. In the simplest form of this model, it is further assumed that there are no leading hadron fragments, and the number of pions produced in an interaction i.e. pion multiplicity, M, remains constant. The number of muons reaching the ground is equal to the number of charged pions in the shower with $E \lesssim E_c^\pi$.

If a single cosmic ray proton with energy E_0 enters the atmosphere, the total number of charged pions after n interactions will be

$$N_\pi = \left(\frac{2}{3}M\right)^n \ . \tag{2.10}$$

Taking into account the energy that has gone into the electromagnetic shower, every charged pion will possess energy

$$E_\pi = \frac{E_0}{M^n} \tag{2.11}$$

The number of pion generations, n_c, required to reach E_p is obtained thus:

$$\begin{aligned} E_p &= \frac{E_0}{M^{n_c}} \\ M^{n_c} &= \frac{E_0}{E_p} \\ n_c &= \frac{\ln(E_0/E_p)}{\ln M} \end{aligned} \tag{2.12}$$

Substituting Eqn. 2.12 into Eqn. 2.10

$$\ln N_\pi = \ln N_\mu = n_c \ln\left(\frac{2}{3}M\right) = \frac{\ln\left(\frac{2}{3}M\right)}{\ln M} \ln\left(\frac{E_0}{E_p}\right) \ . \tag{2.13}$$

The Eqn. 2.13 may be written in the form

$$N_\mu = \left(\frac{E_0}{E_p}\right)^\beta \tag{2.14}$$

where β simplifies to 0.9. Studies based on detailed Monte Carlo simulations [58,59] have reported values of β in the range $0.85 - 0.92$. Since β changes logarithmically with M, a simplifying assumption that treats β as approximately constant does not alter the results significantly.

For the conservation of energy, the energy in the electromagnetic and hadronic subshowers should add up to the energy of the primary particle: $E_0 = E_c^e N_{\max} + E_c^\pi N_\mu$. The fraction of the electromagnetic component, and hence the energy contained in it at the n^{th} generation can therefore be obtained from:

$$\frac{E_{EM}}{E_0} = 1 - \left(\frac{2}{3}\right)^n \ . \tag{2.15}$$

An expression for the depth of shower maximum similar to Eqn. 2.8 can be obtained for the pion cascade as:

$$X_\pi = n\lambda_p \ln 2 = \frac{\ln(E_0/E_p)}{\ln M}\lambda_p \ln 2 = \lambda_p \frac{\ln 2}{\ln M} \lg E_0 - \lambda_p \frac{\ln 2 \ln E_p}{\ln M} \tag{2.16}$$

which is of the form

$$X_\pi = D_\pi \ln E - C_\pi \ . \tag{2.17}$$

A comparison of Eqns. 2.8 and 2.16 gives the ratio of the elongation rate of electromagnetic subshower to that of proton or pion subshower as $\frac{\lambda \ln M}{\lambda_p \ln 2}$. Given that the interaction lengths in air for photons and pions are approximately 37 g/cm^2 and 120 g/cm^2 respectively and $M \sim 10$ [57], the elongation rate of the electromagnetic subshower should therefore be higher than that of the pion subshower by a factor given by this fraction.

2.3.2.4 Nuclear-initiated showers

A nuclear shower is treated by applying the principle of the superposition model: a nucleus with A nucleons and total energy E_0 is approximated to produce a total of A showers, each with initial energy E_0/A and starting at the same point. The number of muons observed at the ground sums up to:

$$N_\mu = A^{1-\beta} \left(\frac{E_0}{E_p}\right)^\beta \ . \tag{2.18}$$

If the value of β is taken, as previously mentioned, to be 0.9, iron showers are expected to produce 50% more muons than proton showers of the same energy. The depth of the first interaction for a proton being $\lambda_p \ln 2$ leads to the shower maximum for a nuclear shower being

$$X^A_{\max} = \lambda_p \ln 2 + D_p \ln(E_0/A) - C_p \tag{2.19}$$

or

$$X^A_{\max} = X^p_{\max} - (D_p \ln A + \xi) \tag{2.20}$$

where p refers to proton and $\xi = C_p - \lambda_p \ln 2$. Equation 2.20 suggests that the depth of a shower initiated by a nucleus will be shallower than that initiated by a proton of the same energy. Hence the depth of shower maximum can be used to discriminate between a heavy composition and a light composition, although the ER remains unaltered in the two cases.

2.3.3 Shower Universality

The determination of the evolution of primary cosmic ray mass with energy normally involves the use hadronic interaction models to interpret the $X_{\max}$ data collected experimentally. The fact that hadronic interactions at the highest energies are quite uncertain renders it difficult to make firm conclusions regarding mass composition. An approach in analysis which is independent of the hadronic interaction model used

would greatly reduce the systematic uncertainties in the conclusions regarding mass composition. In this regard, the works of Chou *et al.* [60] and Schmidt *et al.* [61] among others, have concluded that it is possible to find a function which relates the $X_{\max}$ to one or possibly two other SD shower observables, independent of the type of primary particle and energy. This is referred to as shower universality. Among the air shower observables that have been used to characterize extensive air showers are the energy of the primary cosmic ray, E_0, the overall normalization of the muon component, N_μ, and the shower age, s, which is a function of the depth of individual showers, X, and the $X_{\max}$.

The concept of universality has been studied with regard to the electron component of both electromagnetic and hadronic showers, and applied to data collected by surface detectors of cosmic rays, in particular those that are designed to detect UHECRs [62–64]. Although limitations have been noted, especially with the use of the shower age, s, as a universality in hadronic showers [65], generally good levels of accuracy have been reported for shower observables that were parametrized using universality.

Air shower universality is made possible due to the large statistics on shower properties resulting from the enormous amount of particles involved in an extensive air shower, especially that initiated by an UHECR. The EAS thus behaves like an ensemble of thermodynamic particles, and the parameters describing it can be extracted through simulation [15]. By examining the complete set of EAS parameters, the primary cosmic ray parameters e.g. composition can potentially be derived.

In determining the SD primary energy and $X_{\max}$ used in this study, the parametrized shower observables were the primary energy, E_0, the depth of shower maximum, and the normalization of the muon content, N_μ. Only a brief description of the procedure is given here. Full details are found in [14, 15, 66].

2.3.3.1 N_μ Reconstruction

As a starting point, showers with zenith angles smaller than 60° detected simultaneously by the the FD and SD (hybrid events) are used to estimate the corresponding muon content. Since the distance in grammage from the ground to the electromagnetic or muonic $X_{\max}$ (DX) is more useful in the fit procedure involved, the available monthly molecular atmospheric profiles are used to convert $X_{\max}$ to DX. The signals are recorded in units of vertical equivalent muons (VEM), where a VEM is defined as the charge deposited in an SD detector when a muon that is moving vertically downwards traverses the detector through its center. The signals are fitted using a likelihood method to the parametrization [13]

$$S_0(DX, E) = S_{ref} \left(\frac{E}{10^{19}\,\mathrm{eV}} \right)^{\gamma} \left(\frac{DX - DX_0}{DX_{ref} - DX_0} \right)^{\frac{DX_{max} - DX_0}{\lambda(E)}} \exp\left(\frac{DX_{ref} - DX}{\lambda(E)} \right) \tag{2.21}$$

where S_0 is the signal in an ideal detector for a zenith angle of zero, and which can be estimated from signals in a real detector, while $\lambda(E) = \lambda_0 + f_\lambda \log(E/10^{19}\,\mathrm{eV})$. The value of f_λ is non-zero only in the case of electromagnetic component of a shower, S_{ref} is the value of the signal at DX_{ref} and the value of DX_0 is fixed to

different values for muonic, pure, electromagnetic, electromagnetic from muon and electromagnetic from low energy hadron components.

Two steps are involved in the reconstruction process. As a first step, all the stations which are triggered and have signals larger than 3 VEM are used, while keeping the $X_{\max}$ and energy fixed to the FD values. In the next step, an additional restriction on the SD detectors, that their distance from the shower core should be less than 2500 m, is imposed. Additionally, the shower direction is obtained from hybrid reconstruction.

The muon content in the vertical showers is estimated using the universality ansatz viz. the electromagnetic component at the ground level is a function of $X_{\max}$ and energy, with small deviations that can be parametrized using N_μ. The parametrization is represented in Eqn 2.22,

$$\langle N_\mu \rangle (E, \theta) = N_{\mu\,0} + A_{N_\mu}(\lg E - \lg E_0) + B_{N_\mu}(\sec\theta - 2) \tag{2.22}$$

where $E_0 = 10^{19}$ eV. By using a binned χ^2 fit, the parameters $N_{\mu\,0}$, A_{N_μ} and B_{N_μ} are obtained, from which $\langle N_\mu \rangle$ may be calculated. Events which can be reconstructed from either the SD data alone or FD data alone, referred to as "Golden Hybrid", provide the values of E and θ for use in Eqn 2.22. The value of N_μ is obtained from the correlation between $N_\mu/\langle N_\mu \rangle$ and $X_{\max} - \langle X_{\max} \rangle$ given by Eqn 2.23, which has been confirmed to be true by MC studies [15].

$$N_\mu/\langle N_\mu \rangle = 1.0 + \frac{0.5}{\pi} \arctan\left(\frac{X_{\max} - \langle X_{\max} \rangle}{40}\right) \tag{2.23}$$

2.3.3.2 Energy Reconstruction

With the exception that N_μ and $X_{\max}$ are fixed to the average obtained from Golden Hybrid calibration, a similar procedure to the one described in the previous section is followed for getting the energy. The bias and resolution of the energy are estimated from the Golden events.

2.3.4 Gumbel Parametrization of X_{max} Distributions

In cases where the simulation of the variation of $\langle X_{\max} \rangle$ or its moments with energy for a pure composition has been necessary, the Gumbel-expected values have been adopted in this work. The generalized Gumbel probability density function is defined as

$$\mathcal{G}(z) = \frac{1}{\sigma}\frac{\lambda^\lambda}{\Gamma(\lambda)}\left(e^{-\lambda z - \lambda e^{-z}}\right), \quad z = \frac{x-\mu}{\sigma}, \tag{2.24}$$

where μ and σ represent a location and scale parameter which are related to the mean and spread of the distribution respectively, Γ is the Euler gamma function and λ is a shape parameter. An analytical model based on this function has recently been developed [67], which describes the distribution of $X_{\max}$.

The model is based on the observation that the fluctuations in the position of $X_{\max}$ for a given particle and energy, which is caused by stochastic fluctuations in the position of the first point of interaction in the atmosphere and in the secondary

interactions [68], can be treated as a special dissipative stochastic system. Such a system, consisting of many particles, is kept from thermodynamic equilibrium by energy injection and dissipation. Recent studies have shown that the sum of correlated or non-identically distributed variables or a combination of both, which are in a special dissipative stochastic system, follow a generalized Gumbel distribution [69, 70]. Furthermore, the generalized Gumbel distribution is found to describe the distribution of $X_{\max}$ resulting from any primary nucleus or photon [67].

A slightly modified Heitler model for the description of the development of an air shower (see section 2.3.2) may be applied when treating the development of a shower as a stochastic process. In this more realistic description, some particles do not interact or decay at the i-th step of development of the shower. This causes a fluctuation in the number of particles expected to be produced. Also, some energy is lost between the i-th and $(i+1)$-th steps of the shower development.

The Gumbel function was thus parametrized as a function of UHECR primary particle type and energy using several hadronic interaction models in [67]. Estimates of the parameters μ, σ and λ obtained have been used in this work. In the generalized Heitler model, x represents the $X_{\max}$ and $\mathcal{G}(X_{\max})$ may thus be used to calculate the mean value of $X_{\max}$ of a primary particle at a given energy:

$$\langle X_{\max} \rangle = \int_0^\infty X_{\max} \mathcal{G}(X_{\max}) dX_{\max} \quad . \tag{2.25}$$

3 Mass Composition of UHECR

One of the main characteristics of ultra high energy cosmic rays required for studying their origin is their mass composition. Depending on the evolution of the mass composition of UHECR, we may be able to make conclusions on whether they are extragalactic or not, and also understand better the steepening of the flux observed at around 4×10^{19} eV [71]. However, for one to make firm conclusions from the data regarding mass composition, it is imperative that the data be precise [49, 72]. To this end, the Pierre Auger Observatory has measured both the flux of the primary cosmic rays and the depth of shower maximum with an unprecedented accuracy.

Due to the fact that at the highest energies the intensity of the cosmic rays is very low, it is not feasible to measure the primary mass directly [49]. Instead, the mass can only be inferred from the properties of the particle cascade initiated when a primary cosmic ray interacts with nuclei at the upper atmosphere. In the case where the primary particle is an UHECR, a number of coherent secondary particles from the cascade may be observed at the ground level over a large area, a phenomenon known as an extensive air shower. Thus ground-based detectors spread out over a large area are a convenient means of observation of air showers by for example recording the particle densities at the ground level. Coupled with the observed longitudinal development of the cascade, the properties of the primary particles may be inferred.

Several explanations for the evolution of the mass composition observed by Auger and other experiments have been proposed. A few of them are discussed in this chapter, after a brief discussion of $X_{\max}$ that is more realistic than the simple treatment based on the Heitler model.

3.1 Depth of Shower Maximum

The evolution of the electromagnetic component of the air shower fed by photon and electron contributions, and which dominates the energy deposited in the atmosphere by the shower particles [73], is well described as a function of the slant depth X.

The slant depth is a measure of the traversed air mass, and is obtained by [49]:

$$X = \int_s^\infty \rho(\mathbf{r}(s'))\mathrm{d}s' \tag{3.1}$$

where $\mathbf{r}(s)$ is the trajectory of the shower through the atmosphere and $\rho(\mathbf{r})$ is the density of the air at the point $\mathbf{r}$.

Since the properties of the first few hadronic interactions at the beginning of an air shower fluctuate, it is not possible to measure the primary mass on an event-by-event basis, but it can be statistically inferred from the distribution of shower maxima. The average depth of shower maximum for an electromagnetic shower of energy E is approximately [2]

$$\langle X^{\mathrm{em}}_{\mathrm{max}} \rangle \approx X_0 \ln E/\varepsilon^{\mathrm{em}}_c \tag{3.2}$$

where X_0 is the radiation length in air. At the critical energy $\varepsilon^{\mathrm{em}}_c \approx 84$ MeV, the ionization energy losses in air are equal to the losses due to bremsstrahlung. The semi-superposition theorem [74], which is based on more realistic assumptions than the superposition ansatz used in the simple Heitler model, still leads to the same conclusion that a nucleus of mass A and energy E can be treated as a superposition of A nucleons, each of energy E/A. Considering the total multiplicity of the hadrons produced in an interaction to be N with an average energy E/N, a primary proton produces the shower maximum

$$\langle X^{p}_{\mathrm{max}} \rangle \approx \lambda_p + X_0 \ln \left(\frac{E}{2N\varepsilon^{\mathrm{em}}_c} \right). \tag{3.3}$$

Taking into account the elasticity of the first interaction (κ_{ela}), where κ_{ela} is the ratio of the energy possessed by the most energetic secondary particle produced in the interaction to the hadron energy,

$$\langle X^{p}_{\mathrm{max}} \rangle \approx \lambda_p + X_0 \ln \left(\frac{\kappa_{\mathrm{ela}} E}{2N\varepsilon^{\mathrm{em}}_c} \right). \tag{3.4}$$

3.2 Elongation Rate

The elongation rate (ER), symbolized as D, is defined as [75] the rate of change of the shower maximum with logarithm of energy.

$$D = \frac{\mathrm{d}\langle X_{\mathrm{max}} \rangle}{\mathrm{d} \ln E} \tag{3.5}$$

Assuming constant elasticity for protons and considering the results of hadronic interaction models which predict an approximately logarithmic decrease of λ_p with energy, the ER of a pure proton composition, D_p, is approximately constant and from Eqn. 3.4 may be shown to be given by

$$\langle X^{p}_{\mathrm{max}} \rangle = D_p \ln \left(\frac{E}{E_0} \right) + c_p, \tag{3.6}$$

where $D_p = d\langle X^p_{\max}\rangle / d\ln E$, E_0 is a reference energy and c_p is the average depth of shower maximum from a proton with a reference energy E_0 [76]. Using a superposition model, for any primary composition with average logarithmic mass $\langle \ln A \rangle$

$$\langle X_{\max}\rangle = D_p[\ln(E/E_0) - \langle \ln A\rangle] + c_p = \langle X^p_{\max}\rangle - D_p\langle \ln A\rangle. \tag{3.7}$$

The above equations still hold true in the semi-superposition model in which the fragmentation of the nucleus is taken into account [74]. Assuming that the proton ER does not vary with energy, which is confirmed with a good precision by Monte Carlo (MC) simulations, the general expression for ER of a mixed composition from Eqn. 3.7 is:

$$D = \frac{d\langle X_{\max}\rangle}{d\ln E} = D_p\left(1 - \frac{d\langle \ln A\rangle}{d\ln E}\right). \tag{3.8}$$

The ER has one parameter (c_p) less, compared to $\langle X_{\max}\rangle$, making its analysis more robust with respect to uncertainties in interaction models. For a constant composition, it follows from Eqn. 3.8 that $D = D_p$ (and it is true with a good accuracy for the complete shower simulations). Consequently, D becomes smaller (larger) for the composition getting heavier (lighter) with energy.

For the ER per decade, D_{10}, one has:

$$D_{10} = \frac{d\langle X_{\max}\rangle}{d\lg E} = D_p\left(1 - \frac{d\langle \ln A\rangle}{d\ln E}\right)\cdot \ln(10). \tag{3.9}$$

The derivative of Eqn. 3.9 with respect to $\lg E$ is:

$$\frac{dD_{10}}{d\lg E} = -D_p\frac{d^2\langle \ln A\rangle}{d(\ln E)^2}\cdot(\ln 10)^2. \tag{3.10}$$

The value of $dD_{10}/d\lg E$ can only be zero if the composition is not changing at all or if it changes at a constant rate, i.e. $d\langle \ln A\rangle/d\ln E = \text{const}$. In realistic astrophysical scenarios with differing slopes of energy spectra and cut-offs at different energies for different primary masses, it is a natural expectation that $dD_{10}/d\lg E$ can be non-zero for certain energy ranges.

3.3 Fluctuations of $X_{\max}$

The RMS of the distribution of the depth of shower maximum, $\sigma(X_{\max})$, being an independent observable in itself and varying with energy, can serve as a cross-check for measurements of the evolution of average UHECR mass with energy obtained using $X_{\max}$. Unlike the depth of shower maximum which at a given energy increases as $\ln A$ decreases and vice versa (see Eqn. 2.20), the case of $\sigma(X_{\max})$ is not as straightforward. Simulated individual events of pure proton and pure iron at energy 10^{19} eV [77] is shown in Fig. 3.1. It can be seen in this figure that the width of distribution of $X_{\max}$ for proton is larger than that of iron at the same energy. Given an ensemble of nuclear species, one can obtain its $\sigma(X_{\max})$ [78]:

$$\sigma^2_{\text{tot}} = \sum_{A=1}^{56}\left(f_A\sigma^2_A + f_A(X_{\max,\text{A}} - \langle X_{\max}\rangle)^2\right). \tag{3.11}$$

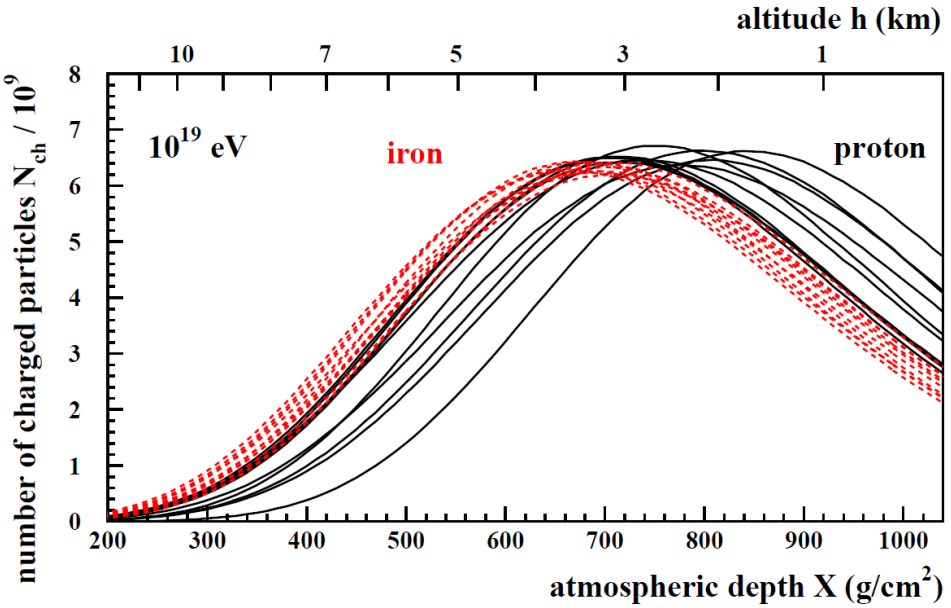


Figure 3.1: *A set of simulated $X_{\max}$ profiles of individual air shower events induced by 10^{19} eV pure proton and pure iron primaries. The extent of horizontal spread of each set is an indication of the magnitude of $\sigma(X_{\max})$ [77].*

Here, f_A represents the fraction of the total population whose atomic mass is A and σ_A is the $\sigma(X_{\max})$ expected for cosmic rays of the same atomic mass. It is clear from Eqn. 3.11 that if the measured values of $\sigma(X_{\max})$ are low, then the arriving nuclei at that particular energy are heavy. This is because the fraction of protons or light nuclei in this case must be small, as dictated by the $f_i(X_{\max,\mathrm{i}} - \langle X_{\max}\rangle)^2$ terms in the sum, which prevent large contributions from the light nuclei.

One can infer from Eqn. 3.11 that there is no one-to-one correspondence between the value of $\sigma(X_{\max})$ and the mean mass, except in the special case of a pure composition. A method that is based on the Heitler model and statistical theory has recently been proposed for interpreting $\sigma(X_{\max})$ to obtain mass composition [79]. Using this approach, the variance of $\ln A$ distribution is given by:

$$\sigma^2_{\ln A} = \frac{\sigma^2(X_{\max}) - \sigma^2_{\mathrm{sh}}(\langle \ln A\rangle)}{b\,\sigma_p^2 + f_E^2} \tag{3.12}$$

where $\sigma^2_{\mathrm{sh}}(\langle \ln A\rangle)$ is the $X_{\max}$ variance for mean mass A, b is a parameter, σ_p^2 is the variance of a pure proton primary composition and f_E depends on energy according to:

$$f_E = \xi - \frac{D}{\ln 10} + \delta \lg\left(\frac{E}{E_0}\right) \ . \tag{3.13}$$

The parameters E_0, D, ξ and δ have values that depend on the hadronic interaction model used.

3.4 Current Experimental Results

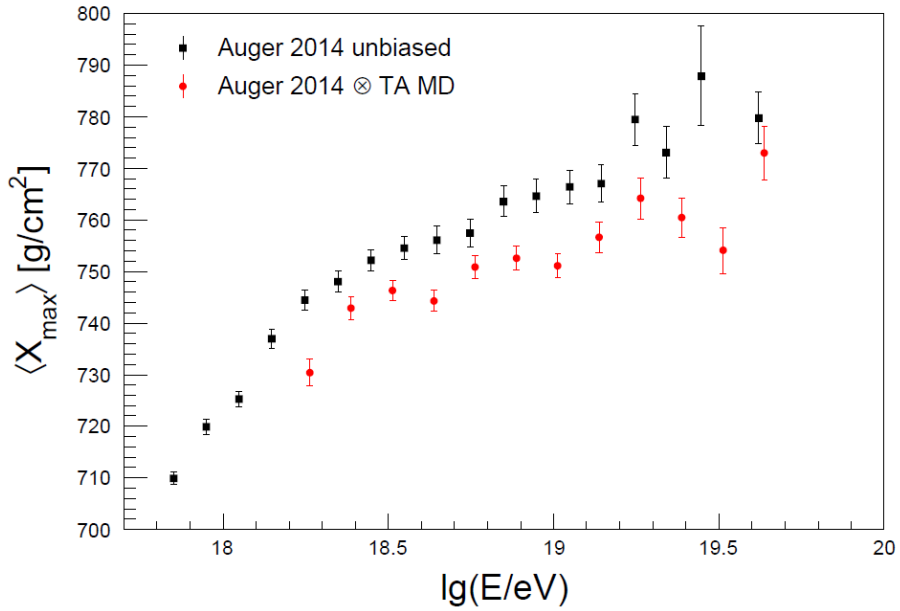
The most recent results of mass composition based on the analysis of $\langle X_{\max} \rangle$ have been released by the two on-going experiments: Telescope Array (TA) and Pierre Auger. The claims by the two collaborations concerning the composition of UHECRs have been apparently contradictory, although they use the same detection procedure. According to the TA, the composition of UHECRs up to the highest energy measured is compatible with a predominantly proton composition [80]. The Pierre Auger experiment on the other hand has consistently reported a composition that initially gets lighter up to $\lg(E/\text{eV}) \approx 18.3$, before beginning to get heavier beyond this point [49]. At $\lg(E/\text{eV}) = 19.6$, the Auger results might indicate that the trend of increasing mass comes to an end [49, 81].

The difference in the conclusions by the two collaborations should be viewed in light of the difference in the analysis procedures used in each case. The values of $\langle X_{\max} \rangle$ as published by TA result from distributions which are folded with detector resolution and efficiency, and interpreted using MC predictions, also folded with detector resolution and efficiency [82]. On the other hand, the Auger $X_{\max}$ distributions used are a result of selection based on shower geometry as described in detail in section 5.1.2. The Auger approach allows an interpretation that does not need to take into account the detector properties.

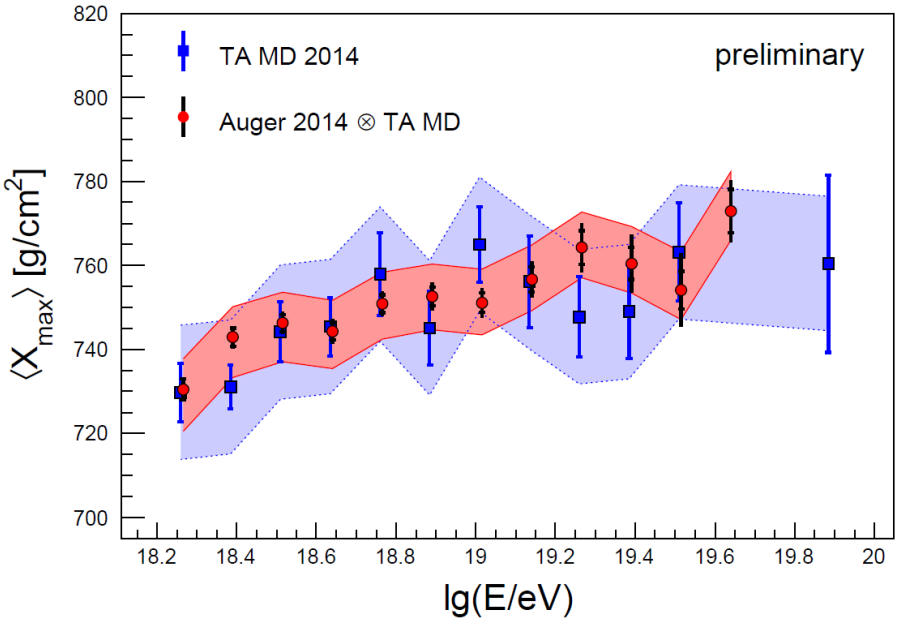
In a recent joint study by TA and Auger collaborations, the distributions of $X_{\max}$ published by the two have been compared [82]. The Auger distributions were fit using simulated events resulting from primary nucleons of proton, helium, nitrogen and iron. The primary fractions giving the best fit were simulated through the FD and SD of the TA and analyzed using the TA procedure for data analysis. The results are as shown in Fig. 3.2.

Although the possible effects of the differences in the energy scales of the two experiments was not taken into account, the small average difference in the data points of $2.9 \pm 2.7(\text{stat.}) \pm 18(\text{syst.})$ g/cm^2 suggests a very good agreement between them. The energy scale effect is not expected to introduce any significant change in the differences. It should be noted that in the TA analysis, MC simulation was done using the QGSJetII03, which has long been upgraded. Furthermore, the number of events observed by the Pierre Auger Observatory exceed those of the TA by a factor of ~ 10 and the sensitivity of the Pierre Auger detector is higher than that of the TA.

A study of composition based on the distributions of $X_{\max}$ measured by the Pierre Auger Observatory [81] has reported results that suggest that the primary composition at $\lg(E/\text{eV}) \sim 19.0$ can neither be dominated by protons nor contain a large proportion of iron nuclei, unless major adjustments are made to the hadronic interaction models used in the study. The best quality fits of simulated values to the measured data yielded the fractions of different elements shown in Fig. 3.3. The fraction of protons is observed in Fig. 3.3 to rise to a maximum of over 60% in the ankle region, before dropping to near zero just above 10^{19} eV. This large fraction of protons just below the ankle is unexpected if the ankle is produced as a result of a transition from galactic to extragalactic cosmic rays. At the same time, a low proton fraction is not consistent with the expectations of the dip scenario.



(a) Effect of TA detector acceptance on Auger $X_{\max}$



(b) Comparison of Auger and TA $\langle X_{\max} \rangle$

Figure 3.2: *(a) The change in $\langle X_{\max} \rangle$ of the Auger $X_{\max}$ distribution after being simulated through the TA detector. (b) A comparison of $\langle X_{\max} \rangle$ of simulated Auger $X_{\max}$ distribution as measured by the TA detector with the $\langle X_{\max} \rangle$ of the actual Auger data after folding with the TA acceptance [82]. The systematic uncertainties of the $X_{\max}$ scales of each experiment lies within the respective colored band.*

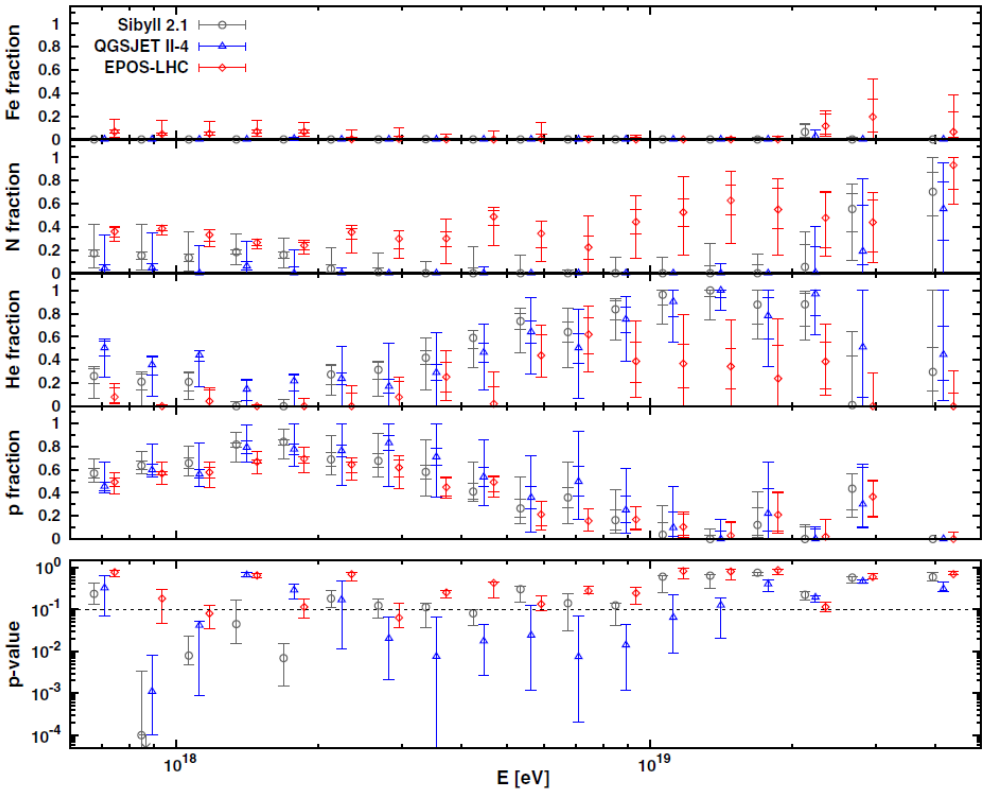


Figure 3.3: *Fractions fit of p, He, N and Fe nuclei obtained for the best fit of simulated distribution to Auger data [81]. The fit quality is represented by the p-values in the bottom panel.*

4 The Pierre Auger Observatory

In this chapter, a brief description of the Pierre Auger Observatory [83] is presented. The observatory has been in operation from January 2004 and has collected data continually since then. It is made up of two main parts: the surface detector (SD) and the fluorescence detector (FD). A combination of the two detection techniques is known as a hybrid detector. Since this study was more involved with the FD, more weight is given to the description of the FD than the SD. However, the data acquisition procedure of the SD is given in more details in Sec. 5.2.

The Pierre Auger Observatory is situated close to the town of Malargüe within the province of Mendoza in the West of Argentina (69°W, 35°S, 1400 m a.s.l). It covers a ground surface area of 3000 km^2 at the base of the Andes mountains in the flat *Pampa Amarilla* (prairies), with an overall variation in altitude which is less than 300 m. The layout of the detector is depicted in Fig. 4.1.

4.1 The Surface Detector

The surface detector consists of an array of 1660 water Cherenkov detectors together with their associated electronics deployed on a 1500 m hexagonal grid [83]. An additional infilled array of 61 detectors with a spacing of 750 m cover an area of 23.5 km^2. Each of the surface detector stations includes a circular 3.6 m diameter tank, the inside of which is lined with a reflective material, and which contains 12 tonnes of pure water. By exploiting the Cherenkov effect, the water is used as the medium for detecting the Cherenkov light produced when charged relativistic air shower particles pass through it. To collect the Cherenkov light, three nine-inch diameter photomultiplier tubes are symmetrically installed at the top of the tank and at a distance of 1.20 m from the center of the tank facing downwards [83, 84]. Windows of clear polyethylene provide optical coupling between the photomultiplier tubes and the water. Each surface detector station is self-powered by the use of a solar-powered system. Communication between each SD station and the central data acquisition system is by means of wireless technology. A photograph of a surface detector and its schematic representation are shown in Fig. 4.2.

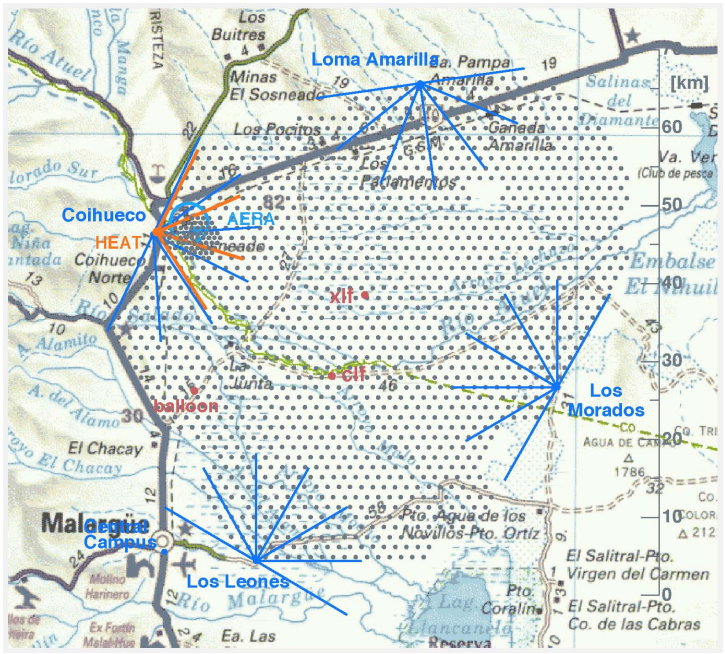
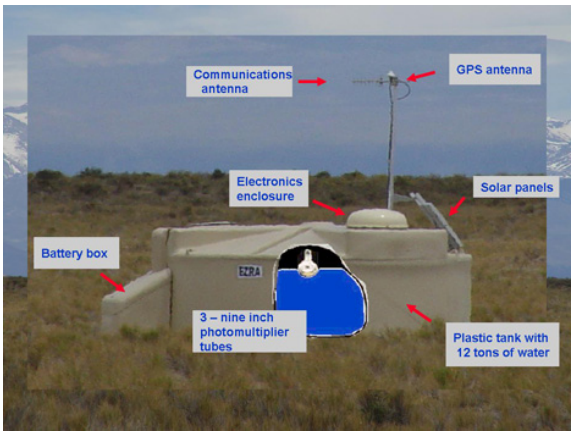


Figure 4.1: *The geographical location of the Pierre Auger Observatory. The black dots represent the detector stations of the surface detector (SD). The fluorescence detector (FD) is housed in four buildings in the locations called 'Los Morados', 'Loma Amarilla', 'Coihueco' and 'Los Leones' on the periphery of the SD array, marked by blue dots. The field of view of the FD cameras are represented by the blue lines. A later enhancement, called the infilled array, is represented by more closely packed SD stations close to 'Coihueco' [83].*

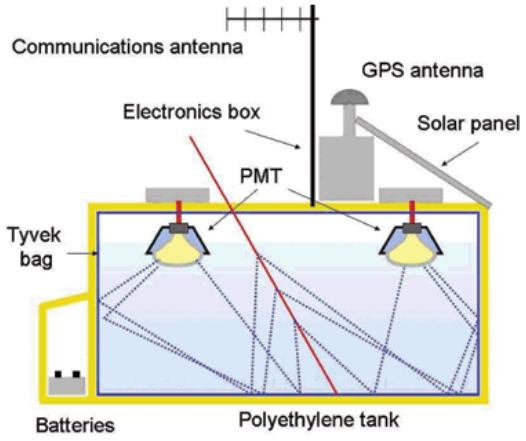
4.2 The Fluorescence Detector

The fluorescence detector (FD) is located atop four hills on the perimeter of the SD, and overlooking it [83]. It consists of 24 fluorescence telescopes which are housed in 4 buildings on top of the hills at the locations named Los Morados, Loma Amarilla, Coihueco and Los Leones as shown in Fig. 4.1. An additional three telescopes with adjustable vertical elevations overlook the SD infilled array. Each building is clean and climate-controlled, in order to prevent accumulation of dust particles, which have the potential of reducing the reflectivity of the telescope mirrors. Every building houses a cluster of 6 independent fluorescence telescopes, each of which has a field of view of $30° \times 30°$ in azimuth and elevation. Only a summary of the main features of the detector are presented here. More details can be found in [83, 85].

The combination of the 6 telescopes in each cluster enables an azimuth view into the SD totaling to 180°. The arrangement of four FD sites overlooking the SD stations ensures that the whole region covered by the SD stations is fully viewed by the FD. This reduces the corrections that have to be made in order to account for uncertainties in atmospheric transmission of nitrogen fluorescence light.



(a) An image of a surface detector station



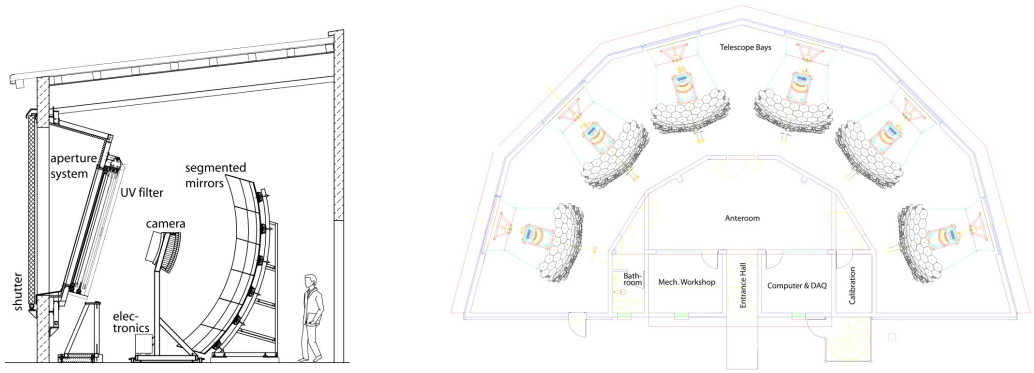
(b) A schematic layout of an SD station

Figure 4.2: *The major components of a surface detector (SD) station. The physical location of each component is indicated in the top panel, while the bottom panel depicts the schematic layout of such a Cherenkov detector [83].*

4.2.1 Design of a Fluorescence Telescope

The Auger fluorescence telescopes are designed to measure the longitudinal profile of the cascade of the electromagnetic component of an extensive air shower. This design, which had already been used successfully in the Fly's Eye experiment [86], relies on the collection of nitrogen fluorescence light which is emitted isotropically during the propagation of an air shower. The fluorescence light is produced during the de-excitation phase of nitrogen molecules, which had initially been excited as a result of their interaction with charged particles contained in the air shower.

A schematic view of a Pierre Auger telescope, together with the arrangement of such telescopes comprising an FD station is shown in Fig. 4.3. Nitrogen fluorescence light is let into the telescope through a circular diaphragm of radius 1.1 m, covered with a filter, which restricts the range of the wavelength of light allowed in. This results in a reduction of the background light flux and an improved signal-to-noise ratio of the measured signal. Visible light is absorbed, while only



(a) Fluorescence telescope (b) Layout of an FD station

Figure 4.3: *(a) A schematic view of an FD telescope of the Pierre Auger Observatory showing its major components and (b) a schematic layout of a building housing the telescopes [83, 85].*

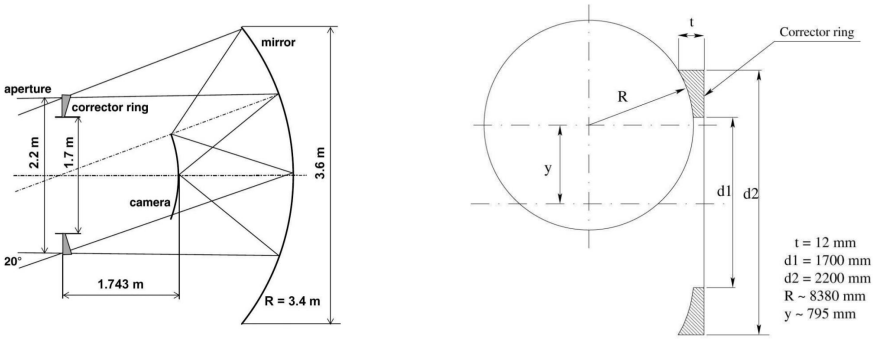
UV photons of wavelength $\sim 290 - 410$ nm, which is almost the full spectrum of nitrogen fluorescence, are transmitted [83, 85]. The filter also shields the telescope and its electronics, thus keeping it clean and climate controlled. The shutters close automatically during day light and at night when it becomes too windy or when rain begins to fall.

Various means are used to minimize light losses and reduce the cost of a telescope. Coma aberration is eliminated by the use of a simplified annular ring which is mounted on the aperture. The design, which in addition to the corrector ring also includes the aperture, the mirror and the PMT camera, partially corrects spherical aberration. It is a modified form of a Schmidt camera i.e. it has a large aperture. Since its area is large, the spherical mirror is segmented to reduce its overall cost and weight.

The telescope mirror elements used at Los Leones and Los Morados sites are produced by covering aluminium blocks with a sheet of $AlMgSiO_5$ reflective material, which is in turn coated with aluminium oxide for protection. The rest of the telescopes at Coihueco and Amarilla are made using mirrors with segments of light-weight glass. This type of glass, called SIMAX, remains stable both thermally and mechanically at different temperatures. A layer of aluminium, coated with SiO_2 for mechanical protection, constitutes the reflecting surface. The geometrical parameters of an FD telescope, as well as the profile of a corrector ring are shown in Fig. 4.4. The results of simulation of the optical system using Geant4 [87, 88] have been found to agree with all the design specifications.

4.2.2 The Telescope Camera

The camera is made up of hexagonal pixels arranged in a matrix of 22 rows by 20 columns. The pixels lie on the focal surface of the optical system, and hexagonal photomultiplier tubes (PMTs) are fixed such that they fit in the pixels. A cosmic ray event is recorded on the camera as a pattern of activated pixels with a clear



(a) FD telescope geometry (b) Corrector ring design

Figure 4.4: *(a) The geometrical layout of a fluorescence detector and (b) a cross-sectional view of a corrector ring. Spherical approximation is used in the design of the corrector ring. The lower horizontal dash-dotted line in (b) is the optical axis of the telescope while d2 is the diameter of the aperture [83, 85].*

time sequence.

It is unavoidable to leave some space between the photocathodes for the convenience of their safe packaging on the focal surface. In addition to this, the effective cathode area is smaller than the area covered by the PMT glass envelope. Light collectors are therefore used to compliment the PMTs so as to maximize the amount of light collected, and at the same time ensure a sharp transition between neighbouring pixels. The light collectors take the form of a simplified Winston cone consisting of a combination of six "Mercedes stars". A Mercedes star has three arms, which are separated from each other by an angle of 120°. Such a "star" is positioned on every pixel vertex. The geometrical arrangement of the light collectors around a pixel is shown in Fig. 4.5. The use of Monte Carlo ray tracing has shown that on the average, the Mercedes stars light collectors increase the efficiency of light collection from 70% to 94% [83, 85].

4.2.3 Operation of the FD

The FD stations are operated securely and remotely from the central campus in Malargüe. The fraction of operational time during which the FD actually acquires data is limited by factors such as the observed sky being too bright due to the presence of sunshine or bright moonlight and dangerous weather conditions such as high wind and snow. During such times, no data is collected. In this way, the quality of the recorded data is assured, and at the same time, the detector components are safeguarded against damage. The fraction of operational time is referred to as up time. For the whole period of time during which the data analyzed in this study was collected, the average up time fraction was approximately 15% of the total time of data collection.

Background conditions are quantified in terms of ADC counts. The ADC counts are derived from an analysis of the variation in ADC signals during each night of data taking. Typical values of background signals in units of $(\text{ADC counts})^2$ range

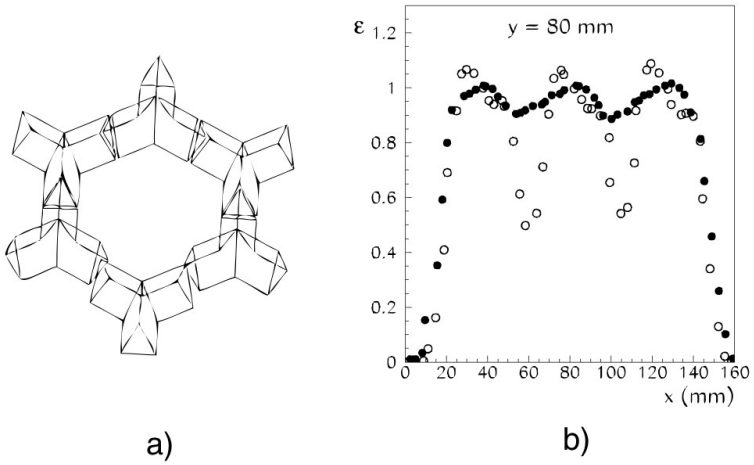


Figure 4.5: *(a) An arrangement of six "Mercedes stars" around a pixel. Each arm of a Mercedes star is separated from the other by an angle of 120°. Note a slight displacement of the bottom "star", deliberately done for clarity. (b) Results of the measurement of light collection efficiency, which was performed with a light spot moving along a line passing through three adjacent pixels. The full (open) dots show the results obtained with (without) Mercedes stars. Note the marked increase in efficiency in the region between two pixels when Mercedes stars are used [83, 85].*

from 3 for electronic background to 60 for clear moonless nights and several hundred when the moon is above the horizon. The optimal conditions for the observation of an event which can be reconstructed with an uncertainty in energy less than 20% range from 25 to 60 ADC counts. Such high quality events, for which the $X_{\max}$ may be determined with an uncertainty better than 40 g/cm^2, are the ones that were used in this work.

4.2.3.1 Event Reconstruction

The reconstruction of events is carried out by taking into account the timing information from the SD stations and the FD pixels. An extensive air shower event that is detected simultaneously by the FD and at least one SD is referred to as a hybrid event. A hybrid event which triggers a sufficient number of SD stations such that it can be reconstructed by using only SD techniques independent of the FD, is called a "golden hybrid". The FD detects cosmic shower events as a series of triggered pixels in the camera, an example of which is shown in Fig. 4.6. From the pulse of light registered by an event, one may obtain the time of the pulse and its uncertainty. The best fit obtained for the differences between the expected and actual times gives the shower axis. In order to perform the reconstruction of the shower event, one requires to determine the plane which passes through the shower axis and the eye of the telescope, called the shower-detector plane (SDP). It is within this plane (see Fig. 4.7) that the shower axis is reconstructed, using the times of pixel trigger.

The shower axis may be characterized by the parameters indicated in Fig. 4.7

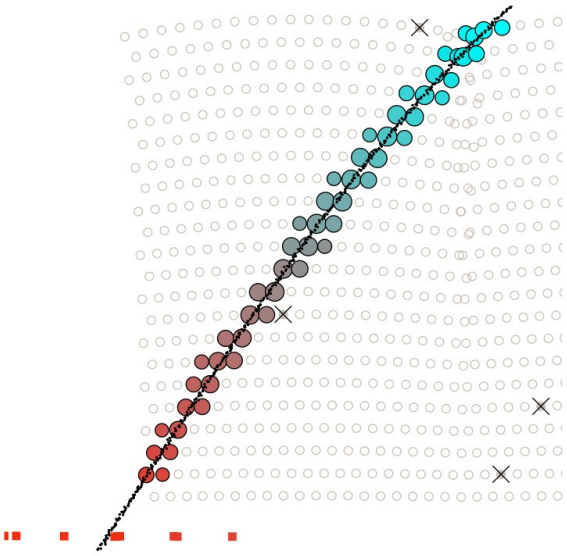


Figure 4.6: *A fluorescence telescope view of a hybrid event, also detected by surface detector stations indicated by red squares at the bottom. The pixels are triggered in the sequence indicated by the colors and the continuous line is the fitted shower detector plane (SDP). The pixels that had a signal within the time of trigger but were rejected because of being too far from the SDP either in time or distance, are marked with crosses [83, 85].*

as R_p and χ_0. They represent the shortest distance between the camera and the track and the angle between the track and the horizontal respectively. If the shower front passes the line R_p at time t_0, the resultant light is registered at the i^{th} pixel at the time

$$t_i = t_0 + \frac{R_p}{c} \tan\left(\frac{\chi_0 - \chi_i}{2}\right). \tag{4.1}$$

A fit of data points to this function yields the shower parameters. However, if the measured angular speed $d\chi/dt$ changes very little over the observed track length, many good fits are possible, hence uncertainty in other shower parameters. This degeneracy is solved by using the timing data from SD stations. By following this procedure, a hybrid event is reconstructed. To determine the reconstruction uncertainties, light scattered from laser pulses with known geometries are used.

With the knowledge of the shower geometry, one can obtain the energy deposited in the shower as a function of slant depth from the light deposited at the aperture as a function of time. However, corrections have to made for light attenuation from the shower to the telescope and light originating from other sources. An estimate of the calorimetric energy of the shower is obtained by integrating the log-likelihood fit of a Gaisser-Hillas function [89], f_{GH},

$$f_{\text{GH}}(X) = \left(\frac{\mathrm{d}E}{\mathrm{d}X}\right)_{\text{max}} \left(\frac{X - X_0}{X_{\text{max}} - X_0}\right)^{(X_{\text{max}} - X_0)/\lambda} e^{(X_{\text{max}} - X)/\lambda} \tag{4.2}$$

to the number of photoelectrons registered in the PMTs of an FD. The energy that

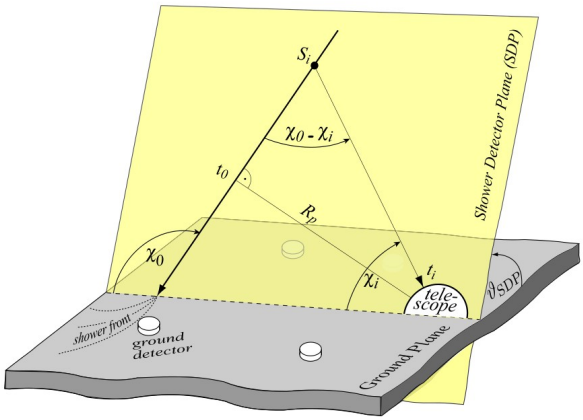


Figure 4.7: *Geometrical drawing illustrating the use of the shower observables of the fluorescence detector telescope to reconstruct the shower-detector plane (SDP) [83, 85].*

is carried away by neutrinos and high energy muons has to be corrected for in order to obtain the total shower energy. An estimate of the required correction has been made using hybrid data [90]. The depth of shower maximum is obtained from the same fit, as a parameter of the fit. An example of a measurement of photoelectrons at a telescope aperture including the reconstructed light contributions is shown in Fig. 4.8. The energy profile obtained from this measurement is shown in Fig. 4.9.

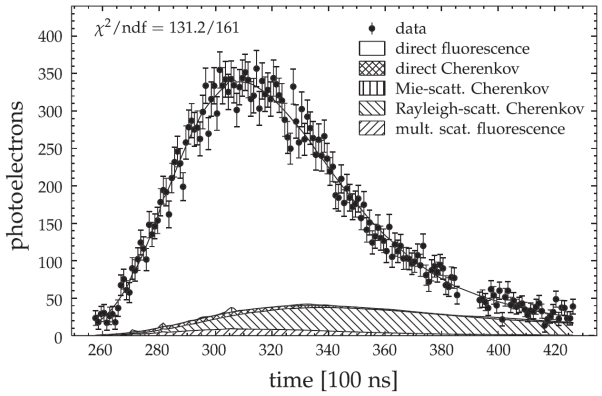


Figure 4.8: *An example of the result obtained for the measurement of light at the aperture of an FD telescope (dots). The hatched areas show other sources, also contributing to the recorded light, as determined by reconstruction [49, 83].*

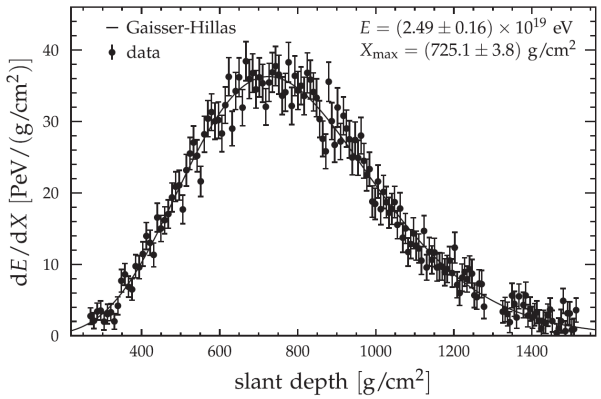


Figure 4.9: *The energy deposit profile reconstructed from the results in Fig. 4.8. The line represents a Gaisser-Hillas fit of the profile. The indicated calorimetric energy of this event is obtained by integrating the fit [49, 83].*

4.3 Enhancements to the Observatory

4.3.1 HEAT

An additional three telescopes with fields of view that are more elevated compared to the original 24 were installed and commissioned in September 2009 [91]. They are called the High-Elevation Auger Telescopes (HEAT). These three telescopes, whose design is similar to the original 24 except that they can be tilted upwards by 29°, are located 180 m from the Coihueco FD building in the north-east direction. A schematic view of such a telescope is shown in Fig. 4.10.

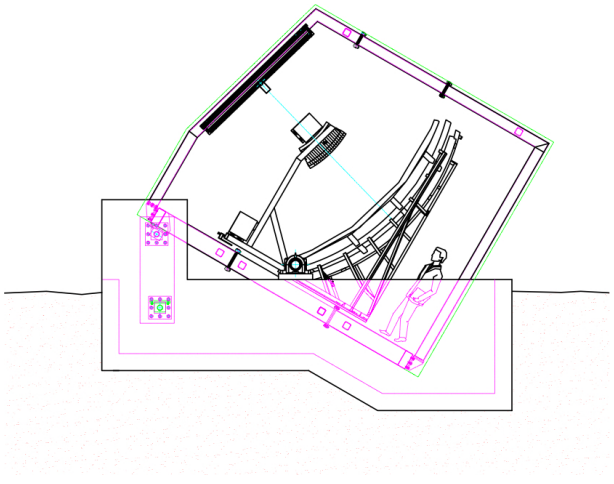


Figure 4.10: *A schematic view of one of the HEAT telescopes in a tilted position, ready to take data . The telescope may be positioned horizontally during servicing and cross-calibration [83, 85].*

Differences between the first FD telescopes and HEAT telescopes arise due to

the necessary tilting mechanism required in HEAT. Tilting is done by an electrically driven hydraulic system and the HEAT telescopes are individually housed in pivot-mounted enclosures. The data acquisition in HEAT has been improved, including a doubling of the sampling rate of the digitizing system from 10 MHz to 20 MHz. HEAT acts effectively as a fifth telescope site, independent of the first four.

4.3.2 AMIGA

Alongside the installation of HEAT, 61 extra SD stations covering an area of 24 km^2 have been deployed as an infilled array among the already operational SD stations, reducing the spacing of the triangular grids in this region to 750 m. The additional SD stations are part of the Auger Muons and Infilled for the Ground Array (AMIGA). An incorporated feature of AMIGA are scintillation muon detectors, installed close to the AMIGA water Cherenkov detectors, and designed to measure directly the muon content of the detected extensive air showers [92]. The distance of separation of SD stations within the infilled is 433 m close to HEAT and 750 m further away (see Fig. 4.11). The infilled array is located close to HEAT and is in its field of view. The combined operation of AMIGA and HEAT allows the lowering of the Auger hybrid detector energy threshold to 10^{17} eV. Coupled with this, the combination now brings about the possibility of coincidentally detecting an event using the independent techniques of the FD, SD, and muon detector, hence obtaining higher quality data. A direct comparison of the Auger data with results of other experiments e.g. KASCADE-Grande becomes realizable, due to their overlapping operating energy ranges.

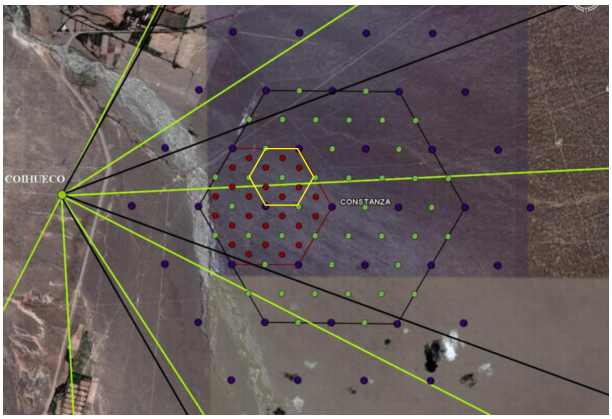


Figure 4.11: *The infilled array of the Pierre Auger Observatory. The largest hexagon encloses the SDs with a spacing of 750 m (blue dots: first installed, green dots: infilled). The second largest hexagon encompasses the 433 m spaced SDs, where the ones shown in red are yet to be deployed. The yellow hexagon highlights a unitary cell, while the the black and green lines represent the fields of view of the Coihueco and HEAT telescopes respectively [93].*

4.3.3 AERA

Another enhancement which is under development is the Auger Engineering Radio Array (AERA) . It is aimed at the detection of the radio emission which accompanies the evolution of showers in the atmosphere in the 30 to 80 MHz frequency range. The technique used was first proposed in 1965 [94], but has recently been revived and used in the LOPES [95] and CODALEMA [96] collaborations, following improvement in technology.

Already, some events have been detected by AERA [97], which were simultaneously and independently detected by the FD, SD, and AMIGA muon counters. After reconstruction methods for shower parameters have been developed, it will be possible to make more accurate measurements of the shower characteristics using AERA, in conjunction with the other detectors in the Pierre Auger Observatory.

4.4 Atmospheric monitoring

Since the FD measurement of $X_{\max}$ uses the atmosphere as a calorimeter, a knowledge of the atmospheric conditions during the time at which an event is recorded is essential in order to not only reconstruct air showers accurately, but also estimate the exposure of the detectors accurately. This follows since atmospheric variables e.g. the presence of aerosols, clouds, temperature, pressure and humidity affect the longitudinal development of extensive air showers as well as the transmission of the isotropic fluorescent light initiated by air showers. Some conditions such as strong winds and heavy snow pose a danger to the scientific equipment and their detection prompts an orderly shutdown of the FD stations. A full description of the atmospheric monitoring devices used in the Pierre Auger Observatory is found in [83], a brief summary of which is presented here.

The equipment used for monitoring the troposphere above the Pierre Auger Observatory (see Fig.4.12) may be broadly categorized into three viz. those used for measuring atmospheric state variables, those for determining atmospheric transmission conditions and those for detecting clouds and extinction. The data recorded by the installations in the first category are required to determine the necessary adjustments to the SD observations and for an accurate reconstruction of the air showers by the FD. Likewise, the data acquired using the second and third category of instruments has to be taken into account in reconstructing the light generated along the shower axis.

4.4.1 Atmospheric state variables

The variables that show the state of the atmosphere namely temperature, pressure and humidity are measured in ground-based weather stations while their vertical profiles up to about 20-25 km a.s.l. are obtained from balloon-borne equipment. The air density and atmospheric depth are calculated from these directly measured values. The weather stations are equipped with wind speed sensors, while those at Los Leones and Coihueco additionally have sensors for wind direction. Up to December 2010, the meteorological radiosondes were launched regularly, but have

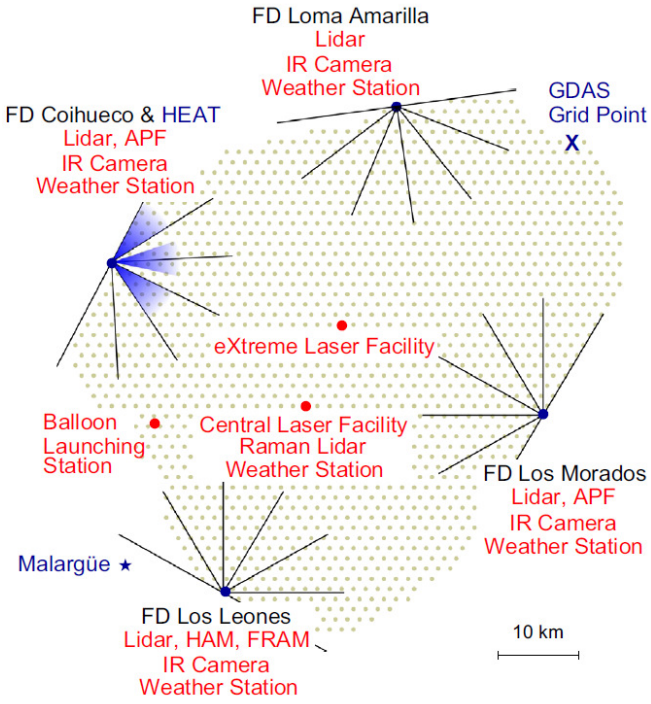


Figure 4.12: *A schematic representation of the atmospheric monitoring devices used by the Pierre Auger Observatory, showing where each device is located within the observatory [83].*

since been replaced by validated data from the Global Data Assimilation System (GDAS). The position of the GDAS grid point used is shown in Fig. 4.12.

4.4.2 Atmospheric transmission

At the wavelength of the fluorescence light induced by extensive air showers, scattering is the dominant cause of attenuation [83]. From the vertical profiles of atmospheric temperature, humidity and pressure, an analytical description of the scattering of photons, $T_{mol}(\lambda, s)$, based on molecular theory, can be calculated. In this analytical description, λ is dependent on the Rayleigh scattering cross-section along the line of sight s. Measurements of the vertical aerosol optical depth $\tau_{aer}(h)$ are applied in the Mie scattering theory to describe the scattering of photons by aerosols using the aerosol transmission factor $T_{aer}(\lambda, s)$. The parameters λ and s in τ_{aer} are equivalent to those in T_{mol}.

The aerosol optical depth profiles are obtained at hourly intervals from the laser tracks of the Central Laser Facility (CLF) and the eXtreme Laser Facility (XLF) [98]. In addition, the Aerosol Phase Function is measured by dedicated measuring devices located at Coihueco and Los Morados FDs. These quantities are required for the reconstruction of every FD air shower event. Another required variable, the angular distribution (phase function) for the aerosols, is obtained using a collimated xenon flash lamp. Light from the lamp is directed horizontally across the fields of view

of the FDs at Los Morados and Coihueco. The results are recorded in the form of light measured by the FD as a function of scattering angle. The heights of clouds directly above the CLF and the XLF can be obtained from an analysis of the laser shots from each one, as detected by the FD in each case [98].

The total horizontal atmospheric attenuation is obtained using the Horizontal Attenuation Monitor (HAM), located close to the Coihueco FD. Horizontal attenuation reveals the dependence of aerosol attenuation on the wavelength of light.

4.4.3 Clouds and extinction

The presence of clouds can distort the detectable signatures of an EAS, thus creating gaps in reliably measured profiles. It also reduces the aperture of the FD. The measurements from infrared cameras and LIDAR systems, together with the FD measurements of CLF and XLF tracks are used by the Auger Observatory to measure cloud cover and height as well as aerosols. The values of $X_{\max}$ recorded during periods when the average cloud cover is more than 25% are excluded from analysis.

The dependence on wavelength of Rayleigh and Mie scattering is measured using the ph(F)otometric Robotic Atmospheric Monitor (FRAM), located near the Los Leones FD station. It is an optical telescope which measures star light, and can operate in the field of view of an FD, since it is passive. Additionally, a Charge Coupled Device (CCD) camera with a large field of view is used to measure atmospheric extinction along the shower detector plane.

5 Data Selection

An overview is given in this chapter of the procedure that was followed in selecting the data that has been analyzed in this work, full details of which can be found in [15,49,66,99]. Mention is made of the pre-selection done on the raw data collected by the FD of the Pierre Auger Observatory as well as the standard procedures followed by the Auger collaboration in specifically selecting the $\langle X_{\rm max} \rangle$ data to be used for analysis. Similarly, a summary of the procedure for SD data selection is given.

5.1 FD Data

5.1.1 Preliminary Selection

The analysis procedure relies heavily on shower parameters which are obtained by reconstruction. Certain precautions are therefore necessary in order to ensure that the shower parameters are reconstructed in a reliable way. Only events satisfying a strict criteria are selected for analysis from the data collected. A summary of the cuts and the percentage of events remaining after each cut, also known as selection efficiency, is given in Table 5.1.

Only the data that was collected during time periods when the conditions allowed for collection of reliable data as indicated by the databases of the measuring instruments described in section 4.4 is selected. Any data collected during a period when calibration constants are not available, or that which is recorded in a telescope with one or more bad pixels is rejected. Similarly, data for which the correction for aerosol attenuation is not available or that which was collected when the atmospheric conditions were very poor e.g. during storms are excluded from analysis. Clouds obscure a part of the profile, which can bias the energy as well as the $X_{\rm max}$ measured by the FD. The presence of clouds is indicated by cloud camera measurements taken at the ground level and from images taken by the Geostationary Operational Environmental Satellites (GOES). An event is accepted if either of the two detectors record no cloud along the direction to the shower. It is also accepted if both the LIDAR and CLF measurements indicate that the base of the cloud lies

at a height which is above the field of view or 400 g/cm^2 above the fiducial depth cut, or if the LIDAR data shows an average cloud fraction of 25% during the data taking period [100].

The event must also satisfy geometrical conditions such that it can be fully reconstructed, together with a simultaneously detected SD event, as a hybrid event. This implies that the reconstructed profile of the event must be good. The energy threshold is set at $10^{17.8}$ eV.

5.1.2 Quality Selection

A bias in the SD resulting from the dependence of its trigger probability on the mass of the primary particle [101] is bound to be propagated to the overall hybrid measurement of $\langle X_{\rm max} \rangle$. Given that the SD trigger efficiency is a function of energy, the distance from the station to the shower core $d_{\rm tank}$ and the zenith angle θ, fiducial volume cuts are applied in order to keep the trigger probability as close to one as possible. Events are only accepted if the average trigger probability is greater than 95%, given their energy, zenith angle and core location. Using the reconstructed distance to the shower axis, the trigger probability can be determined using the function obtained in [102]. The application of this cut on the data set used in this analysis results in the rejection of only about 5% of the events which survived previous cuts, as can be seen in Table 5.1. The rest have a sufficient number of profile points and statistical precision, which allow a full reconstruction of energy and $X_{\rm max}$.

Since the measurement of energy and shower maximum can only be reliable if $X_{\rm max}$ is in the field of view of the FD telescopes [103], a bias is introduced in the event selection resulting from the exclusion of those showers whose $X_{\rm max}$ lie outside the field of view of the telescopes. Even a shower that falls within the field of view may still be excluded if it is very far away from the telescope, since in this case the amount of light the telescope receives is reduced. A Monte Carlo study of the field of view bias has reported biases of up to 12 g/cm^2 for pure proton composition with an energy of $10^{19.5}$ eV [99]. In order to correct for this bias, the procedure that has been adopted for the data used in this work is to determine for every shower, a depth $[X_{\rm low} - X_{\rm up}]$ window within which it is possible to detect the shower given its geometry as defined by its arrival direction and distance to telescope. The selection of showers according to this window is independent of the observed $X_{\rm max}$ of individual events and reduces under-sampling of the tail in the distribution, hence reducing the bias in $\langle X_{\rm max} \rangle$. The efficiency of data selection based on the $X_{\rm max}$ being in the field of view is 69.4%, as seen in Table 5.1.

The farthest distance up to which a shower can be reconstructed reliably is determined on an event by event basis, depending on the derived expected uncertainty of the $X_{\rm max}$. The expected uncertainty, which is greater closer to the field of view boundaries [99], is limited to less than 40 g/cm^2. Since the uncertainty on the reconstructed geometry increases with a decrease in the zenith angle, showers with viewing angle less than 20° are discarded, based on the results of a study of the evolution of $\langle X_{\rm max} \rangle$ with viewing angle [49, 99]. This procedure effectively defines the allowed field of view in terms of $X_{\rm max}$ resolution and the minimum viewing angle,

and thus avoids cuts on the observed profile. In this way, the optimum values of the selection parameters $X_{\rm low}$ and $X_{\rm up}$ are obtained directly from the data rather than theoretical predictions, which would introduce their own uncertainties. The relative percentage reduction in the number of events resulting from cuts related to the field of view are shown as fiducial field of view cuts in Table 5.1.

In addition to the enforcing of the field of view cuts, quality cuts are applied. In order to ensure that the profiles used are of a good quality, some conditions are imposed on the profiles to be used. If there is any gap in the profile, it must not be more than 20% of the observed length. A poor value of the reduced χ^2 of a Gaisser-Hillas function (Eqn. 4.2) fit to the photoelectrons measured by the FD cameras is an indicator of cloud contamination and horizontal non-uniformities of aerosols. In order to exclude such contaminated events, a standard-normal transformation, $z = (\chi^2 - {\rm Ndf})/\sqrt{2{\rm Ndf}}$, is applied to the χ^2 of the profile fit and the showers in the non-Gaussian tail are rejected at $> 2.2\sigma$. The final requirement for a good quality profile is that the observed track length should be more than 300 g/cm^2, to ensure that the observed $X_{\rm max}$ falls within the observed profile range, such that it can be reliably estimated.

Cut	Events	eff.(%)
pre-selection		
air shower candidates	2573713	–
hardware status	1920584	74.6
aerosols	1569645	81.7
hybrid geometry	564324	35.9
profile reconstruction	539960	95.6
clouds	432312	80.1
$E > 10^{17.8}$ eV	111194	25.7
quality and fiducial selection		
P(hybrid)	105749	95.1
$X_{\rm max}$ observed	73361	69.4
quality cuts	58305	79.5
fiducial field of view	21125	36.2
profile cuts	19947	94.4

Table 5.1: *Summary of the number of events remaining after the application of event selection criteria to the Auger data [49]. The selection efficiency in each case is calculated relative to the previous cut.*

The implementation of the cuts described above is expected to result in a good $X_{\rm max}$ resolution, the value of which needs to be known in order to correct for detector effects when comparing the measured and simulated $X_{\rm max}$ fluctuations. The evolution of $X_{\rm max}$ resolution with energy, including the contributing factors, is shown in Fig. 5.1. A suitable tool for determining the resolution contributed by the fluorescence detectors is the REALMC [104]. The simulations done using this code reproduce the data-taking conditions of the Pierre Auger Observatory, which are

time dependent. The resolution obtained using this simulator compares excellently with measured data [100,104]. The contribution of atmospheric aerosols to the $X_{\max}$ resolution is incorporated into the total resolution by adding it in quadrature to the detector resolution. The final value obtained can be cross-checked using events which are measured independently by more than one FD, since the distribution of the differences between two $X_{\max}$ measurements of the same shower has a width which is proportional to detector resolution [99].

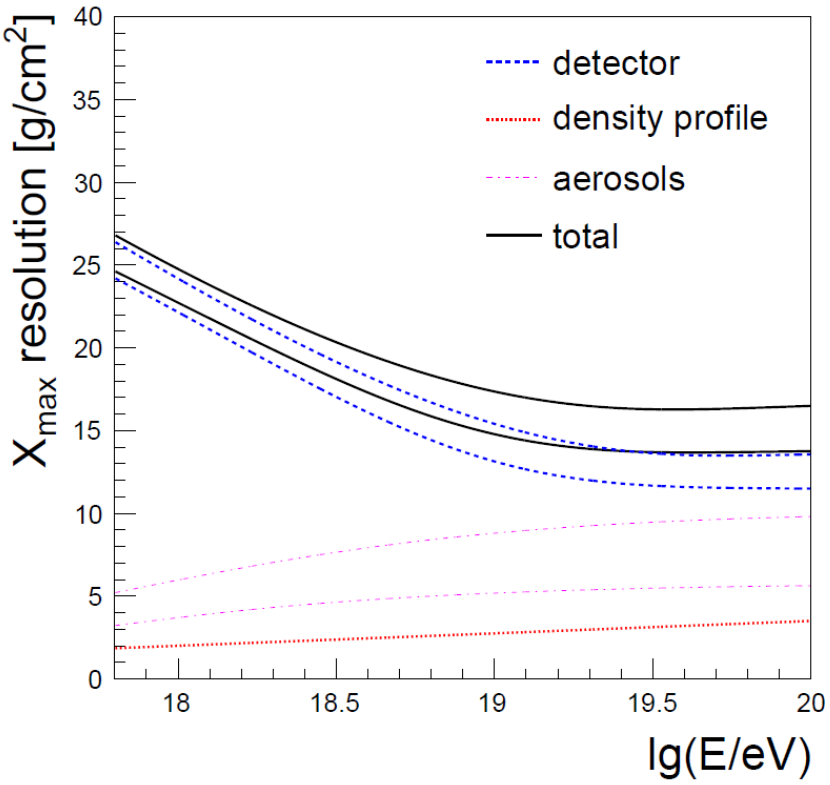


Figure 5.1: *Evolution with energy of the sources of the resolution of $X_{\max}$ measured by the Auger FD detector [100].*

5.1.3 Bias Correction

5.1.3.1 Effect of acceptance

The probability of detecting an air shower at an arbitrary $X_{\max}$ is a function of energy and the magnitude of the $X_{\max}$ itself. To obtain the $X_{\max}$ acceptance, a detailed simulation of the atmosphere, the FD and SD has been performed for the Auger collaboration as reported in [49]. Using a graph that is normalized to give a maximum acceptance of 1, a flat region exists for the intermediate values of $X_{\max}$, while the ends fit a parametrization of exponentially rising and exponentially falling

functions respectively, described by Eqn. 5.1

$$\mathrm{acc}(X_{\max}) = \begin{cases} e^{+\frac{X_{\max}-x_1}{\lambda_1}} & , X_{\max} \leq x_1 \\ 1 & , x_1 < X_{\max} \leq x_2 \\ e^{-\frac{X_{\max}-x_2}{\lambda_2}} & , X_{\max} > x_2 \end{cases} \tag{5.1}$$

where the parameters x_1, λ_1, x_2 and λ_2 depend on the energy. They may be described by a polynomial, $p_i = p_{i0} + p_{i1} \lg(E/\mathrm{EeV}) + p_{i2} \lg(E/\mathrm{EeV})^2$, in which case i represents the respective parameter subscript. A fit of the $X_{\max}$ acceptance with Eqn. 5.1 yields the graph in Fig. 5.2. In the normalized plot in Fig. 5.2, the acceptance expressed

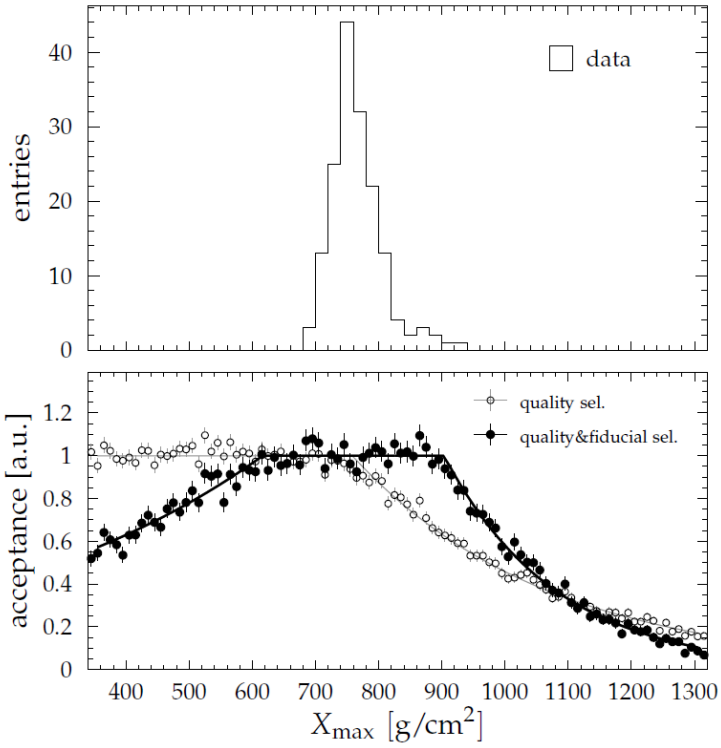


Figure 5.2: Top panel: *The distribution of measured* $X_{\max}$ *at energy* $19.0 < \lg(E/eV) < 19.1$. Bottom panel: *The corresponding relative acceptance after quality cuts (open circles) and after addition fiducial cuts (filled circles) [49].*

in arbitrary units may be considered as the probability of detecting an event at a given energy. It is then obvious that some events outside the flat acceptance region will go undetected as a result of the $X_{\max}$ acceptance of the detector. In order to take these events into account, the mean $X_{\max}$ and weighted standard deviation, σ_w, were calculated as follows:

$$\langle X_{\max} \rangle = \frac{\sum w_i X_{\max,i}}{\sum w_i} \tag{5.2}$$

$$\sigma_w = \sqrt{\frac{\sum w_i}{\left(\sum w_i\right)^2 - \sum w_i^2} \cdot \sum w_i (X_{\max,i} - \langle X_{\max}\rangle)^2} \tag{5.3}$$

where

$$w_i = 1/\text{acceptance}(X_{\max\, i}, \text{Energy}_i) \tag{5.4}$$

is the weight of the i-th event. The weight is one within the flat acceptance region but larger than one outside this region.

5.1.3.2 Reconstruction bias

It has been shown that biases are possibly present in the algorithms used to estimate the hybrid geometry, Cherenkov and fluorescence light contributions as well as the parameters of the longitudinal energy deposit profile [99]. Hence the bias in the reconstruction of every event has been corrected for before beginning the binning process. The full details of the procedure are found in [100]. In this procedure, the results of studies of detector resolution performed using realMC, weighted to match the $X_{\max}$ distribution, yields the correction

$$X^{'}_{\max} = X_{\max} - \mu \tag{5.5}$$

where μ has been determined to be given by $\mu = -3.4 + 0.93 z_{18}$ and $z_{18} = \lg(E/\text{EeV})$. Additionally, the light falling outside an optimal angular region around the shower detector plane within which light at the aperture is collected has to be corrected for by taking into account the lateral width of both fluorescence and Cherenkov light [105]. This is known as the lateral width correction. Based on the results of a recent study [106], the bias introduced by using the lateral width correction is given by

$$\text{b}_{\text{LWcorr}} = 6.5\ \text{g/cm}^2 \left(e^{\frac{\lg(E/\text{eV})-18.23}{0.41}} + 1\right)^{-1}. \tag{5.6}$$

Taking into account equations (5.5) and (5.6), a total bias correction was made for every data event such that the true $X_{\max}$ in each case is given by:

$$X^{\text{true}}_{\max} = X^{\text{data}}_{\max} + 3.4 - 0.93 \lg(E/\text{EeV}) - 6.5 \left(e^{\frac{\lg(E/\text{eV})-18.23}{0.41}} + 1\right)^{-1}. \tag{5.7}$$

5.1.4 Binning of the data

Starting at the lowest bin boundary of $\lg(E/\text{eV}) = 17.8$, the bin widths are set at $0.1\lg(E/\text{eV})$ except for the last one where all the events with energy $\lg(E/\text{eV}) \geq 19.5$ are all binned. The number of events in the bins for combined data varies from over 3000 in the first two bins to about 40 in the last two bins, as shown in appendix A.

5.2 SD Data

The signal registered by the PMTs when Cherenkov light is produced in a ground station is digitized by a Flash Analog-to-Digital Converter (FADC) at the rate of 40

MHz (time bins of width 25 ns). The recorded signal is measured in units of VEM. A minute by minute on-line calibration that is done by identifying the maximum in the atmospheric background muon signal histograms is used to calculate the VEM. The total signal, S, at each SD station within a given time interval, is obtained by integrating the signal traces over that interval.

5.2.1 SD Trigger levels

The SD data acquisition system employs a hierarchical trigger system to filter out background muons and record quality events [107,108]. The two lowest trigger levels, T1 and T2, are performed autonomously by each station.

T1 is implemented in two independent modes to ensure the detection of both the electromagnetic and muonic components of an air shower. The first mode is defined by a signal above a simple threshold of 1.75 VEM in each PMT in an SD station and is denoted as TH-T1. The second mode, referred to as the Time-over-Threshold (ToT) and designated as ToT-T1, captures the contribution to the signal of the electromagnetic component of the shower, which is more spread out in time. The logic chain of the trigger modes is illustrated in Fig. 5.3.

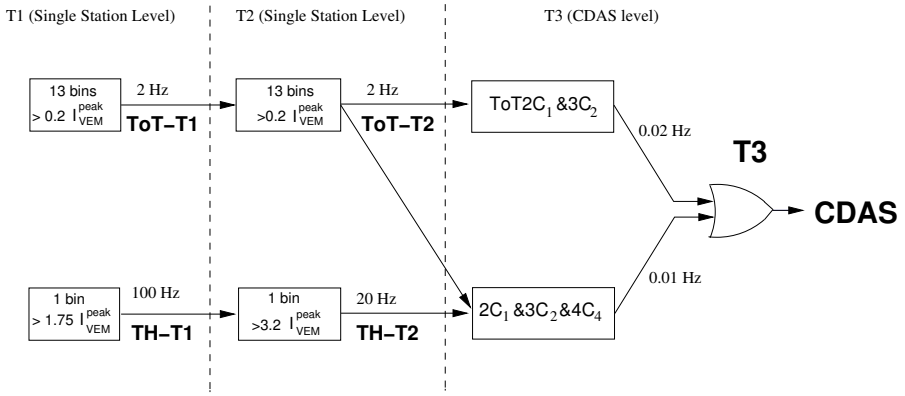


Figure 5.3: *A schematic view of the hierarchy of the Auger SD trigger system [107]. The figures indicate the approximate rates. See Fig. 5.4 for explanation of C_1, C_2, C_3, C_4 and text for explanation of the other notations.*

T2 limits the rate of events that are eventually to be transmitted to the Central Data Acquisition System (CDAS) to a frequency of 24 Hz, which the band width of the communication system can cope with. All ToT-T1 triggers are allowed to move to the ToT-T2 level, while in the case of TH-T1 signals, a simple threshold above 3.2 VEM in each PMT, or 13 time bins above 0.2 VEM in a sliding window of 3 μs in at least 2 PMTs is required.

In the third level, T3, the output of T1 and T2 including the time stamp are transmitted to the CDAS where they are combined. At T3, groups of stations that are clustered in time and space are identified. The requirement for coincidence in time is that the rest of the selected detectors, which must be at least two in number, must have been triggered within a time interval of ± 25 μs after the first one.

The system of identification is based on concentric hexagons centered at each station. Coincidence in space is satisfied in two modes. In the first mode, at least three detectors that passed the ToT condition must coincide. One of the fired detectors must be positioned such that a second fired detector is in its first hexagon of neighbours while the third one is no further than the second hexagon.

In the second, more liberal mode, any type of T2 four-fold coincidence is required. Among the four fired detectors, a station in up to the fourth hexagon is allowed, so long as a station is in the first hexagon and another does not lie beyond the second hexagon. The geometrical arrangement of the second trigger mode is illustrated in Fig. 5.4.

After spatial coincidence is confirmed, it is required that each T2 must be within $(6 + 5n)$ μs of the central station, where n is the hexagon number. This assures the selection of predominantly physics events. For showers whose zenith angles are below 60°, the application of this rule results in 90% of selected events being real showers [107]. At this stage, the event is recorded.

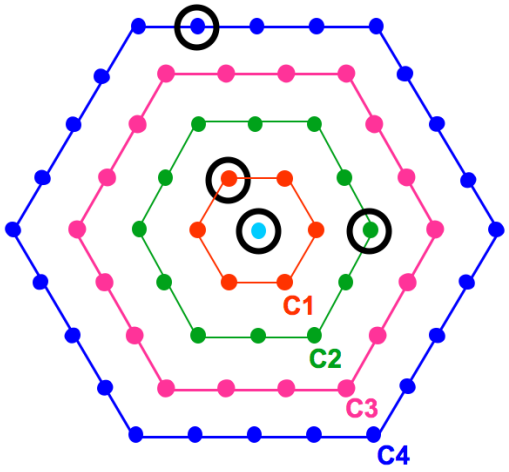


Figure 5.4: *An illustration of a T3 configuration of SD detectors showing the 4-fold mode [107]. The first, second, third and fourth sets of neighbours (hexagons) relative to the detector at the center are indicated respectively by C1, C2, C3 and C4. The distance from a given detector to a neighbouring set increases by 1.5 km as one moves from one set to the next.*

An additional trigger, T4, is applied to the T3 events to select actual physical events and reject accidental ones. A distribution of events selected by the T4 trigger is shown in Fig. 5.5. The final trigger is the quality selection, T5, at which the events which are usable for building up an energy or angular spectrum are selected from the physical events. The events whose shower cores cannot be reconstructed well are rejected at this point. A particular requirement at T5 is that the detector from which a selected event was triggered must have all its 6 closest neighbours working at the time of the event (see the central hexagon in Fig. 5.4). This condition is abbreviated as 6T5.

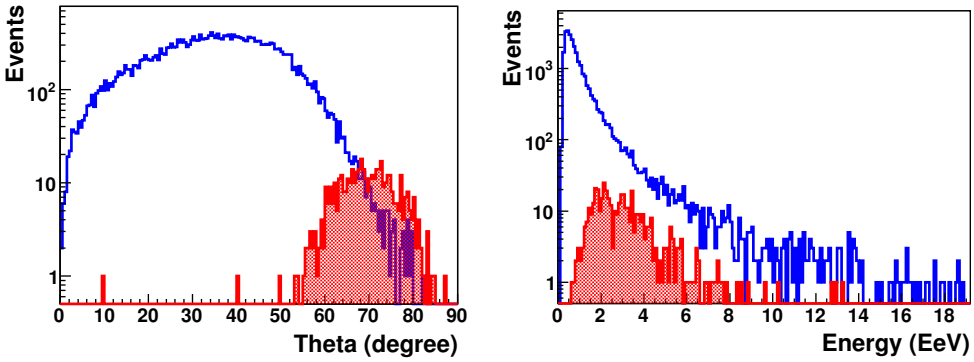


Figure 5.5: *Angular and energy distribution of events selected by the T4 triggers [107]. The filled histograms in both panels are for 4C1 events that are not 3ToT, while the unfilled ones are for 3ToT.*

5.2.2 Cuts in SD data

An SD event qualifies for use in $X_{\max}$ estimation if it is a T4 event that meets the 6T5 condition and at the same time it induces signals larger than 5 VEM in at least four stations [15]. This corresponds to a minimum trigger probability of 0.8 for a single station [107]. The zenith angle is reliable in the range $0° - 60°$, although showers up to 80° are used. If the value of χ^2 obtained in the fit to determine the direction of the shower is compatible with a lightning event, the event is rejected [15]. After reconstruction, some further cuts are applied in order to be certain of the quality of the reconstruction. The distribution used in the time variance model for the SD stations must be well defined in all stations. The time variance $V[T_s]$ used is given by [109]:

$$V[T_s] = a^2 \left(\frac{2\ T_{50}}{n}\right)^2 \frac{n-1}{n+1} + b^2 \tag{5.8}$$

where T_{50} is the time interval within which the first 50% of the total signal is measured, n is the effective number of particles contributing to the shower and a and b are parameters. Further, there must be no time quantities with deviations from the expected values larger than 4σ and the error in the estimated $X_{\max}$ must be smaller than 200 g/cm^2. The final selection efficiencies are shown in Fig. 5.6. Full explanations of the cuts are found in [15].

5.2.3 X_{max} data from SD

The $X_{\max}$ data from the SD that has been used in this study is estimated from parametrization based on shower universality [13]. The data is only for vertical events i.e. events whose zenith angles do not exceed 60°. Use is made of the calculated number of muons, N_μ, arriving at the detector as described in sections 2.3.3.1 and 2.3.3.2.

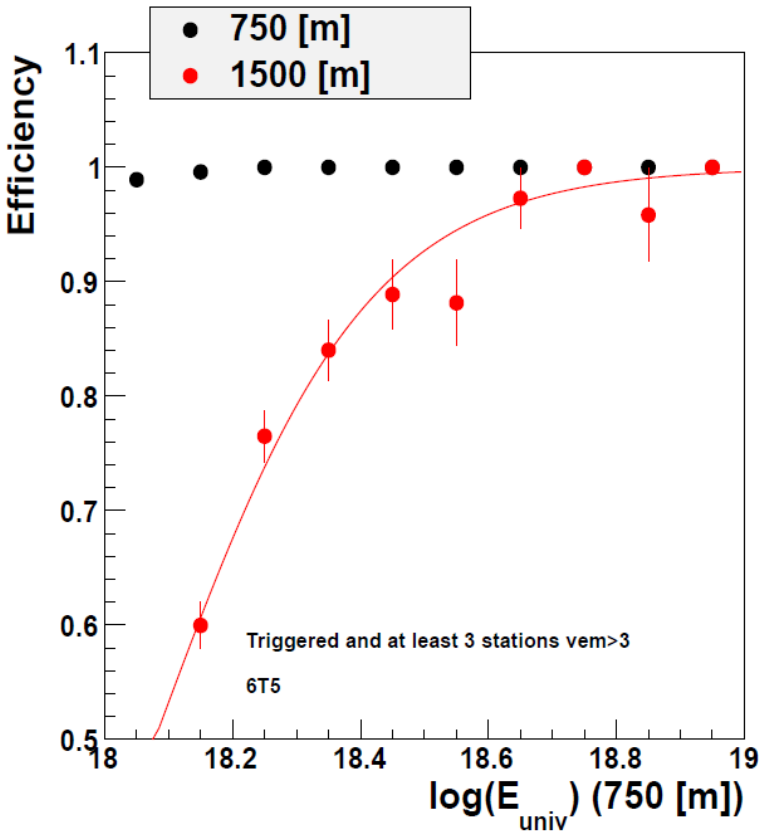


Figure 5.6: *Evolution of the selection efficiency for the 750 m and the 1500 m arrays with energy [15].*

There are two stages involved in the reconstruction of $X_{\max}$. The first stage begins with the signals satisfying the above defined cuts being fitted to the function in Eqn. 2.21 using a likelihood method. The first reconstruction step uses all triggered stations with signals larger than 3 VEM while the energy and $X_{\max}^{\mathrm{SD}}$ [1] are fixed to the average obtained from "golden hybrid" calibration. The calibration of SD energy using "golden hybrid" events is shown in Fig. 5.7.

In the second step, the fit is done using signals from all detectors within a distance of 2500 m from the shower core and whose predicted signal is larger than 3 VEM, while the energy and $X_{\max}^{\mathrm{SD}}$ remain fixed as in the first step.

In the second stage, the start time of the signal or the 1% time quantile (t_{1vem}), together with the 10% (t_{10}) and 50% (t_{50}) time quantiles are used additionally to complete the reconstruction of $X_{\max}^{\mathrm{SD}}$ for every event [15]. The t_{1vem}, t_{10} and t_{50} of a station are flagged as usable if their signals are more than 30, 5 and 5 VEM respectively and the distance from the core of the shower to the detector is less than 2000 m. These time quantities, together with the signals recorded by the SD

[1] This denotation of $X_{\max}$ is used to distinguish the results obtained using SD data from those obtained using the FD data. Where no confusion of the two is likely, the expressions are used interchangeably.

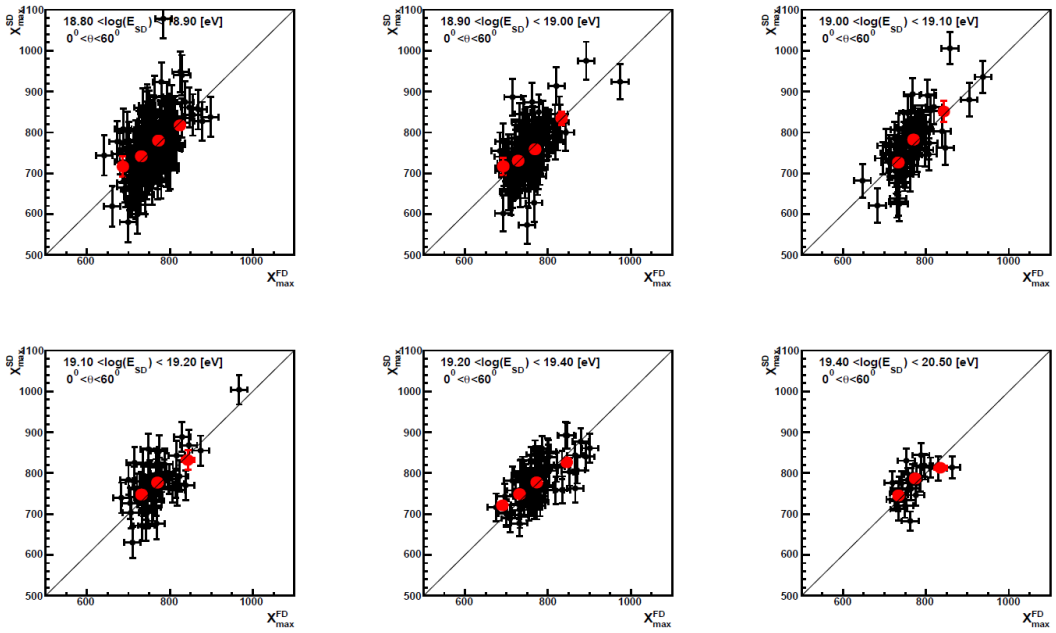


Figure 5.7: *Correlation between $X^{SD}_{\max}$ and $X^{FD}_{\max}$ in different energy ranges for "golden hybrid" events [15]. A black marker represents an event in each case while a red marker represents the average $X^{SD}_{\max}$ in four bins of $X^{FD}_{\max}$.*

detectors, are used in a likelihood minimization to reconstruct energy and $X^{\mathrm{SD}}_{\max}$ of individual events (see [15, 66] for details).

6 Data Analysis and Results

In this chapter, an analysis of ER based on the Auger data of $X_{\max}$ as a function of energy, which was performed in this study, is presented. The data was collected by the Pierre Auger Observatory during the period 1$^{\text{st}}$ Dec 2004 to 31$^{\text{st}}$ Dec 2012 [49]. After all the preselection and quality cuts (see Table 5.1), a total of 19759 events remained for further analysis. The tabulated values of $X_{\max}$ and $\sigma(X_{\max})$ that were published in [49] are shown in appendix A. The binned $X_{\max}$ values that have been calculated in this work are shown in appendix B.

The results of the analysis of the actual data is first presented. As a part of the analysis, the data was truncated into two components, designated as "light" and "heavy". The purpose of this division was to form a data subset that is sensitive to changes in the proton fraction ("light" subset) such that any such changes can be investigated, if they are present. As a part of the data analysis, a description of the procedure that was followed in estimating the bias in the truncated data, and hence the true distribution, is given.

6.1 ER of the Whole Data Set

The $X_{\max}$ data was stored in bins of width $\lg(E/\text{eV}) = 0.1$ from $\lg(E/\text{eV}) = 17.8$ up to $\lg(E/\text{eV}) = 19.5$. All the events with energy higher than $10^{19.5}$ eV were put in the last bin. Thus a total of 18 bins were filled. The data in the bins ranged from about 4000 events in the first bin to about 40 in the last two. A plot of the mean of $X_{\max}$ in the bins as a function of energy is shown in Fig. 6.1.

In this analysis, the acceptance, which is a function of $X_{\max}$, was corrected for by upweighting the number of events where the relative acceptance is less than one. A description has been given in section 5.1.3.1 of the procedure for determining the weight of each event. The errors in the binned data were calculated by using the

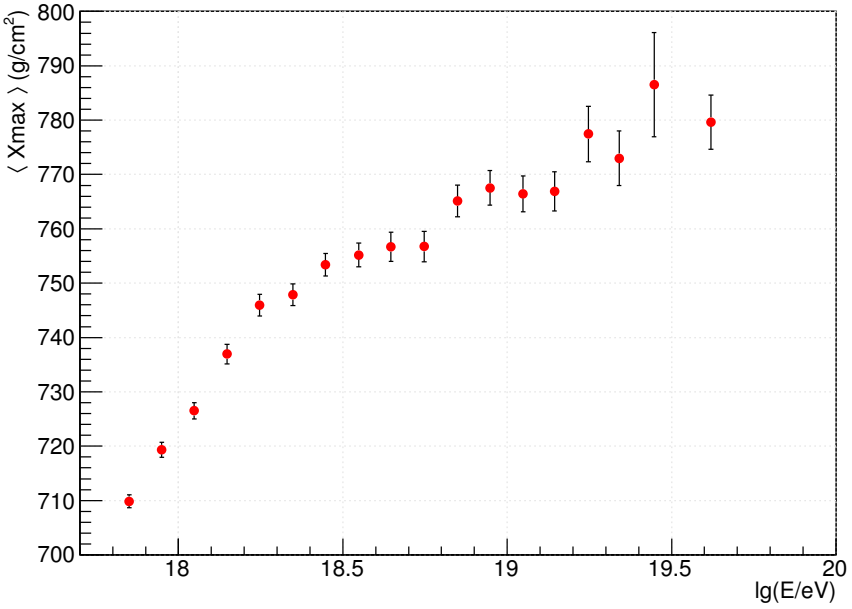


Figure 6.1: *The evolution of $\langle X_{\max} \rangle$ with energy for events remaining after application of selection cuts to the Auger data [49].*

formula suggested by Cochran [110]:

$$\begin{aligned}\sigma^2 = \frac{n}{(n-1)(\sum w_i)^2} \Big[& \sum (w_i X_{\max,i} - \overline{w}\overline{X}_w)^2 \\ & - 2\overline{X}_w \sum (w_i - \overline{w})(w_i X_{\max,i} - \overline{w}\overline{X}_w) \\ & + \overline{X}_w^2 \sum (w_i - \overline{w})^2 \Big]\end{aligned} \tag{6.1}$$

where n is the size of the data sample, w_i is the weight of $X_{\max,i}$, $\overline{w}$ is the mean of the weights and $\overline{X}_w$ is the weighted mean of the sample. The reason for using this formula is discussed further in section 6.3.1.

On the other hand, the published Auger data corrected for acceptance by approximating the distribution of shower maxima of the events in the tails to be exponential. By considering the fraction of events in the tail of a distribution to be η, the exponential slope of the tail, Λ_η, was obtained from a fit to the data in each bin. Using an unbinned likelihood method, the observed damped exponential was hence corrected to be an exponential function, which was then used as the true distribution in the bin. Such a fit and correction are shown in Fig. 6.2. The errors in the binned data published by Auger in [49] were obtained using the bootstrap method [111].

A comparison of the estimations of errors on $\langle X_{\max} \rangle$ obtained by following the procedure in [49], the use of the Cochran formula and standard approach $(\sigma(X_{\max})/\sqrt{N})$ for the full data set and standard binning $\Delta \lg(E/\mathrm{eV}) = 0.1$ is shown in Fig. 6.3. One can see that Cochran estimation compares well with the

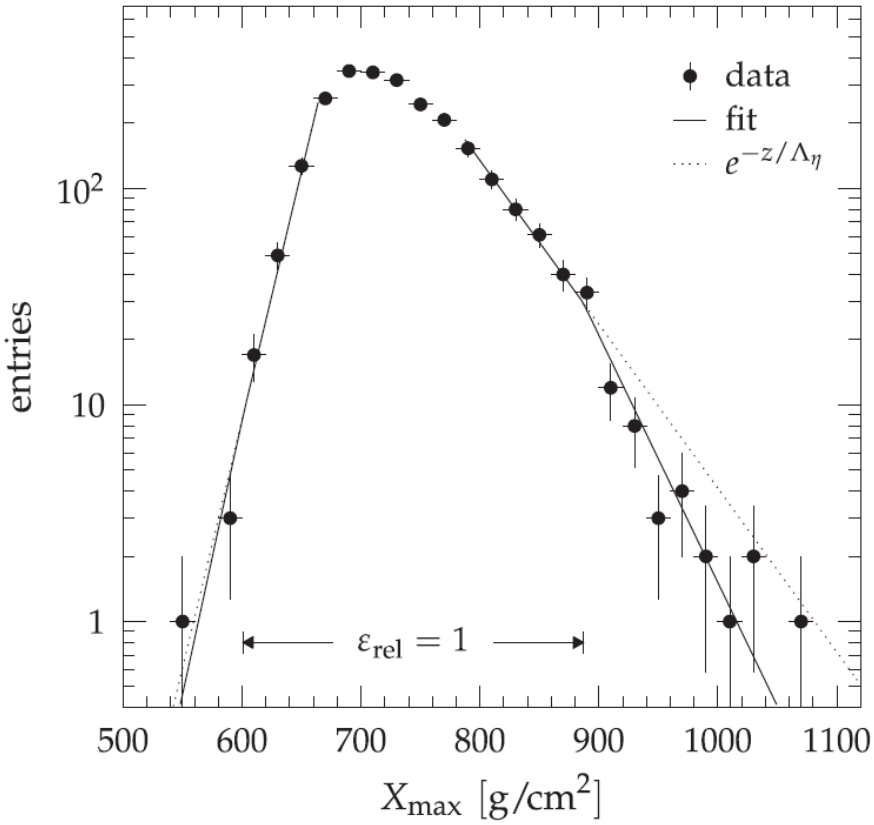


Figure 6.2: *A fit to the tails of the $X_{\rm max}$ distribution in the data bin* ($18.1 < \lg(E/\mathrm{eV}) < 18.2$). *The corrected exponential function is indicated by the dotted lines, while the region where the acceptance is* 1 *is indicated by the arrows [49].*

published error estimation. Since the Cochran approach is technically simpler to use than the bootstrap method, it was adopted for further use in the estimation of errors in the light and heavy subsets of data (see Sec. 6.3).

A fit with a single slope does not describe the data ($\chi^2/\mathrm{Ndf} = 152.7/16$). However, a fit of the data with a broken line function

$$\langle X_{\rm max} \rangle = \begin{cases} X_0 + D_1 \lg(E/E_0), & E < E_0 \\ X_0 + D_2 \lg(E/E_0), & E > E_0 \end{cases} \tag{6.2}$$

gives a good description, with $\chi^2/\mathrm{Ndf} = 7.7/14$, $P = 0.90$, as seen in Fig. 6.4. Included in the top panel of Fig. 6.4 are the ER of proton and iron primaries for different hadronic interaction models, obtained using Gumbel parameterization of $X_{\rm max}$ distributions. The fit parameters obtained in this analysis as well as those reported in [49], by the use of Eqn. 6.2, are given in Table 6.1, where a good agreement between the two is observed. The slight difference in the values of ER can be explained by the different approaches used in the two cases for correcting for acceptance of the detector and also for determining the error in the binned data.

Although the broken line fit provides a good description of the data (Fig. 6.4), a closer look suggests that a structure beyond a simple two-line fit might take place

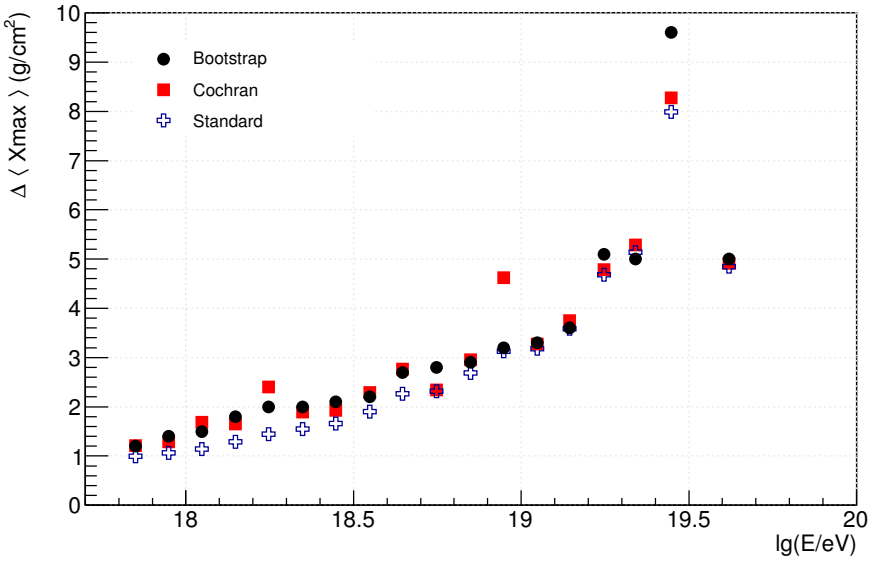


Figure 6.3: *Comparison between the statistical errors on $\langle X_{\mathrm{max}} \rangle$ in the published Auger analysis (bootstrap), Cochran formula and standard approach ($RMS/\sqrt{N}$) for the full data set and standard binning $\Delta \lg(E/\mathrm{eV}) = 0.1$.*

		D_{10} (g/cm^2/decade)		
	$\lg(E_0/\mathrm{eV})$	D_1	D_2	X_0 (g/cm^2)
This analysis	18.26 ± 0.03	90.0 ± 5.0	26.0 ± 2.6	747.5 ± 2.0
[49]	18.27 ± 0.04	86.4 ± 5.0	26.4 ± 2.5	746.8 ± 2.1

Table 6.1: *Fit parameters for $\langle X_{\mathrm{max}} \rangle$ data [49] with the broken line function from Eqn. (6.2). The errors are only statistical.*

(Fig. 6.5). The curve in in Fig. 6.5 is a 5-order polynomial function. The change from a composition which is getting lighter to that which is getting heavier may not be taking place at an instant, but rather gradually over a range of energy. In order to investigate the nature of the structure, ER was estimated in narrow energy ranges. To estimate the slopes D_{10} in given energy ranges, straight line fits were performed using 3-bin sliding windows (Fig. 6.6 and Fig. 6.7).

The same was done for the MC predictions of $\langle X_{\mathrm{max}} \rangle$ from Gumbel parametrizations (Table 6.3) and slopes for all models, primary particles and energies were found to be within $54-62$ g/cm^2/decade range. All these ERs are summarized in Fig. 6.8.

It should be noted in Fig. 6.8, that points with ordinal numbers $1, 4, 7 \ldots$; $2, 5, 8 \ldots$; $3, 6, 9 \ldots$ form three sets, each point in which does not have common data with the neighboring ones. In other words, slopes D_{10} at points such as $1, 4, 7 \ldots$ are not correlated with each other. The magnitude of D_{10} in data decreases from $\approx$ 80 g/cm^2/decade at $E \approx 10^{18.0} - 10^{18.2}$ eV to $\approx 10 - 30$ g/cm^2/decade for

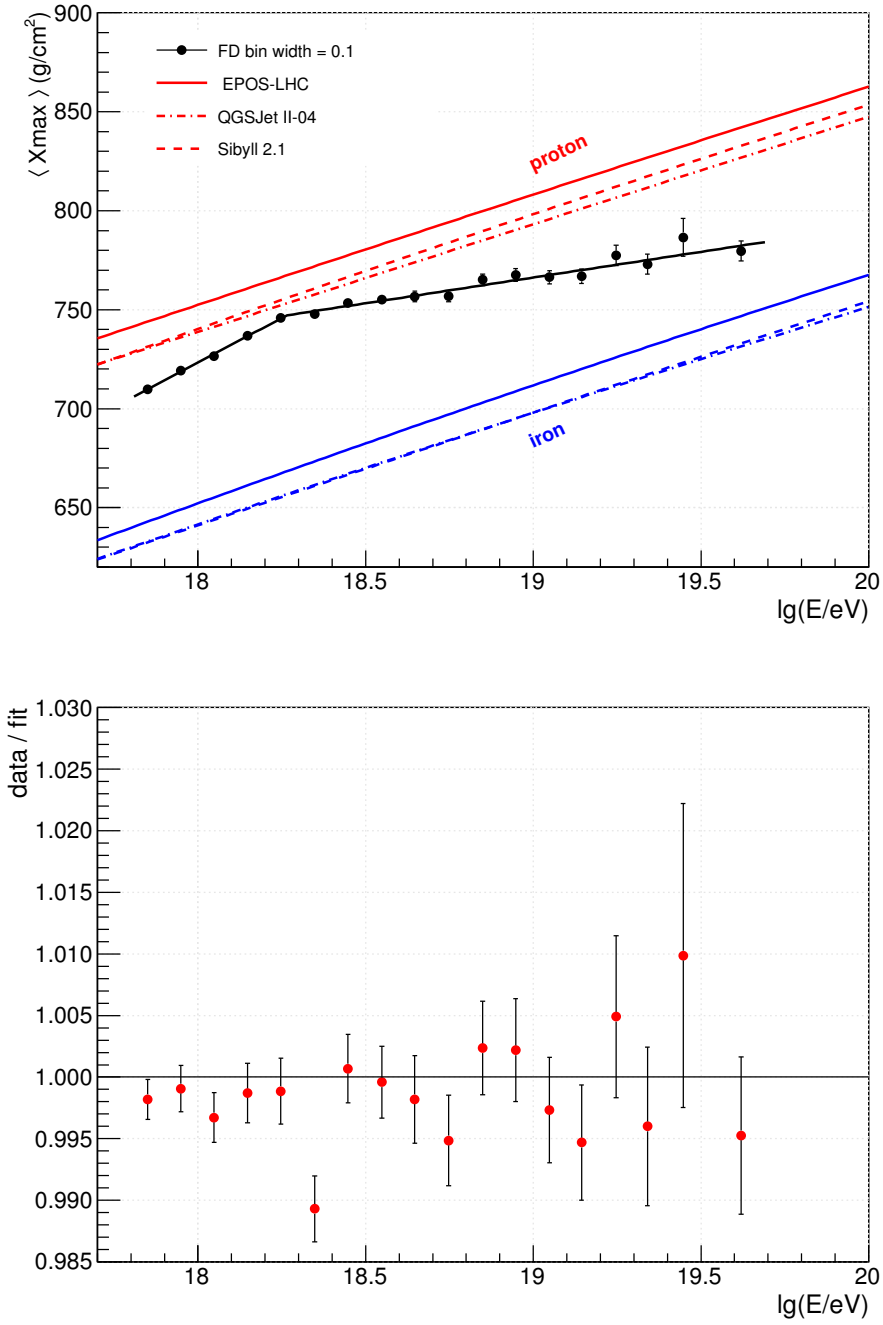


Figure 6.4: Top panel: *Auger* $\langle X_{\mathrm{max}} \rangle$ *data used in this work together with predictions for protons and iron obtained using different interaction models and the Gumbel parametrizations [67]. Parameters of the fit function (black line) are given in Table 6.1.* Bottom panel: *The ratio of* $\langle X_{\mathrm{max}} \rangle$ *in data to the fit function.*

$E \approx 10^{18.3} - 10^{19.4}$ eV. From Fig. 6.8, one can see that within statistical errors, the ER remains constant at these two values within the respective energy ranges. This is an indication that the composition becomes lighter at around 10^{18} eV and then

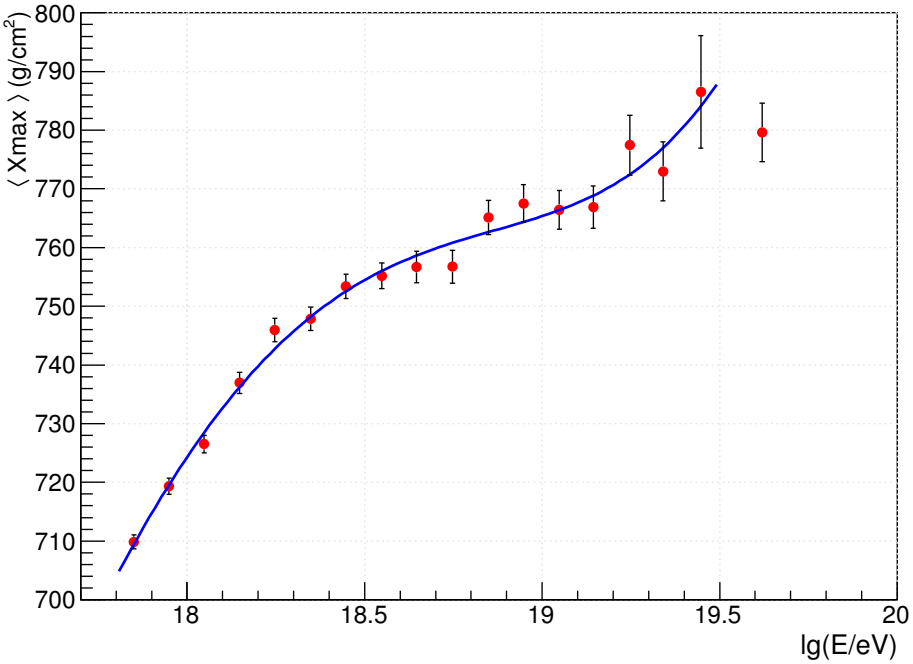


Figure 6.5: *Auger data [49] (zoom-in of Fig. 6.4(a)) with a fit of a 5-order polynomial function, to guide the eye.*

progressively heavier at higher energies. This observation from the analysis done using sliding windows not only confirms the analysis already done by the Auger collaboration for the whole data set using ER, but it also offers new insight into how quickly the transition from light to heavy occurs. A fairly narrow transition region between the energy range where the composition is getting lighter and that in which it gets heavier occurs within the interval of primary energy $\sim 1.3-2.0$ EeV. If the composition around the "ankle" is getting heavier with increase in energy, that implies that it is not pure, but rather mixed. This is in agreement with the results of [115].

An estimation of the systematic errors in the ER before and after the break in the fit, as well as the energy at the point of break, has been carried out. As a first rough estimate, the recorded primary energy was changed by $\pm 14\%$, which is equivalent to the systematic error in the Auger energy scale [116], and the events were rebinned. This is expected to cause a systematic change in the ER, as a result of the migration of events from one bin to another. The ERs before and after the break are both found to be shifted from the data position in the same direction. Before the break, the systematic shift in ER is -6 g/cm^2, while after the break it is -1 g/cm^2. The systematic shift in energy at the point of break is $\Delta \lg(E/\mathrm{eV}) = {}^{+0.09}_{-0.04}$. The respective systematic effect on the energy at the point of break of fit line, the first and second ER are shown in Fig. 6.9. A further check on the possible effect of binning, done by altering the width of bin from $\Delta \lg(E/\mathrm{eV}) = 0.1$ to $\Delta \lg(E/\mathrm{eV}) = 0.05$ (Fig. 6.10), showed marginal systematic shifts of -0.78 g/cm^2/decade, -0.47 g/cm^2/decade and

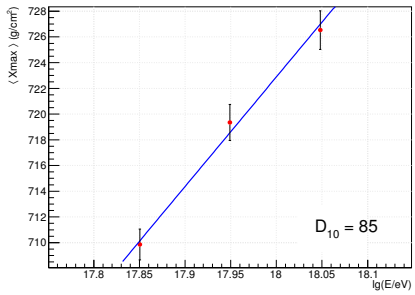
Central value $\lg(E/\text{eV})$	χ^2/Ndf	Prob	D_{10} (g/cm^2/decade) Data	MC
17.949	0.5/1	0.49	85 ± 10	$56 - 62$
18.048	0.7/1	0.40	87 ± 11	$56 - 62$
18.148	0.1/1	0.76	98 ± 12	$56 - 62$
18.247	2.3/1	0.13	55 ± 13	$56 - 61$
18.348	0.6/1	0.45	37 ± 14	$56 - 61$
18.447	0.6/1	0.46	37 ± 15	$56 - 60$
18.548	0.02/1	0.96	17 ± 17	$56 - 60$
18.646	0.1/1	0.82	8 ± 18	$55 - 60$
18.747	1.4/1	0.23	41 ± 20	$55 - 59$
18.849	0.6/1	0.42	55 ± 21	$55 - 59$
18.947	0.2/1	0.65	7 ± 22	$55 - 59$
19.048	0.04/1	0.85	-4 ± 24	$55 - 59$
19.144	1.1/1	0.30	48 ± 30	$55 - 58$
19.247	1.5/1	0.22	38 ± 31	$55 - 58$
19.340	1.4/1	0.23	21 ± 51	$54 - 58$
19.447	1.2/1	0.28	22 ± 25	$54 - 58$

Table 6.2: *The ERs in sliding windows of three bins for the Auger data [49] (see Figs. 6.6 and 6.7). The MC ERs for pure proton, helium, oxygen and iron for SIBYLL-2.1 [112], EPOS-LHC [113] and QGSJetII-04 [114] hadronic interaction models fall within the ranges indicated in the 'MC' column (see also Table 6.3).*

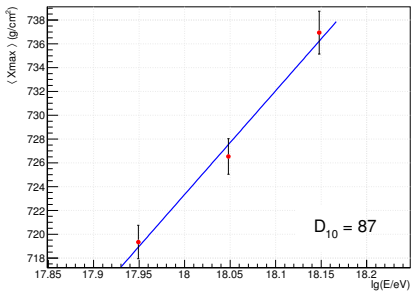
D_{10} (g/cm^2/decade)					
proton			iron		
SIBYLL 2.1	EPOS-LHC	QGSJetII-04	SIBYLL 2.1	EPOS-LHC	QGSJetII-04
$57 - 62$	$57 - 59$	$57 - 58$	$57 - 59$	$56 - 62$	$54 - 59$

Table 6.3: *Bounds of MC ERs for pure proton and iron obtained using SIBYLL-2.1, EPOS-LHC and QGSJetII-04 models in the energy range $10^{17.8}$ eV to $10^{19.8}$ eV. Gumbel parametrizations of $X_{\max}$ distributions are used to determine $\langle X_{\max} \rangle$ values, which are then fitted with straight lines in 3-bin sliding windows.*

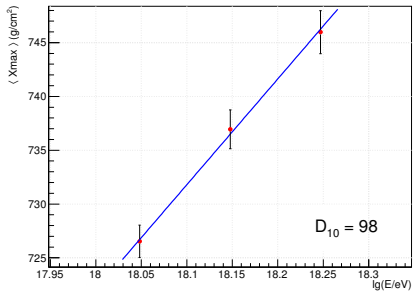
$+0.01$ in D_1, D_2 and $\lg(E/\text{eV})$ respectively.



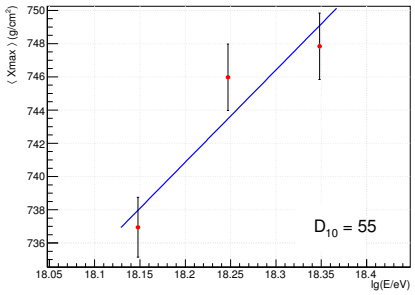
(a) 1–3



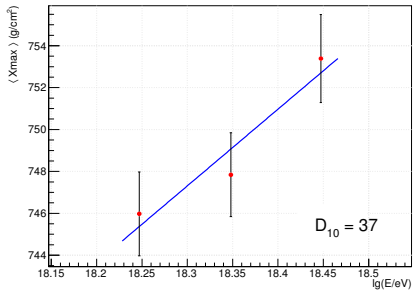
(b) 2–4



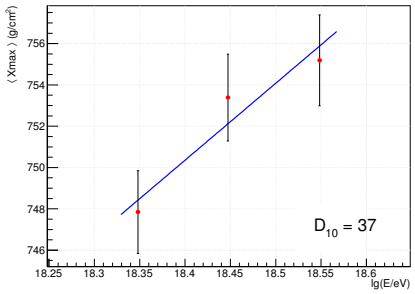
(c) 3–5



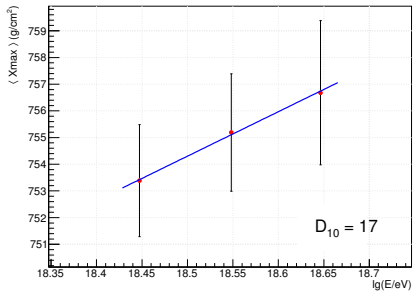
(d) 4–6



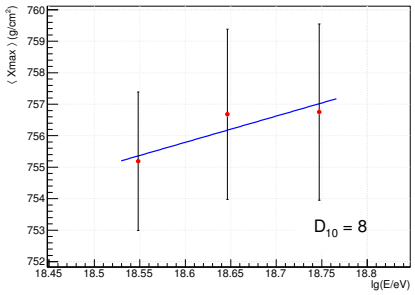
(e) 5–7



(f) 6–8

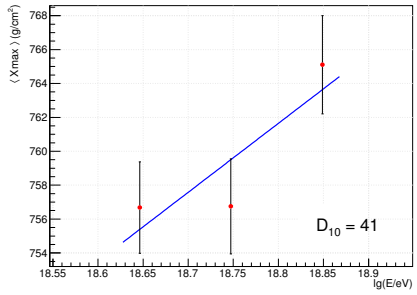


(g) 7–9

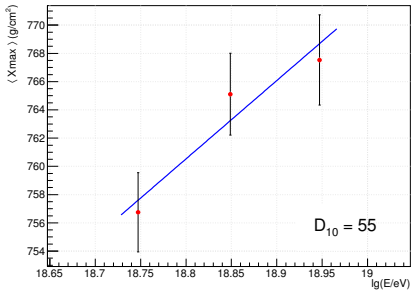


(h) 8–10

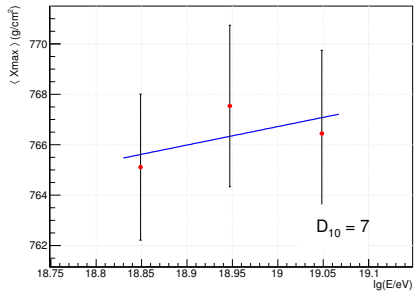
Figure 6.6: *Linear fits of Auger* $\langle X_{\max} \rangle$ *data [49] from Fig. 6.4 in 3-bin sliding windows. Ordinal numbers of bins used for fitting are indicated under each panel, while the values of* D_{10} *are in* $\mathrm{g/cm^2/decade}$. *The fit results are summarized in Table 6.2. Note that the mid-points are not correlated, although the ERs are.*



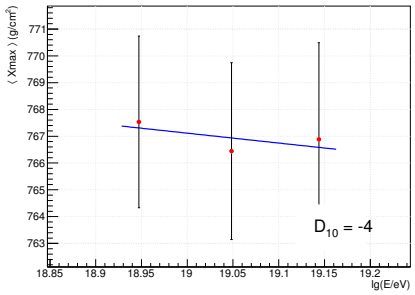
(a) 9–11



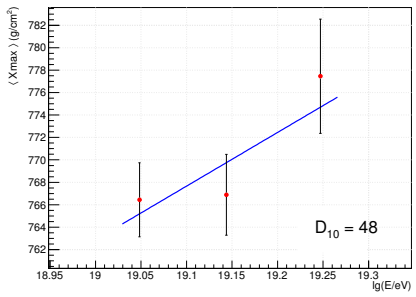
(b) 10–12



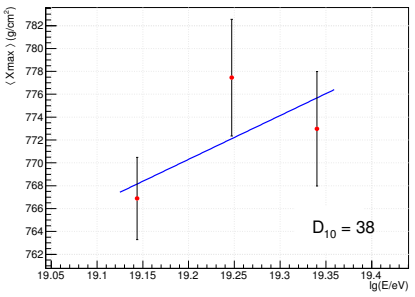
(c) 11–13



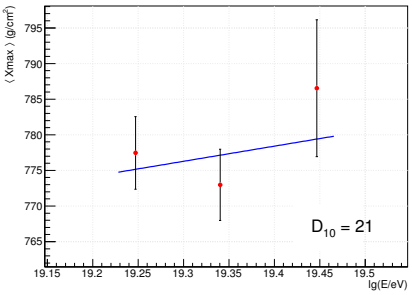
(d) 12–14



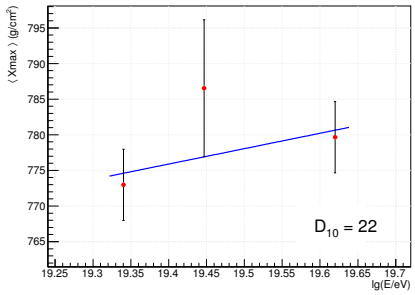
(e) 13–15



(f) 14–16



(g) 15–17



(h) 16–18

Figure 6.7: *Same as Fig. 6.6 but for bins of ordinal numbers* $9 - 18$.

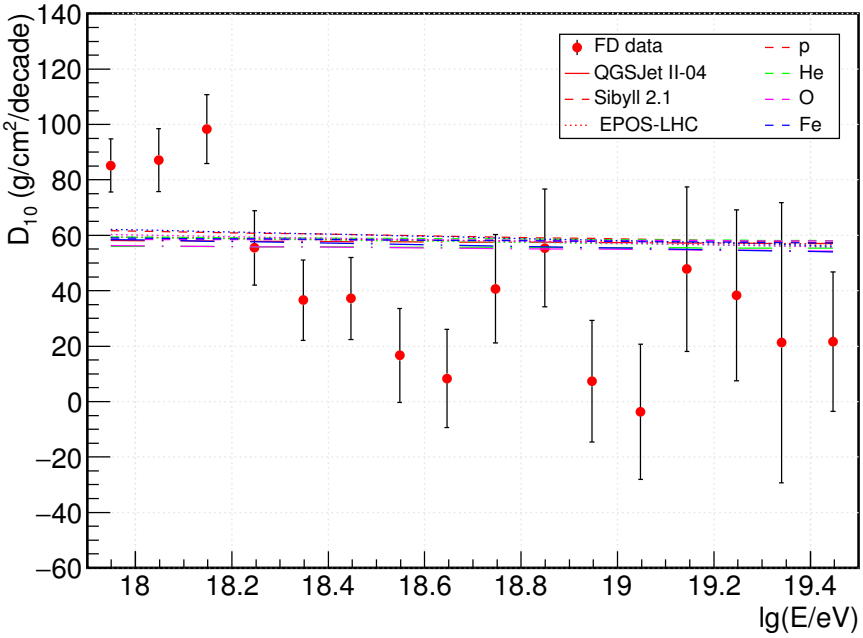


Figure 6.8: *ERs in sliding windows of three bins for the Auger* $\langle X_{\mathrm{max}} \rangle$ *data [49] together with MC values for different primary particles and interaction models (see Tables 6.2, 6.3). Data values larger (smaller) than the MC values indicate a composition getting lighter (heavier). Each data point is correlated with four neighboring points: two on either side.*

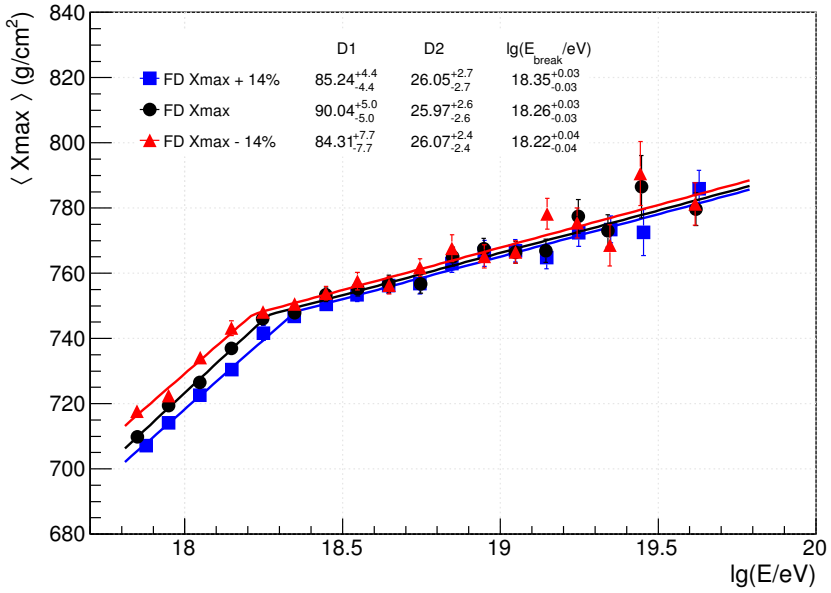


Figure 6.9: *Fits to the Auger FD data, assuming a change of 14% in the primary energy, compared to the unaltered data. The difference between the corresponding fit parameters are used as a rough estimate of the systematic error in each parameter.*

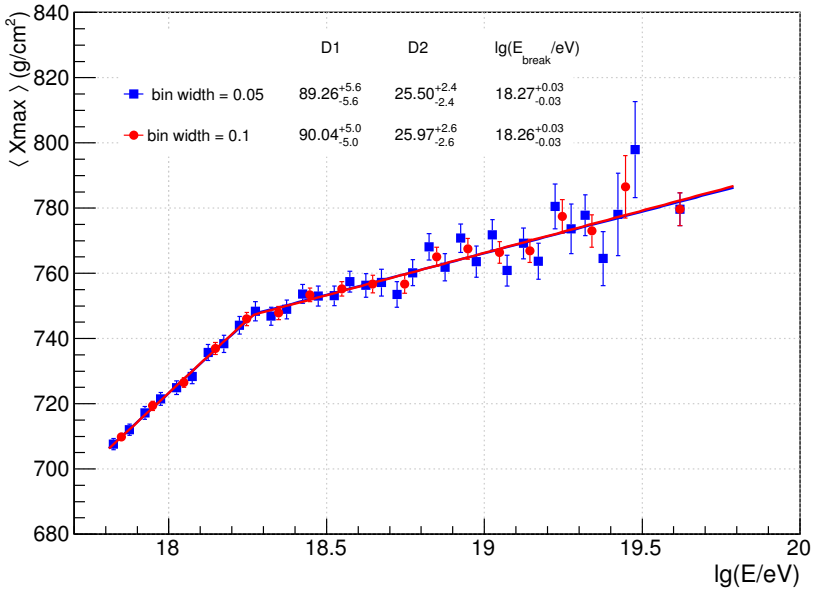


Figure 6.10: *Fits to the Auger data with the bin width changed from* $\Delta \lg(E/\mathrm{eV}) = 0.1$ *to* $\Delta \lg(E/\mathrm{eV}) = 0.05$*. The corresponding change in the fit parameters is used as an estimate of the systematic error in each parameter.*

6.2 Rate of change of ER with energy

As was explained in Sec. 6.1, slopes D_{10} in the sets of points with ordinal numbers 1, 4, 7 ...; 2, 5, 8 ...; 3, 6, 9 ... in Fig. 6.8 are not correlated with each other. If the points in such a set are in turn fitted in 3-bin sliding windows, an estimate of $dD_{10}/d\lg E$ at the energy center of each window is obtained. The rate of change of ER obtained using this approach is given in Fig. 6.11 and Table 6.5. Two of the values of $dD_{10}/d\lg E$ in the first set – the first and the last points in this set – are not correlated. A look at these two points seems to suggest that the rate of change of ER may not be constant. The rest of the points are from overlapping regions in the $\langle X_{\max}\rangle$ data as shown in Table 6.4.

Central value	$dD_{10}/d\lg E$ bin range		
$\lg(E/\mathrm{eV})$	Set 1	Set 2	Set 3
18.247	2-8		
18.348		3-9	
18.447			4-10
18.548	5-11		
18.646		6-12	
18.747			7-13
18.849	8-14		
18.947		9-15	
19.048			10-16
19.144	11-17		

Table 6.4: *The range of ordinal numbers of energy bins covered by the sliding windows of* $dD_{10}/d\lg E$.

Even from these latter data points of $dD_{10}/d\lg E$, one is able to get a rough idea of how fast the ER is changing. From Fig. 6.11 and Table 6.5, one can see that $dD_{10}/d\lg E \approx -(130 \div 110)$ g/cm^2/decade^2 and differs from zero by around 3σ for the energies $\lg(E/\mathrm{eV}) \lesssim 18.3$. Thus, in this region, the rate of change of the primary mass $d\langle\ln A\rangle/d\ln E$ might not be constant (see Eqn. (3.10)). Since most of the points fitted to obtain the slope are correlated as has been mentioned earlier, this conclusion must be treated with caution. The change from the trend of decrease of the primary mass with increasing energy to that of increase of the primary mass with energy, a point where $d^2\langle\ln A\rangle/d(\ln E)^2$ is compatible with zero, happens for $\lg(E/\mathrm{eV}) \gtrsim 18.3$.

To investigate the behavior of $dD_{10}/d\lg E$ for $\lg(E/\mathrm{eV}) \lesssim 18.2$ and also obtain more independent points for D_{10}, the Auger data was re-binned using $\Delta\lg(E/\mathrm{eV}) = 0.05$. The errors were estimated as $\Delta\langle X_{\max}\rangle_{0.05} \approx \Delta\langle X_{\max}\rangle_{0.1}\sqrt{N_{0.1}/N_{0.05}}$, where the numerical subscripts represent the respective binnings in energy and $\Delta\langle X_{\max}\rangle_{0.1}$ is the statistical error from the published Auger analysis [49]. In the same way as had earlier been done for the standard binning, the obtained ER, shown in Fig. 6.12, was split into three subsets with independent points and each subset was fitted in 3-bin sliding windows to get $dD_{10}/d\lg E$. From Fig. 6.13 and Table 6.6, one can see

Central value $\lg(E/\mathrm{eV})$	χ^2/Ndf	Prob	$dD_{10}/d\lg E$ ($\mathrm{g/cm^2/decade^2}$)
First Set			
18.247	0.1/1	0.79	-112 ± 31
18.548	3.4/1	0.07	-20 ± 40
18.849	0.7/1	0.40	68 ± 54
19.144	0.1/1	0.79	-55 ± 55
Second Set			
18.348	0.4/1	0.54	-137 ± 34
18.646	0.4/1	0.53	-55 ± 43
18.947	0.3/1	0.57	39 ± 56
Third Set			
18.447	3.0/1	0.09	-114 ± 37
18.747	0.97/1	0.32	-55 ± 46
19.048	0.88/1	0.35	-81 ± 76

Table 6.5: *Values of $dD_{10}/d\lg E$ for Auger data [49], obtained using linear fits of the ERs for $\Delta\lg(E/\mathrm{eV}) = 0.1$ in 3-bin sliding windows (see Fig. 6.8).*

that the statistical errors are so large that it is not possible to make any statistically significant inference on the behavior of $dD_{10}/d\lg E$. To confirm this, fits of the first four points of each set, which all have negative $dD_{10}/d\lg E$ values, were additionally performed with a constant function $f(x) = p_0$. The fits were done first for p_0 fixed to zero, and then for free p_0 (Table 6.7). In both cases, the fits are good and p_0 values do not differ from zero by more than 1σ for the fits when p_0 was left to vary. The same applies to the fit of the last 4 points. These results are thus consistent with a primary mass that is changing at a constant rate.

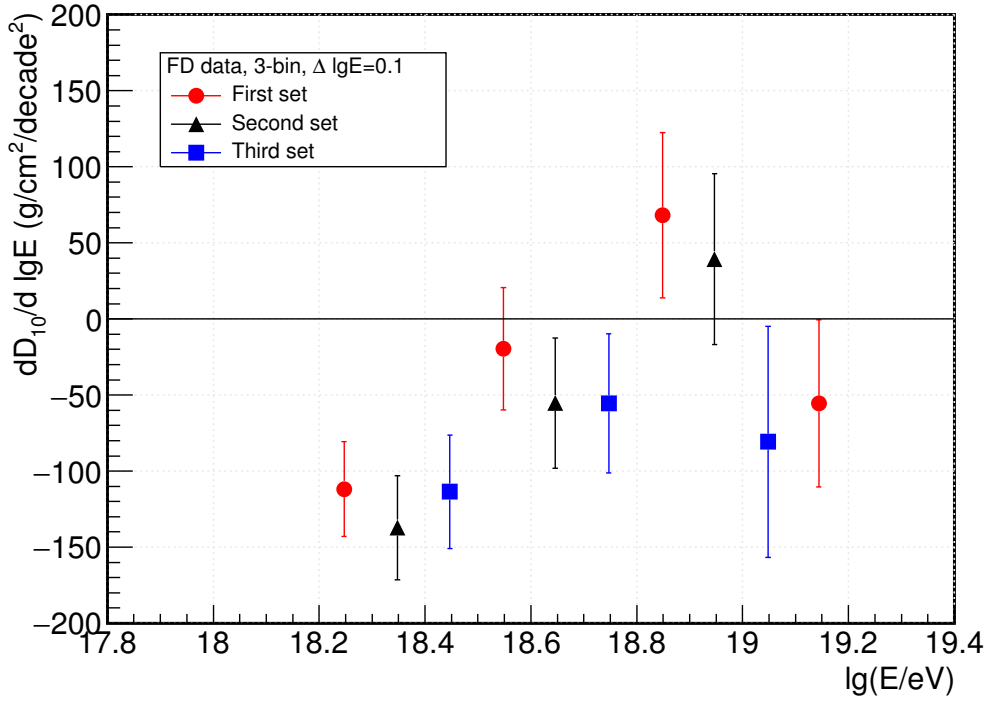


Figure 6.11: *Variation of* $dD_{10}/d\lg E$ *with* $\lg E$ *for Auger data [49], obtained using linear fits of the ERs for* $\Delta\lg(E/\mathrm{eV}) = 0.1$ *in 3-bin sliding windows (see Fig. 6.8). A value of zero indicates a constant rate of primary mass change. Only the first and last data points in the first set are not correlated.*

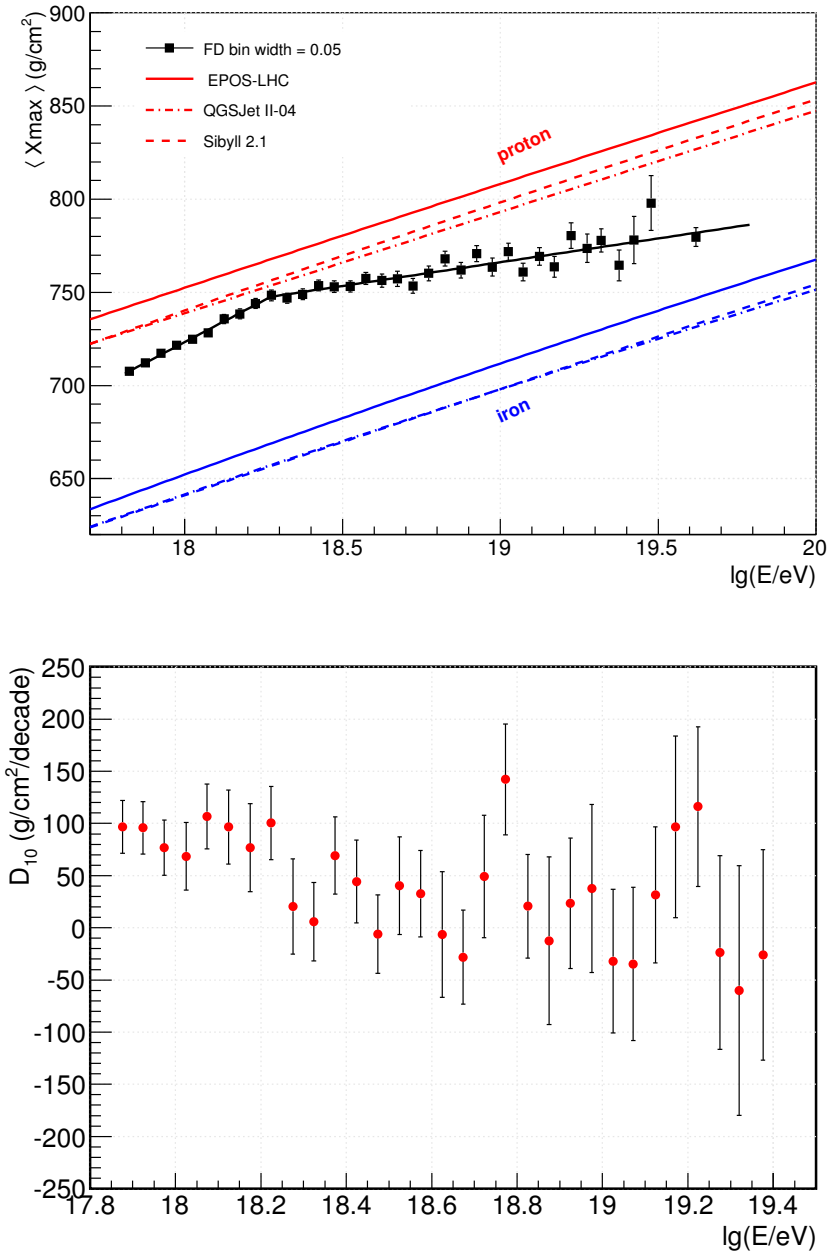


Figure 6.12: Top panel: *Auger* $\langle X_{\max} \rangle$ *data [49], binned using a bin width of* $\Delta \lg(E/\mathrm{eV}) = 0.05$*, together with predictions for protons and iron for different interaction models obtained with the use of the Gumbel parametrizations [67].* Bottom panel: *the (correlated) ERs obtained using fits in 3-bin sliding windows of data from the upper panel.*

Central value $\lg(E/\mathrm{eV})$	Reduced χ^2	Prob	$dD_{10}/d\lg E$ ($\mathrm{g/cm^2/decade^2}$)
First Set			
18.025	0.2	0.65	-88 ± 158
18.175	0.65	0.42	-199 ± 166
18.324	0.39	0.53	-268 ± 189
18.474	0.013	0.91	-48 ± 226
18.625	1.2	0.27	456 ± 214
18.773	3.8	0.051	109 ± 289
18.925	0.14	0.71	-614 ± 297
19.072	1.4	0.24	269 ± 328
Second Set			
18.074	0.053	0.82	21 ± 142
18.224	0.084	0.77	-121 ± 160
18.374	0.00069	0.98	-200.5 ± 192
18.525	0.13	0.72	-317 ± 193
18.674	1.1	0.3	-74 ± 228
18.825	0.051	0.82	248 ± 292
18.975	0.017	0.9	41 ± 272
19.124	0.071	0.79	-198 ± 414
Third Set			
18.124	1.2	0.28	-132 ± 168
18.275	0.9	0.34	-184 ± 178
18.424	0.13	0.72	37 ± 205
18.574	0.068	0.79	2 ± 230
18.723	0.28	0.6	-94 ± 281
18.875	0.051	0.82	-274 ± 298
19.025	0.68	0.41	348 ± 399
19.171	1.7	0.2	91 ± 446

Table 6.6: *Values of $dD_{10}/d\lg E$ for Auger data [49], obtained using straight line fits in 3-bin sliding windows of the ERs for $\Delta\lg(E/\mathrm{eV}) = 0.05$.*

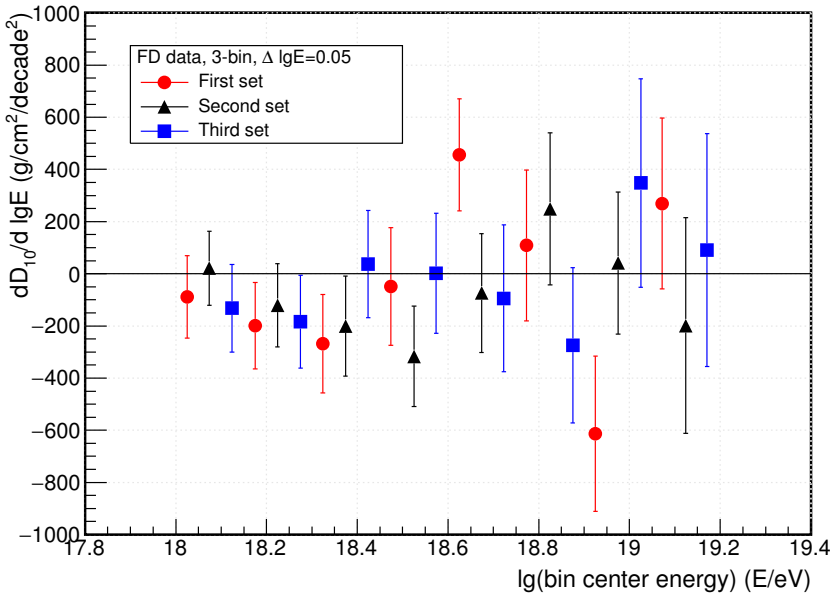


Figure 6.13: *The first derivative of ER ($dD_{10}/d\lg E$) for the Auger data, obtained using linear fits of the ERs for $\Delta\lg(E/\mathrm{eV}) = 0.05$ in 3-bin sliding windows (see Fig. 6.12). A value of zero indicates a constant rate of primary mass change, and points in a given set are correlated.*

	$p_0 = 0$		Free p_0		
	χ^2/Ndf	Prob	p_0	χ^2/Ndf	Prob
First Set					
All points	13.42/8	0.10	-70 ± 75	12.54/7	0.08
First 4	3.83/4	0.43	-155 ± 90	0.83/3	0.84
Last 4	9.59/4	0.05	126 ± 136	8.73/3	0.03
Second Set					
All points	5.46/8	0.71	-86 ± 72	4.02/7	0.78
First 4	4.37/4	0.36	-123 ± 84	2.21/3	0.53
Last 4	1.08/4	0.90	18 ± 141	1.07/3	0.78
Third Set					
All points	3.49/8	0.90	-78 ± 83	2.61/7	0.92
First 4	1.72/4	0.79	-88 ± 96	0.88/3	0.83
Last 4	1.76/4	0.78	47 ± 169	1.68/3	0.64

Table 6.7: *Fit with a constant function $f(x) = p_0$ of three subsets of data for $dD_{10}/d\lg E$ (Fig. 6.13). First, all points in each subset were fitted, then only first four points, and finally the last 4 points. Fits were performed with p_0 fixed to zero and also with free p_0.*

6.3 Division of data into "light" and "heavy" components

6.3.1 Definition of the cut

In most of the astrophysical scenarios, it is anticipated that changes in mass composition with the growth of energy should happen via a successive reduction of abundances of nuclei, starting from the lightest ones, protons. This might be, for example, due to the maximum attainable energy of particles at the acceleration sites being proportional to their charge or, in a less standard scenario including possible Lorenz invariance violation, due to energy loss by vacuum Cherenkov radiation for energies above a certain threshold proportional to the particle mass [117–120]. Disappearance of medium mass nuclei (helium, CNO) from the primary beam, without anything happening to the lightest (protons) and heaviest (iron) ones, is not easy to justify.

With this in mind, one can split hybrid events into lighter and heavier parts according to their $X_{\rm max}$ values and check the ER behavior for each of the parts. Since the lighter component will be dominated by protons, a significant change in the proton fraction should manifest itself in a change in the ER in this component. This change might be obscured in the full data ER due to simultaneous changes in the heavy subset, which can average out to result in no net change in ER. Two criteria for splitting can be considered, each of which in this case leads to the same cut on $X_{\rm max}$. The first one is to keep approximately the same number of events below and above the cut, in order to have comparable statistics in the first energy bin, and then to increase the $X_{\rm max}$ cut value with energy using the ER for a constant composition ($\approx$ 60 g/cm^2/decade). The second way is to define the cut based on the probability of a shower to be produced by helium. The latter approach has been used in this work.

The $X_{\rm max}$ cut was selected such that 50% of simulated helium events for EPOS-LHC have deeper depth of shower maximum than the cut value, $X^{cut}_{\rm max}(E)$. Thus, in data, events with $X_{\rm max}$ deeper than the cut value should be dominated by the proton and helium nuclei, in case they are both present in the primary radiation; the heavier part will be less influenced by the presence of protons. For EPOS-LHC, $\approx$ 70% of protons, $\approx$ 17% of oxygen and 0% of iron showers are deeper than $X^{cut}_{\rm max}(E)$ (see Fig. 6.14(b)). These values are almost independent of the energy. For the purpose of being brief, the two subsets will be referred to simply as 'light' and 'heavy' from this point onwards.

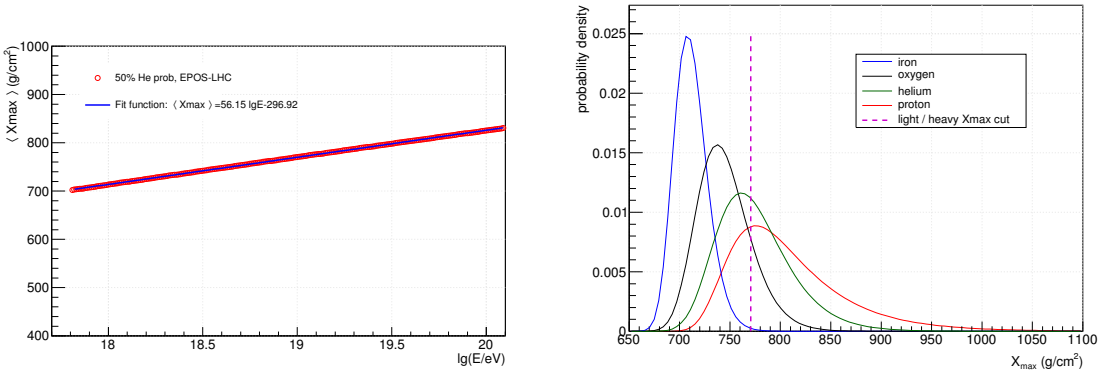
It is important to note two features of such a division. First, if the increase of the primary mass, which one observes in the ER for the whole Auger data set is in a large part due to a reduction of the helium fraction, then the light part would show some lightening and the mass of the heavy part would seem to increase relative to the combined data set. The second feature is that if one would use some other model instead of EPOS-LHC to define the cut, that would change the relative helium fraction e.g. for SIBYLL 2.1 and QGSJet II.04, $\approx 33-36$% of helium showers are deeper than $X^{cut}_{\rm max}(E)$. However, the change would be by a constant value for all

energies, thus not affecting much the relative changes in the composition and the ER.

In Fig. 6.14(a), the function that is used to apply a cut on $X_{\max}$ is shown. It is given by

$$f(\lg E) = 56.15 \lg E - 296.92 \, . \tag{6.3}$$

The events whose value of $X_{\max}$ are above the line of cut are the 'light' events, while those which lie below the line of cut are the 'heavy' ones.



(a) Line of cut

(b) Gumbel-parametrized distributions

Figure 6.14: Left panel: *The line dividing $X_{\max}$ data into 'light' and 'heavy' events, obtained with a cut at 50% probability for helium to produce a deeper shower using EPOS-LHC model.* Right panel: *The Gumbel-parametrized distributions of iron, oxygen, helium and proton at energy* 10^{19} eV*, obtained using the EPOS-LHC interaction model. The line of delimitation of $X_{\max}$ into light and heavy events at this energy is included.*

The cut does not only result in a reduction in the statistics for the two parts, but also gives rise to two parts whose behaviors are qualitatively different: the lighter events have larger $X_{\max}$ fluctuations than the heavier ones. Hence the errors on $\langle X_{\max} \rangle$ cannot be estimated by relating the standard errors ($\sigma_M = \sigma/\sqrt{N}$) of the original data to those in the light and heavy subsets. It is for this reason that the estimation of the error on the weighted mean given by Eqn. 6.1 has been used in this study.

6.3.2 Bias in the $X_{\max}$ of light and heavy events

In this section, the bias in the $\langle X_{\max} \rangle$ values of light and heavy subsets in simulated distributions, resulting from detector resolution, is investigated. This is done with a view to quantifying any possible bias in the mean of the light and heavy subsets resulting from truncation. Starting with a smeared distribution, the procedure which is described below was used to first estimate the true full distribution, and then the biases in the truncated sets.

The approach that was used to estimate the bias introduced in the truncated mean $X_{\max}$ of the light and heavy subsets of simulated data involved the use of the exponentially modified Gaussian function (EMG). The EMG is a convolution of the normal and exponential probability density functions, a property which makes it useful for deconvoluting a measured distribution which fits it, but whose true distribution is Gaussian. It is defined as [121]:

$$f(x;\mu,\sigma,\lambda) = \frac{\lambda}{2} e^{\frac{\lambda}{2}(2\mu+\lambda\sigma^2-2x)} \mathrm{erfc}\left(\frac{\mu+\lambda\sigma^2-x}{\sqrt{2}\sigma}\right) \tag{6.4}$$

where *erfc* is the complimentary error function defined as $\mathrm{erfc}(x) = 1 - \mathrm{erf}(x)$, μ and σ^2 are the mean and variance of the Gaussian component and λ is the rate of the exponential component.

The suitability of the EMG was tested with simulations that were performed using the CONEX [122] code with EPOS-LHC as the hadronic interaction model, at the energy $10^{19.0}$ eV. Using a composition made up of proton-helium, proton-oxygen and proton-iron in percentages varying from 100% of proton and 0% of the second nucleus to 0% proton and 100% of the second nucleus in intervals of 20%, a total of 100000 events were simulated in each case. A sample of the distributions is shown in Fig. 6.15 and Fig. 6.16. To bring about the resolution effect, the MC distributions were smeared with Gaussians with $\sigma = 20$ and 50 g/cm^2 corresponding to the resolutions of the FD and SD respectively. A shift in the value of $X_{\max}$ in both the light and heavy parts of the distribution can be seen in comparing the true and the "measured" distributions. This is more marked when the width of smear is 50 g/cm^2. In the case of proton mixture with iron, a significant change in the nature of the distribution is evidently introduced by the smearing process, as the two peaks are replaced by just a single one.

If a reasonably good fit is obtained on the smeared data, then the variance, σ, of the underlying normal distribution and hence its standard deviation may be obtained from the relation

$$\sigma^2_{\mathrm{EMG}} = \sigma^2 + 1/\lambda^2. \tag{6.5}$$

The mean of the EMG distribution is given by

$$\overline{x}_{\mathrm{EMG}} = \mu + 1/\lambda\,. \tag{6.6}$$

For a large sample of data, it is expected that the smearing of the sample should not cause its mean to deviate from that of the parent distribution, hence $\overline{x}$ in Eqn. 6.6 should remain the same for the sample as for the parent distribution. Furthermore, it has been shown that for a large variety of primaries, λ remains the same after smearing [123]. Hence one expects that μ should also remain invariable.

To begin the procedure of EMG analysis, the smeared distribution was fitted with an EMG for various p-He, p-Fe, and p-O compositions, and hence λ_{fit}, σ_{fit} and μ_{fit} were obtained. In the meantime, the true mean of the light and heavy subsets were obtained from two separate histograms, filled for light and heavy events respectively. The standard deviation of the EMG from Eqn. 6.5 is given by

$$\sigma^{\mathrm{EMG}}_{\mathrm{fit}} = \sqrt{\sigma^2_{\mathrm{fit}} + 1/\lambda^2_{\mathrm{fit}}}\,. \tag{6.7}$$

Then

$$\sigma_{\rm true}^{\rm EMG} = \sqrt{(\sigma_{\rm fit}^{\rm EMG})^2 - \sigma_{\rm res}^2} \tag{6.8}$$

where $\sigma_{\rm res}$ is the resolution due to smearing, and the standard deviation of the unsmeared underlying Gaussian is given by

$$\sigma_{\rm true} = \sqrt{(\sigma_{\rm true}^{\rm EMG})^2 - 1/\lambda_{\rm fit}^2}\,. \tag{6.9}$$

The fit parameters $\sigma_{\rm true}$, $\lambda_{\rm fit}$ and $\mu_{\rm fit}$ were then substituted in the EMG, and the function integrated to obtain the deconvoluted means of the 'light' and 'heavy' subsets of the data; the two subsets being delimited by $X_{\rm max}^{\rm cut}$. The mean of $X_{\rm max}$ for the light and heavy components are given respectively by Eqns. 6.10 and 6.11,

$$\langle X_{\rm max}^{\rm light}\rangle = \frac{\int_{X_{\rm max}^{cut}}^{+\infty} X_{\rm max} f(X_{\rm max};\mu_{\rm fit},\sigma_{\rm true},\lambda_{\rm fit})dX_{\rm max}}{\beta_{\rm light}} \tag{6.10}$$

$$\langle X_{\rm max}^{\rm heavy}\rangle = \frac{\int_{-\infty}^{X_{\rm max}^{cut}} X_{\rm max} f(X_{\rm max};\mu_{\rm fit},\sigma_{\rm true},\lambda_{\rm fit})dX_{\rm max}}{\beta_{\rm heavy}} \tag{6.11}$$

where $\beta_{\rm light}$ and $\beta_{\rm heavy}$ are the normalization factors given respectively by

$$\beta_{\rm light} = \int_{X_{\rm max}^{cut}}^{+\infty} f(X_{\rm max};\mu_{\rm fit},\sigma_{\rm true},\lambda_{\rm fit})dX_{\rm max} \tag{6.12}$$

and

$$\beta_{\rm heavy} = \int_{-\infty}^{X_{\rm max}^{cut}} f(X_{\rm max};\mu_{\rm fit},\sigma_{\rm true},\lambda_{\rm fit})dX_{\rm max}\,. \tag{6.13}$$

By subtracting the mean in the light and heavy subsets of the true distribution from the corresponding mean in the deconvoluted distribution, an estimate of the bias is obtained.

A summary of the results of this procedure, showing all the fits to selected compositions together with the fit parameters in each case, is contained in Figs. 6.17 and 6.18. Superposed plots of the actual true distributions give one an idea of how well this procedure unfolds the smeared distribution.

It was observed that generally, when the primary distribution contains two distinct peaks, the smeared distribution does not reflect this, i.e. it has only one peak. In such cases, the true distribution obtained by the use of the EMG differs significantly from the actual true distribution. A look at the quality of fit represented by the values of χ^2/Ndf for different compositions shown in Fig. 6.19 suggests that compositions containing only iron and proton generally have the lowest quality fits. This is especially so when the width of smear is 20 g/cm^2. However, the procedure worked generally well, resulting in a bias in truncated $\langle X_{\rm max}\rangle$ close to zero as expected. A summary of the biases in $\langle X_{\rm max}\rangle$ for the resolutions $\sigma_{\rm res} = 20$ and $\sigma_{\rm res} = 50$ g/cm^2 are presented in Fig. 6.20.

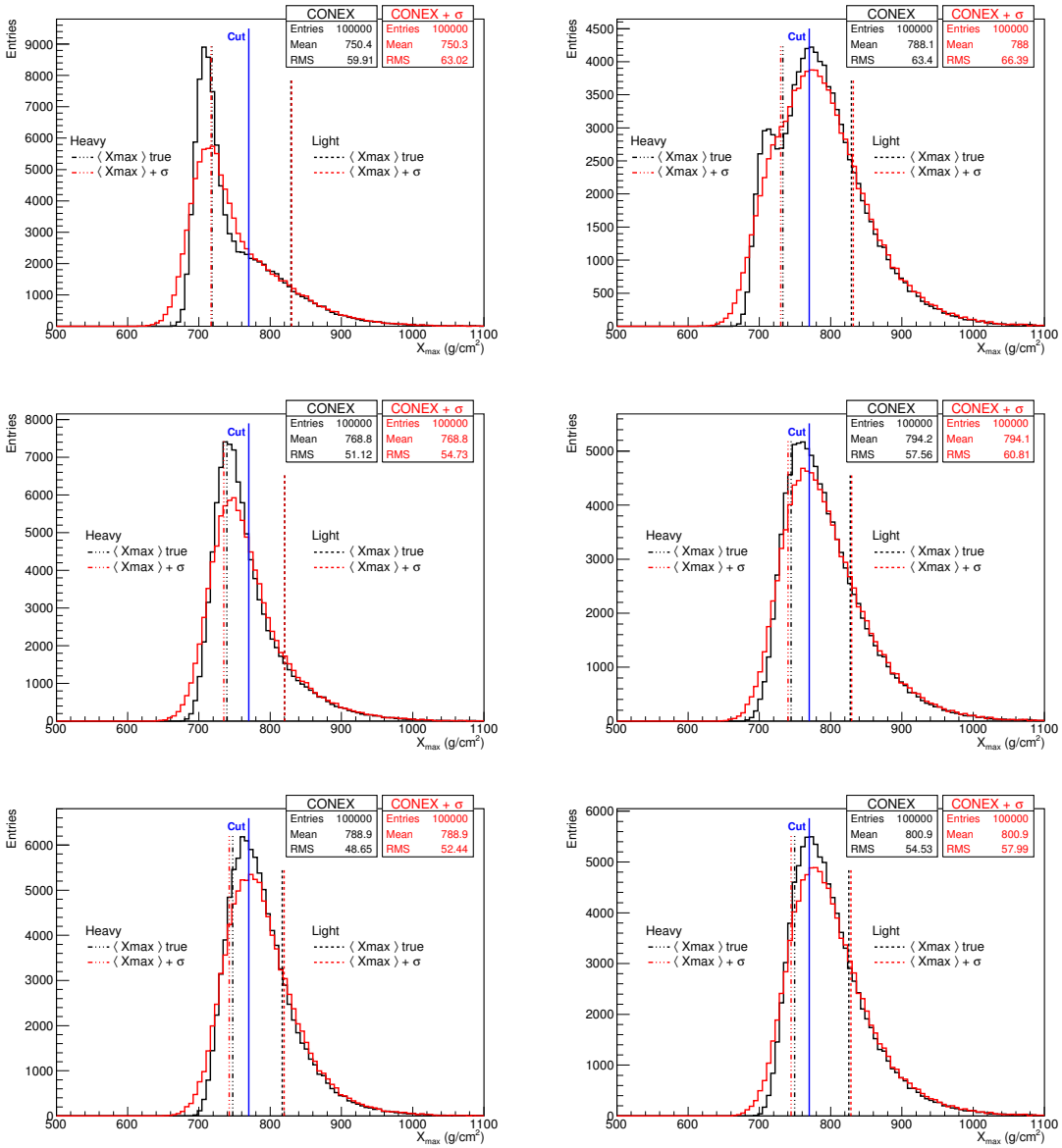


Figure 6.15: *Simulated compositions of p-Fe* (top)*, p-O* (middle) *and p-He* (bottom) *mixtures. A smear of each distribution with* $\sigma_{\text{res}} = 20$ g/cm^2 *is shown in red. In the first column, the proton fraction is 40% while in the second column it is 80%. The mean of the light and heavy subsets of data before and after smearing are represented by the vertical dashed lines.*

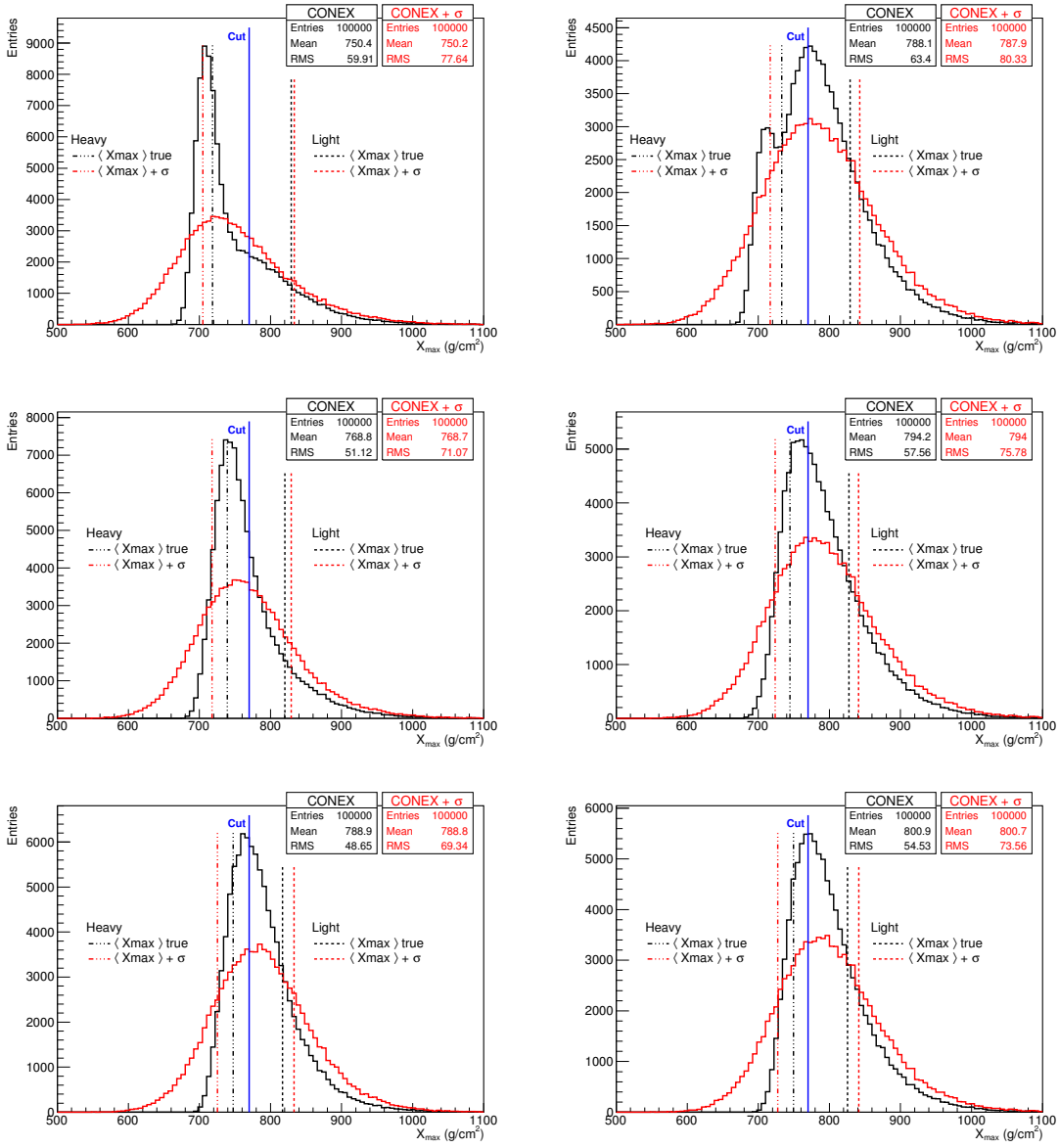


Figure 6.16: *Same as in Fig. 6.15, but for* $\sigma_{\mathrm{res}} = 50\ \mathrm{g/cm^2}$.

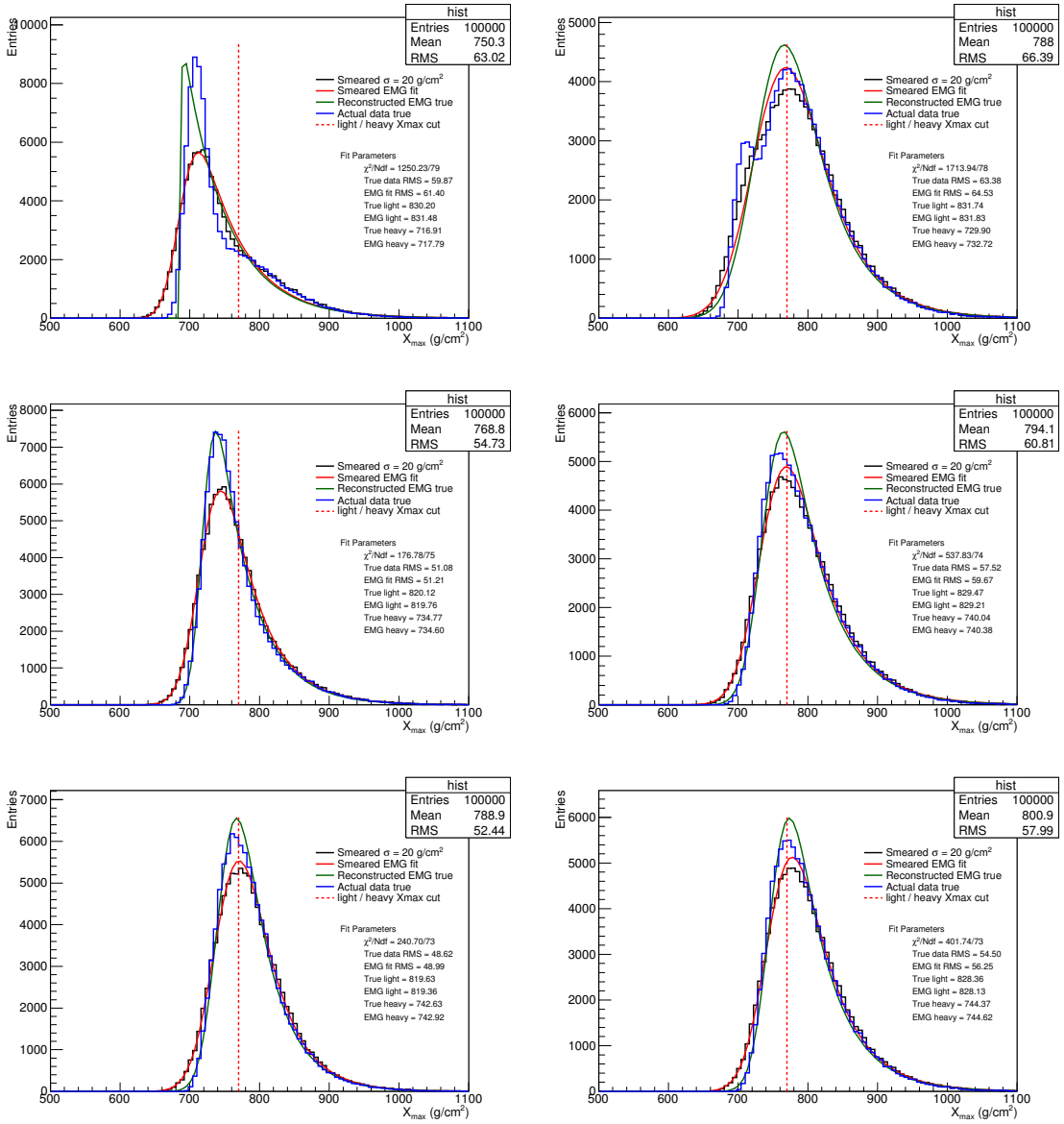


Figure 6.17: *Smeared ($\sigma_{\mathrm{res}} = 20\ \mathrm{g/cm^2}$) simulated p-Fe, p-O, and p-He mixtures fitted with an EMG. The composition of the distribution in every row and column is as described under Fig. 6.15. The reconstructed true distributions as predicted by the EMG are also shown. Alongside this, the distributions of the original data (blue) are included for comparison. The unit of $\langle X_{\mathrm{max}} \rangle$ shown on the r.h.s. of every plot for both light and heavy subsets is in* $\mathrm{g/cm^2}$.

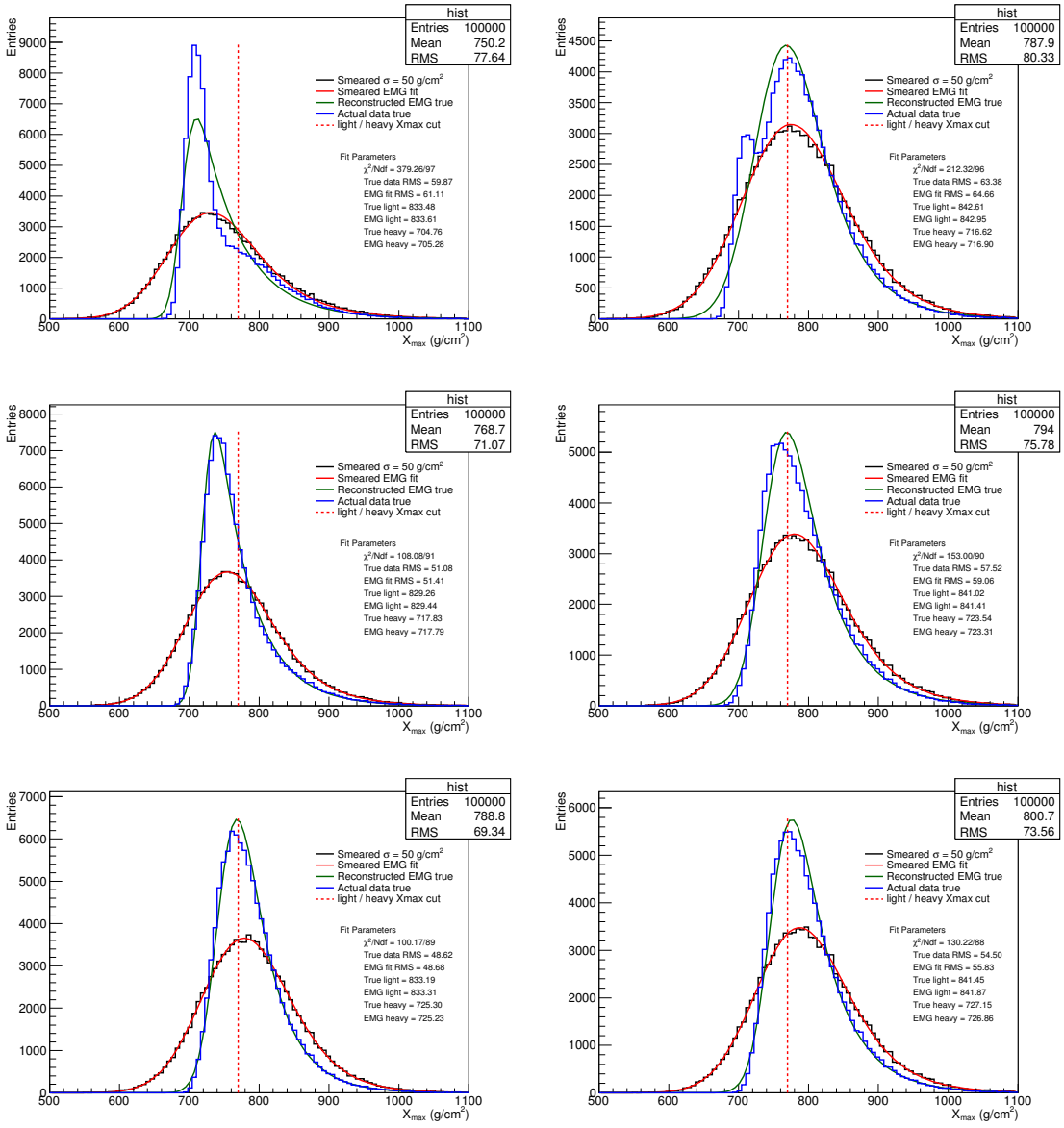


Figure 6.18: *Same as in Fig. 6.17, but for* $\sigma_{\mathrm{res}} = 50\ \mathrm{g/cm^2}$.

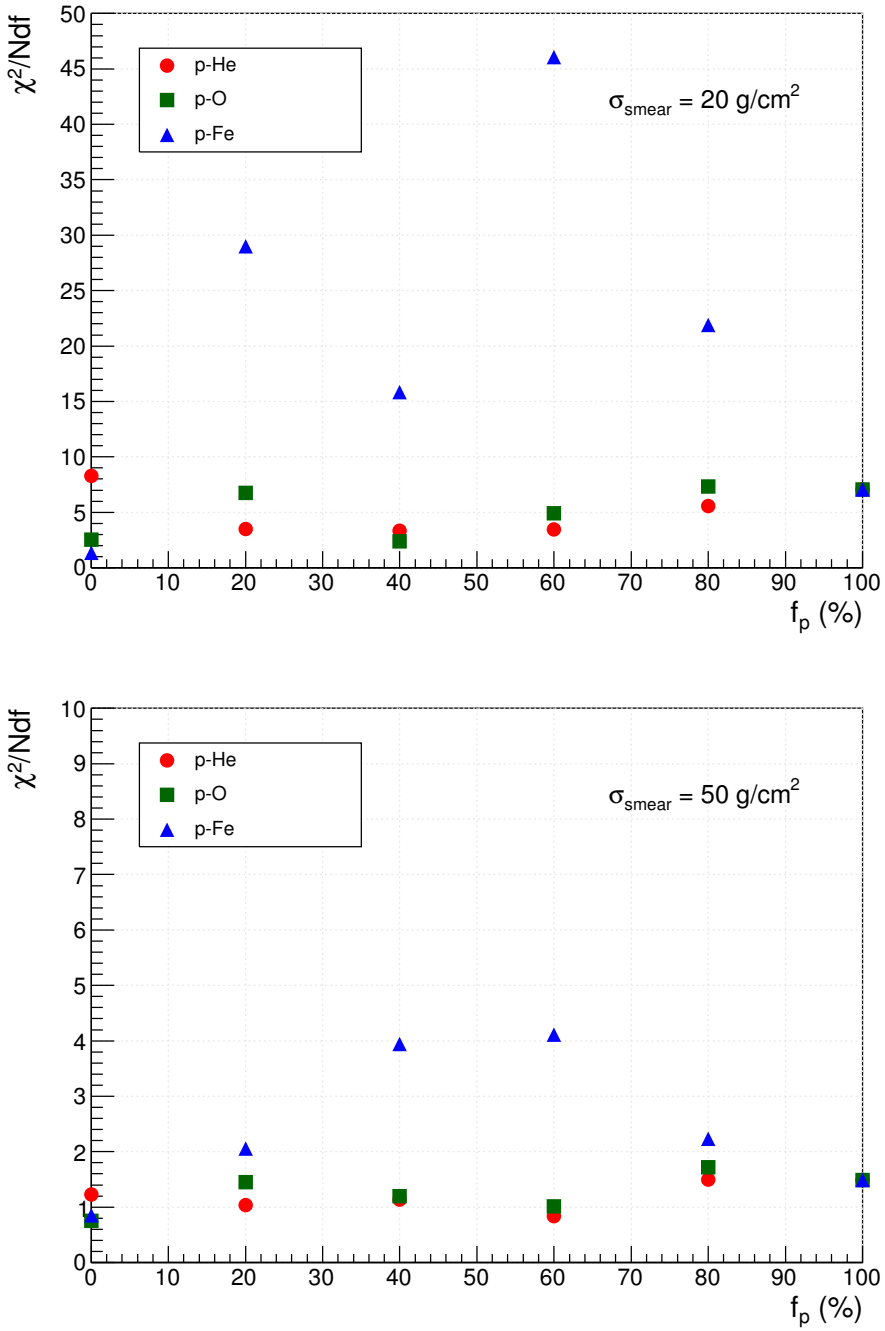


Figure 6.19: *Evolution of the quality of EMG fit with proton fraction for p-He, p-O and p-Fe mixtures for widths of smear* $\sigma_{\text{res}} = 20$ g/cm^2 (top panel) *and* $\sigma_{\text{res}} = 50$ g/cm^2 (bottom panel). *The horizontal axis shows the percentage of proton in each mixture.*

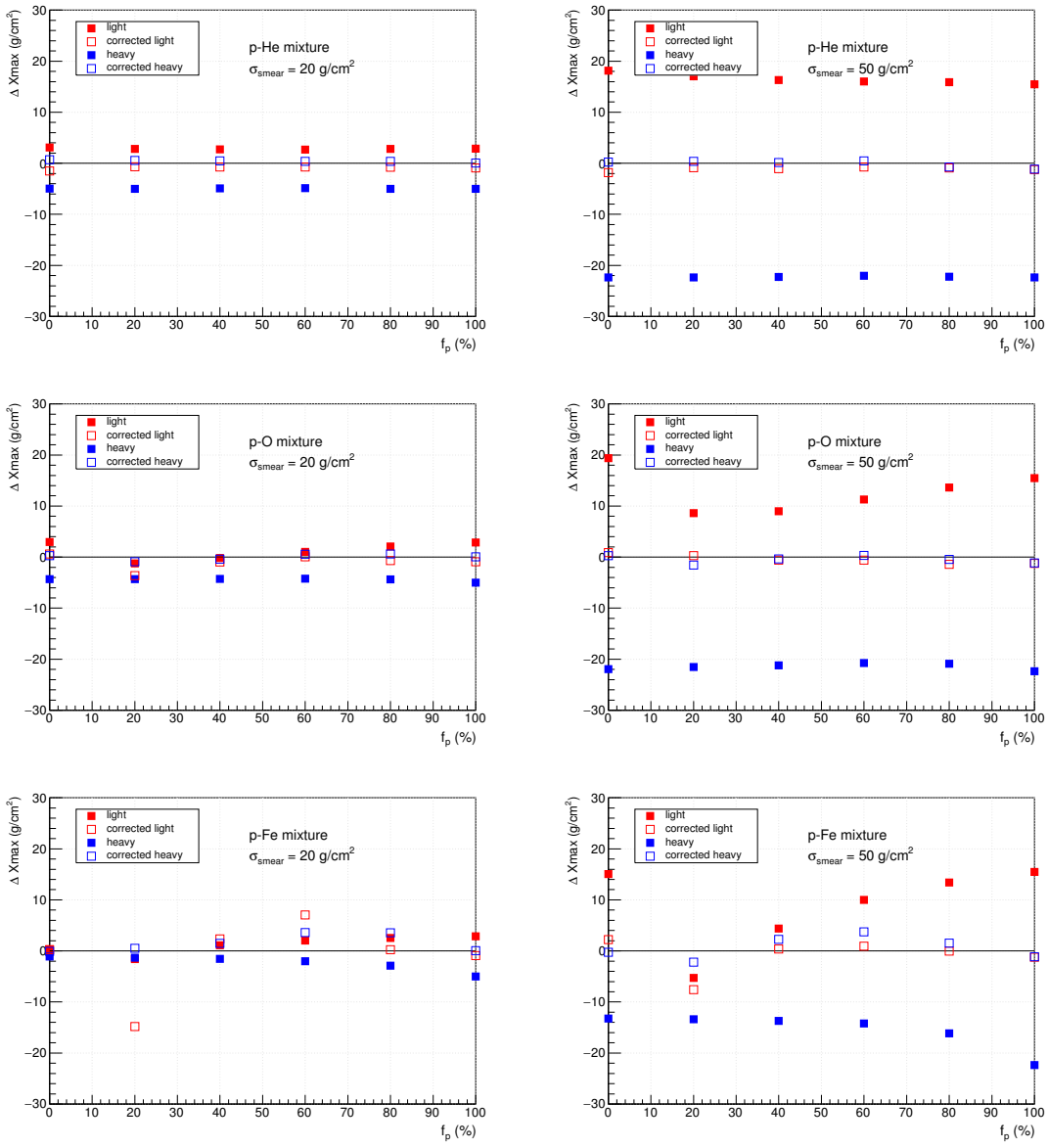


Figure 6.20: *Bias in the light and heavy $\langle X_{\max} \rangle$ due to resolution effects (closed symbols) and results for the corrected $\langle X_{\max} \rangle$ values (open symbols).*

6.3.3 Application of cut on FD data

An event is classified in this study as 'light' or 'heavy' according to the criteria explained in section 6.3.1. A study of the ER of the two subsets of the Auger data is presented in this section.

In Fig. 6.21 one can see the cut on the Auger FD $X_{\max}$ data and the behavior of the fraction of events in the light part. The fraction of light events stops growing at $\lg(E/\text{eV}) \approx 18.3$. Shortly after this point, it begins to actually decline. This is the same energy at which the ER in the full data changes, as seen in Fig. 6.8, indicating a change in trend from lighter to heavier composition. It is therefore interesting to investigate whether a decrease in the fraction of the lighter component of the primary UHECRs is responsible for the observed overally increasing mass.

To this end, the evolution of $\langle X_{\max} \rangle$ with energy for 'light' and 'heavy' parts of events has been investigated. Initially, fits to the respective subsets of data with only one break were used, and following observations in the light subset, two-break fits were used as well. The results obtained using a single-break fit are discussed first. The resolution effect on $\langle X_{\max} \rangle$ of the 'light' and 'heavy' subsets of data and the corresponding ERs are also examined.

It was noted that just like the simulated distributions, the Auger data in all the energy bins takes a form that is close to that of an EMG. The distribution in every bin was fitted with an EMG and the procedure as earlier described for simulated events was followed to find the shape of the true distributions. The value of the fits as indicated by the value of reduced χ^2 is good: it ranges from 0.45 in bin 17 to 2.24 in bin 7.

In order to assure conformity to the published Auger analysis [49], the distribution for every bin that was obtained in this work was plotted alongside the published distributions, for comparison. Contrasting the light component reconstructed from data with that reconstructed from the EMG fit in Figs. 6.22, 6.23 and 6.24, one notes that there is only a small change in the value of $\langle X_{\max} \rangle$ of magnitude $2 - 5$ g/cm^2 in all the bins except the 16^{th} and 17^{th} bins where the deviation is significant at ~ 16 g/cm^2. In these two bins, either the distribution of the data is irregular such that a significant amount of the data is left out of the fit curve, or the fit curve extends to regions where there are no data events. However, the exclusion of these two points from the fit does not affect the fit results. The unbiased ERs are statistically consistent with the results for the uncorrected data. A further analysis is thus done for the $X_{\max}$ distributions which are directly measured by the FD.

The resolution is slowly changing with energy (Fig. 5.1), therefore biases in $\langle X_{\max} \rangle$ of 'light' and 'heavy' subsets are similar and small throughout the whole energy range, and the ERs are practically unaffected. This can be seen in Fig. 6.25. It is noted that the ERs before and after the break, as well as the energy at the point of break, before correction for bias, are compatible with the corresponding values after correction for bias. In the light subset, the application of the resolution correction changes the ERs from 69.8 ± 6 and 33.0 ± 3 g/cm^2/decade to 67.5 ± 6 and 28.0 ± 3 g/cm^2/decade respectively. The energy at the point of break remains unaltered at $\lg(E/\text{eV}) = 18.16 \pm 0.05$. In the heavy subset, the ERs before and after the break change from 67.1 ± 1 and 48.6 ± 5 g/cm^2/decade to 68.8 ± 1 and

45.5 ± 4 g/cm^2/decade respectively. In this data subset, the energy at the point of break changes from $\lg(E/\text{eV}) = 18.85 \pm 0.10$ to $\lg(E/\text{eV}) = 18.80 \pm 0.07$.

The fit results of using a single break defined by Eqn. 6.2 are shown in Figs. 6.26 and 6.27. In Table 6.8, one can see the results for these fits and results for the fits with straight lines (which in both cases are bad). In the light subset, the features are similar to those in the full data set (see Fig. 6.5) i.e. the primary mass is initially decreasing and then at $\lg(E/\text{eV}) \approx 18.2$ it begins to increase. Considering the probability and χ^2/Ndf, this fit appears to be necessary; the ER values differ significantly below and above break at $\lg(E_0/\text{eV}) = 18.16$. It is worth noting, as already shown in Fig. 6.21, that the fraction of light events stops to grow at this energy as well. A possible cause of the changing primary mass beyond $\lg(E/\text{eV}) \approx 18.2$, if the primary beam is dominated by protons, could be due to the loss of protons with increase in energy. Alternatively, there could be an accelerated increase in proton interaction cross-section as the energy increases, leading to a decrease in the $X_{\max}$ for protons.

Light subset

	D_{10} (g/cm^2/decade)				
$\lg(E_0/\text{eV})$	D_1	D_2	X_0 (g/cm^2)	χ^2/Ndf	Probability
$18.16^{+0.05}_{-0.04}$	70^{+6}_{-6}	33^{+3}_{-3}	783^{+2}_{-2}	1.30	0.20
no break	43 ± 1			3.54	0.00

Heavy subset

	D_{10} (g/cm^2/decade)				
$\lg(E_0/\text{eV})$	D_1	D_2	X_0 (g/cm^2)	χ^2/Ndf	Probability
$18.55^{+0.08}_{-0.08}$	69^{+1}_{-1}	56^{+2}_{-2}	713^{+5}_{-5}	1.04	0.41
$18.85^{+0.09}_{-0.08}$	67^{+1}_{-1}	49^{+4}_{-5}	733^{+5}_{-5}	1.04	0.41
no break	64 ± 1			2.53	0.00

Table 6.8: *Fit parameters for 'light' and 'heavy' subsets of Auger $\langle X_{\max} \rangle$ data with a broken line from Eqn. 6.2 and also with a straight line. For the 'heavy' subset, two possible fits with broken lines are presented.*

For the heavier part, the break is weakly pronounced and fit results depend on the initial value of the position of the break. Five possible fits with the break positions between $\lg(E/\text{eV}) = 18.55$ and $\lg(E/\text{eV}) = 18.95$ were found. In Fig. 6.27, two examples of fits with break positions at $\lg(E_0/\text{eV}) = 18.55$ and $\lg(E_0/\text{eV}) = 18.85$ are shown. Of all the 5 possible fits, these two cases have the smallest errors on $\lg(E_0/\text{eV})$. The ER below break is compatible with the MC expectations for a constant composition; above break, the ER for the fit with $\lg(E_0/\text{eV}) = 18.85$ might indicate an increase of the primary mass. The existence of the break at $\lg(E_0/\text{eV}) = 18.55$ can be explained by the reduction of the fraction of protons. The change in ER at the break is not so pronounced in the heavy subset since in this case, the relative contribution of protons should be small, assuming there is a

significant number of heavier nuclei. The application of the cut $X_{\max}^{cut}$ reduces the influence of protons in the heavy subset.

As a way of carrying out a cross-check, the effect of altering the position of the cut on the fits to the data has been investigated. In the first approach, the ER of the line of cut was shifted positively and negatively by 20 g/cm^2 at each energy. The effect of this is seen in the first row of Fig. 6.28. In an alternative approach, the ER of the cut remained fixed at 70 and 45 g/cm^2 respectively instead of 56 g/cm^2 throughout the energy range studied. The lines of cut are shown on the left hand side of the lower panel in Fig. 6.28, while the fits obtained using single breaks in the ER are shown on the right hand side. The obtained fit parameters are given in Table 6.9.

Cut	lg(E_0/eV)	X_0(g/cm^2)	D_{10} (g/cm^2/decade) D_1	D_2	Prob	χ^2/Ndf
Default	$18.16^{+0.05}_{-0.04}$	783^{+2}_{-2}	70^{+6}_{-6}	33^{+3}_{-3}	0.20	1.30
shift						
+20 g/cm^2	$18.15^{+0.08}_{-0.08}$	804^{+4}_{-4}	74^{+2}_{-1}	30^{+4}_{-4}	0.28	1.19
−20 g/cm^2	$18.19^{+0.05}_{-0.05}$	773^{+2}_{-3}	75^{+8}_{-8}	24^{+3}_{-3}	0.33	1.12
Fixed						
70 g/cm^2/dec	$18.15^{+0.08}_{-0.08}$	791^{+4}_{-4}	89^{+9}_{-9}	42^{+7}_{-7}	0.47	0.98
45 g/cm^2/dec	$18.17^{+0.06}_{-0.06}$	785^{+2}_{-2}	67^{+13}_{-9}	16^{+3}_{-3}	0.34	1.11

Table 6.9: *Fit parameters for single-break line fits to the 'light' part of Auger data when the position of cut is shifted by* ±20 g/cm^2 *and when it is fixed at 45 and* 70 g/cm^2/decade. *The parameters for the fit with the defined position of cut are included for comparison.*

Since there is no dramatic departure from the ER as originally defined, one can conclude that the inferences made concerning the trend in the 'lighter' component of Auger data are not altered, even if there is a slight change in the position of cut.

The possibility that a fit line with two breaks could describe the data in the 'light' subset well was investigated. The same was not done for the heavy subset of the $X_{\max}$ data since in this case, as has already been mentioned earlier, the break in the line of fit is weakly pronounced. The fit function for a two-break fit is given by:

$$\langle X_{\max}\rangle(E) = \begin{cases} X_0 + D_1 \lg(E/E_0), & E < E_0 \\ X_0 + D_2 \lg(E/E_0), & E_0 < E < E_1 \\ X_0 + D_2 \lg(E_1/E_0) + D_3 \lg(E/E_1), & E > E_1 \end{cases} \tag{6.14}$$

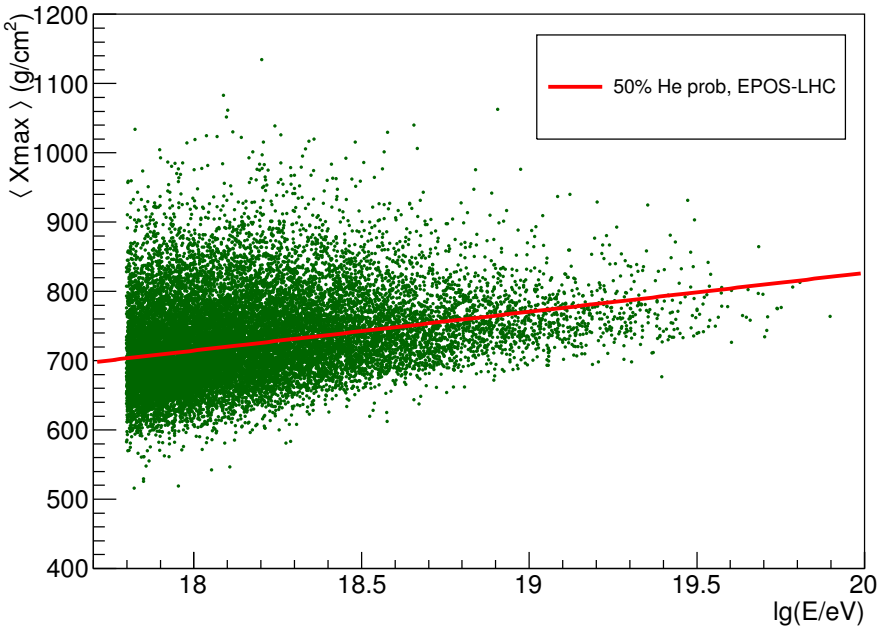
where the parameters X_0, E_0, D_1 and D_2 are the same as in Eqn. 6.2, while E_1 is the energy at the second break and D_3 is the gradient of the last part of the broken line. From Table 6.10 and as has been shown in Fig. 6.26(b), it can be seen that the fit is good, and actually has a higher probability compared to the one with only one break. A look at the ER in the last straight part of the fit, as shown more clearly

$\lg(E/\mathrm{eV})$ at break		D_{10} (g/cm^2/decade)					
E_0	E_1	D_1	D_2	D_3	X_0 (g/cm^2)	χ^2/Ndf	Prob
$18.20^{+0.05}_{-0.04}$	$19.02^{+0.12}_{-0.20}$	70^{+6}_{-6}	28^{+4}_{-4}	64^{+19}_{-17}	785^{+2}_{-2}	0.97	0.48

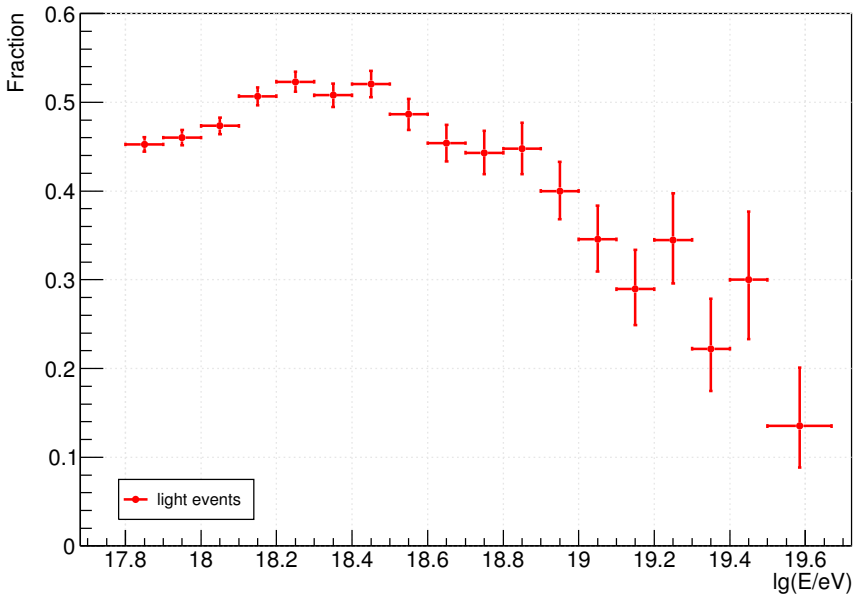
Table 6.10: *Fit parameters for light subset of Auger $\langle X_{\mathrm{max}} \rangle$ data with a two-break line from Eqn. 6.14.*

by comparison to the helium showers with $X_{\mathrm{max}} > X_{\mathrm{max}}^{cut}$ in Fig. 6.29(a), suggests that the ER of the 'light' subset in this energy region is compatible with constant composition (see Table 6.3). This is brought out more clearly by a corresponding fit with two breaks in the ER, in Fig. 6.29(b). Since by construction the 'light' subset is composed mainly of proton and helium, it appears as if protons disappear in this range of energy, leaving behind a composition dominated by helium. The energy at which this change takes place, which is $\lg(E/\mathrm{eV}) \sim 19.0^{+0.12}_{-0.20}$, is slightly above the ankle of the cosmic ray spectrum, but is compatible, within errors.

In the same way as for the full data set, fits of the light and heavy subsets in 3-bin sliding windows were performed (Fig. 6.30). In the light subset, one observes a decrease of the ER relative to pure composition (increase of the primary mass) in the interval $\lg(E/\mathrm{eV}) \approx 18.3 - 19.0$. The change is however not steady, probably due to the low statistics. Afterward, the ER becomes compatible with a constant composition, though the errors are rather too large to make any claim. The heavy part has no explicitly expressed trends or structures and ER is compatible with a constant composition. In the plots for $dD_{10}/d\lg E$ (Fig. 6.31), the situation is similar. In the light subset, $dD_{10}/d\lg E$ changes significantly – from negative values at $\lg(E/\mathrm{eV}) \lesssim 18.5$ to values compatible with zero (the last two points deviate from zero to the positive side). On the other hand, in the heavy subset, the derivative of ER is mostly compatible with zero.

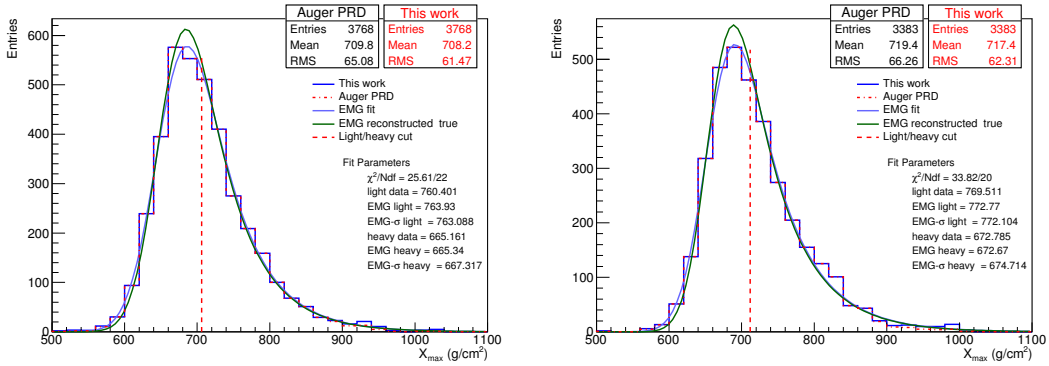


(a) Separation of light/heavy events

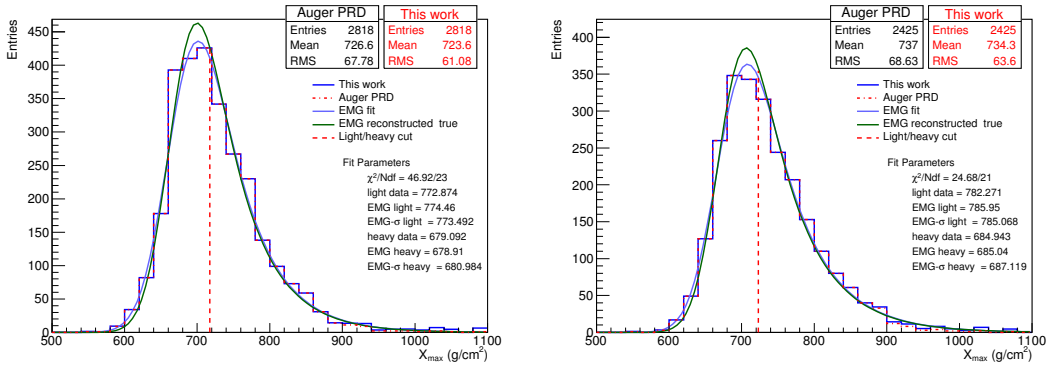


(b) Fraction of events in "light" subset

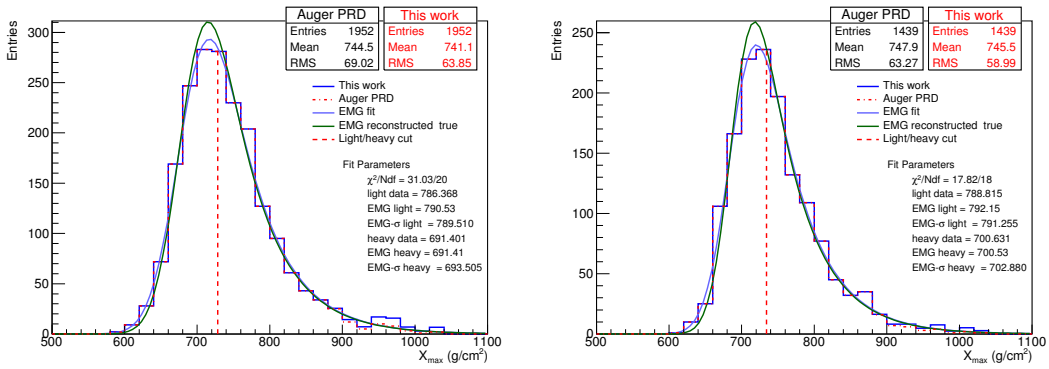
Figure 6.21: *(a) Auger hybrid data [49] with a cut at 50% probability for helium to produce a deeper shower using EPOS-LHC model. (b) the fraction of the events with* $X_{\max}$ *deeper than cut value (the light subset). Errors on the fraction are obtained using Wilson interval [124] (see* `TEfficiency` *class of* `ROOT`*).*



(a) 17.8 – 17.9 (b) 17.9 – 18.0

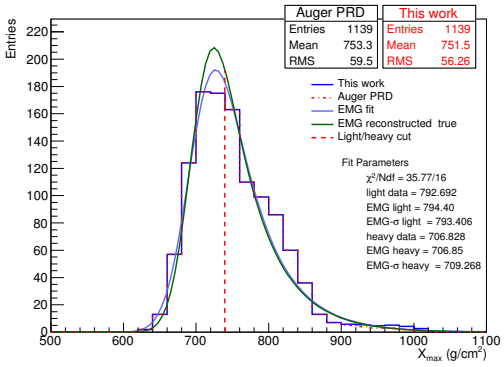


(c) 18.0 – 18.1 (d) 18.1 – 18.2

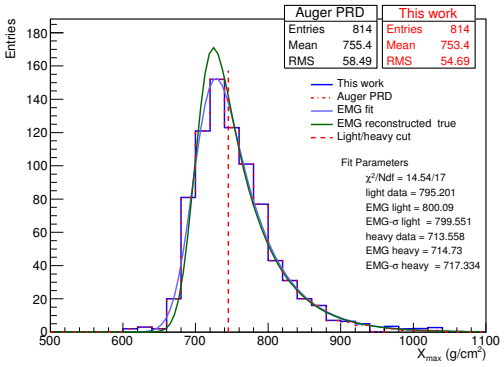


(e) 18.2 – 18.3 (f) 18.3 – 18.4

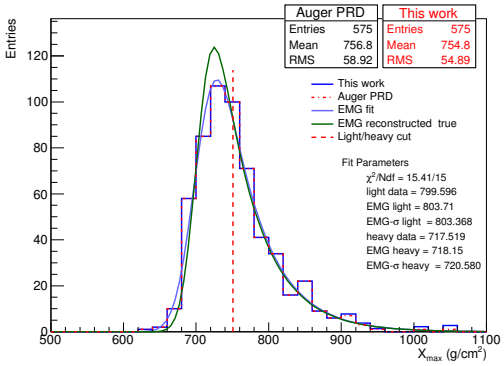
Figure 6.22: *Fits of the exponentially modified Gaussian function to the Auger data in bins with ordinal numbers* $1-6$ *(blue curve). The deconvoluted distribution is represented in green and the mean values of* $X_{\max}$ *for the combined data, ‘light’ and ‘heavy’ subsets are as shown on the lower right hand side of each panel. The cut on the* $X_{\max}$ *values at the mean bin energy is represented by the vertical line in each case and the numbers indicate energy ranges in* $\lg(E/\mathrm{eV})$.



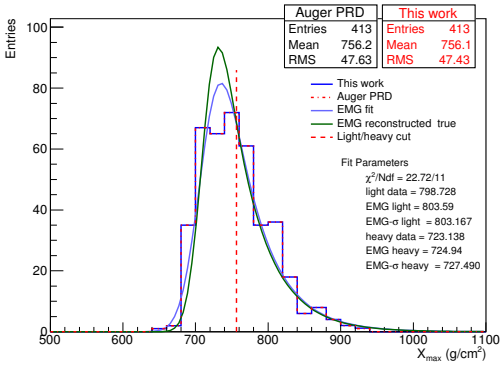
(a) 18.4 – 18.5



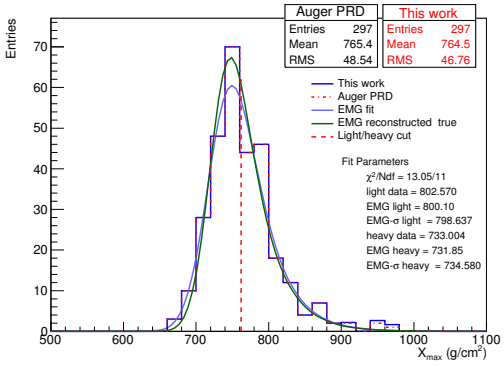
(b) 18.5 – 18.6



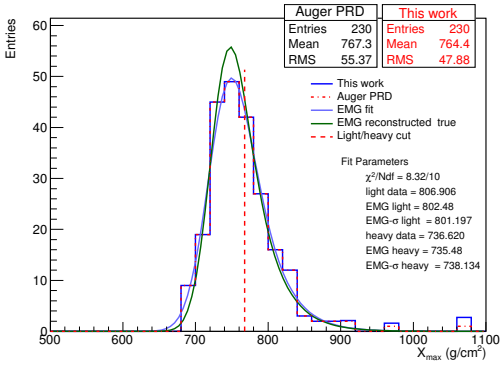
(c) 18.6 – 18.7



(d) 18.7 – 18.8

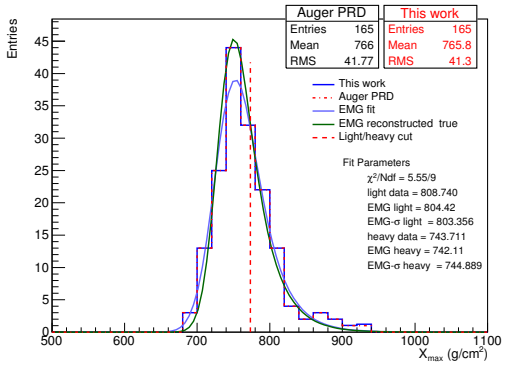


(e) 18.8 – 18.9

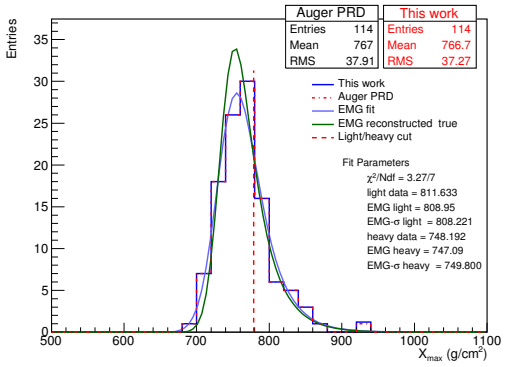


(f) 18.9 – 19.0

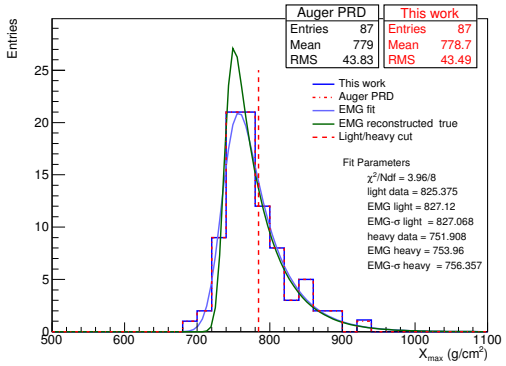
Figure 6.23: *Same as in Fig. 6.22 but for bins with ordinal numbers* $7 - 12$.



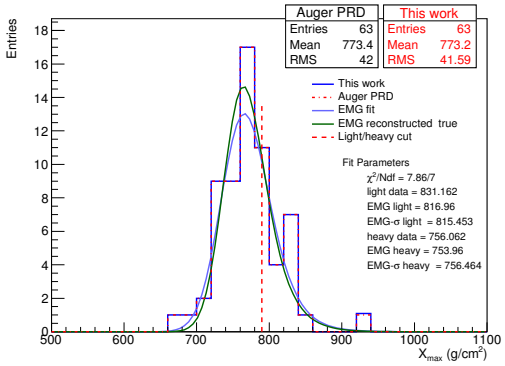
(a) 19.0 – 19.1



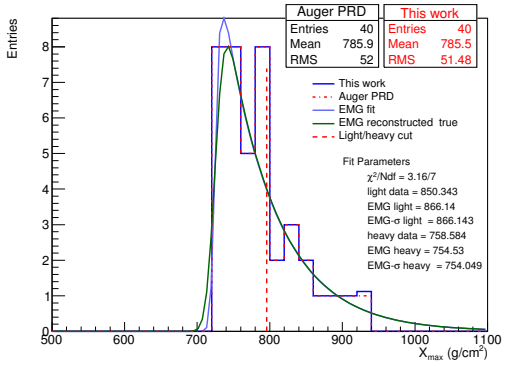
(b) 19.1 – 19.2



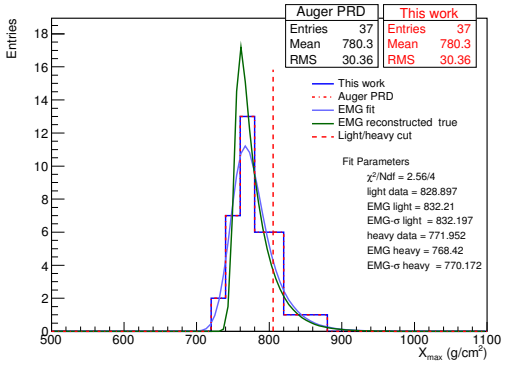
(c) 19.2 – 19.3



(d) 19.3 – 19.4



(e) 19.4 – 19.5



(f) 19.5 – ∞

Figure 6.24: *Same as in Fig. 6.22 but for bins with ordinal numbers* 13 – 18.

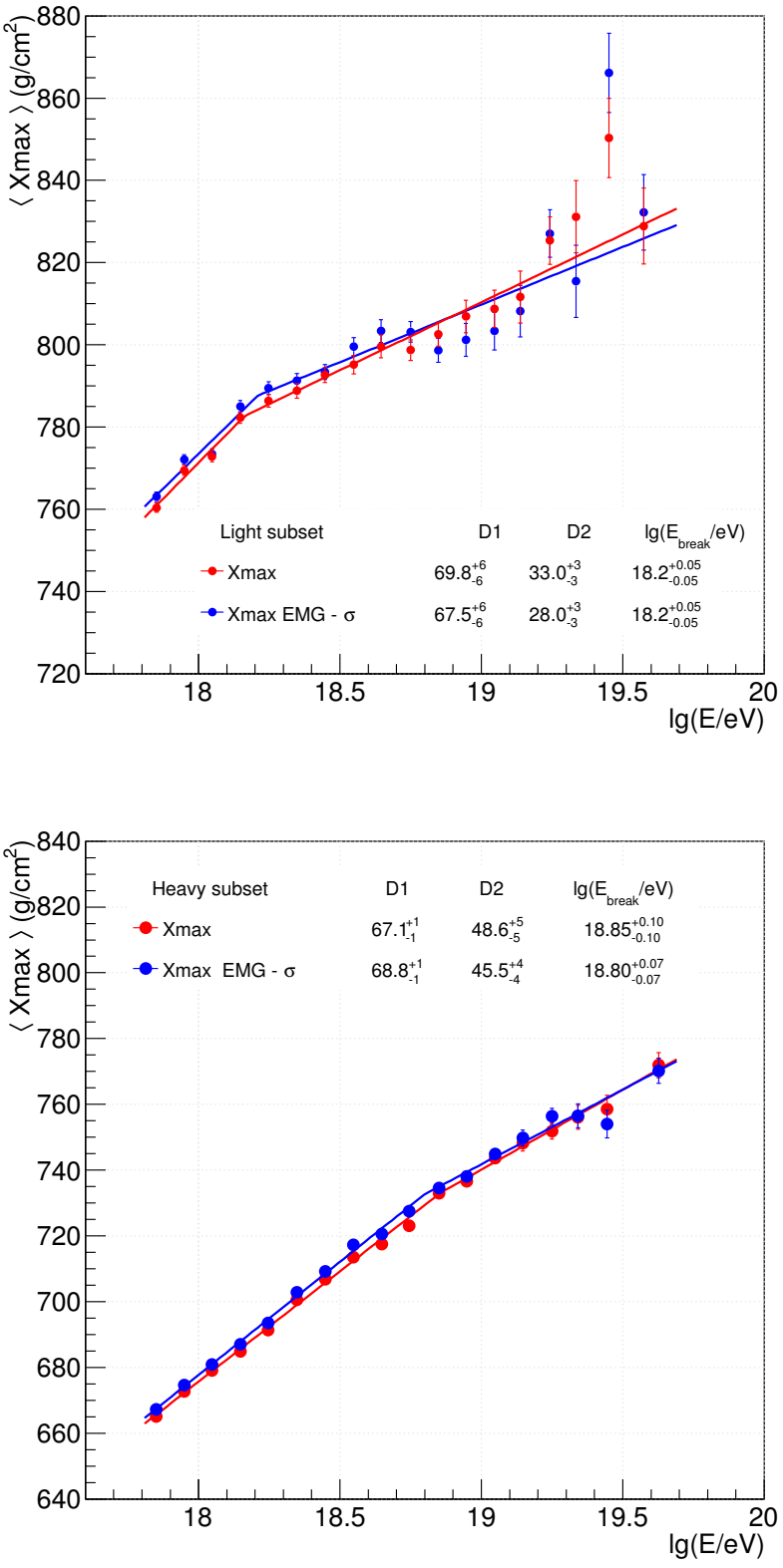
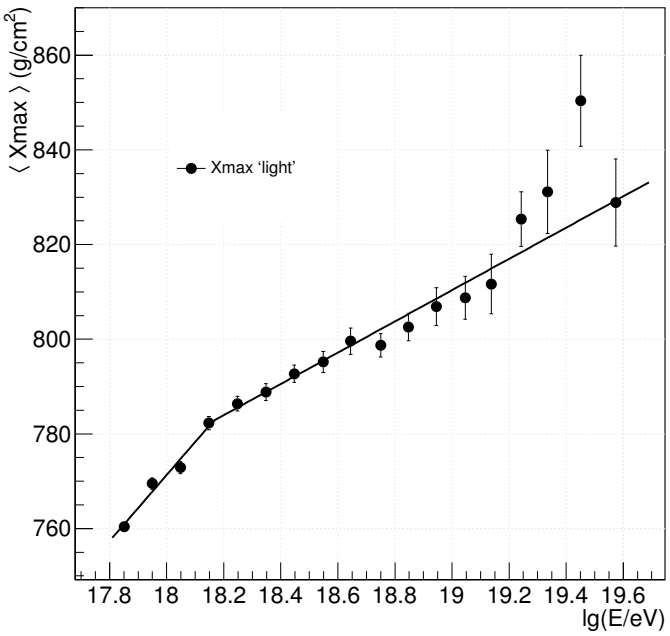
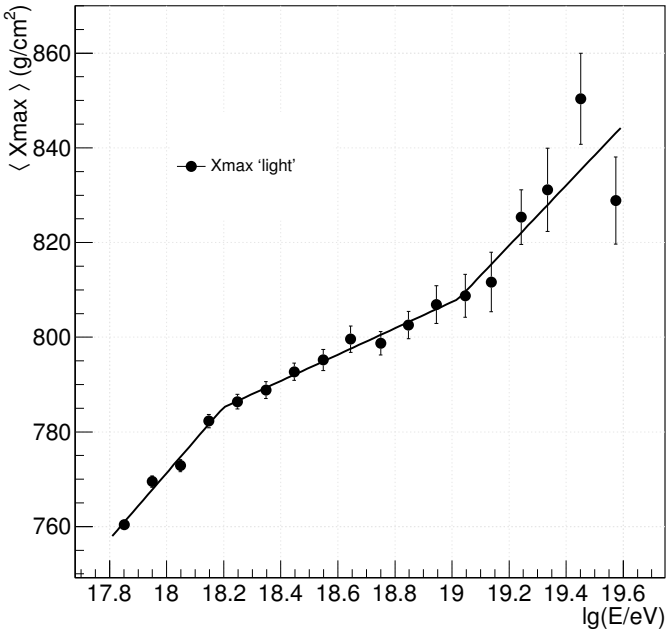


Figure 6.25: Top panel: *The ER fits obtained before and after correcting for resolution effect in the values of FD $X_{\max}$ of the 'light' subset.* Bottom panel: *Similar fits for the 'heavy' subset.*

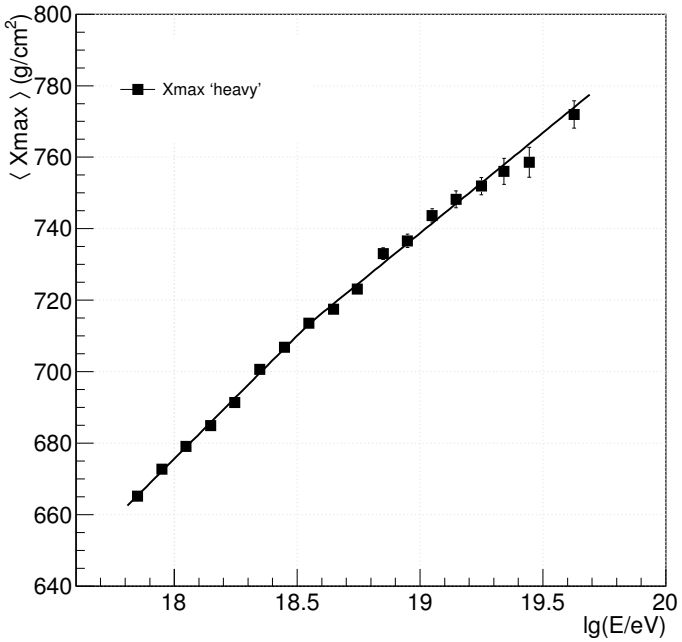


(a) Break at one point

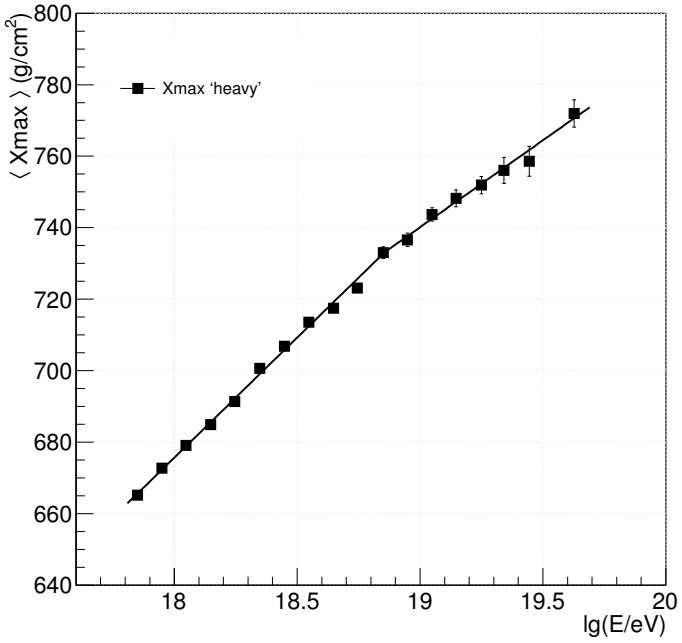


(b) Breaks at two points

Figure 6.26: Top panel: *Fit to the 'light' subset of Auger* $\langle X_{\max} \rangle$ *data, with a single-break line. Parameters of the fit function (solid line) are given in Table 6.8.* Bottom panel: *Fit to the same subset of Auger data, but with two breaks in the line of fit. The fit parameters are found in Table 6.10.*



(a) Break at $\lg(E_0/\mathrm{eV}) = 18.55$



(b) Break at $\lg(E_0/\mathrm{eV}) = 18.85$

Figure 6.27: *Fits to Auger data $\langle X_{\mathrm{max}} \rangle$ of events in the heavy subset. The parameters of the fit functions (solid lines in each case) are given in Table 6.8.*

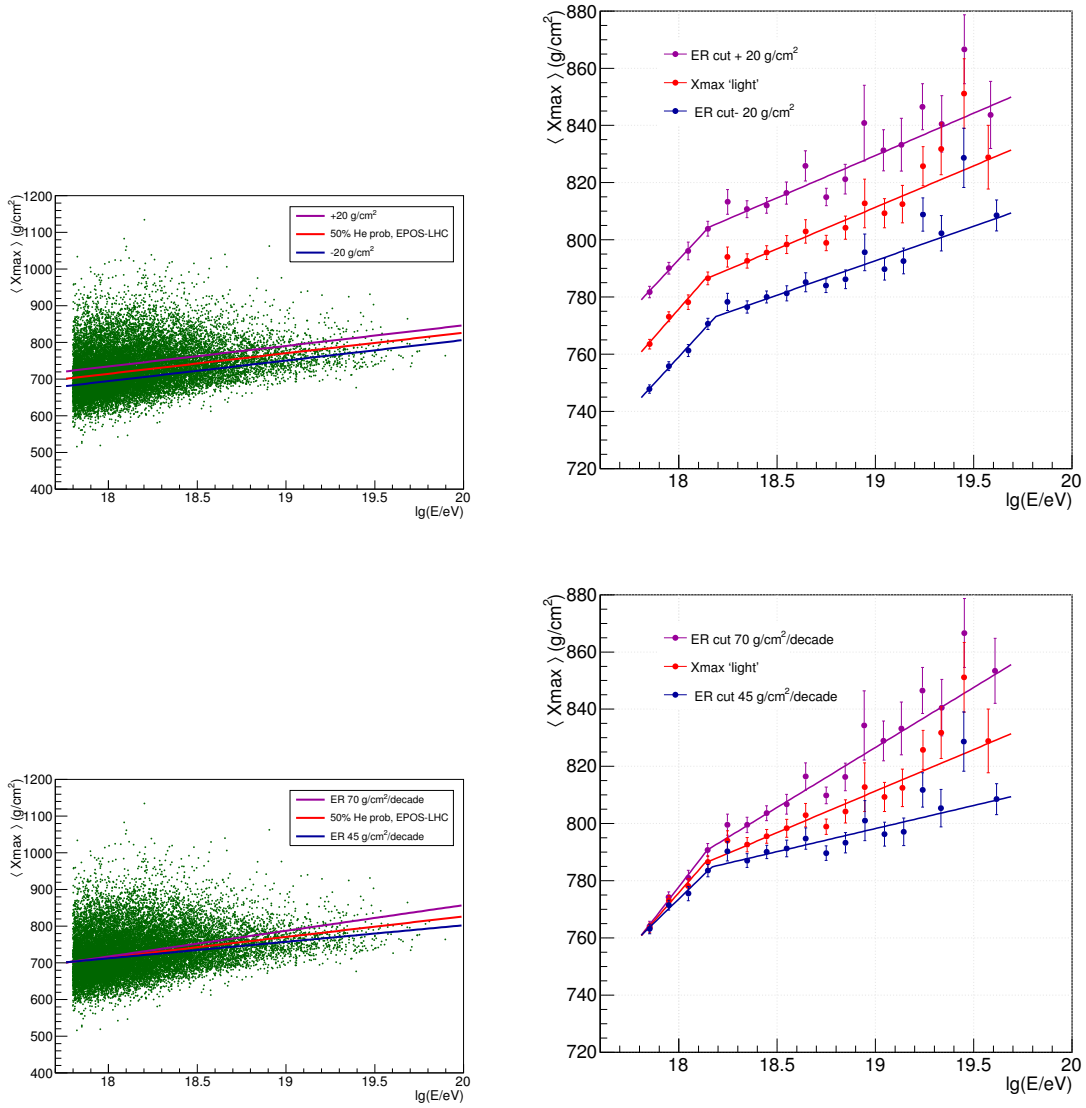
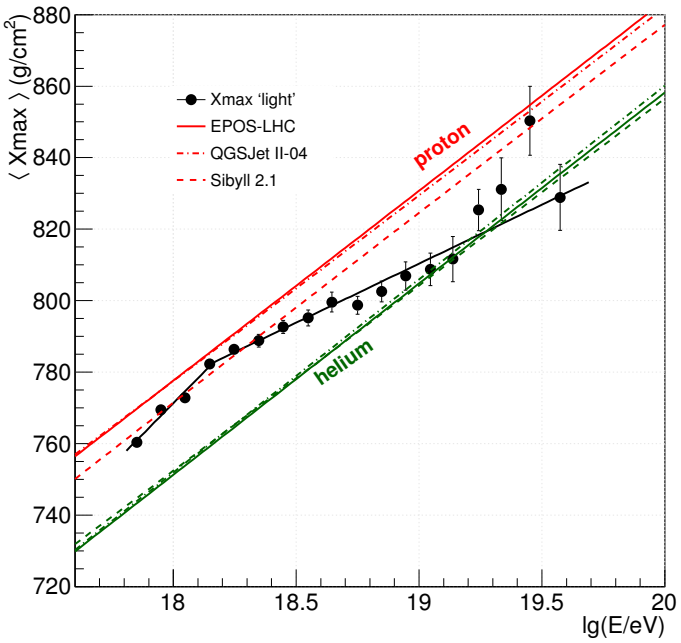
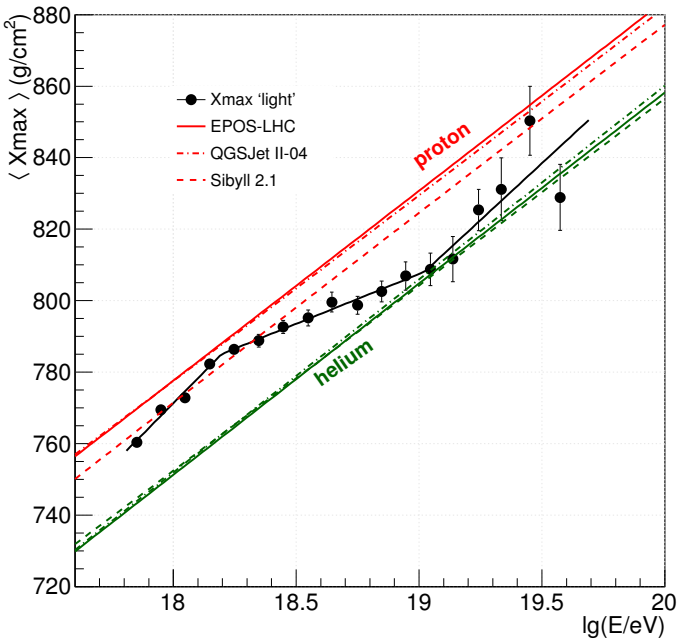


Figure 6.28: Top row: *Cuts on $X_{\max}$ data shifted by 20* g/cm^2 *from the originally defined value* (left) *and the resultant single-break ER curves for the 'light' subset* (right). Bottom row: *ER when the lines of cut are fixed at 70 and 45* g/cm^2 (left) *and the corresponding single-break fits to the data* (right). *The fit parameters are given in Table 6.9.*

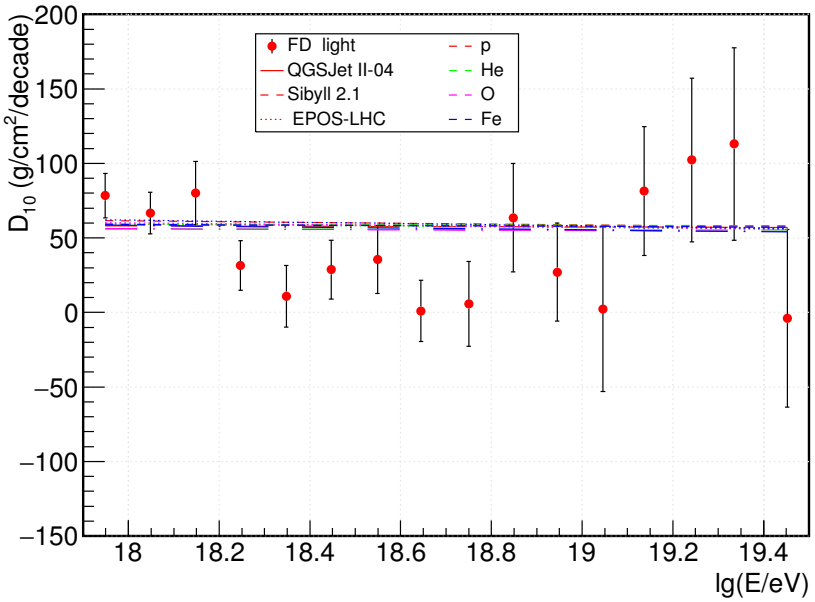


(a) Single break

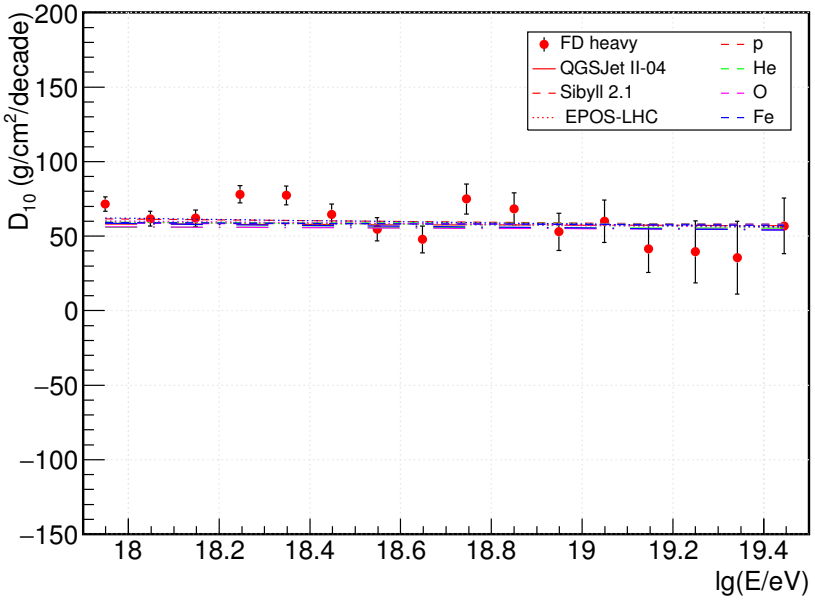


(b) Double break

Figure 6.29: *Fits to Auger data $\langle X_{\max} \rangle$ of events in the light subset. The pure compositions are for parts of proton and helium distributions with $X_{\max} > X_{\max}^{cut}$ and the parameters of the fit functions (solid lines in each case) are given in Tables 6.8 and 6.10.*

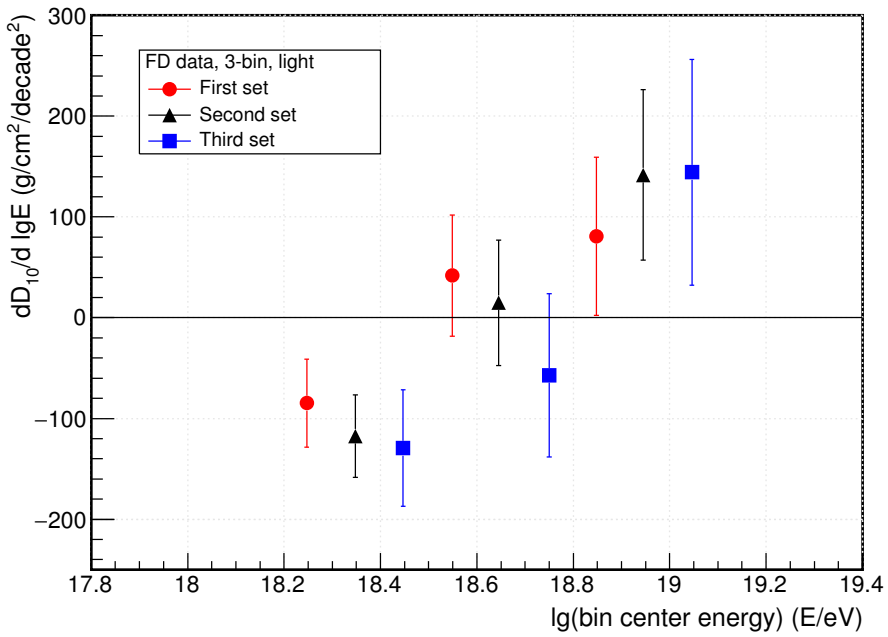


(a) light subset

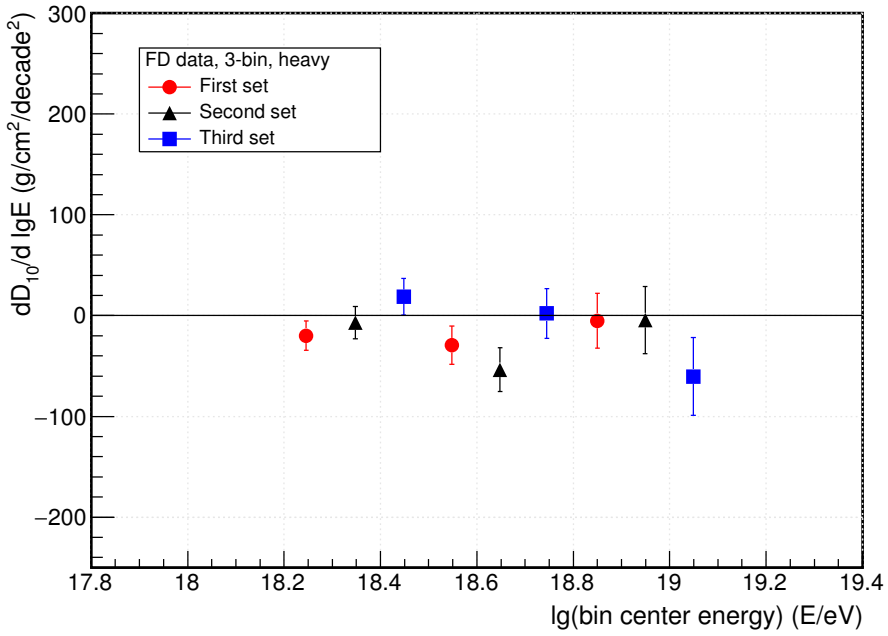


(b) heavy subset

Figure 6.30: *ERs in sliding windows of three bins for the Auger $\langle X_{\mathrm{max}} \rangle$ data in the light (a) and heavy (b) subsets together with MC values for different primary particles and interaction models. Data values larger (smaller) than the MC values indicate a composition getting lighter (heavier). Note that the data points are correlated.*



(a) light subset



(b) heavy subset

Figure 6.31: *Evolution of* $dD_{10}/d\lg E$ *with* $\lg E$ *for the Auger data, in the light (a) and heavy (b) subsets, obtained using linear fits of the ERs for* $\Delta\lg(E/\mathrm{eV}) = 0.1$ *(see Fig. 6.30) in 3-bin sliding windows. A value of zero indicates a constant rate of primary mass change. The data points are correlated because of the overlap of the windows used.*

6.3.4 Application of cut on SD data

The effect of applying the cut as defined in section 6.3.1 to the SD data [125] at $\lg(E/\mathrm{eV}) \geq 18.8$ has been investigated. This is natural, since the SD has a much larger duty cycle and hence has collected more data than the FD, and the possibility of applying the procedure of the cut on the SD data is interesting to look at. The procedure for acquiring the SD data has been described in section 5.2. Due to the events whose energy is below 10^{19} eV and which additionally arrive at zenith angles above 38° being outside the full efficiency for the universality approach, such events have been excluded from analysis. This should eliminate an obvious bias from the reconstructed X_{max}.

The combined SD X_{max} data at the highest energy is very compatible with the FD data, as seen in Fig. 6.32. In this plot, a broken line fit to the FD data is shown alongside a straight line fit to the SD data. One can observe a very good agreement

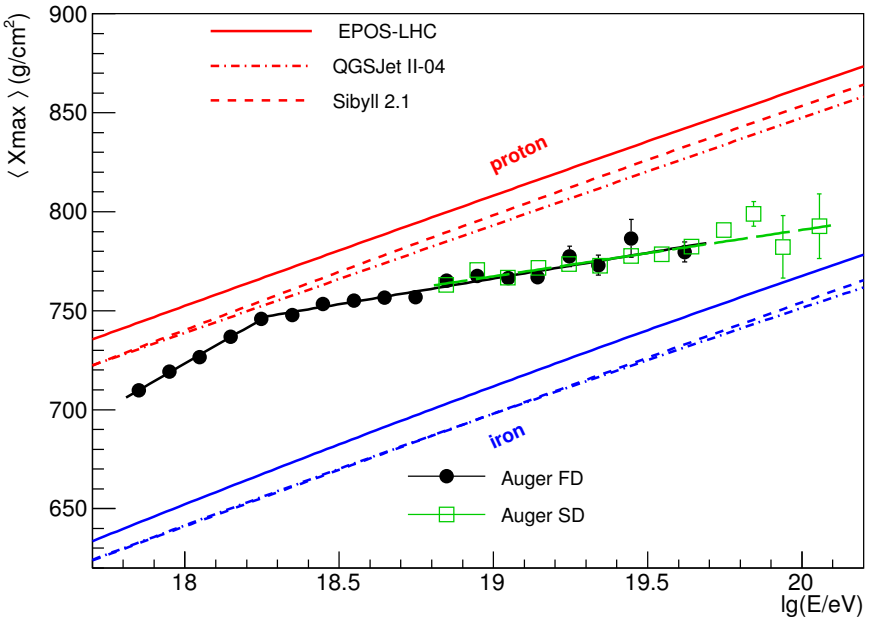
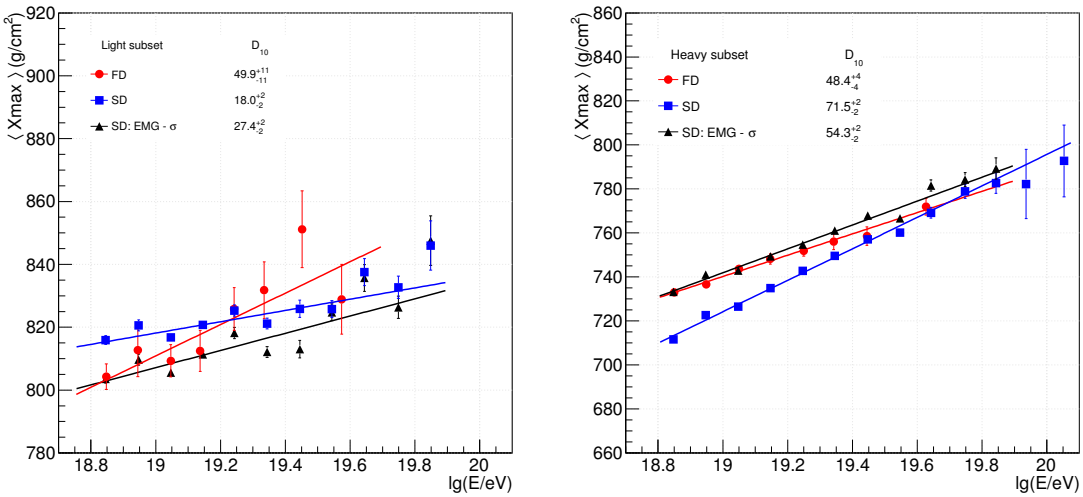


Figure 6.32: *Fits to the full energy range of the Auger FD data and SD data in the range* $\lg(E/\mathrm{eV}) \geq 18.8$.

between the ERs of FD and SD data. The gradient of the second part of the broken line fit for the FD data is 26 ± 3 g/cm^2/decade, while that of the SD fit is 23 ± 2 g/cm^2/decade.

After the application of the cut separating the SD data into light and heavy components, the elongation rates obtained differ significantly from those of the FD data (Fig. 6.33). In part, this is an expected behaviour since the SD X_{max} data had not yet been corrected for resolution effects, while that of the FD had been corrected.

As for the FD data, the procedure of EMG analysis was applied to the SD data.



(a) light components (b) heavy components

Figure 6.33: *Straight line fits for* $\lg(E/\mathrm{eV}) \geq 18.8$ *to the (a) light and (b) heavy FD and SD Auger data. The circles are for the FD data while the squares are for the SD data. The effect of applying EMG correction for the removal of resolution bias is shown by the triangles, and the values of the gradients of the straight line fits are as indicated, in* g/cm^2/decade.

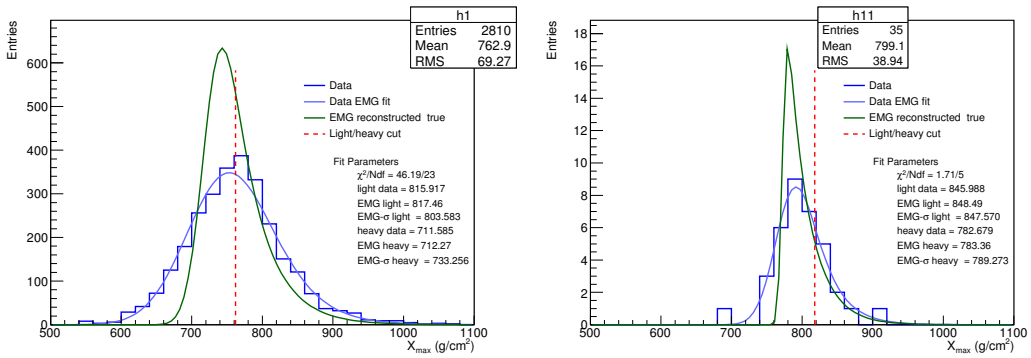
A typical result is shown in Fig. 6.34(a) for the bin $\lg(E/\mathrm{eV}) = 18.8 - 18.9$. In the last two bins, the fit does not converge due to a scarcity of data. Apart from these two cases, the fit parameters obtained in the course of EMG analysis were generally good, with the shape of fit curves conforming to that expected of an EMG. In the last bin in which the fit converged (Fig. 6.34(b)), it can be seen that the EMG has a rather sharp peak. Nevertheless, even in this case, good fit parameters are obtained.

The SD resolution was estimated to vary according to [15, 125]

$$\sigma_{\mathrm{res}}(\lg(E/\mathrm{eV})) = 15 + be^{(-c(\lg(E/\mathrm{eV})-19))}\ \mathrm{g/cm^2}, \tag{6.15}$$

where $b = 28$ and $c = 1.35$. The resolution for the SD data, compared to that of the FD, is quickly decreasing with energy, from ~ 50 g/cm^2 to ~ 25 g/cm^2 within the range $\lg(E/\mathrm{eV}) = 17.8 - 20.0$ (Fig. 6.35). Thus the bias in truncated X_{max} is larger at the lower energies. At energy $\lg(E/\mathrm{eV}) = 18.8 - 18.9$ which corresponds to a resolution of ≈ 45 g/cm^2, for example, the EMG analysis yields a bias in the light subset of -12 g/cm^2, and $+22$ g/cm^2 for the heavy subset. At a higher energy $\lg(E/\mathrm{eV}) = 19.8 - 19.9$ where the resolution is ≈ 25 g/cm^2, the biases in the light and heavy subsets reduce to -2 g/cm^2 and -7 g/cm^2 respectively.

The procedure reduces the differences between ERs of FD and SD data in the 'light' and 'heavy' subsets to levels that are compatible with each other. Although a fit of the light subset using a free parameter gives an ER in FD data of 50 ± 11 g/cm^2/decade compared to that in SD data of 27 ± 2 g/cm^2/decade, keeping D_{10} for the FD fixed at the SD value results in a fit probability of 24.4%. At the same



(a) High statistics (b) Low statistics

Figure 6.34: *(a) Deconvolution of SD* $X_{\max}$ *in the bin defined by* $\lg(E/eV) = 18.8 - 18.9$ *in which statistics are high. The results of EMG analysis are indicated on the right hand side. (b) Results for* $\lg(E/eV) = 19.8 - 19.9$*, a typical bin with low statistics.*

time, in the heavy subsets, the ER in the FD data is 48 ± 4 g/cm^2/decade while in the SD data it is 54 ± 2 g/cm^2/decade, after correcting for the bias using the EMG procedure.

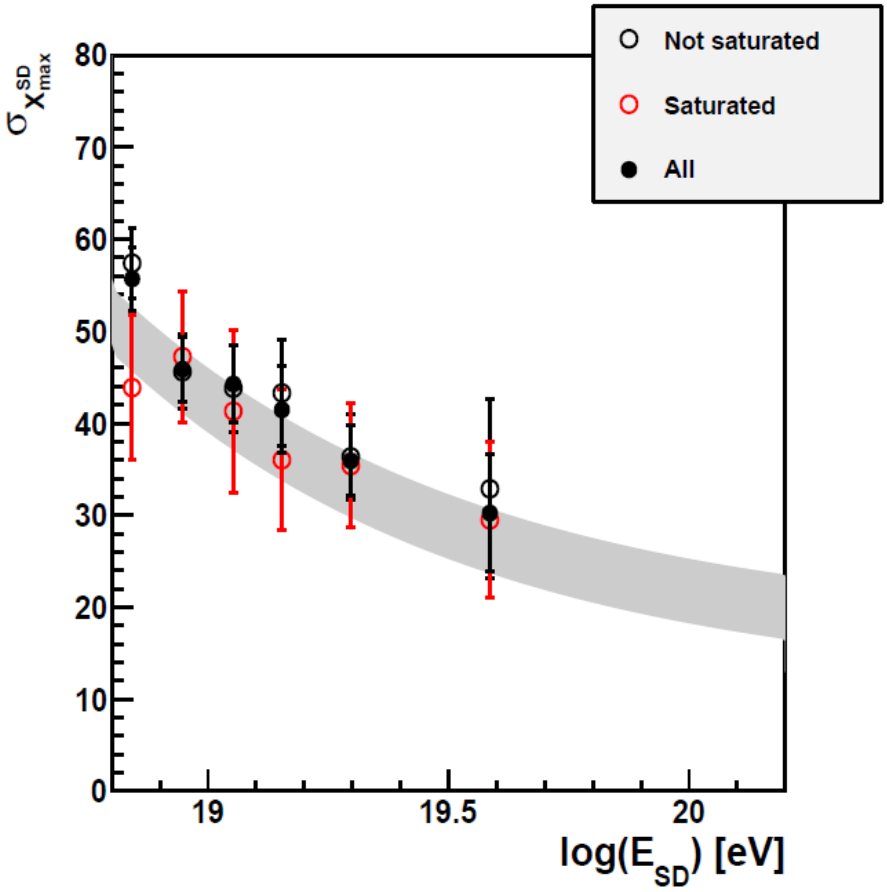


Figure 6.35: *An estimate of the growth of resolution of $X^{\mathrm{SD}}_{\mathrm{max}}$ with energy. The shaded band shows a MC estimation of resolution [15].*

6.4 Varied proton-air cross-section

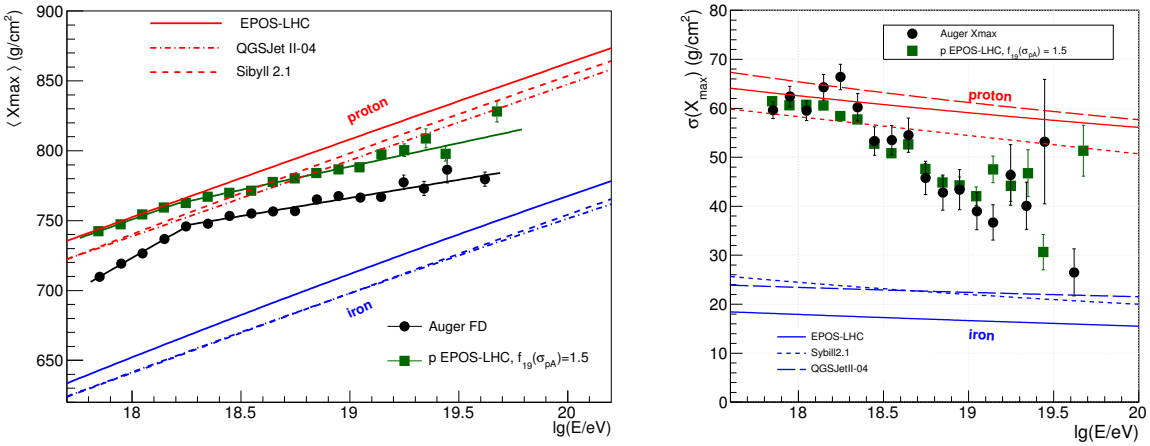
A fit to the light component of the Auger data as defined in section 6.3.1 does not give good results for a single straight line fit, but for a broken line fit as seen in Fig. 6.26, the fit is reasonable. This reinforces the observation in Fig. 6.21(b), that the fraction of 'light' events contained in the combined data seems to rise to a maximum at around $\lg(E/\mathrm{eV}) = 18.2$ before beginning to fall. However, in the context of the dip model in which all the primary cosmic rays are protons, the transition from the composition getting lighter to getting heavier can only occur as a result of the interaction cross-section of protons increasing. In regard to this, a simulated pure proton composition has been investigated for the possibility that at energy $\lg(E/\mathrm{eV}) = 18.2$, the cross-section of proton-air interaction, $\sigma_{\mathrm{p-Air}}$, begins to change with increase in energy. At the same time, a significant fraction of protons in the 'heavy' part would manifest itself in a notable change in ER in this subset. The simulation was carried out using the code CONEX, with EPOS-LHC [126] as the hadronic interaction model. A total of 160000 events were simulated. The value of $\sigma_{\mathrm{p-Air}}$ was treated as changing by a factor [127]:

$$f(E, f_{19}) = 1 + (f_{19} - 1)\frac{\lg(E/\,\mathrm{eV} - 18.2)}{\lg(10^{19}\,\mathrm{eV}/10^{18.2}\,\mathrm{eV})} \tag{6.16}$$

where f_{19} is 1.5. A comparison of the first two moments of X_{max} with those in the Auger data shows that whereas the RMS is described quite well (Fig. 6.36(b)), the mean is not (Fig. 6.36(a)). The probability of broken line fit of MC $\langle X_{\mathrm{max}} \rangle$ is 0.02 while $\chi^2/\mathrm{Ndf} = 27.8/14$. It seems therefore, that some energy-dependent composition adjustments are necessary in order to bring the MC values into agreement with the experimentally observed data. A statistical fluctuation is seen in the last two points of MC $\sigma(X_{\mathrm{max}})$. These points should therefore be interpreted with caution.

The behavior of the light and heavy subsets are shown in Fig. 6.37. Unlike the case of the combined Auger data, it is observed that there is no break at all in the heavy part, while the light part has a break that is much less pronounced than the corresponding data part. The ER before the break in the light subset differs appreciably from the corresponding ER in data as can be seen in comparing Fig. 6.26(a) with Fig. 6.37(a). Whereas the data gives a value of 70 ± 6 g/cm^2/decade as seen in Table 6.8, the MC result is only 55 ± 2 g/cm^2/decade. After the break, the MC light fit ER results are compatible with data, with the values being 34 ± 1 and 33 ± 3 g/cm^2/decade respectively. In the case of the heavy subset, the ER of the single line fit in MC is not compatible with that in the region of data before break. The data results range from 64 ± 1 to 69 ± 1 g/cm^2/decade (Fig. 6.27, Table 6.8) while the outcome for MC is 57 ± 0 g/cm^2/decade. It is noteworthy that a check on the ER for $f_{19} = 1.5$ reveals that just like in the case of data, there is no break in the heavy part but only in the light part.

In a related study [128], an investigation of the parameters in the QGSJetII-04 hadronic interaction model was carried out by MC simulation of primary protons using CONEX. The aim of this study was to adjust the model, such that the MC X_{max} and $\sigma(X_{\mathrm{max}})$ distributions get as close to data as possible. It would then be possible to compare the behaviour of the ER in the light and heavy subsets of the



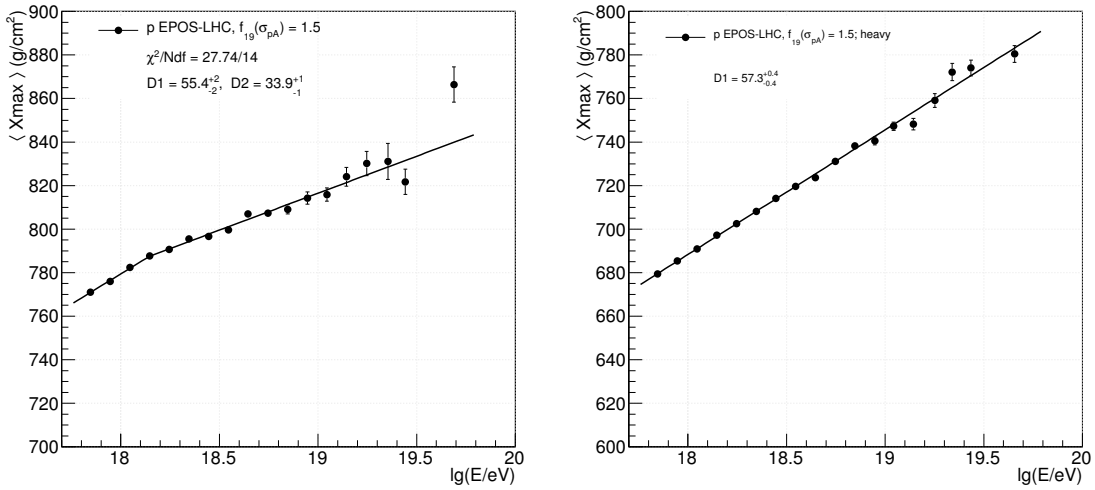
(a) Elongation rates

(b) RMS of $X_{\max}$

Figure 6.36: *(a) Elongation rates for a proton-dominated scenario with EPOS-LHC in which the proton-air cross-section is altered, beginning at* $\lg(E/eV) = 18.2$ *according to Eqn. 6.16 (*$f_{19} = 1.5$*), contrasted to the ER in [49] and (b) the distribution of the second* $X_{\max}$ *moments.*

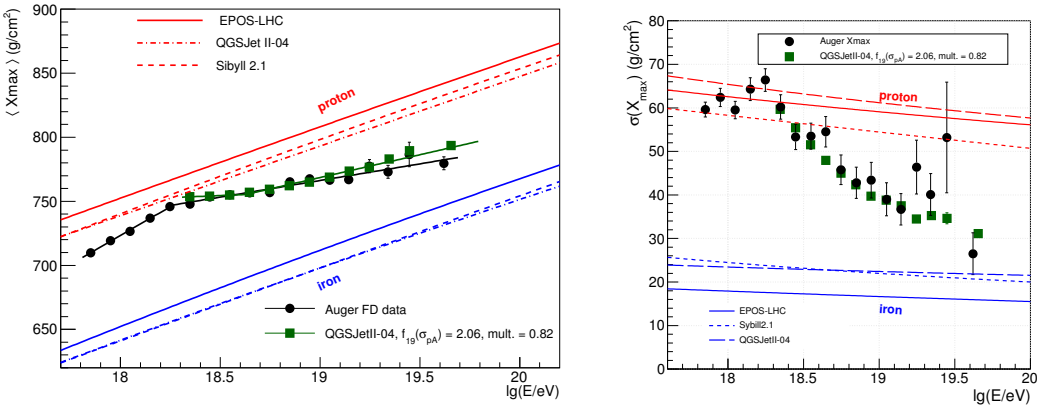
resulting distribution with that of the data. The $\sigma_{\mathrm{p-Air}}$ was made to change linearly according to Eqn. 6.16, starting at $\lg(E/\mathrm{eV}) = 18.3$. From the scan of a grid of multiplicity and cross-section factors obtained from a number of productions, an optimum combination of the two was found to be 0.82 and 2.06 respectively. The evolution with energy of $X_{\max}$ from simulations performed using these multiplicity and cross-section factors, binned in a similar manner to the Auger data, is shown in Fig. 6.38. It can be seen that the first two moments of $X_{\max}$ agree with the Auger data quite well. The implication of this is that a scenario where the primary CRs are dominated by protons throughout the energy range measured could be feasible if there is linear variation in cross-section, multiplicity and possibly other factors with $\lg(E/\mathrm{eV})$.

Nevertheless, the ER of the light part for this modification of QGSJetII-04 differs from the ER in FD data. In particular, the ER is flatter below the break (Fig. 6.39); its value is only 11 ± 2 g/cm^2/decade compared to 27 ± 5 g/cm^2/decade for the FD data. Likewise, above the break, the ER is 36 ± 2 g/cm^2/decade in the modified cross-section, while in the FD data it is 64 ± 16 g/cm^2/decade. Hence a proton-dominated scenario, which is a basic assumption of the dip model, is questioned. This is in agreement with the analysis of the full Auger $X_{\max}$ data [49]. One has to note, however, that the amount of data that is currently available at the highest energy is quite small. As more $X_{\max}$ data becomes available, possibly from the Auger SD, a confirmation of the existence of the second break would strongly disfavor the hypothesis that changes in interaction cross-sections are responsible for the observed increase in mass above $\lg(E/\mathrm{eV}) = 18.3$.



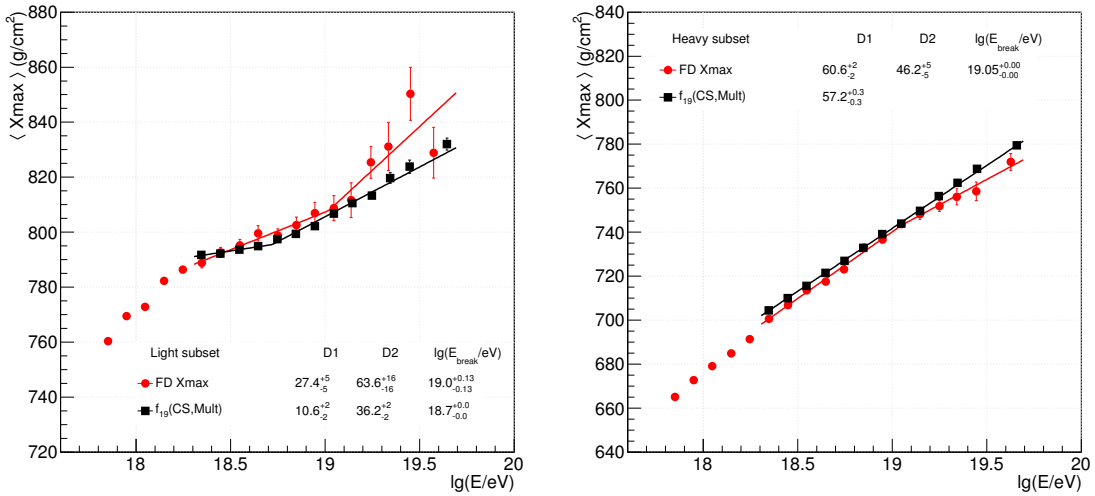
(a) light part (b) heavy part

Figure 6.37: *Elongation rates for pure proton events in which the proton-air cross-section from* $\lg(E/eV) = 18.2$ *is altered by a factor given by Eqn. 6.16 (*$f_{19} = 1.5$*). The values of the fit parameters are as indicated (see text for a discussion).*



(a) Fit from $\lg(E/\mathrm{eV}) = 18.2$ (b) $\sigma(X_{\max})$ for $\lg(E/\mathrm{eV}) \geq 18.3$

Figure 6.38: *(a) A broken line fit to pure proton composition simulated with QGSJetII-04, whose cross-section* $f_{19}(\sigma_{\mathrm{p-Air}}) = 2.06$ *and multiplicity* $f_{19}(\sigma_{\mathrm{mult}}) = 0.82$ *are changing from* $\lg(E/\mathrm{eV}) = 18.3$ *according to Eqn. 6.16 [128], together with a fit to the Auger FD data. (b) The corresponding* $\sigma(X_{\max})$ *distribution.*



(a) Light fit from $\lg(E/\mathrm{eV}) = 18.3$ (b) Heavy fit in the same energy range

Figure 6.39: *A broken line fit to (a) light and (b) heavy subsets of pure proton composition, simulated with QGSJetII-04, whose cross-section* $f_{19}(\sigma_{\mathrm{p-Air}}) = 2.06$ *and multiplicity* $f_{19}(\sigma_{\mathrm{mult}}) = 0.82$ *are changing from* $\lg(E/eV) = 18.3$ *[128]. Broken line fits to Auger light and heavy subsets are included in each case for comparison. The elongation rates and energy at the point of break are as indicated.*

7

Summary and Outlook

In this work, a study of the mass composition of ultra-high energy cosmic rays based on the data from the Pierre Auger Observatory has been carried out. The main focus has been the FD $X_{\rm max}$ data reported in [49]. In addition to this, the study has taken a first look at the SD $X_{\rm max}$ data. The results of ER analysis of the FD data has led to conclusions that are similar to those in [49]. However, some finer details regarding the changes in mass composition, which are difficult to observe in [49], are additionally reported in this study. This has been made possible by the calculation of ER in sliding windows, and the truncation of the data into 'light' and 'heavy' subsets. The behavior of $\langle X_{\rm max}\rangle$ is analyzed in terms of change in mass vis-a-vis a growth in interaction cross-section.

This study has taken advantage of the considerable statistics on $\langle X_{\rm max}\rangle$ as a function of energy that has been accumulated by the Pierre Auger Observatory since it started collecting data in January 2004. Compared to the analysis presented in [129], for example, one now has statistics that are larger by a factor of almost three. Due to the increased statistics, coupled with a lower energy threshold, the ER can be studied in narrower energy bins. The results obtained in this study using 3-bin sliding windows ($\Delta \lg(E/{\rm eV}) = 0.3$) yields ER at $\lg(E/{\rm eV}) \approx 18.0$ that varies from 82 to 100 g/cm^2/decade, with a typical error of ± 10 g/cm^2/decade. This is larger than MC values for pure compositions ($55 - 62$ g/cm^2/decade), an indication that the primary mass is getting lighter. For $\lg(E/{\rm eV}) \approx 18.2 - 18.8$, the value of D_{10} is less than that of pure composition, reaching a minimum at $\lg(E/{\rm eV}) \approx 19.05$ where $D_{10} = -4 \pm 24$ g/cm^2/decade, an indication that the composition is getting heavier. A relatively narrow transition in the sliding windows is observed in the interval $\sim 1.3 - 2.0$ EeV between the regions where the composition is getting lighter and where it is getting heavier. These observations still remain the same when the bin width in the data is altered to $\Delta \lg(E/{\rm eV}) = 0.05$. The statement that the composition is getting heavier implies, in particular, that it is not pure. In other words it should be mixed, an observation which supports the results of [115].

The rate of change of ER with energy, $dD_{10}/d\lg E \sim -d^2\langle \ln A\rangle/d(\ln E)^2$, is negative and differs from zero for energies $\lg(E/{\rm eV}) \approx 18.2 - 18.5$ by up to 3σ. This indicates that the rate of change of the primary mass, $d\langle \ln A\rangle/d\ln E$, might not be

constant and that $\langle \ln A \rangle$ is a convex function of the logarithm of energy in this range (cf. Fig. 4 in [79]).

The separation of the full data set into 'light' and 'heavy' parts using a cut on $X_{\max}$ has shown that the ER for the 'light' part changes significantly with energy. It indicates that the primary mass is initially decreasing up to $\lg(E/\mathrm{eV}) = 18.2$, beyond which it begins increasing. It is observed that at $\lg(E/\mathrm{eV}) \approx 19.0$, the increase in primary mass ceases and the primary mass becomes consistent with MC predictions of a constant composition. The 'heavy' part of the data has an ER that is essentially compatible with a constant composition for all energies, although the weakly pronounced broken line fit might indicate some increase of the primary mass for energies above $\lg(E/\mathrm{eV}) \approx 18.9$.

The 'light' part of the Auger data, by construction, should be dominated by protons and helium nuclei. The observed trend in ER, that it is getting heavier as the energy increases, can come about as a result of the fraction of protons decreasing, thus leaving a more helium-dominated composition or due to the increase of the proton interaction cross-section. The latter scenario is challenged by the results of this study due to the observation that at $\lg(E/\mathrm{eV}) \approx 19.0$, the light part begins to be compatible with constant composition. The results for simulated distributions where the proton interaction cross-section is increasing from $\lg(E/\mathrm{eV}) = 18.3$ and $f_{19}(\sigma_{\mathrm{p-Air}}) = 2.06$, $f_{19}(\sigma_{\mathrm{mult}}) = 0.82$ do not give any hint of the composition beginning to look similar to a constant composition at any later point. Contrary to what would be expected if the primary composition was proton-dominated but with a varying $\sigma_{\mathrm{p-Air}}$ or σ_{mult}, the ER of the 'light' part of this modification of QGSJetII-04 is found to differ from the ER in the 'light' subset of data. A broken line fit gives D_{10} in the first part to be 27 ± 5 and 11 ± 2 g/cm^2/decade for the FD and modified cross-section respectively; the second part has the respective values of D_{10} as 64 ± 16 and 36 ± 2 g/cm^2/decade.

Using the Auger data, this study has shown that the mass of the 'lighter' component increases from $\sim 10^{18.2}$ eV, and might be getting constant from $\approx 10^{19.0}$ eV. The second break is however still subject to confirmation, as the little data beyond the ankle that is currently available make the statistical errors to be large.

The behavior of ER for the 'heavy' part, which indicates that composition changes there are small, can be in a large part just due to the decrease of the proton fraction: the contribution of protons to the heavy part is relatively small. It is worth noting that in [81], an analysis of $X_{\max}$ produced fits to the Auger data of fractions of proton and iron that were reasonable for all hadronic interaction models only in the cases where intermediate nuclei such as helium were also present in the primary cosmic rays. Thus a decrease in the proton fraction would leave the intermediate and heavier nuclei dominating the CR composition in the 'heavy' subset.

This analysis shows that the increase of the primary mass, indicated by the values of the ER for the whole Auger FD data set in the energy range $\lg(E/\mathrm{eV}) \approx 18.4 - 19.0$, might happen mostly due to the reduction of fraction of protons. For the sum of other components, the behavior of the ER appears consistent with expectations for a constant primary mass.

The ER in the SD data of 23 ± 2 g/cm^2/decade is compatible with that in the FD, of 26 ± 3 g/cm^2/decade. After correction for bias in the mean values of

the truncated X_{max} distributions measured by the SD, ERs in the light and heavy subsets of the SD data are compatible with those in the FD, in the same energy range. However, unlike the case of the FD where the ER in the light subset suggests that the growth in mass ceases at $\sim \lg(E/\mathrm{eV}) = 19.0$, the light subset of the SD data shows a continuous increase in mass.

In this study, some useful tools have been developed for the analysis of X_{max} data. They can be applied for example, based on the trends in measured $\langle X_{\mathrm{max}} \rangle$, in discriminating between scenarios of mass composition changes and those of cross-section changes. An application of the tools in an initial look at the Auger SD X_{max} data has shown promising results. With the implementation of the *AugerPrime* upgrade [130], more data on the depth of shower maximum will be collected. This should make the picture clearer concerning the trend in mass at the highest energies. This should in turn help to constrain the proposed models of UHECR origin currently available in literature.

A

First two moments of $X_{\max}$

$\langle \lg(E/\mathrm{eV}) \rangle$	N	$\langle X_{\max} \rangle$	Stat. error	Syst. error	$\sigma(X_{\max})$
17.850	3768	709.9	1.2	+7.6–10.2	$59.6 \pm 1.7^{+1.9}_{-1.7}$
17.949	3383	719.9	1.4	+7.5–10.2	$62.4 \pm 2.1^{+2.1}_{-1.8}$
18.048	2818	725.2	1.5	+7.4–10.2	$59.5 \pm 2.0^{+2.2}_{-1.9}$
18.148	2425	736.9	1.8	+7.3–10.1	$64.3 \pm 2.6^{+2.4}_{-2.1}$
18.247	1952	744.5	2.0	+7.3–9.9	$66.4 \pm 2.6^{+2.6}_{-2.2}$
18.348	1439	748.0	2.0	+7.3–9.7	$60.2 \pm 2.8^{+2.3}_{-2.0}$
18.448	1139	752.2	2.1	+7.3–9.4	$53.3 \pm 2.9^{+2.1}_{-1.8}$
18.549	814	754.5	2.2	+7.3–9.1	$53.5 \pm 3.0^{+1.9}_{-1.7}$
18.647	575	756.1	2.7	+7.4–8.8	$54.5 \pm 3.5^{+1.7}_{-1.6}$
18.747	413	757.4	2.8	+7.5–8.5	$45.8 \pm 3.4^{+1.5}_{-1.5}$
18.849	297	763.6	2.9	+7.7–8.1	$42.8 \pm 3.6^{+1.4}_{-1.4}$
18.947	230	764.6	3.2	+7.8–7.8	$43.4 \pm 4.1^{+1.3}_{-1.4}$
19.048	165	766.4	3.3	+8.0–7.6	$39.0 \pm 3.8^{+1.3}_{-1.4}$
19.146	114	767.0	3.6	+8.2–7.4	$36.7 \pm 3.6^{+1.3}_{-1.4}$
19.245	87	779.5	5.1	+8.5–7.2	$46.4 \pm 6.2^{+1.2}_{-1.3}$
19.340	63	773.1	5.0	+8.7–7.1	$40.1 \pm 4.8^{+1.3}_{-1.4}$
19.446	40	787.9	9.6	+8.9–7.0	$53.2 \pm 12.7^{+1.3}_{-1.4}$
19.620	37	779.8	5.0	+9.4–6.9	$26.5 \pm 4.8^{+1.5}_{-1.6}$

Table A.1: *Auger published values of the first two moments of the $X_{\max}$ distributions [49]. The units of $\langle X_{\max} \rangle$ and $\sigma(X_{\max})$ are $\mathrm{g/cm^2}$. The statistical and systematic uncertainties are given after the values of $\langle X_{\max} \rangle$ and $\sigma(X_{\max})$. The numbers in the fifth column represent the positive and negative systematic errors respectively, and the number of selected events are in the second column.*

B

Binned $\langle X_{\mathrm{max}} \rangle$ data calculated in this work

Bin #	$\langle \lg(E/\mathrm{eV}) \rangle$	N	$\langle X_{\mathrm{max}} \rangle$	Stat. error
1	17.850	3768	709.9	1.2
2	17.949	3383	719.3	1.3
3	18.048	2818	726.5	1.7
4	18.148	2425	736.9	1.7
5	18.247	1952	746.0	2.4
6	18.348	1439	747.8	1.9
7	18.447	1139	753.4	1.9
8	18.548	814	755.2	2.3
9	18.646	575	756.7	2.8
10	18.747	413	756.8	2.3
11	18.849	297	765.1	3.0
12	18.947	230	767.5	4.6
13	19.048	165	766.4	3.3
14	19.144	114	766.9	3.7
15	19.247	87	777.4	4.8
16	19.340	63	773.0	5.3
17	19.446	40	786.5	8.3
18	19.620	37	779.6	4.9

Table B.1: *The mean energy and $\langle X_{\mathrm{max}} \rangle$ in binned data as calculated in this work. The number of events per bin and statistical error in $\langle X_{\mathrm{max}} \rangle$ are also included. The unit of $\langle X_{\mathrm{max}} \rangle$ is* $\mathrm{g/cm^2}$.

List of Figures

List of Tables

List of Abbreviations

AERA	Auger Engineering Radio Array
AMIGA	Auger Muons and Infilled for the Ground Array
APF	Aerosol Phase Function
CCD	Charge Coupled Device
CDAS	Central Data Acquisition System
CIB	Cosmic Infrared Background
CLF	Central Laser Facility
CMB	Cosmic Microwave Background
CODALEMA	COsmic ray Detection Array with Logarithmic ElectroMagnetic Antennas
EMG	Exponentially Modified Gaussian
ER	Elongation Rate
FADC	Flash Analog-to-Digital Converter
FD	Fluorescence Detector
FRAM	ph(F)otometric Robotic Atmospheric Monitor
GDAS	Global Data Assimilation System
GOES	Geostationary Operational Environmental Satellites
GZK	Greisen, Zatsepin and Kuz'min effect
HAM	Horizontal Attenuation Monitor

HEAT	High-Elevation Auger Telescopes
KASCADE	KArlsruhe Shower Core and Array DEtector
LHC	Large Hadron Collider
LIDAR	Light Detection And Ranging
LOPES	Low-frequency array Prototype Station
MC	Monte Carlo
Ndf	Number of degrees of freedom
RMS	Root Mean Square
SD	Surface Detector
SNR	Supernova Remnant
TA	Telescope Array
ToT	Time-over-Threshold
UHECR	Ultra-High Energy Cosmic Rays
VCV	Véron-Cetty and Véron catalogue
VEM	Vertical Equivalent Muon
XLF	eXtreme Laser Facility

Bibliography

[1] J. Matthews. Measuring the highest energy cosmic rays. Pierre Auger Collaboration internal note (GAP-NOTE 97-007).

[2] K.-H. Kampert and M. Unger. Measurements of the cosmic ray composition with air shower experiments. *Astropart. Phys.*, 35:660–678, May 2012.

[3] H. Bhabha and W. Heitler. The Passage of Fast Electrons and the Theory of Cosmic Showers. *Proc. Roy. Soc. A*, 159:432–458, 1937.

[4] L. Bonolis. Walther Bothe and Bruno Rossi: The birth and development of coincidence methods in cosmic-ray physics. *arXiv*, 1106.1365v2, 2011.

[5] J. Linsley. Evidence for a Primary Cosmic-Ray Particle with Energy 10^{20} eV. *Phys. Rev. Lett.*, 10:146–148, February 1963.

[6] M. Nagano and A. A. Watson. Observations and implications of the ultrahigh-energy cosmic rays. *Rev. Mod. Phys.*, 72(3):689–732, July 2000.

[7] D. J. Bird et al. Changes in the Spectrum and Composition of Cosmic Rays at Extremely High Energies. *Phys. Rev. Lett.*, 71:3401, 1993.

[8] T. K. Gaisser, T. Stanev, and S. Tilav. Cosmic Ray Energy Spectrum from Measurements of Air Showers. *arXiv*, 1303.3565v1, 2013.

[9] R. Engel, D. Heck, and T. Pierog. Extensive Air Showers and Hadronic Interactions at High Energy. *Annu. Rev. Nucl. Part. Sci.*, 61:467–489, 2011.

[10] J. Linsley. The cosmic ray spectrum above 10^{19} eV at Volcano Ranch and Haverah Park. In 8^{th} *ICRC, Jaipur*, volume 77, 1963.

[11] K. Greisen. End to the Cosmic Ray Spectrum? *Phys. Rev. Lett.*, 16(17), 1966.

[12] G.T. Zatsepin and V.A. Kuzmin. Upper Limit of the Spectrum of Cosmic Rays. *JETP Lett*, 4, 1966.

[13] M. Ave. Ground signal parameterization for water Cherenkov and scintillator detectors and performance of ground arrays that combine both techniques. Pierre Auger Collaboration internal note (GAP-2012-141).

[14] I. Allekotte, M. Ave, X. Bertou, R. Engel, and M. Roth. Preliminary results for the $X_{\max}$ reconstruction using SD vertical events. Pierre Auger Collaboration internal note (GAP-2013-075).

[15] M. Ave, X. Bertou, R. Engel, D. Maurel, M. Roth, and A. Schulz. (Energy, N_μ, X_{max}) from the Bariloche reconstruction. Pierre Auger Collaboration internal note (GAP-2013-109).

[16] G. van Aar and C. Timmermans. Data driven event-by-event primary mass information from the SD. Pierre Auger Collaboration internal note (GAP-2014-031).

[17] The High Resolution Fly's Eye Collaboration. Analysis of large-scale anisotropy of ultra-high energy cosmic rays in HiRes data. *arXiv*, 1002.1444v1, 2010.

[18] The Pierre Auger Collaboration. Correlation of the highest-energy cosmic rays with the positions of nearby active galactic nuclei. *Astropart. Phys.*, 29:188–204, 2008.

[19] The Pierre Auger Collaboration. Constraints on the origin of cosmic rays above 10^{18} eV from large scale anisotropy searches in data of the Pierre Auger Observatory. *ApJ.*, 762:L13, 2013.

[20] The Pierre Auger Collaboration. Searches for anisotropies in the arrival directions of the highest energy cosmic rays detected at the Pierre Auger Observatory. *ApJ*, 804(15), 2015.

[21] The Pierre Auger Collaboration. Update on the correlation of the highest energy cosmic rays with nearby extragalactic matter. *Astropart. Phys.*, 34:314–326, 2010.

[22] K.-H. Kampert for the Pierre Auger Collaboration. Highlights from the Pierre Auger Observatory. In *32nd International Cosmic Ray Conference, Beijing, China*, 2011.

[23] P. Abreu et al. for the Pierre Auger Collaboration. Search for First Harmonic Modulation in the Right Ascension Distribution of Cosmic Rays Detected at the Pierre Auger Observatory. *Astropart. Phys.*, 34:627–639, 2011.

[24] The Pierre Auger Collaboration. Large scale distribution of arrival directions of cosmic rays detected above 10^{18} eV at the Pierre Auger Observatory. *Astrophys. J. Suppl.*, 203:34, 2012.

[25] E. Fermi. On the Origin of the Cosmic Radiation. *Phys. Rev.*, 75(8):1169–1174, 1949.

[26] P.O. Lagage and C. J. Cesarsky. The maximum energy of cosmic rays accelerated by supernove shocks. *Astron. Astroph.*, 125:249–257, 1983.

[27] A.M. Hillas. Can diffusive shock acceleration in supernova remnants account for high-energy galactic cosmic rays? *J. Phys. G*, 31:R95–R131, 2005.

[28] A. M. Hillas. The Origin of Ultra-High-Energy Cosmic Rays. *Ann. Rev. Astron. Astrophys.*, 22:425–444, 1984.

[29] D. F. Torres and L. A. Anchordoqui. Astrophysical origins of ultrahigh energy cosmic rays. *Rep. Prog. Phys.*, 67:1663–1730, 2004.

[30] H. J. de Vega and N. Sanchez. Extreme Energy Cosmic Rays: Bottom-up vs. Top-down scenarii. *arXiv*, 0301039v2, 2003.

[31] V. S. Berezinsky, M. Kachelriess, and A. Vilenkin. Ultrahigh Energy Cosmic Rays without Greisen-Zatsepin-Kuzmin Cutoff. *Phys. Rev. Lett.*, 79:4302, 1997.

[32] G. B. Gelmini, O. E. Kalashev, and D. V. Semikoz. GZK Photons as Ultra High Energy Cosmic Rays. *JETP*, 106:1061, 2008.

[33] V. Berezinsky. Ultra High Energy Cosmic Rays. *Nucl. Phys. B Proc. Suppl.*, 70:419–430, 1999.

[34] P. Bhattacharjee and G. Sigl. Origin and propagation of extremely high-energy cosmic rays. *Phys. Rep.*, 327:110, 2000.

[35] T. Weiler. Resonant Absorption of Cosmic-Ray Neutrinos by the Relic-Neutrino Background. *Phys. Rev. Lett.*, 49(3), 1982.

[36] T. Weiler. Cosmic-ray neutrino annihilation on relic neutrinos revisited: a mechanism for generating air showers above the Greisen-Zatsepin-Kuzmin cutoff. *Astropart. Phys.*, 11:303–316, 1999.

[37] A. Benson, A. Smialkowski, and A.W. Wolfendale. Ultra high energy cosmic rays and the Galactic halo . *Astropart. Phys.*, 10:313–320, 1999.

[38] C. Bleve For the Pierre Auger Collaboration. Update of the neutrino and photon limits from the Pierre Auger Observatory. In *The Pierre Auger Observatory: Contributions to the 34^{th} International Cosmic Ray Conference (ICRC 2015)*, 2015.

[39] R. Aloisio. Ultra High Energy Cosmic Rays: A Short Review. *Nucl. Phy. B Proceedings Supplement*, 00:1–7, 2013.

[40] D. Allard, N.G. Busca, G. Decerprit, A.V. Olinto, and E. Parizot. Implications of the cosmic ray spectrum for the mass composition at the highest energies. *JCAP*, 10:033, 2008.

[41] R. Aloisio, V. Berezinsky, P. Blasi, and S. Ostapchenko. Signatures of the transition from galactic to extragalactic cosmic rays. *Phys. Rev. D*, 77:025007, 2008.

[42] E. G. Berezhko and H. J. Völk. TeV gamma rays expected from supernova remnants in different uniform interstellar media. *Astropart. Phys.*, 14:201–206, 2000.

[43] D. de Marco and T. Stanev. On the shape of the UHE cosmic ray spectrum. *arXiv*, 0506318v2, 2005.

[44] M.A. Lawrence, R.J.O. Reid, and A. A. Watson. The cosmic ray energy spectrum above 4×10^{17} eV as measured by the Haverah Park array . *J. Phys. G*, 17:733–757, 1991.

[45] M. Nagano et al. Energy spectrum of primary cosmic rays above $10^{17.0}$ eV determined from extensive air shower experiments at Akeno. *J. Phys. G*, 18:423–442, 1992.

[46] A.A. Ivanov, S.P. Knurenko, and I.Ye. Sleptsov. Measuring extensive air showers with Cherenkov light detectors of the Yakutsk array: The energy spectrum of cosmic rays. *New J. Phys.*, 11:065008, 2009.

[47] A. Aab et al. Measurement of the cosmic ray spectrum above 4×10^{18} eV using inclined events detected with the Pierre Auger Observatory. *JCAP*, 08:049, 2015.

[48] V. Berezinsky. UHECR: Signatures and Models. *arXiv*, 1307.4043v1, 2013.

[49] A. Aab et al. Depth of Maximum of Air-Shower Profiles at the Pierre Auger Observatory: Measurements at Energies above $10^{17.8}$ eV. *Phys.Rev.D*, 90:122005, 2014.

[50] V. S. Berezinsky and S.I. Grigorieva. A bump in the ultra-high energy cosmic ray spectrum. *Astron. Astroph.*, 199:1, 1988.

[51] V. Berezinsky, A. Z. Gazizov, and S. I. Grigorieva. On astrophysical solution to ultrahigh energy cosmic rays. *Phys. Rev. D*, 74:043005, 2006.

[52] V. Berezinsky, A. Z. Gazizov, and S. I. Grigorieva. Dip in UHECR spectrum as signature of proton interaction with CMB. *Phys. Lett. B*, 612:147–153, 2005.

[53] A. Schulz For the Pierre Auger Collaboration. The measurement of the energy spectrum of cosmic rays above 3×10^{17} eV with the Pierre Auger Observatory. In *33rd ICRC, Rio de Janeiro, Brazil*, 2013.

[54] K.H. Kampert and A. A. Watson. Extensive Air Showers and Ultra High-Energy Cosmic Rays: A Historical Review. *arXiv*, 1207.4827v2, 2012.

[55] http://upload.wikimedia.org/wikipedia/commons/thumb/2/2c/AirShower.svg/.

[56] W. Heitler. *The Quantum Theory of Radiation*. Oxford University Press, 1954. third ed.

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

[58] R. Engel et al. Anti-protons as probes of solar modulation . In *26^{th} ICRC conference, Salt Lake City*, volume 1, page 415, 1999.

[59] J. Alvarez-Muniz et al. Hybrid simulations of extensive air showers. *Phys. Rev. D*, 66:033011, 2002.

[60] A. S. Chou et al. An Universal Description of the Particle Flux Distributions in Extended Air Showers. In *29th International Cosmic Ray Conference Pune*, pages 101–104, 2005. 29th International Cosmic Ray Conference Pune.

[61] F. Schmidt, M. Ave, L. Cazon, and A. Chou. A model-independent method of determining energy scale and muon number in cosmic ray surface detectors. *Astropart. Phys.*, 29:355–365, 2008.

[62] D. Góra et al. Universal lateral distribution of energy deposit in air showers and its application to shower reconstruction. *Astropart. Phys.*, 24:484–494, 2006.

[63] F. Nerling, J. Blümer, R. Engel, and M. Risse. Universality of electron distributions in high-energy air showers–Description of Cherenkov light production. *Astropart. Phys.*, 24:421–437, 2006.

[64] M. Ave, N. Busca, L. Cazon, F. Schmidt, and T. Yamamoto. Determining the SD Energy Scale and Number of Muons Using Air Shower Universality. Pierre Auger Collaboration internal note GAP 2007-021., 2007.

[65] P. Lipari. Concepts of "age" and "universality" in cosmic ray showers. *Phys. Rev. D*, 79:063001, 2009.

[66] M. Ave and X. Bertou. Alternative shower reconstruction for SD vertical events. Pierre Auger Collaboration internal note (GAP-2013-067).

[67] M. De Domenico, M. Settimo, S. Riggi, and E. Bertin. Reinterpreting the development of extensive air showers initiated by nuclei and photons. *JCAP*, 7:50, July 2013.

[68] C. J. T. Peixoto, V. de Souza, and J. Bellido. Comparison of the moments of the Xmax distribution predicted by different cosmic ray shower simulation models. *arXiv*, 1301.5555v2, 2013.

[69] E. Bertin. Global Fluctuations and Gumbel Statistics. *Phys. Rev. Lett.*, 95:170601, 2005.

[70] E. Bertin and M. Clusel. Generalized extreme value statistics and sum of correlated variables. *J. Phys. A: Math. Gen.*, 39:7607, 2006.

[71] R. Abbasi et al. First Observation of the Greisen-Zatsepin-Kuzmin Suppression. *Phys. Rev. Lett.*, 100:101101, 2008.

[72] R. Aloisio et al. A dip in the UHECR spectrum and the transition from galactic to extragalactic cosmic rays. *Astropart.Phys*, 27:76–91, 2007.

[73] M. Risse and D. Heck. Energy release in air showers. *Astropart.Phys.*, 20:661, 2004.

[74] J. Engel, T. K. Gaisser, P. Lipari, and T. Stanev. Nucleus-nucleus collisions and interpretation of cosmic-ray cascades. *Phys. Rev. D*, 46(11):5013–5025, 1992.

[75] J. Linsey and A. A. Watson. Validity of scaling to 10^{20} eV and High-Energy Cosmic-Ray Composition. *Phys. Rev. Lett.*, 46(7):459–463, 1981.

[76] M. Unger. Study of the Cosmic Ray Composition above 0.4 EeV using the Longitudinal Profiles of Showers observed at the Pierre Auger Observatory . *arXiv*, 0706.1495, 2007.

[77] M. Risse. Properties of Extensive Air Showers. *arXiv:astro-ph/0402300v1*, 2004.

[78] D. Hooper and A. M. Tylor. On the heavy chemical composition of the ultra-high energy cosmic rays. *Astropart. Phys.*, 33:151–159, 2010.

[79] P. Abreu et al. Interpretation of the Depths of Maximum of Extensive Air Showers Measured by the Pierre Auger Observatory . *JCAP*, 26(02), 2013.

[80] R. U. Abbasi et al. Study of Ultra-High Energy Cosmic Ray Composition Using Telescope Array's Middle Drum Detector and Surface Array in Hybrid Mode. *arXiv*, 1408.1726v3, November 2014.

[81] A. Aab et al. Depth of Maximum of Air-Shower Profiles at the Pierre Auger Observatory: Composition Implications. *Phys. Rev. D*, 90:122006, 2014.

[82] R. Abbasi et al. Report of the Working Group on the Composition of Ultra High Energy Cosmic Rays. *arXiv*, 1503.07540v1, 2015.

[83] The Pierre Auger Collaboration. The Pierre Auger Cosmic Ray Observatory. *Nucl. Instrum. Meth. A*, 798:172–213, 2015.

[84] I. Allekotte et al. The Surface Detector System of the Pierre Auger Observatory. *Nucl. Instrum. Meth. A*, 586:409–420, 2008.

[85] J. Abraham et al. The Fluorescence Detector of the Pierre Auger Observatory. *Nucl. Instrum. Meth. A*, 620:227–251, 2010.

[86] R.M. Baltrusaitis, R. Cady, G.L. Cassiday, R. Cooper, J.W. Elbert, et al. The Utah Fly's Eye Detector. *Nucl. Instrum. Meth. A*, 240:410–428, 1985.

[87] S. Agostinelli et al. GEANT4: A simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250–303, 2003.

[88] M.G. Pia. The Geant4 Toolkit: Simulation capabilities and application results. *Nucl. Phys. Proc. Suppl.*, 125:60–68, 2003.

[89] T. K. Gaisser and A. M. Hillas. Reliability of the method of constant intensity cuts for reconstructing the average development of vertical showers. *International Cosmic Ray Conference*, 8:353–357, 1977.

[90] M. J. Tueros et al. For the Pierre Auger Collaboration. Estimate of the non-calorimetric energy of showers observed with the fluorescence and surface detectors of the Pierre Auger Observatory. In 33^{rd} *ICRC, Rio de Janeiro* , 2013.

[91] H.-J Mathes For the Pierre Auger Collaboration. The HEAT Telescopes of the Pierre Auger Observatory: Status and First Data. In *32^{n}d International Cosmic Ray Conference, Beijing, China*, 2009.

[92] A. Bueno. The physics of ultra-high energy cosmic rays and the Pierre Auger observatory: present and future directions. An internal report for the Pierre Auger collaboration.

[93] N. Accialini et al. AMIGA Status Report. Pierre Auger Collaboration internal report (GAP2011_120).

[94] H.R. Allan, 1971. Prog. in Elem. part. and Cos. Ray Phys., ed. J.G. Wilson and S.A.Wouthuysen.

[95] H. Falcke et al. Detection and imaging of atmospheric radio flashes from cosmic ray air showers. *Nature*, 435:313–316, 2005.

[96] D. Ardouin et al. Radioelectric field features of extensive air showers observed with CODALEMA. *Astropart. Phys.*, 26:341–350, 2006.

[97] Q. Dorosti, F. Sanchez, and F.G Schröder. First quadruple hybrid events (SD+FD+MD+RD). Pierre Auger Collaboration internal note (GAP-2014-016), 2014.

[98] P. Abreu et al. Techniques for measuring aerosol attenuation using the Central Laser Facility at the Pierre Auger Observatory. *Journal of Instrumentation*, 8:P04009, 2013.

[99] E. J. Ahn, J. Bellido, S. BenZvi, R. Engel, F. Schüssler, R. Ulrich, and M. Unger. Measurement of the Depth of Shower Maximum of Cosmic Rays above 10^{18} eV. Pierre Auger Collaboration internal note (GAP-2009-078).

[100] M. Unger and J. Bellido. Supplementary Material for the 'long Xmax Paper'. Pierre Auger Collaboration internal publication., 2014.

[101] Pierre Auger Collaboration. *Aperture calculation of the Pierre Auger Observatory surface detector*, 2005.

[102] The Pierre Auger Collaboration. The Lateral Trigger Probability function for the Ultra-High Energy Cosmic Ray showers detected by the Pierre Auger Observatory. *Astropart. Phys.*, 35:266–276, 2011.

[103] J. Bellido. Systematics in the Reconstructed $X_{\max}$ as a function of the observed Profile Length. Talk at the Auger Analysis Meeting, Chicago, September 2006.

[104] F. Schüssler. Measurement of the Energy Spectrum of Ultra-High Energy Cosmic Rays using Hybrid Data of the Pierre Auger Observatory. Pierre Auger Collaboration internal note (GAP-2008-155).

[105] M. Unger, J. Bäuml, R. Engel, and R. Ulrich. Simulation of the Point Spread Function of the FD Telescopes and its Effect on the Reconstructed Energy and $X_{\max}$. Pierre Auger Collaboration internal note (GAP-2013-069), 2013.

[106] M. Unger et al. Lateral shower light distributions in the Cherenkov-Fluorescence-Matrix Profile Reconstruction. Pierre Auger Collaboration internal note (GAP-2008-052), 2008.

[107] J. Abraham et al. Trigger and aperture of the surface detector array of the Pierre Auger Observatory. *Nucl. Instrum. Meth. A*, 613:29–39, 2010.

[108] E. Parizot, I. Lhenry, D. Allard, P. Ghia, and G. Navarra. First steps towards the definition of a "quality trigger" (T5) for the SD acceptance calculations. Pierre Auger Collaboration internal note (GAP-2004-023).

[109] C. Bonifazi. The angular resolution of the Pierre Auger Observatory. *Nucl. Phys. B Proc. Suppl.*, 190:20–25, 2009.

[110] D. F. Gatz and L. Smith. The standard error of a weighted mean concentration—I. Bootstrapping vs other methods . *Atmospheric Environment*, 29:1185–1193, 1995.

[111] B. Efron. Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7(1):1–26, 1979.

[112] 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*, 50:5710, 1994.

[113] T. Pierog, Iu. Karpenko, J. M. Katzy, E. Yatsenko, and K. Werner. EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider. *Phys. Rev. C*, 92:034906, 2015.

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

[115] M. Risse, A. Yushkov, F. Guarino, and R. Colalillo. Correlation between depth of shower maximum and signal in ground stations: pure or mixed composition? Pierre Auger Collaboration internal note (GAP-NOTE 2014-006).

[116] V. Verzi for the Pierre Auger Collaboration. The Energy Scale of the Pierre Auger Observatory. In *33rd ICRC, Rio de Janeiro* , 2013.

[117] F. Klinkhamer and M. Risse. Ultra-high-energy cosmic-ray bounds on nonbirefringent modified-Maxwell theory. *Phys. Rev. D*, 77(016002), 2008.

[118] F. Klinkhamer and M. Risse. Addendum: Ultrahigh-energy cosmic-ray bounds on nonbirefringent modified-Maxwell theory. *Phys. Rev. D*, 77(117901), 2008.

[119] M. Risse, D. Boncioli, A. di Matteo, et al. Auger observations and modified Maxwell theory. Pierre Auger Collaboration internal note (GAP-NOTE 2013-064).

[120] F. Klinkhamer. Potential sensitivities to Lorentz violation from nonbirefringent modified-Maxwell theory of Auger, HESS and CTA . *Phys. Rev. D*, 82:105024, 2010.

[121] E. Grushka. Characterization of Exponentially Modified Gaussian Peaks in Chromatography. *Anal. Chem.*, 44(11):1733–1738, 1972.

[122] T. Bergmann, R. Engel, D. Heck, N. Kalmykov, S. Ostapchenko, et al. One-dimensional Hybrid Approach to Extensive Air Shower Simulation. *Astropart. Phys.*, 26:420, 2007.

[123] M. Werner, P. Heimann, M. Risse, and A. Yushkov. Sensitivity of the exponential slope of the $X_{\max}$ distribution tail to the primary mass composition. Pierre Auger Collaboration internal note (GAP-2014-059).

[124] L. D. Brown, T. Cai, and A. DasGupta. Interval Estimation for a Binomial Proportion . *Statistical Science*, 16:101–103, 2001.

[125] X. Bertou. Private communication.

[126] K. Werner, I. Karpenko, and T. Pierog. The "Ridge" in Proton-Proton Scattering at 7 TeV. *Phys. Rev. Lett.*, 106:122004, 2011.

[127] R. Ulrich and M. Unger. Hadronic multiparticle production at ultrahigh energies and extensive air showers. *Phys. Rev. D*, 83:054026, 2011.

[128] M. Unger. Fitting the $X_{\max}$ Distributions with Modified p-air Interactions. Pierre Auger Collaboration meeting, February, 2015.

[129] P. Facal for the Pierre Auger Collaboration. The distribution of shower maxima of UHECR air showers. In *32nd International Cosmic Ray Conference, Beijing, China*, 2011.

[130] The Pierre Auger Collaboration. The Pierre Auger Observatory Upgrade "AugerPrime" Preliminary Design Report. *arXiv*, 1604.03637v1 [astro-ph.IM], 2016.