Measurement of the proton-air cross section using hybrid data of the Pierre Auger Observatory

Zur Erlangung des akademischen Grades eines
DOKTORS DER NATURWISSENSCHAFTEN
von der Fakultät für Physik der Universität (TH)
Karlsruhe

genehmigte
DISSERTATION

von Dipl.-Phys. Ralf Matthias Ulrich
aus Basel (Schweiz)

Tag der mündlichen Prüfung: 21.12.2007
Referent: Prof. J. Blümer
Korreferent: Prof. G. Quast

LaTeX

Abstract

The subject of this thesis is the measurement of the proton-air cross section at ultra high energy with hybrid data of the Pierre Auger Observatory. Based on a critical review of the shortcomings of previous air shower measurements, a new analysis method is developed. This analysis method takes into account the most important and relevant experimental and air shower physics effects. The impact of a changed cross section extrapolation on the resulting air shower development is considered in addition to its more obvious effect on the distribution of shower starting points. Furthermore, detector acceptance effects are explicitly included in the reconstruction ansatz, which allows us to use almost the complete data set in the analysis. Systematic uncertainties on the resulting cross sections are thoroughly studied and quantified. Assuming a proton dominated composition, the analysis is applied to hybrid data of the Pierre Auger Observatory. The obtained cross section is, within the statistical and systematic uncertainties, in agreement with predictions from hadronic interaction models up to $10^{18.4}$ eV. At higher energies the resulting cross section increases rapidly. Finally the proton-air cross section is converted to a proton-proton cross section using Glauber theory and limits on the elastic scattering slope are derived.

Bestimmung des Wechselwirkungsquerschnittes von Protonen mit Luft mittels Hybrid-Daten des Pierre Auger Observatoriums

In dieser Arbeit werden Hybrid-Daten des Pierre Auger Observatoriums verwendet, um den Wechselwirkungsquerschnitt von Protonen mit Kernen der Luft bei ultra-hoher Energie zu bestimmen. Basierend auf einer Analyse der Schwachpunkte früherer Luftschauer-Messmethoden wird ein neuer Rekonstruktionsansatz entwickelt. Dieser berücksichtigt explizit die wichtigsten experimentellen sowie physikalischen Effekte. Der Einfluss einer veränderten Extrapolation des Wechselwirkungsquerschnittes auf die Luftschauer-Entwicklung wird ebenso berücksichtigt wie der Einfluss auf die Verteilung der Luftschauer-Startpunkte. Darüber hinaus wird die Akzeptanz des Detektors direkt in der Rekonstruktionsmethode berücksichtigt, wodurch fast der komplette Datensatz für die Analyse verwendet werden kann. Systematische Unsicherheiten des resultierenden Wechselwirkungsquerschnittes werden im Detail untersucht und quantifiziert. Die Hybrid-Daten des Pierre Auger Observatoriums werden unter der Annahme einer proton-dominierten Zusammensetzung der kosmischen Strahlung analysiert. Der resultierende Wechselwirkungsquerschnitt ist, innerhalb der statistischen und systematischen Unsicherheiten, bis zu einer Energie von $10^{18.4}$ eV mit den Vorhersagen von hadronischen Wechselwirkungsmodellen kompatibel. Bei höherer Energie nehmen die resultierenden Wechselwirkungsquerschnitte jedoch sehr schnell zu. Abschliessend wird der Proton-Luft-Wechselwirkungsquerschnitt mittels der Glauber-Theorie in den Proton-Proton-Wechselwirkungsquerschnitt konvertiert und es werden Einschränkungen des Steigungsparameters der elastischen Streuung abgeleitet.

Contents

1 Introduction

In the early years, elementary particle physics began with the observation of cosmic rays and cosmic ray induced air showers. Many discoveries are directly linked to the analysis of cosmic ray interactions, like the finding of positrons [1], muons [2] and pions [3, 4, 5]. Following these early discoveries based on cosmic rays, accelerators were built for precision studies. This was leading to huge achievements in understanding fundamental particle physics and to the formulation of the extremely successful *standard model* of elementary particle physics. All physics so far discovered at accelerators can be described very well within this framework (for example [6, 7]). Mainly the difficulty to unify the coupling constants of all elementary forces at high energy and the general problem to describe gravity are theory-motivated arguments that require physics beyond the standard model. Again, it is only astro- and astroparticle physics, which are currently providing experimental evidence for the existence of such physics. Namely, the observation of the non-vanishing mass of neutrinos [8, 9] and the required existence of dark matter [10].

Furthermore man-made accelerators are limited in their maximum energy, on the other hand astroparticle physics can provide observations of the most extreme environments in the universe and even probe ultra-high energy physics directly. For example, the extragalactic propagation of cosmic rays at ultra-high energies is sensitive to violations of the Lorentz invariance of space-time [11, 12] as well as new particle physics. Moreover, by the observation of the highest energy cosmic rays it is also possible to study hadronic interactions directly at energies far larger, than any Earth-based accelerator is able to reach.

With the Pierre Auger Observatory there is the first time a detector available, which allows us to do high precision studies of cosmic rays at ultra-high energies [13]. This is mainly due to the hybrid detector design, combining a ground based air shower array with the data of telescope detectors, which observe the shower development within the atmosphere. The enormous size of the Auger detector of more than $3000\,\mathrm{km}^2$ allows us to gather high event statistics already after the first few years of operation despite of the extremely small cosmic ray flux at these energies. This is ideal for studying cosmic ray properties like their origin, but also to study hadronic interactions at ultra-high energies.

The aim of this thesis is the first measurement of the proton-air cross section based on the hybrid data of the Pierre Auger Observatory collected so far.

As a starting point methodical studies of previous attempts to measure the proton-air cross section using air shower data are studied, unveiling the common scheme, on which they are all founded on. The principal shortcoming, which is inherent to all of them, is pointed out.

Based on these fundamental studies a novel method to derive the proton-air cross section from air shower data is developed. Compared to previous methods to measure the proton-air cross section, two main improvements are to be pointed out. Firstly, not only the obvious exponential dependence of the fluctuation of the shower starting point from the cross section are considered, but also the equally important impact of a changing cross section on the subsequent air shower development. Secondly, the acceptance of the detection and reconstruction process is studied carefully and is incorporated directly within the presented cross section analysis.

This could only be achieved by developing a detailed detector simulation for hybrid events that allows us to study the impact of the detection and reconstruction process on the final observables. Special attention is paid to enhance the general understanding of the Auger detector and to improve the reconstruction algorithms, with an emphasis on the propagation of uncertainties into the finally used quantities.

The fundamental necessity to utilize Monte Carlo simulations to relate any air shower based observation to properties of the primary cosmic ray particle introduces an inherent model-dependence. The new method to measure the proton-air cross section developed in this thesis fully accounts for the existing model-dependence and reduces it as much as possible. Estimates of the systematic uncertainties induced by the model-dependence are presented, which apply also to other air shower based cross-section measurements.

Other general sources of systematic uncertainties are thoroughly studied. First and foremost this is the unknown mass composition of cosmic rays at ultra-high energies. This includes the impact of a possible contribution of very high energy gamma-rays.

After application of the new method to the data of the Pierre Auger Observatory and the extraction of the proton-air cross section at ultra-high energy, the results are converted into predictions for the corresponding proton-proton cross section by utilizing Glauber theory. The implications of the measurement, for example, on the expected interaction characteristics at the LHC, but also on the interpretation of cosmic ray data in terms of mass composition are discussed.

2 Cosmic rays, extensive air showers and high energy hadronic interactions

2.1 Overview of cosmic rays

After almost 100 years of research cosmic ray physics is remaining to be an exciting field and the current activity is probably larger than ever before [14]. Recent high quality observations are uncovering more and more of the mysteries associated to the existence of cosmic rays. It seems very likely that cosmic rays up to the highest energies will fit seamlessly into the framework of astro and particle physics. This implies important improvements in understanding our cosmic environment as well as ultra-high energy interactions during the next years. Since these cosmic rays are particles with macroscopic energies of up to several Joules (16 EeV$\sim$1 J), this will be a major step for particle physics.

The currently favored theories to explain the phenomena of cosmic rays are founded on the assumption of charged particle acceleration at collisionless magnetic shock fronts and the propagation as well as confinement of charged particles within the galactic and extra-

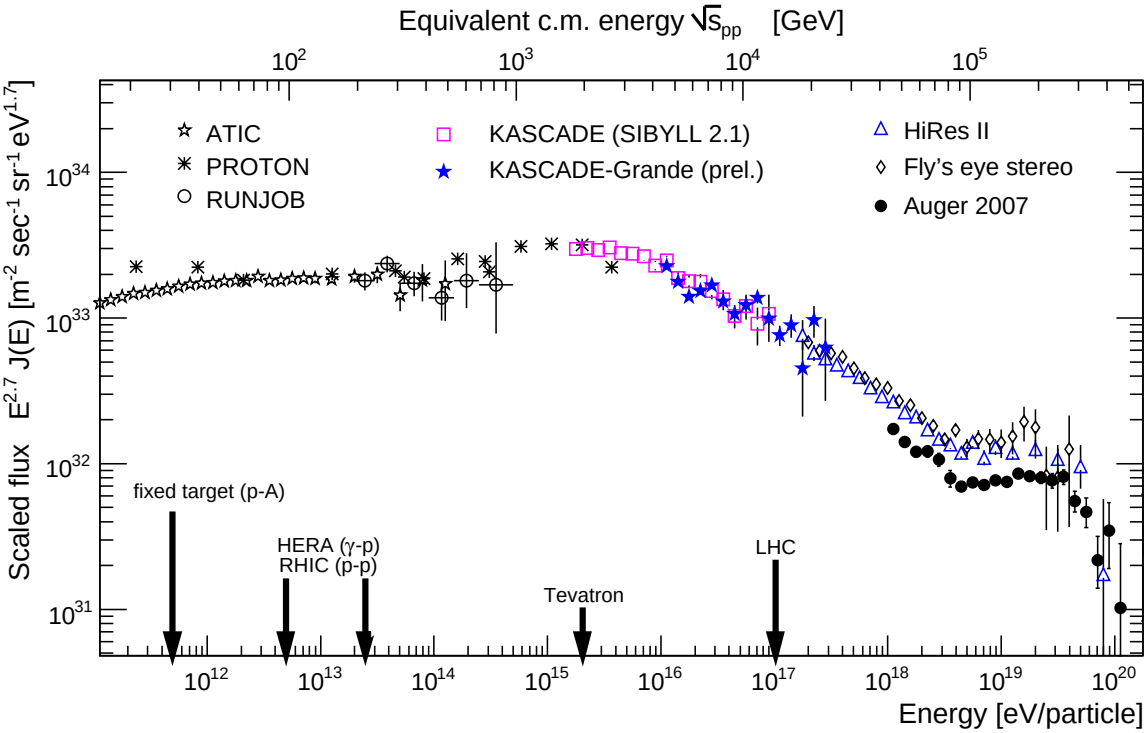


Figure 2.1: Cosmic ray energy spectrum as measured by many experiments over a wide range in energy [15, 16, 17, 18, 19, 20, 21]. The energy spectrum is multiplied by $E^{2.7}$ to remove the enormous slope over $\sim$15 orders in magnitude of the flux. In this representation the structures of the energy spectrum can be seen very clear: the *knee* at several PeV, the *ankle* around $\sim$3 EeV and the *cutoff* above $\sim$50 EeV. While the lower axis reflects the energy of the primary cosmic ray nuclei E_{lab}, the upper axis denotes the corresponding center-of-mass energy per nucleon. Some typical energies, which can be accessed by accelerators, are emphasized.

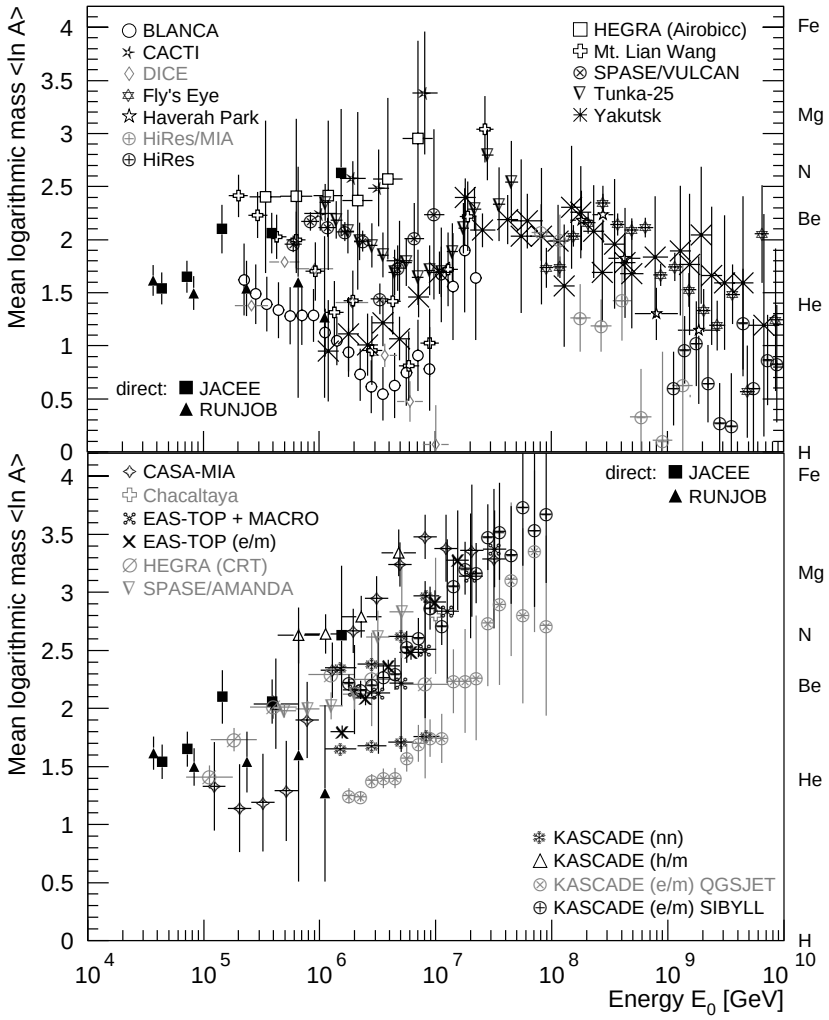


Figure 2.2: Mean logarithmic mass of the primary cosmic ray composition as derived from observations of the shower maximum (top panel) or of electrons, muons and hadrons at ground level (bottom panel). Direct measurements are included in both plots. For the references see [22].

galactic environment. The stochastic nature of the acceleration and propagation processes is generating the power-law-like shape of the energy spectrum as it is observed on Earth, starting from the solar modulation cutoff at ∼GeV up to the high energy cutoff around $\sim 50\,\mathrm{EeV}$. Figure 2.1 shows the cosmic ray particle flux, multiplied by $E^{2.7}$ in order to compensate for the enormous steepness of the spectrum. Several distinct features of the energy spectrum are instantly visible, commonly associated by the eye-catching analogy to the anatomy of a human leg. In addition, Figure 2.2 summarizes the available data on the mass composition of the high energy cosmic ray flux arriving at Earth. The uncertainties in the associated analyses are becoming larger with increasing energy and therefore there are large systematic errors associated to any absolute measurement of the mass composition at ultra-high energy. However, there are indications for two transitions within the data, where the mass composition changes rapidly, first around ∼PeV from light to heavier and then at ∼EeV back from heavier to lighter primary particles.

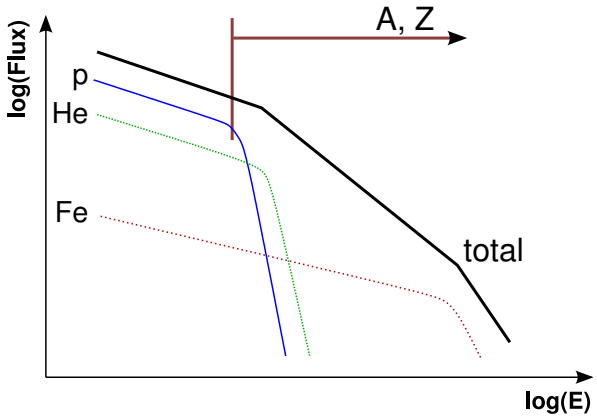


Figure 2.3: Possible origin of the knee structure in the cosmic ray energy spectrum. The successive cutoff of proton, helium and heavier elements is already observationally confirmed [24]. However, the precision of the measurement is not yet high enough to distinguish between the ∝A or ∝Z scenarios.

At lower energies the flux is well described by $\Phi \propto E^{-2.7}$ and is therefore appearing flat in Figure 2.1. At the energy of several PeV a sudden change of the power-law slope to $\Phi \propto E^{-3}$ appears. This feature is called the *knee* [23]. It coincides with a rapid increase of the mean primary mass. The knee structure is most probably caused by the galactic cosmic ray accelerators reaching their maximum energy. Since acceleration depends on the charge $q = eZ$ of particles, the knee energy gets $E_{\text{knee}} \propto Z$. Also the increasing leakage of higher energy cosmic rays out of the magnetic confinement of our galaxy may contribute. This effect is governed by the rigidity $R = pc/eZ$ of the trajectory of a particle with charge eZ in the galactic magnetic field, also leading to $E_{\text{knee}} \propto Z$. Indeed it was shown by the KASCADE experiment [24] that the first component which reaches their knee energy are protons, followed by helium. This is a very strong observation to support the idea of cosmic rays at these energies being of galactic origin. On the other hand, it still can not be ruled out that a sudden change of the hadronic interaction characteristics at E_{knee} may be responsible for the knee feature [25]. This would cause the knee energy to shift according to the energy per projectile nucleon and thus being proportional to the nucleon number A. Figure 2.3 summarizes the different possible scenarios for the knee.

At some energy above E_{knee} the spectral index changes to $\Phi \propto E^{-2.7}$. This is typically associated with the transmission to a cosmic ray component of extragalactic origin. The break in the energy spectrum is called *ankle* and was discovered by the Fly's eye Collaboration [26]. Also data on the depth of the shower maximum, X_{max}, indicate a changing composition of cosmic rays in the transition region from heavy to light particles [27, 28]. However, the mass composition and also the shape of the all-particle energy spectrum above the knee are still not very well known. There is a debate on the transition from the galactic to the extragalactic cosmic ray component going on. Figure 2.4 illustrates the ultra-high energy cosmic ray spectrum together with the two most popular models explaining the features.

It is possible that the ankle itself is the transition point [29, 30]. These models typically assume a mixed composition of the extragalactic component. The composition of the extragalactic flux, as it arrives at Earth, is strongly altered during propagation over extragalac-

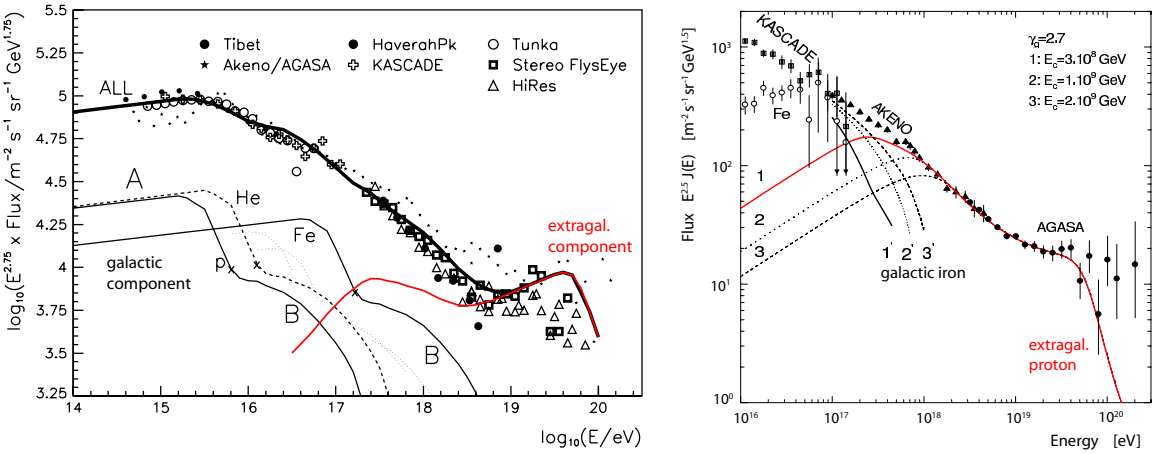


Figure 2.4: The upper end of the cosmic ray spectrum. Left panel: the extragalactic mixed composition scenario (taken from [34] and modified). Right panel: The *dip* model with the transition to a proton dominated extragalactic component at lower energies [35].

tic distances (∼Mpc), and thus must be very different from the original composition at the sources. Because of the required low flux of the extragalactic component, a new galactic component needs to be introduced (component "B"), which fills the gap between the vanishing galactic flux from component "A" and the observed all particle spectrum in the energy interval from the knee up to the ankle (cf. Figure 2.4, left). It can be argued that the component "B" may be associated with special classes of supernova explosions [31, 32].

Very differently the ankle can also be explained by a propagation effect of the extragalactic proton dominated component [33]. While traveling through the *cosmic microwave background* (CMB) over extragalactic distances, protons loose energy due to Bethe-Heitler e^+e^- pair production. This causes a suppression of the flux and a pile up at slightly lower energies. Figure 2.4 (right) shows the resulting characteristic *dip* in the proton flux, which is then identified with the ankle. In this model, the transition from galactic to extragalactic cosmic rays does occur at much lower energies. The dip model requires a strong proton dominance of the cosmic ray flux, reaching down at least to 1 EeV ($\geq$ 80% protons and $<$ 20% helium).

There are two plausible explanations for the spectrum cutoff at ultra-high energies. Firstly, the extragalactic cosmic ray accelerators run out of power. Secondly, the cutoff is due to a propagation effect of cosmic ray nucleons through the CMB photon field. The latter is much more popular within the astroparticle community and seems to be favored by recent observations. The *Greisen-Zatsepin-Kutsmin* (GZK) effect [36, 37] was predicted more than 40 years ago as a consequence of the discovery of the 3 K microwave photon field. Ultra-high energy nucleons ($E_{\mathrm{GZK}} \gtrsim 50$ EeV) traveling through extragalactic space interact with photons of the CMB photon field via the formation of resonances (i.e. $\mathrm{N} + \gamma_{3\mathrm{K}} \rightarrow \Delta \rightarrow \mathrm{N} + \pi$), which leads to the emission of pions. This corresponds to an energy loss of cosmic ray nucleons via pion emission, effectively leading to a strong flux suppression above the GZK-threshold energy $E_{\mathrm{GZK}} \sim 50$ EeV (see Figure 2.5). Observational claim for the cutoff was published by the HiRes Collaboration [17], being in contradiction to the observation of missing cutoff by the AGASA Collaboration [38]. This observational discrepancy was settled recently by the new results of the HiRes [39] as well as the Auger [21] Collaboration. Thus, after many

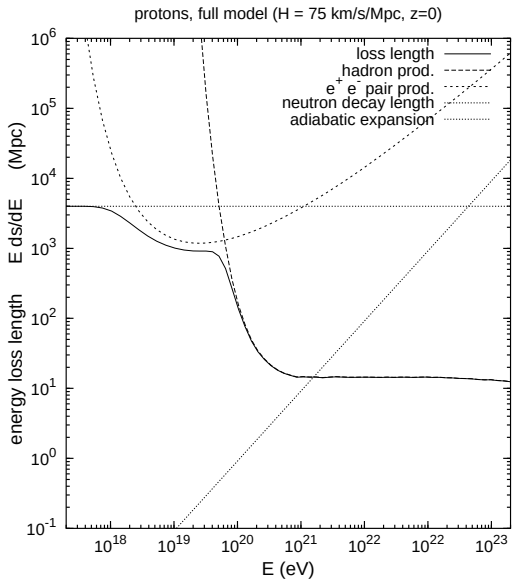


Figure 2.5: Energy loss length for extragalactic cosmic ray propagation [40]. The hadron production curve corresponds to the GZK-effect, while the e^+e^- pair production is causing a *dip*, which might be responsible for the *ankle* feature within the cosmic ray energy spectrum [33, 41, 35].

years of discussion and unclear observational results, a GZK-like feature can now be clearly identified within the HiRes ($\sim 5\,\sigma$) and Auger ($\sim 6\,\sigma$) energy spectra. However, it is still not clear, whether the observed flux suppression is due entirely to the GZK-effect.

While there is almost complete isotropy of the arrival direction of cosmic rays, very recent results from the Pierre Auger Observatory show a correlation of the arrival direction of cosmic rays at energies above $E > 57\,\mathrm{EeV}$ with close-by *active galactic nuclei* (AGN) [42]. Figure 2.6 depicts the published arrival direction distribution together with AGN locations. Active galactic nuclei are supermassive black holes with a mass of $M_{\mathrm{BH}} > 10^6\,M_\odot$ located in the center of galaxies. If supermassive black holes accrete large amounts of matter, they release enormous energies in the formation of extragalactic jets, which often are several times larger than their host galaxy. These jets are very bright from radio up to X-ray emission and are suspected to be the prime location for charged particle acceleration beyond energies of 100 EeV. Several models have been proposed that allow extremely efficient particle acceleration in the vicinity of AGNs. For example the acceleration of particles in the central region close to the supermassive black hole [43], or in the very hot spots, where the AGN-jets are finally stopped by the intergalactic medium ($\rightarrow$*"galactic beam-dump"*) [44]. A recent proposal can even explain ultra-high energy particle acceleration in the remains of former AGN jets (*fossil* jets) [45].

Within the current statistics available for such analysis, this does not yet mean the detection of a cosmic ray point source, but just the possible identification of a source population. Unfortunately, the cosmic ray flux above the considered energy is so small that the southern Pierre Auger Observatory can only measure a handful of events within a year of operation. For a high statistics analysis of this correlation, it might be necessary to wait for the northern Pierre Auger Observatory, which must be especially enlarged to enhance the detection

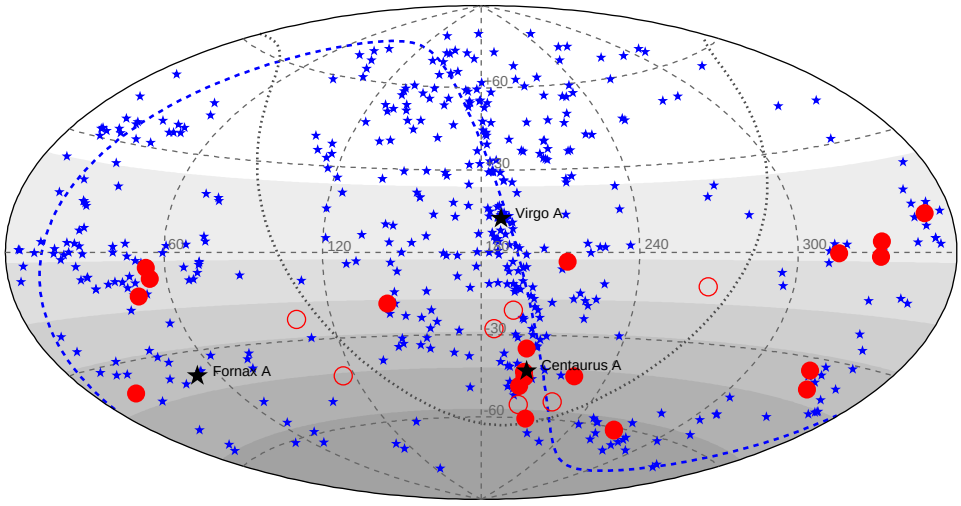


Figure 2.6: Arrival direction of the 27 events of the Pierre Auger Observatory with energies above 57 EeV (circles) and the position of close-by AGN objects (stars). The plot is given in Earth-bound equatorial coordinates, while the projection of the galactic plane is the dotted and of the supergalactic plane is the dashed line. The shaded areas denotes the integrated exposure of the Auger detector. The full circles denote the events, which are located within an angular distance of 3.1° from the next AGN. Events not correlating are shown as open circles. All the open circles are located not far from the galactic plane, where the used AGN catalog is known to be incomplete and the expected cosmic ray deflection are the largest [42]. The three strongest AGNs are marked as black stars.

efficiency at these ultra-high energies.

Cosmic ray particles are reaching energies that are far larger than any Earth-based accelerator will ever be able to achieve. To illustrate this, it is demonstrated what kind of machinery is needed to accelerate particles up to the highest cosmic ray energies of $E_{\text{lab}} \sim 100\,\text{EeV}$, using current technology.

- With the proposed ILC technology [46], which has a mean acceleration gradient of 35.1 MV/m, a linear collider with a length of $l \approx 2.8 \cdot 10^9$ km, roughly corresponding to the diameter of the Saturn orbit around our sun, would be needed.
- A circular accelerator structure with a continuous magnetic field of $B = 8.3\,\text{T}$ (corresponding to the LHC peak magnetic field) must have a radius of $r = \frac{p}{ZeB} \approx 40 \cdot 10^6$ km, which is about the radius of the Mercury orbit around the sun.

Machines of this size will probably never be realized. Therefore the only feasible way to study interactions at these energies is to rely on the cosmic ray particle beam provided by nature for free. The only but nevertheless severe disadvantage, at least from the point of view of a high energy particle physicist, is the extremely low luminosity due to the vanishing flux of cosmic ray particles at ultra-high energies. The observed flux at ultra-high energies is of the order of 1 particle per century and km^2. Every other living being other than high energy physicist is better feeling happy about this vanishing flux, since otherwise natural radiation levels would rise to critical levels.

In order to investigate the cosmic ray beam at the highest energies, it is necessary to build detectors on the scale of at least $1000\,\mathrm{km}^2$ and operate them over several decades. This sounds like instead of building unrealistically large acceleration structures, the problem is shifted to building unrealistically large detectors. Fortunately this is not the case, since over the last decades cosmic ray detector construction has proven to be feasible on a huge scale. Also in terms of cosmic ray detectors, a possible solution is provided by nature for free. It is straightforward to utilize the atmosphere of the Earth as fixed target, absorber material, active detector volume and light propagation medium at the same time. If cosmic ray primary particles hit the atmosphere, they initiate a large cascade of secondary particles which is called *extensive air shower*. This phenomenom was already discovered in the 1930's [47, 48]. Figure 2.7 illustrates how extensive air showers of different energy penetrate the atmosphere. Above a certain energy, air showers are reaching down to the surface of the Earth, where they can be detected by large arrays of particle detectors (*air shower arrays*). Air-fluorescence and especially Cherenkov radiation from air showers can be detected with optical detectors.

The center of mass interaction energy per nucleon of a cosmic ray particle with energy E_{CR} and mass A hitting the atmosphere is

$$E_{\mathrm{cms}} = \sqrt{s} = \sqrt{2\,m_{\mathrm{N}}^2 + 2\,m_{\mathrm{N}}\,E_{\mathrm{CR}}/A}\,, \tag{2.1}$$

where m_{N} is the nucleon mass. For proton primaries with $E_{\mathrm{CR}} = 100\,\mathrm{EeV}$ this results in $E_{\mathrm{cm}} = 4.5 \cdot 10^5\,\mathrm{GeV}$. On the other hand the proton-proton center of mass energy at the LHC is $E_{\mathrm{cm}} = 1.7 \cdot 10^4\,\mathrm{GeV}$ and corresponds to an incident cosmic ray proton with $E_{\mathrm{CR}} \sim 100\,\mathrm{PeV}$.

2.2 Extensive air shower phenomenology

When high energy cosmic ray particles hit the atmosphere, they interact with molecules of the air. These interactions at the highest energies are initializing the formation of *extensive air showers* (EAS). The total energy of the primary cosmic ray particle is distributed to the rapidly growing number of air shower secondaries. Hadronic interactions produce a large number of shortlived mesons (mainly pions) of which many decay into muons, electrons and photons. It is convenient to group the particles of an EAS into three main groups: *hadronic*, *muonic* and *electromagnetic*. In addition, there are particles not contributing much to the total energy balance, i.e. *UV-photons* (fluorescence and Cherenkov) and *radio emission*, or which are not detectable and are therefore called *invisible component* (neutrinos and very low energy particles). Figure 2.8 illustrates the decomposition of an EAS into particle groups.

The differences in the development of the electromagnetic and muon shower components are shown in Figure 2.9. Electromagnetic particles are by far the most abundant and they carry the largest fraction of the total energy. The electromagnetic shower develops fast, mainly by bremsstrahlung interactions and pair production. Below a certain energy E_{crit}, ionization energy losses start to dominate over new particle production and the shower is absorbed by the atmosphere. Thus, the electromagnetic shower shows a pronounced maximum number of particles at some stage of shower development. The location of this point in traversed slant depth of atmospheric matter is called *shower maximum* X_{max}. Since equal energy air showers mainly fluctuate in the resulting X_{max} but hardly in the maximum number of electrons, a pronounced overall minimum of fluctuations of the electron number is

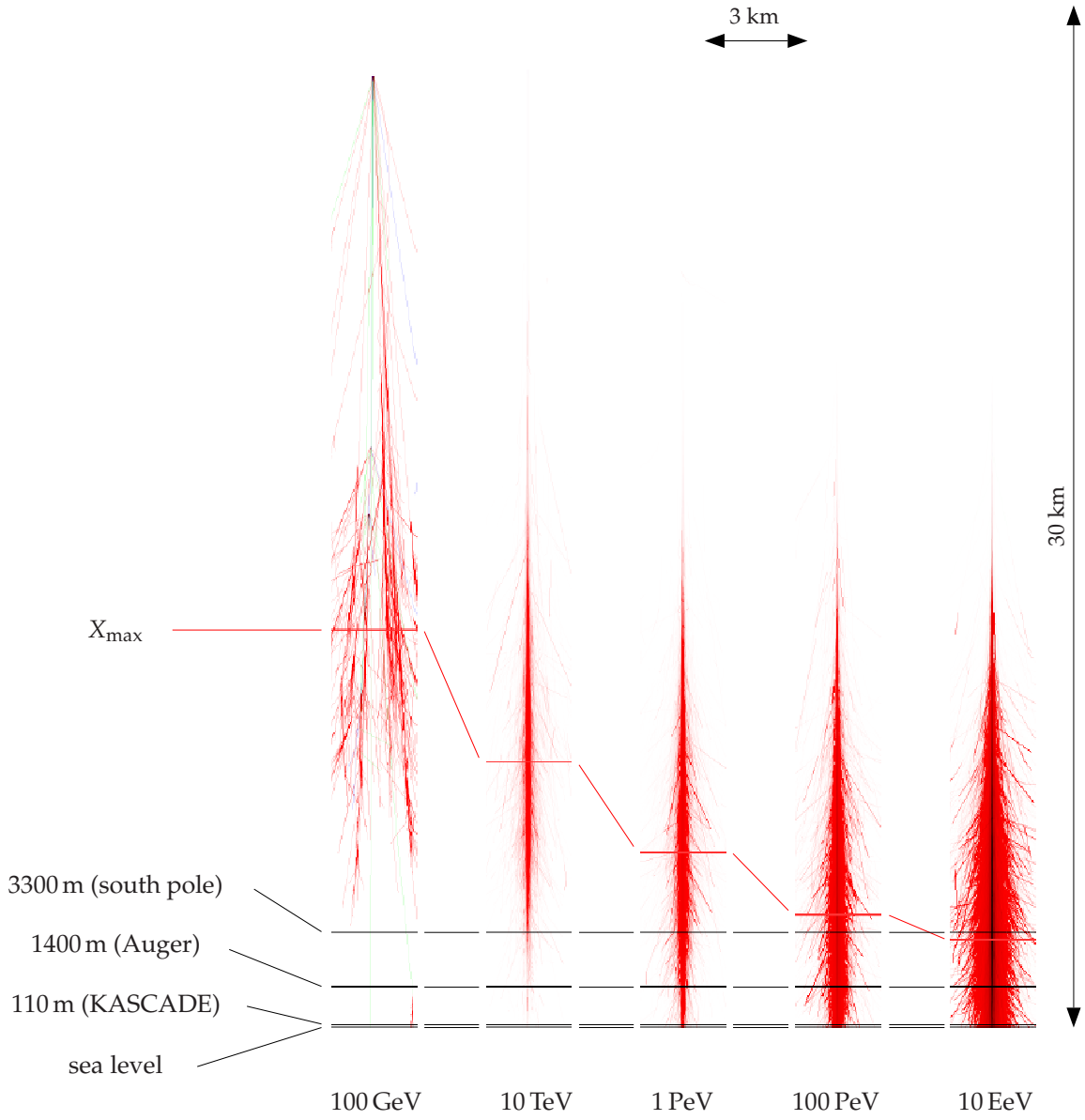


Figure 2.7: Simulated proton induced air showers at different primary energies E with CORSIKA [49] (visualized using COAST[2]). Due to the enormous difference in the total number of particles ($N_{\text{part}} \sim E$) the used intensity scale is different at each energy. The impression of the shower starting point moving deeper into the atmosphere with increasing energy is just an artifact of the changing intensity scale. In fact, higher energy showers tend to start higher in the atmosphere. The line marked by "X_{max}" identifies the shifting depth of the shower maximum, which increases approximately proportional to $\log(E/\text{eV})$. Typical observation levels in altitude above sea level are indicated as horizontal black lines. Low energy air showers do not reach any ground based detectors, so they can only be studied using balloons or satellites. Starting from $\sim$ 1 TeV air showers are penetrating the atmosphere, which makes observations possible at high altitude detectors. With increasing energy X_{max} moves closer to the observation level, reaching it at some point.

[2]See `http://www-ik.fzk.de/~rulrich`.

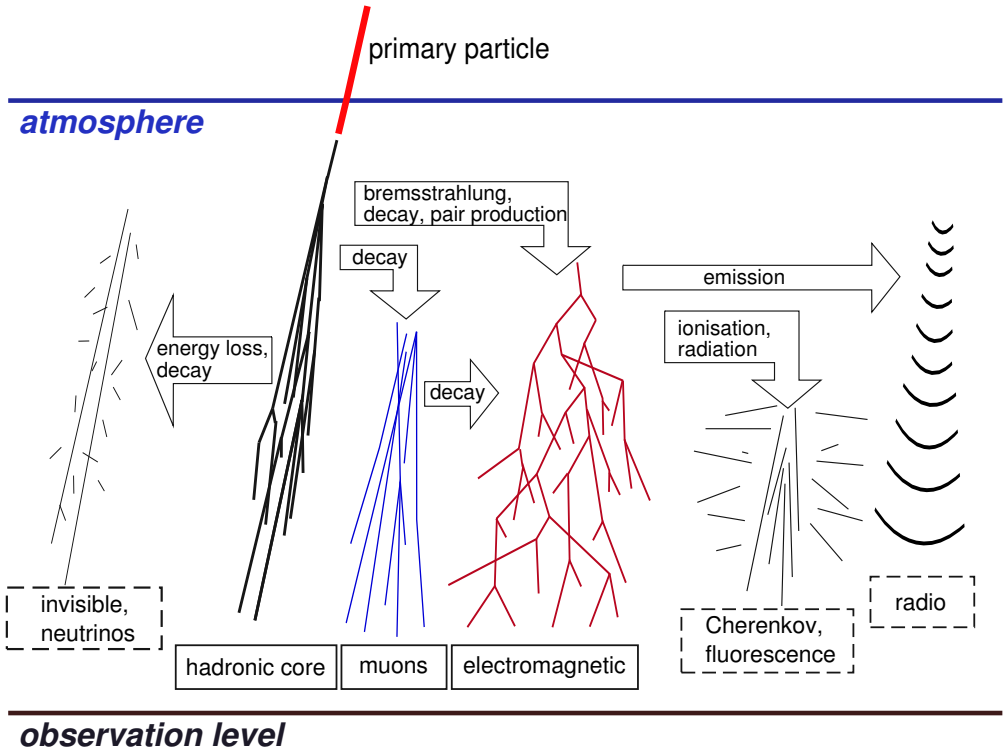


Figure 2.8: A primary cosmic ray nucleus hitting the atmosphere initiates a cascade of secondary hadrons, forming the inner core of an air shower. Muons are produced mostly due to the decay of charged pions. Electromagnetic particles are permanently produced mainly by decaying neutral pions and also decaying muons. In fact, after a few hadronic interaction lengths most of the shower energy is transferred to the electromagnetic shower component.

formed around the location of the mean $X_{\max}$. On the other hand, muons are highly penetrating particles. They hardly interact and loose only slowly energy due mainly to ionization. High energy muons can even penetrate kilometers of rock and reach deep underground detectors. Only lower energy muons may decay before they reach the detection level. The overall fluctuations of the muon number versus slant depth are small and almost constant. This is due to the fact that the position of the maximum development as well as the resulting shape of the muon cascade do hardly fluctuate.

Hadronic particles stay close to the shower axis, which is the direction of motion of the primary cosmic ray particle. After a few hadronic interactions, most of the hadronic energy is transferred into the electromagnetic and muonic shower parts. Since the hadronic shower core is long lived and therefore propagates deep into the atmosphere, it serves as a permanent source of new electromagnetic particles and muons. Therefore, the hadronic shower core and the hadronic interactions therein are paramount for all of the EAS physics and interpretation.

Figure 2.10 illustrates the different stages of shower development. As the progenitor of the air shower, the primary cosmic ray particle enters the atmosphere from the left side. At point X_1 it interacts with an air molecule. The amount of atmospheric matter traversed up to X_1 is determined by the exponential distribution $e^{-X_1/\lambda_{\mathrm{p-air}}}$, given by the interaction

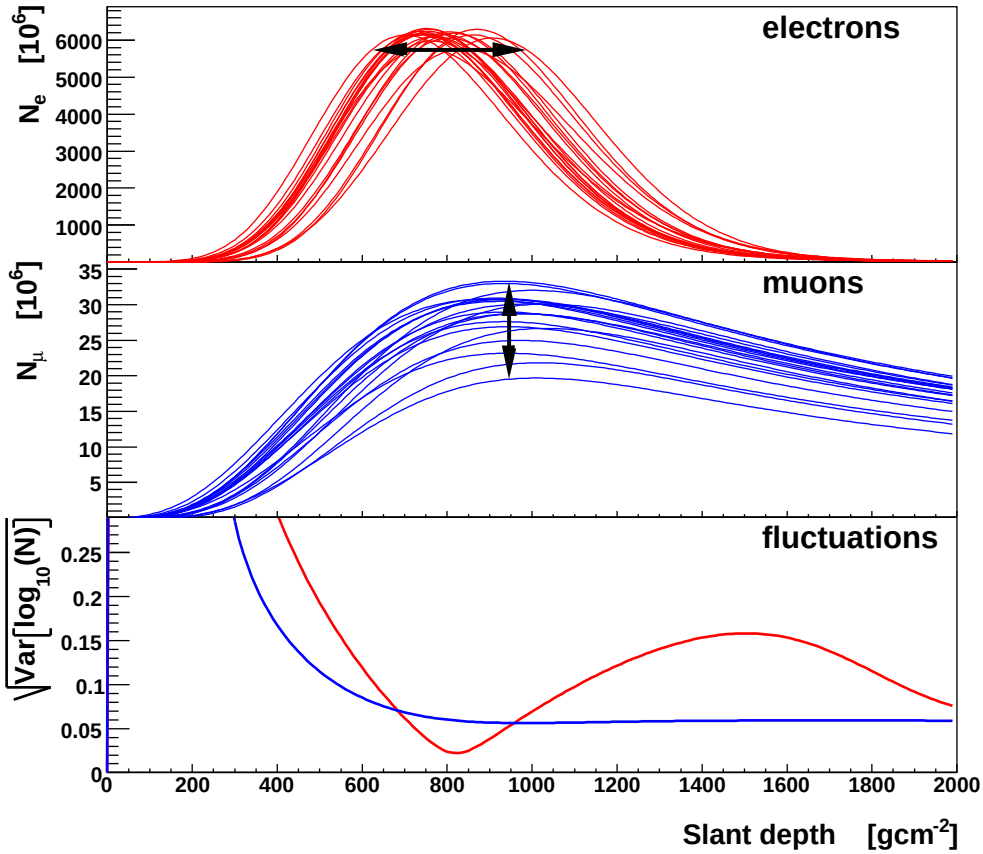


Figure 2.9: Fluctuations of proton initiated air showers at 10 EeV (simulated with CONEX [50]). The electron profile (top) fluctuates strongly, mainly in the depth of the shower maximum X_{max}, while the number of electrons at the shower maximum N_{max} stays almost constant. On the other hand, all of the muon profiles have a very similar shape with some fluctuations in the overall normalization (middle). The electron profile fluctuations in X_{max} are leading to a pronounced minimum of the overall shower fluctuations at the position of the mean X_{max} (bottom).

mean free path $\lambda_{\mathrm{p-air}} = \langle m_{\mathrm{air}} \rangle / \sigma_{\mathrm{p-air}}$, with the mean target mass of air $\langle m_{\mathrm{air}} \rangle \approx 14.5\, m_{\mathrm{p}} \approx 24253.01\,\mathrm{mb\,g/cm^2}$ and the proton-air cross section $\sigma_{\mathrm{p-air}}$. In the first interaction, a set of secondary particles is created that still can have very large energies. In the high energy interactions of the *shower startup phase*, the energy fractions transferred to the electromagnetic shower component are strongly fluctuating. After several high energy interactions a sufficiently large fraction of the primary energy is transferred to the electromagnetic shower component and the *shower cascade phase* is taking over. The cascade phase is entirely dominated by the large number of lower energy electromagnetic particles. The high number statistics results in an almost complete washing-out of additional fluctuations. At this stage, an air shower behaves like a well defined object, with macroscopically observable properties like X_{max}, N_{max}, particle *lateral distribution functions* (LDFs) and total energy E_{tot}. This is why measuring and interpreting EAS is possible in the first place. By looking at Figure 2.9 one can get an impression of how the air shower fluctuations are having very different impact on air shower arrays, as compared to optical (telescope) detectors. Ground based experiments

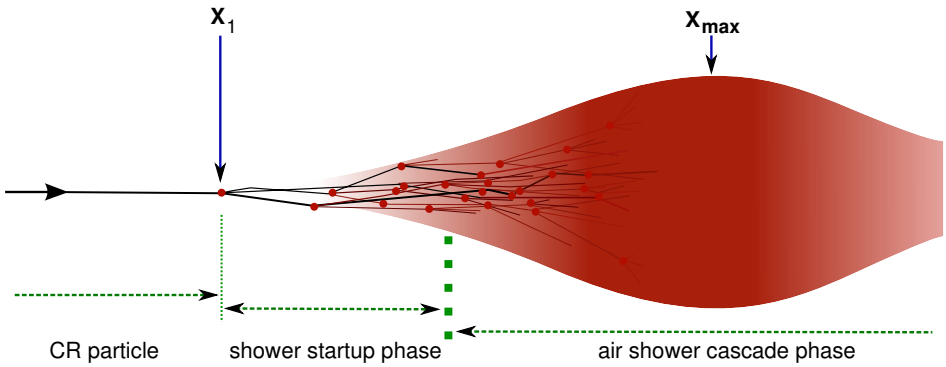


Figure 2.10: Development stages of an extensive air shower. The first interaction takes place at depth X_1. Black trajectories denote high energy hadronic particles that have a significant impact on the resulting shape of the electromagnetic shower bulk. Thus, the *shower startup phase* is extremely important for the properties of the resulting air shower. The red shaded area illustrates the density of electromagnetic particles that mostly have a small energy. Because of the very large number of particles in the shower bulk, the individual fluctuating particles are completely washed out and the resulting *air shower cascade phase* is entirely determined by the startup phase. At X_{max} the air shower reaches a pronounced maximum number of particles after which it is slowly absorbed within the atmosphere.

have no handle to measure X_{max} directly, for which reason this acts as the primary source of fluctuations for air showers of the same energy. In contrast to this, telescope detectors are directly observing X_{max}, thus its fluctuations do not introduce any additional uncertainty in air shower interpretation. In terms of fluctuations this corresponds to aligning all electron profiles (Figure 2.9, top) at their individual X_{max} . Since the electron profile normalization and shape are almost constant, the remaining shower-to-shower fluctuations are very small. Moreover, the fluctuations themselves carry valuable information about the primary mass composition and the ultra-high energy hadronic interaction characteristics. With telescope detectors this can be studied directly.

2.2.1 Heitler models

To learn more about air showers, it is helpful to devise a simplified picture of particle cascades. Figure 2.11 (a) displays the electromagnetic Heitler model [51], where electromagnetic means high energy e^-, e^+ and γ with an energy larger than their critical energy. The critical energy is defined as the energy at which ionization energy losses are getting equal to bremsstrahlung energy losses

$$\left.\frac{\mathrm{d}E}{\mathrm{d}X}\right|^{\mathrm{ion}}_{E=E_{\mathrm{crit}}} = \left.\frac{\mathrm{d}E}{\mathrm{d}X}\right|^{\mathrm{brems}}_{E=E_{\mathrm{crit}}} . \tag{2.2}$$

These particles only undergo pair production and bremsstrahlung interactions. It is assumed that after exactly one *splitting length*

$$\lambda = X_0 \ln 2 \tag{2.3}$$

an interaction occurs and the energy of the initial particle is distributed equally onto the two interaction products. In fact, λ is the grammage distance after which electromagnetic

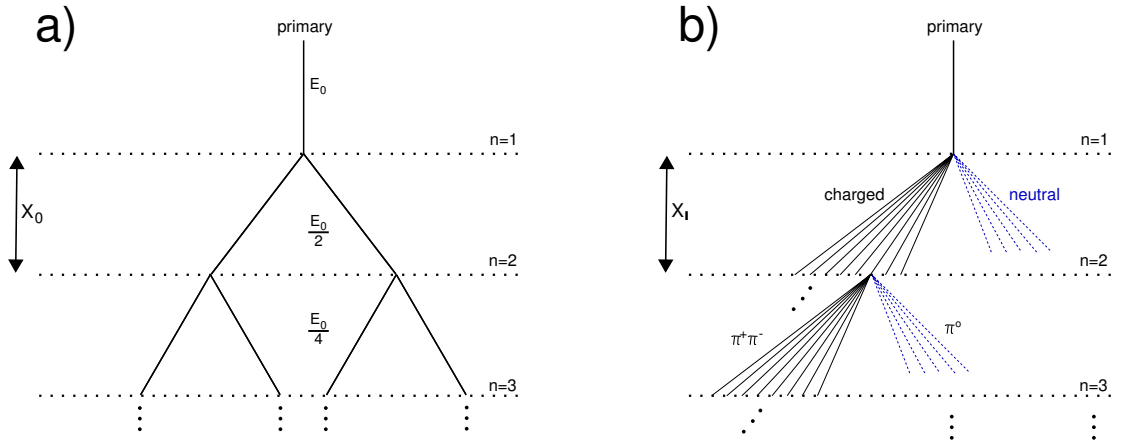


Figure 2.11: Heitler model for an effective electron-photon shower (a) and extended Heitler model for hadronic particles (b).

particles have lost half of their energy. The total number of particles doubles after every λ and the total energy E_0 is subdivided equally onto the particles of the cascade

$$N_n = 2^n \qquad \text{and} \qquad E_n = E_0/2^n \,. \tag{2.4}$$

When the energy per particle drops below the critical energy E_{crit} (85 MeV in air) they are assumed to stop producing new particles, but loose their remaining energy by ionization (collisional energy loss). At this point the cascade reaches its maximum in particle number and suddenly dies out:

$$N_{\text{max}} = E_0/E_{\text{crit}} \qquad \text{and} \qquad X_{\text{max}} = X_0 \, \ln \frac{E_0}{E_{\text{crit}}} \propto \ln E_0 \,. \tag{2.5}$$

Detailed calculations show (for example [61]) that X_0 can be identified with the electromagnetic radiation length, which depends on the medium in which the shower develops and is 36 to 37 gcm^{-2} in air.

Although the assumptions of the Heitler model are quite crude, which is reflected in the predicted unrealistic end of the particle cascade, the logarithmic dependence of X_{max} on the total energy E_0 is indeed observed in data and also in much more sophisticated simulations. Furthermore, the decadic *elongation rate*

$$D_{10} = \frac{\text{d}X_{\text{max}}}{\text{d}\log_{10} E_0} = X_0 \, \ln 10 = 85\,\text{gcm}^{-2} \tag{2.6}$$

can be calculated. Figure 2.12 shows that D_{10} is larger than what is observed in EAS data. This can be explained by the hadronic origin of real air showers.

A more sophisticated Heitler approach is depicted in Figure 2.11 (b). It takes into account hadronic interactions [62]. The hadronic splitting length $\lambda_{\text{I}} = \ln 2 \, X_{\text{I}}$, with $X_{\text{I}} \approx 120\,\text{gcm}^{-2}$ for pions in air, is introduced in analogy to the electromagnetic splitting length λ. In each hadronic interaction N_{mult} particles are produced, of which one third are π^0 and two thirds are $\pi^{\pm}$. The neutral pions decay instantly into two photons, initiating an electromagnetic cascade. The hadronic cascade stops when the pions drop below their critical energy, where

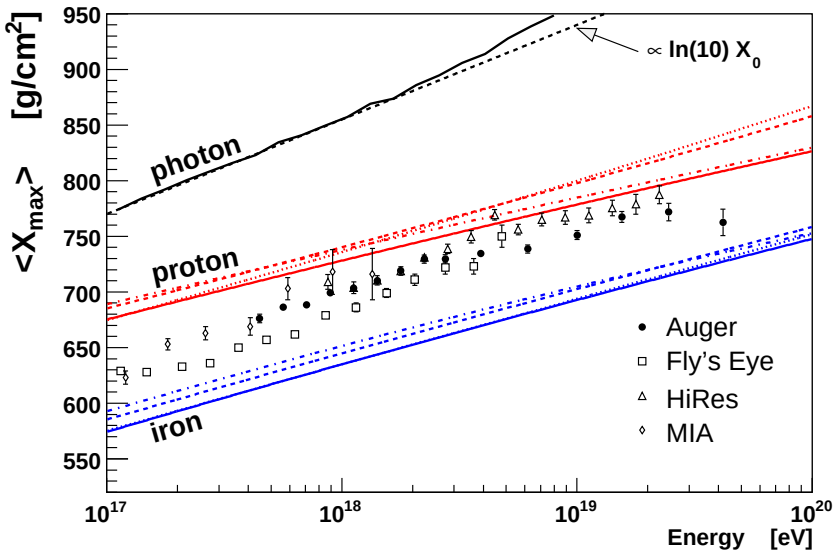


Figure 2.12: The mean $X_{\max}$ versus energy [26, 27, 28, 52] compared to air shower simulations for photon, proton and iron primaries. For hadronic primary particles, the differing model predictions are given as solid line for QGSJET01c [53, 54], dotted line for SIBYLL2.1 [55, 56], dashed line for EPOS1.61 [57] and dashed-dotted line for QGSJETII.3 [58]. The data identify a general trend from heavy cosmic ray primaries at lower energies to light primaries at high energy. The Heitler model prediction coincides very well with the photon simulation, whereas the starting deviation at high energy is related to the Landau-Pomeranchuk-Migdal effect [59, 60].

they are then assumed to decay, yielding muons. The critical energy for pions in air depends on the splitting length, the atmospheric density and the pion decay length. It can be approximated to be 20 GeV [63]. The total number of produced muons is identical with the number of $\pi^{\pm}$ at their critical energy $E^{\mathrm{I}}_{\mathrm{crit}}$

$$N_{\mu} = N_{\pi^{\pm}} = \left(\tfrac{2}{3}\, N_{\mathrm{mult}}\right)^{n_{\mathrm{c}}} = \left(\frac{E_0}{E^{\mathrm{I}}_{\mathrm{crit}}}\right)^{\beta} \quad \text{with} \quad n_{\mathrm{c}} = \frac{1}{\ln N_{\mathrm{mult}}} \ln \frac{E_0}{E^{\mathrm{I}}_{\mathrm{crit}}}\,, \tag{2.7}$$

where $\beta = \ln\left(\frac{2}{3}\, N_{\mathrm{mult}}\right) / \ln\left(N_{\mathrm{mult}}\right) \approx 0.9$.

To compare the depth of the shower maximum predicted by the hadronic Heitler model to that of the electromagnetic Heitler model, only the electromagnetic showers induced by the first generation of π^0 are accounted for. This will somewhat underestimate the real $X_{\max}$, since all the subsequent cascades starting at larger depth are neglected, but it is not possible to properly do the summation over many electromagnetic subshowers starting from different altitudes within this very simple approach. The result is

$$X_{\max} = \lambda_{\mathrm{I}} + X_0 \ln \frac{E_0}{N_{\mathrm{mult}}\, E^{\mathrm{e.m.}}_{\mathrm{crit}}}\,, \tag{2.8}$$

whereas $X_{\max}$ still depends on $\ln E_0$, but here it also depends on the multiplicity $\ln N^{-1}_{\mathrm{mult}}$.

The Heitler approach neglects an important aspect of hadronic interactions which has observable consequences. In hadronic interactions a significant fraction of the energy is carried away by the *leading particle*. This energy is not available for inelastic particle production.

The *inelasticity* is defined as the fraction of the energy that goes into new particle production to the total energy E_{tot}

$$k_{\mathrm{inel}} = 1 - {}^{E_{\mathrm{max}}}/_{E_{\mathrm{tot}}}\,, \tag{2.9}$$

where E_{max} is the energy carried away by the leading particle. The concept of inelasticity is immediately leading to a failure of the Heitler model, since the primary assumption of the equal energy distribution on the secondary particles does not apply any more. However, by simple reasoning one can get an impression on the importance of elasticity. For the assumption of $k_{\mathrm{el}} = 0.5$ only half of the energy is transferred into the electromagnetic cascade in the first interaction compared to the prediction of (2.8). This is producing a cascade with the shower maximum at shallower depth. At the same time the leading particle transfers a considerable amount of energy to the following interactions. The cascades initiated at the following interactions have therefore a much larger importance for the resulting total shower profile than for the case without inelasticity. Their effect is significant and can not be ignored without drawing wrong conclusions. However, there is no simple analytic way to get the result from this complex superposition of cascades. From full Monte Carlo simulations it is known that the net effect of a $k_{\mathrm{inel}} < 1$ is a push of the shower maximum to larger depths. Clearly, the inelasticity is an important parameter for air shower development [62, 64].

After this short excursion into schematic cascade development using Heitler models, it is obvious that air shower observables like X_{max} but also N_{e} and N_{μ} are closely linked to the properties of the hadronic interactions, like particle multiplicity N_{mult}, elasticity k_{inel} and the hadronic interaction length λ_{I}. All of these parameters are not well known at EAS energies and have to be extrapolated from low energy data by hadronic interaction models.

2.2.2 Extensive air shower simulations

The Monte Carlo simulation of extensive air showers needs 3D particle tracking in an atmosphere with changing density and the magnetic field of the Earth as well as a detailed implementation of particle decays and interactions over up to 14 orders of magnitude in energy. Since the number of shower particles grows proportionally to the energy, the number of particles that needs to be simulated increases by the same orders of magnitude as the energy. Therefore the demand in computation resources for a 10 EeV shower is $\sim$10000 times larger than that of a 1 PeV shower. With our contemporary computing infrastructure we are strongly limited in what can be simulated this way.

A standard method of Monte Carlo programs like MOCCA [65], AIRES [66] or CORSIKA [49] to drastically reduce the computation demands is particle *thinning*. By this technique not each individual particle is tracked any more, but particles can get a weight and therefore represent a whole bunch of corresponding particles. At extreme particle densities (several 1000-10000 per m^2), it is very efficient to track only few high weighted particles instead of all the individual particles. A thinning algorithm has to provide a reliable *importance sampling*, keeping the few relevant particles and removing the very abundant particles that are predicted not to give an important contribution to the total extensive air shower. By using a typical weight limitation of 100-10000 the computation demands are reduced almost correspondingly [67]. Unfortunately, thinning does remove relevant information from the shower, leading to artificial fluctuations. Also the reverse procedure of thinning, usually

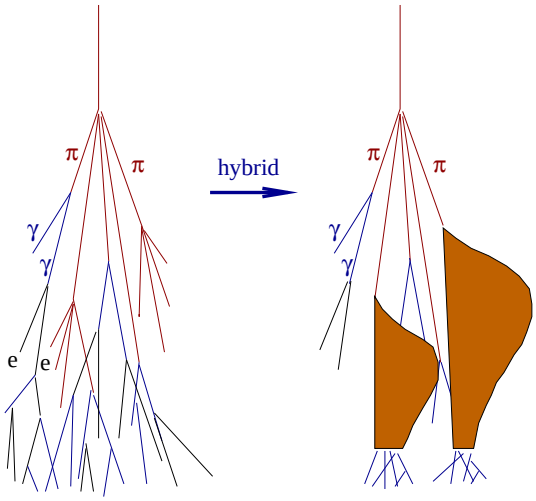


Figure 2.13: Air showers can be simulated in full 3D Monte Carlo, as CORSIKA does, or in combination with cascade equations (called *hybrid*) [70], as done for example by CONEX.

called *un-thinning*, needed for subsequent detector simulation is a delicate process with its limitations [68, 69]. Air shower thinning leads to artifacts, visible even after detector simulation, strongly limiting its usefulness for the interpretation of data. In conclusion, it should be mentioned that thinning is always a compromise between computation speed and simulation quality.

Figure 2.13 illustrates a different approach to reduce computation demands by partly relying on the numerical solution of cascade equations (see for example [71]), which describe propagation of particles through the atmosphere as well as energy losses and conversions from one particle type to another. Cascade equations themselves can not reproduce shower fluctuations. They describe a mean shower development. Therefore, to generate realistic air showers, it is necessary to handle the first few highest energy interactions explicitly and subsequently transfer all particles falling below a defined energy threshold to cascade equations. This approach is implemented in the air shower simulation programs CONEX [50] and SENECA [72]. A limitation of cascade equation based simulation tools is the missing information about particle timing and geometries. There are approaches to regenerate this information (SENECA), or pure 1-dimensional solutions that only calculate the shower profile along the shower axis (CONEX).

To generate the huge number of Monte Carlo events needed for this work, the CONEX program is ideally suited. If needed, the missing particle timing and lateral distributions are added afterwards using additional models and parameterizations. This is described in Section 3.4.8, presenting the chosen approach for Auger surface detector simulation and in some more detail in Appendix A.

2.2.3 Impact of hadronic interaction characteristics on air shower observables

Secondary particle production in the highest energy interactions is the largest source of uncertainty in EAS modeling. All characteristic parameters of hadronic interactions are chang-

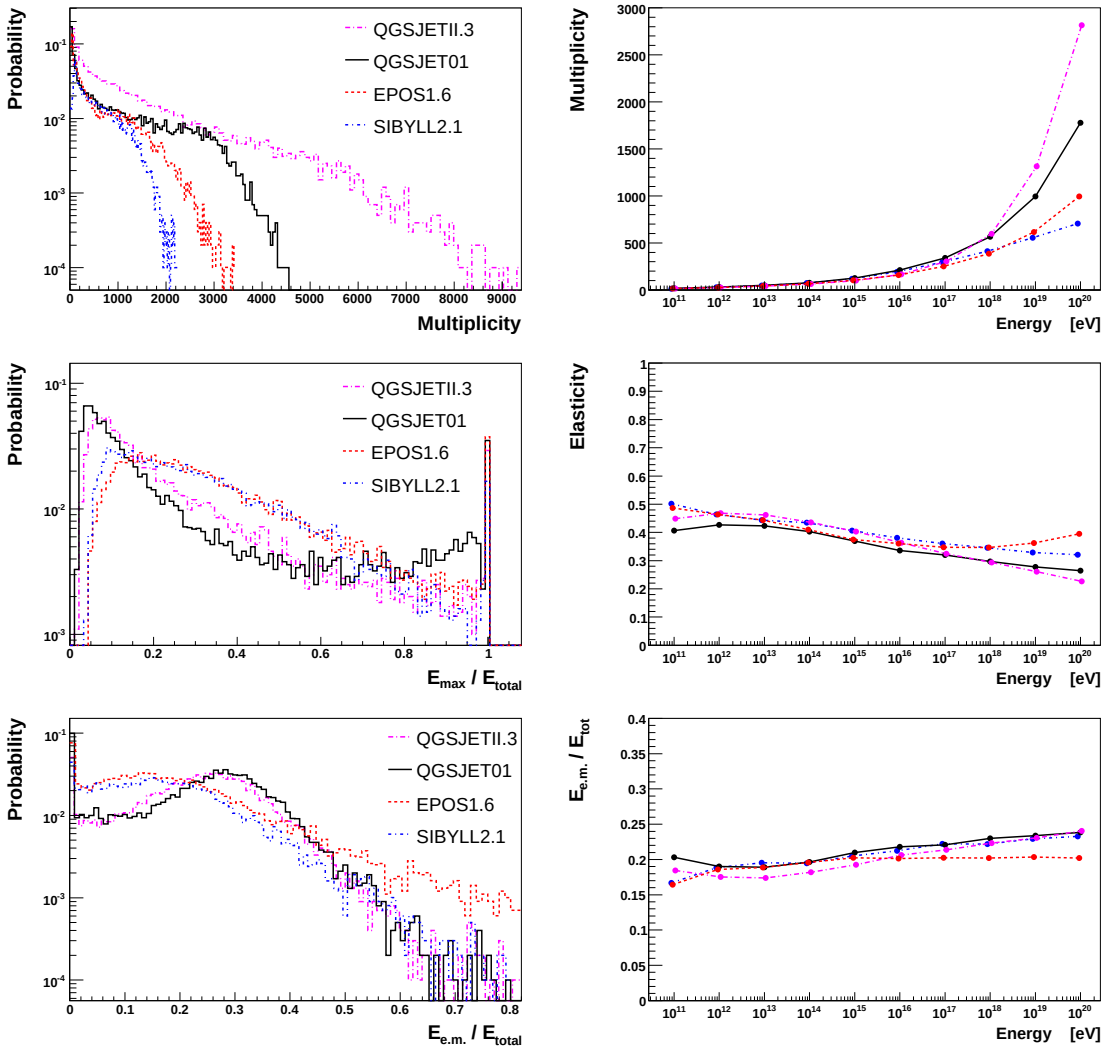


Figure 2.14: The diagrams on the left side show the distributions of features of the secondary particle distributions for proton-air interactions at 10 EeV, which are the most relevant for air shower development. On the right side it is shown how the mean of the distributions is changing with energy. Especially the secondary multiplicity distributions are uncertain by up to a factor of 2 to 4. The predictions of the mean elasticity $k_{\mathrm{el}} = {}^{E_{\mathrm{max}}}/_{E_{\mathrm{tot}}}$ vary only by less than 10%, while the shape of the distributions differs widely. Also note the differences in the diffractive peak at $k_{\mathrm{el}} = 1$. Only QGSJET01 and QGSJETII have their $r_{\mathrm{e.m.}} = {}^{E_{\mathrm{e.m.}}}/_{E_{\mathrm{tot}}}$ peak around 0.3, as expected from the simple $\pi^{\pm}/\pi^{0}$-picture (cf. Figure 2.11 (b)), while the mean $r_{\mathrm{e.m.}}$ of all models is very similar within a few percent.

ing with primary energy and can only be measured at energies much lower, than needed for EAS physics. The resulting extrapolations over many orders of magnitude in energy are a large source of uncertainty. This is well reflected by the different predictions of hadronic interaction models for the high energy behavior, as shown in Figure 2.14. Depending on the hadronic interaction model, there is not much freedom to change its behavior, since all predictions of the model are based on the *microscopic* physics or *phenomenological* implementation that produces a set of consistent interaction characteristics. If one of the underlying

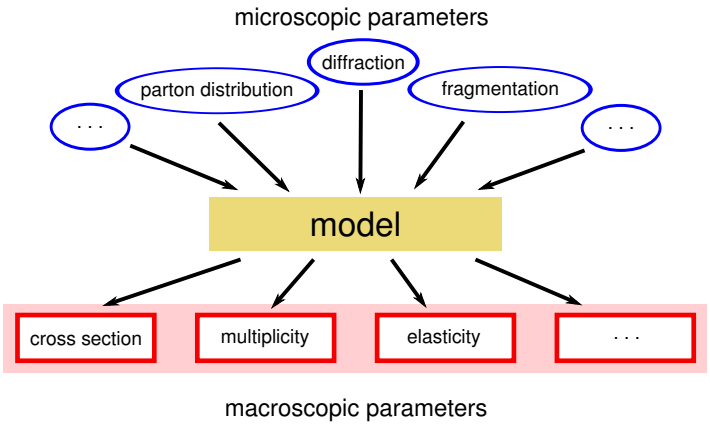


Figure 2.15: Schematic view of high energy hadronic interaction models. The internal implementation of the physics and phenomenology of a model (internal parameters) produces a set of correlated characteristic features of the secondary particle distributions relevant to air shower development (macroscopic parameters).

microscopic processes is changed, usually all model predictions are altered. This correlation is illustrated in Figure 2.15. Important features with a direct impact on the shape of the resulting EAS are:

High energy hadronic cross section, σ_{prod}, which determines the distribution of air shower starting points. At the same time it influences the location of subsequent interactions and therefore the speed of air shower development.

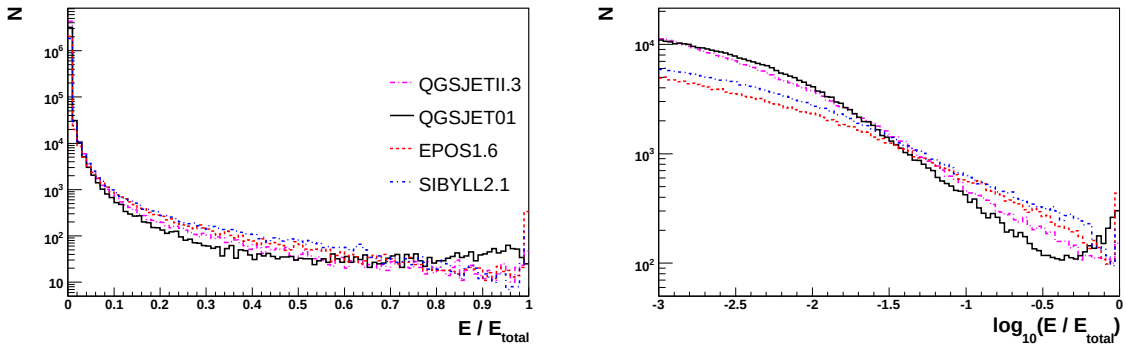
Secondary particle multiplicity, n_{mult}, which is related to the distribution of energy onto the secondaries. An increased number of secondaries leads to smaller overall fluctuations of the subsequent EAS development.

Diffraction is not really a property of the secondary particles, but more a specific type of interaction with a certain probability and therefore cross section σ_{diff}. In diffractive interactions both target and projectile particles may dissociate, producing very few particles. We can classify interactions with an elasticity close to 1 as diffractive with target dissociation, leaving the projectile particle almost unchanged and therefore have minimal impact on EAS development.

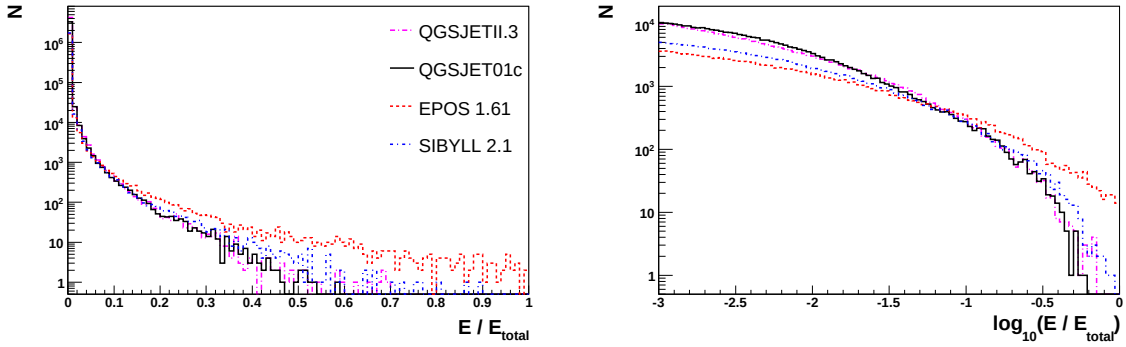
Elasticity, $k_{\text{el}} = {}^{E_{\text{max}}}/_{E_{\text{tot}}}$, quantifies how much of the total energy E_{tot} is carried away by the leading particle E_{max}. Only the inelastic part of the interaction participates in secondary particle production. A large elasticity produces more and higher energy interactions, which increases the overall shower fluctuations.

Fraction of energy going into electromagnetic particles, $r_{\text{e.m.}} = {}^{E_{\text{e.m.}}}/_{E_{\text{tot}}}$, which is a measure of how fast the electromagnetic shower is initiated. The faster this happens the less fluctuations are introduced and the smaller will be the resulting X_{max}.

The above list is far from complete, but it contains the most relevant points. In Figure 2.16 the complexity of differences in the produced secondary energy spectra can be observed. Since the points listed above hardly alter the resulting energy spectrum, these changes are



(a) hadronic secondaries (all hadrons except π^0)



(b) electromagnetic secondaries ($e^{\pm}$, γ and π^0)

Figure 2.16: Energy spectra generated in 10 EeV proton-air collisions by hadronic interaction models. Linear scale on the left and logarithmic scale on the right in E/E_{tot}, where E is the energy of the secondaries and E_{tot} the energy of the primary.

not considered here. Certainly, we can also expect an impact of the shape of the energy spectrum on the resulting air shower cascade.

For the measurement of the proton-air cross section it is certainly of paramount important to study the impact of a changing cross section on the resulting air shower development. Because of the very large differences in model predictions, also the impact of the multiplicity is studied in more detail, since it is a likely source of uncertainty. Also diffraction (see Figure 2.17) is handled later, in the context of *invisible cross section* (in Section 4.5). The elasticity as well as $r_{\text{e.m.}}$ are considered as higher order corrections for the cross section measurement that can be treated in a more general way, as described in Section 4.5.

Since different hadronic interaction models differ significantly about their high energy predictions, there seems to exist some freedom to choose the interaction characteristics within the accessible range opened up by the different models. This is supported by the fact that none of the present models can provide an overall good description of existing EAS data [19, 24, 73, 74, 75, 76, 77]. Presently, in the best case, a single EAS observable can be explained with a given model, leading to inconsistencies when comparing to other observables. This is intensively studied by the KASCADE Collaboration [74, 75, 77, 78]. While each hadronic interaction model describes EAS data more or less reasonably well, at least with respect to one aspect, none of the models gives an overall good description.

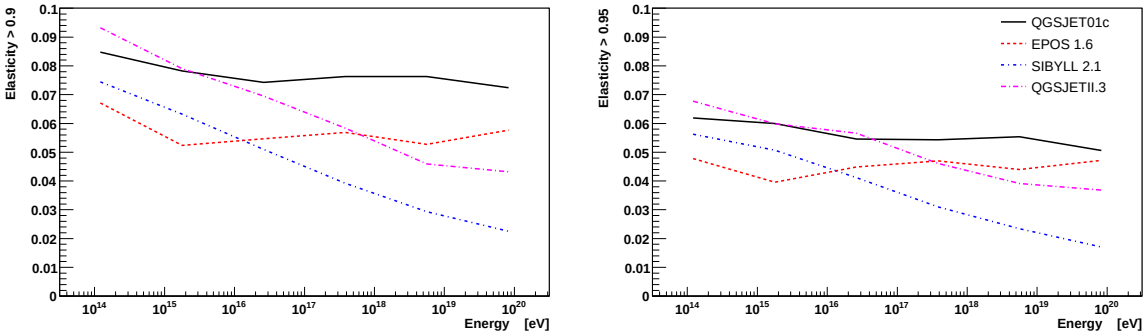


Figure 2.17: Fraction of the diffractive cross section $\sigma_{\mathrm{diff}} = \sigma(k_{\mathrm{el}} > k_{\mathrm{el}}^{\mathrm{min}})$ for the two values of $k_{\mathrm{el}}^{\mathrm{min}}$ 0.90 (left) and 0.95 (right).

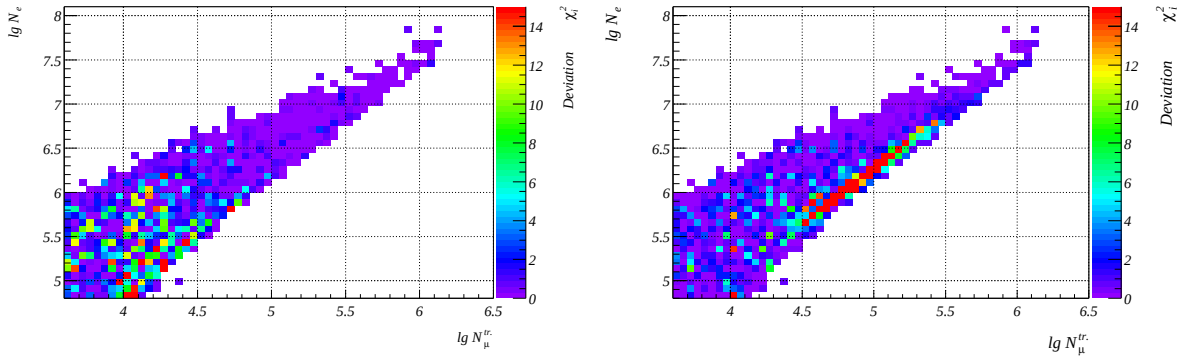
2.2.4 Implications on the interpretation of air shower data

The main purpose of measuring EAS is to determine the flux and composition of cosmic rays at ultra-high energy. Since air showers are the product of primary cosmic ray particles interacting with the atmosphere, it is possible to relate air shower observables to properties of the initial particle. This is possible by means of Monte Carlo simulations of extensive air showers, and is thus strongly depending on our present knowledge of physics on all scales within air showers [79]. The key to understanding air showers is the extrapolation of hadronic interaction characteristics to the highest energies, over many orders of magnitude in energy. The most prominent air shower observables are the slant depth of the shower maximum X_{max} and the number of electrons N_{e} and muons N_{μ} at the slant depth corresponding to the observation level of an air shower array.

KASCADE – frequency spectrum of N_{e} - N_{μ}. An approach to measure the primary composition between 1 and 100 PeV is undertaken by the KASCADE experiment using unfolding of the 2-dimensional N_{e}-N_{μ} frequency distribution [24]. Figure 2.18 shows the resulting χ^2-distributions after unfolding. For all choices of hadronic interaction models, the approach fails to consistently reproduce the full data distribution. Some models require primaries lighter than protons (while photons are excluded) and other models primaries heavier than iron.

The KASCADE data demonstrate that, even at galactic cosmic ray energies (∼PeV), the predictions of current hadronic interaction models are proven to deviate from reality. Extrapolated to higher energies this deviation may get a substantial problem. It seems that first hints of this are already visible within the data of the Auger experiment.

Auger surface detector – muon numbers. The Auger surface detectors by themselves can not distinguish between particle types. With simulations it can be studied what fraction of the energy deposit in an Auger detector at a distance of 1000 m from the shower core, denoted with S_{1000}, has to be attributed to electromagnetic particles or muons [76]. By exploiting the universality of the electromagnetic shower development [68, 80, 81] it is then possible to extract the energy deposit caused by muons S_{1000}^{μ} from Auger data. Figure 2.19



(a) QGSJET01 (b) SIBYLL2.1

Figure 2.18: χ^2-deviation from the original KASCADE N_e-N_μ frequency distribution after the unfolding procedure [19].

(left) compares the reconstructed energy deposit due to muons with simulations made with CORSIKA. The muon energy deposit is plotted over the slant depth distance to $X_{\max}$ and therefore resembles the shape of the longitudinal muon shower development. Apparently, the Auger data, even after considering a possible systematic 30% shift of the energy scale, can not be explained even by iron primaries. The muon number is by a factor of 1.53 ± 0.05 larger than proton simulations with QGSJETII [58]. If real, this would point to primary cosmic ray particles heavier than iron, which is very improbable.

Auger telescopes − longitudinal profiles. Figure 2.19 (right) shows the distribution of the depth of shower maxima $X_{\max}$, as it is measured by high quality Auger hybrid data (see Section 5 for the details on the measurement and selection of $X_{\max}$). The data are superimposed with the corresponding predictions by CONEX + QGSJETII for different air shower primary particles. By comparing just the shape, proton primaries are fitting best, but need a shift of $\sim 30\,\mathrm{gcm}^{-2}$ to lower $X_{\max}$ values, which requires significant changes of the high energy interaction models. Taking the predictions as they are, a mixed composition, starting with some iron to describe the rising edge, a selection of middle mass primaries to obtain the shape around the maximum, a considerable fraction of helium beyond the maximum and a small amount of protons to get the high $X_{\max}$-tail would be needed. Because the composition at these energies is still unknown and at the same time the uncertainties of the hadronic interaction models are large, it is not obvious to what extent the data need interpretation in terms of primary composition or changing hadronic interaction characteristics. However, as demonstrated in Section 2.1, there are strong observational arguments, supported by the expectations of theorists, for a large fraction of protons at ultra-high energy.

2.3 Hadronic interaction models

It is beyond the scope of the this work to attempt a comprehensive description of current hadronic interacting models. In the following only a very brief summary is given, empha-

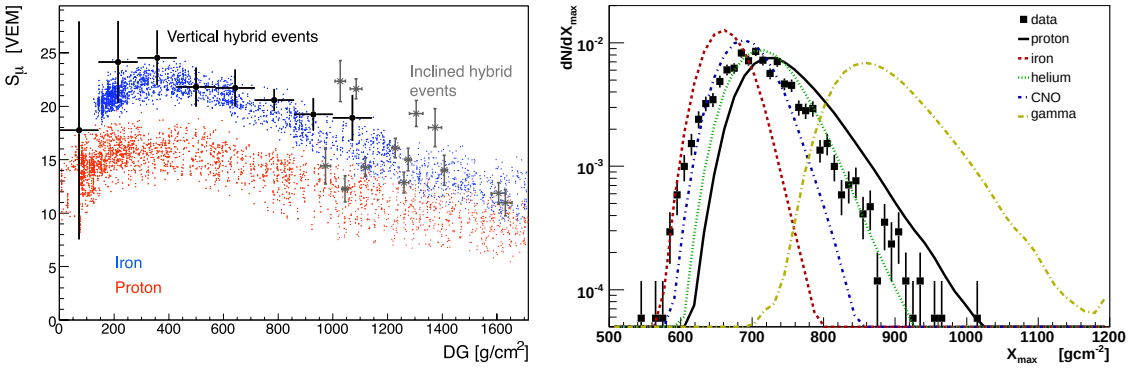


Figure 2.19: Comparison of reconstructed air shower observables of the Pierre Auger Observatory with predictions of simulations based on QGSJETII.3. Left panel: Muon tank signal contribution in dependence on the slant depth distance from $X_{\max}$ to ground DG $= X_{\text{obs}} - X_{\max}$, for vertical and inclined hybrid events. The muon profiles expected from QGSJETII.3 simulations for proton and iron induced showers are indicated by points. The energy scale for the simulations was shifted by $E = 1.3E_{\text{FD}}$ [76]. Right panel: Depth of the shower maximum $X_{\max}$ of high quality hybrid events between 1 and 10 EeV. The best match of the shape would make the proton simulation, but it has to be shifted by $\sim 30\,\text{gcm}^{-2}$ to lower $X_{\max}$ values. Alternatively, many mixed composition scenarios can explain the distribution. The simulations for the different primary particles are performed for exactly the same energy distribution as the one for the Auger data events, also considering the Auger $X_{\max}$-resolution.

sizing some important aspects that are of relevance to air shower predictions. Reviews of recent developments in this field can be found in [82, 82].

Although there is a well-established theory of strong interactions, Quantum Chromodynamics (QCD), we are still not able to calculate the bulk of hadronic particle production processes. Whereas processes with a hard scale (e.g. a large momentum transfer) can be calculated within perturbative QCD, most of the soft interactions with typically very small momentum transfer are still beyond the reach of modern computational techniques. Therefore one has to resort to rather general postulates about scattering amplitudes (i.e. analyticity and unitarity) and model-specific assumptions, whereas internal consistency of the theoretical and phenomenological framework is one of the desired aims in model building. Model assumptions are tested with extensive sets of accelerator measurements, which are also used to tune the free parameters of the models to obtain a good description of the data.

Accelerator measurements cover only a small fraction of the energy and particle production phase space ranges needed for describing interactions in EAS. In addition, very rarely the projectile-target combinations that most frequently occur in EAS, i.e. pion-, kaon- and nucleon-nitrogen, are measured. Therefore it is not surprising that there are large differences between the predictions of the currently available models [83], in particular if they are extrapolated to ultra-high energy [79, 84].

Mainly motivated by different approximation schemes applied for simulating multiparticle production in hadronic interactions, low- and high-energy interaction models are distinguished.

2.4 Low-energy models

Low-energy models typically cover the energy range from the particle production threshold up to about 200 GeV lab. energy. They are mainly parameterizations of measurements (GHEISHA, UrQMD) but also combinations of parameterizations and suitably adapted higher energy string models (FLUKA). Here, low-energy models are not discussed since the influence of the difference between these models is negligible as long as $X_{\max}$ is concerned. This can be understood in terms of the Heitler-type model introduced in Section 2.2.1. Most of the energy initially carried by hadrons is transferred to the electromagnetic shower component already in the first few hadronic interactions. The low-energy secondary particles of the first interactions are unimportant for the shower development since they produce subshowers of very small size.

2.5 High-energy models

The currently most successful high-energy interaction models are all based on the ideas of Gribov's Reggeon Field Theory (RFT) [85]. Within all of these models, perturbative QCD is applied for predicting the cross section of minijet production through parton-parton scattering [86, 87]. Significant differences of the models originate from different implementations of these ideas and from the model-specific treatment of high-parton density phenomena in high-energy scattering.

In Gribov's RFT, interactions are described by the exchange of quasi-particles, called pomerons and reggeons. A reggeon exchange corresponds to the sum over all possible graphs in which hadronic objects are exchanged that have a given set of quantum numbers (charge, parity, etc.) but different angular momenta. Classically, a pomeron is regarded as one particular reggeon, namely the reggeon with vacuum quantum numbers [88]. The amplitude for a single pomeron exchange graph is not known from first principles and parameterized differently in the various models.

The multiplicity of secondary particles is, to a large extent, dominated by the number of minijets produced in the hadronic collisions. Although perturbative QCD is applied in all models, the phase space region for which minijet production is applied is different. Another important source of differences stems from the assumption on the connection between perturbative partons and the soft "underlying event", in particular the spectator partons.

In the following a list of event generators considered in this work is given together with some overall characterization of the models. Detailed information can be found in the references given for each model.

2.5.1 QGSJET01

QGSJET01 [53, 89, 90] is an implementation of the quark-gluon strings model by Kaidalov and collaborators [91, 92]. A minijet component has been added to the quark-gluon strings model in 1997 [53]. Non-diffractive events are characterized by multiple soft and hard interactions.

This model is empirically found to describe data of many EAS detectors reasonably well. There are a number of theoretical shortcomings. For example, the minijets are calculated

with parton densities that stem from pre-HERA times and are now considered outdated and inadequate. The cross section for diffraction dissociation is a constant fraction of the elastic cross section.

The model predictions are characterized by a nearly perfect scaling of the leading particle distributions and a very high number of secondary particles at high energy. Also the fluctuations in multiplicity and inelasticity are very large.

2.5.2 QGSJETII

QGSJETII [54, 93] is the successor to QGSJET01. In addition to the physics already implemented in the previous version, pomeron-pomeron interactions are now taken into account. New parton densities are applied for minijet production and the ratio of diffractive and elastic cross sections decreases with increasing energy.

The implementation of pomeron-pomeron interactions made it possible to use updated parton densities and still keep an energy-independent transverse momentum cutoff for distinguishing between soft and hard interactions. The secondary particle multiplicity is lower than that of QGSJET01 over wide range in energy, however, at ultra-high energy, the new model version predicts even more secondary particles than any other model.

2.5.3 SIBYLL 2.1

SIBYLL [56, 55, 94] was initially constructed as a pure minijet model [86, 87], in which multiple hard but only up to one soft interaction is generated per hadron-hadron collision. In 1999 it was extended to comply with Gribov's RFT (without pomeron-pomeron graphs) and new parton densities were implemented [55]. An energy-dependent transverse momentum cutoff is applied for the minijet cross section calculation. This way the high-density regions of the minijet phase space are avoided, but an additional degree of freedom is added to the high energy extrapolation.

SIBYLL is the model which predicts the highest proton-proton and proton-air cross section at high energy. It is characterized by moderate scaling violation of the leading particle distributions. The ratio of diffractive to elastic cross sections decreases with increasing energy.

2.5.4 EPOS 1.6

In EPOS [95, 96], Gribov's RFT is applied to individual partons instead of hadrons, i.e. pomerons are exchanged between partons and not hadrons as in the other models [97]. A more self-consistent implementation of RFT could be achieved this way [98]. EPOS is optimized to give a very good description of recent RHIC measurements. The many features of hadronic multiparticle production in dense environments found at RHIC could not be described with the conventional approach of interacting and spectator partons. Therefore phenomenological parameterizations of projectile and target remnant excitations were introduced. EPOS is the successor to neXus, which has been declared as obsolete by the authors.

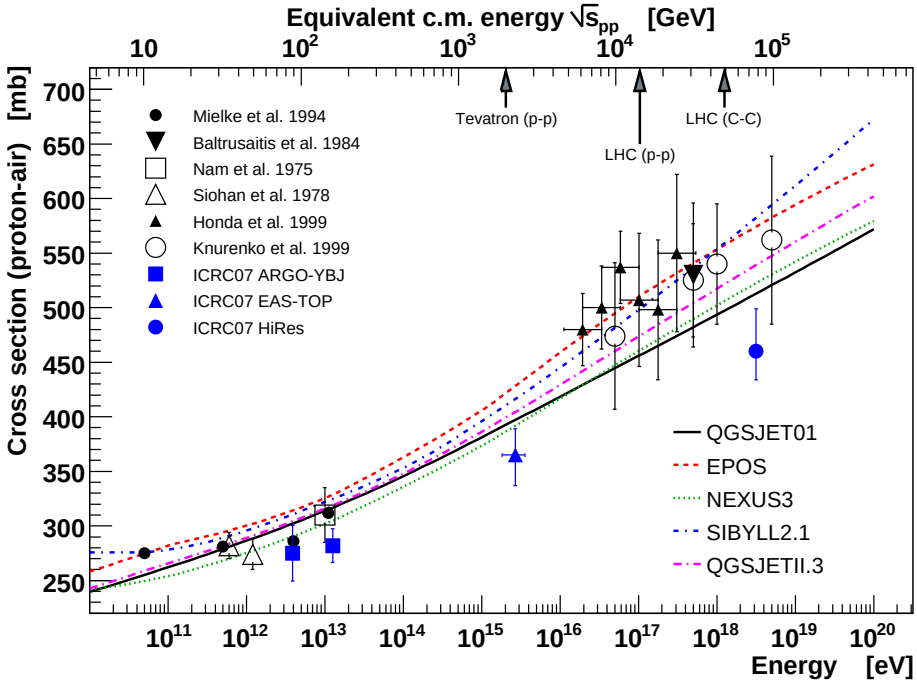


Figure 2.20: Current status of published values of the proton air cross section $\sigma_{\mathrm{p-air}}$ from air shower measurements [73, 99, 100, 101, 102, 103, 104, 105, 106, 107]. The data up to 100 TeV are based on unaccompanied hadrons measurements, with the exception of the ARGO-YBJ data points coming from a high altitude air shower array [107]. All data above this energy come from various air shower arrays, with the exception of the fluorescence telescope based Fly's eye [108] (X_{max}-tail) and HiRes [109] (X_{max}-deconvolution) measurements.

EPOS predicts strong scaling violation of the leading particle distributions. Also a very hard pion spectrum in proton-induced interactions is characteristic for this model. Both scaling violation and hard pion spectra are related to the treatment of remnant excitations and a special fragmentation effect, called popcorn.

At the time this thesis is written, the EPOS model is undergoing a major revision, mainly by re-tuning a number of parameters. It is expected that this re-tuning will lead to significant changes of the model predictions. Therefore EPOS will not be applied in the analysis of the cross section in this work.

2.6 Proton-air cross section measurements using cosmic ray data

Already in the 1960's first estimates of the proton-air cross section $\sigma_{\mathrm{p-air}}$ were performed using cosmic ray data [99]. These early measurements were based on the direct observation of the cosmic ray proton beam with hadron calorimeters. In Figure 2.20 most of the available cosmic ray based experimental results are summarized. Only for cosmic ray energies up to 100 TeV proton-air cross section measurements are based on a direct measurement of primary cosmic ray protons, with the exception of the high altitude air shower array ARGO-YBJ [107], which measures air showers at TeV energies. All other analyses are based on the observation of air showers, and are thus depending on the models used for EAS simulation. The following section summarizes how the proton-air cross section was inferred from

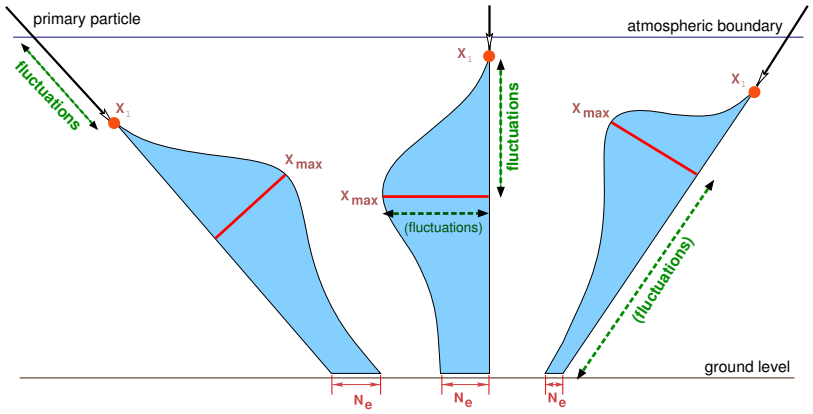


Figure 2.21: Relation of the air shower observables $X_{\max}$ and N_e to the location of the point of the first interaction X_1. The major sources of fluctuations are the exponential distribution of X_1 and the shower startup phase from X_1 to $X_{\max}$. The maximal size of the air shower, $N_{\max}$, is very well correlated to the primary energy and fluctuations are small. The shower development after the shower maximum can be assumed to be almost universal, since the shower at this stage consists mostly of lower energy electromagnetic particles.

measurements of air shower observables by previous analysis attempts. Generally all cross section analyses are based either on the direct measurement of $X_{\max}$ or the measurement of an observable that can be linked to $X_{\max}$, such as N_e. In Figure 2.21 the impact of the slant depth of the first interaction point X_1 on the typical air shower observables $X_{\max}$ and N_e is illustrated.

In Figure 2.22 it is shown, which parts of the total cross section are of relevance for shower development. Only the *production cross section* has a measurable impact on the resulting shower development. The *diffractive target-dissociation cross section* is an exception, since it corresponds to some particle production, but without any importance for the air shower. For cosmic ray primaries heavier than protons, the quasi-elastic *beam-breakup* cross section is contributing modestly to shower development. The ultra-high energy hadronic interaction models, as shown in Figure 2.20, usually use the production cross section, including diffraction.

There is some confusion about which part of the cross section is really measured within previous analyses. In principle, air shower development is not sensitive to interactions without particle production. Therefore we can define the *invisible cross section*, basically comprising elastic, quasi-elastic and target-dissociation. While there is no sensitivity to the magnitude of the invisible cross section, does not mean, it is not included in the results of a cross section analysis. By simply ignoring it, the utilized air shower Monte Carlo simulations will usually implicitly apply a corresponding model dependent correction. This is studied in more detail in Section 4.5.

2.6.1 Primary cosmic ray proton flux

The most straightforward method to measure the proton-air cross section is to measure the primary proton flux at different depths in the atmosphere (Figure 2.23) and relate the flux attenuation directly to the interaction free path length [73, 99, 100, 101, 102]. This measurement

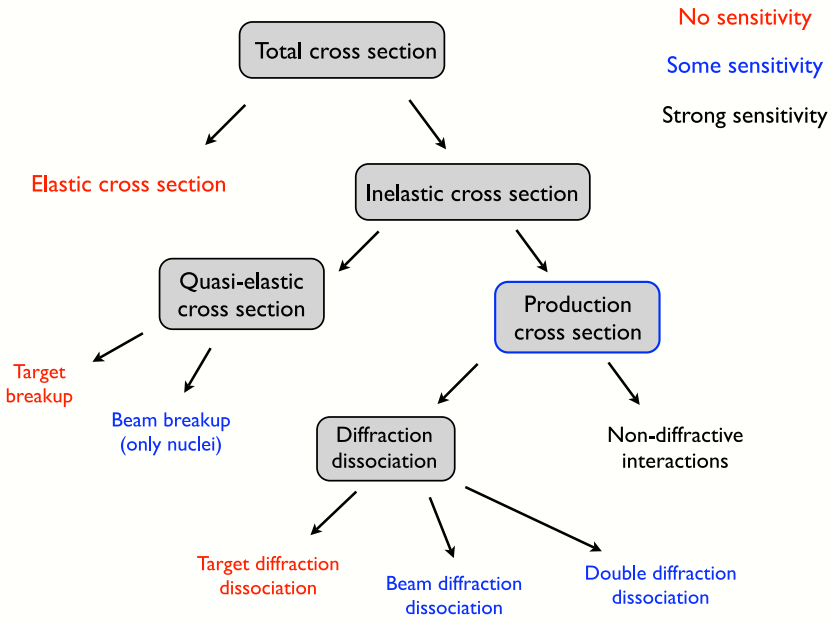


Figure 2.22: Cross section components. Relevant for air shower development is only the *production cross section*, with the exception of the *target diffraction dissociation cross section*, which does not significantly contribute to shower development. For all cosmic ray nuclei larger than protons the *quasi-elastic beam breakup cross section* has a slight additional contribution to EAS development.

technique requires to put hadronic calorimeters very high up on balloons or even outside of the atmosphere on satellites to measure the primary proton flux $\Phi_{\text{top}} = \Phi(X_{\text{top}})$ after a very small amount of traversed atmospheric matter $X_{\text{top}} \sim 1 - 5\,\text{gcm}^{-2}$ [15, 16, 18]. In addition, a measurement of the primary proton flux $\Phi_{\text{bottom}} = \Phi(X_{\text{bottom}})$ after a significant amount of traversed atmospheric matter $X_{\text{bottom}} \sim 600 - 1000\,\text{gcm}^{-2}$ is needed. There are two main limitations on this method, arising from these two measurements. Firstly, the size of a detector on a balloon or satellite is strongly limited, thus limiting the maximum observable energy of primary protons to ∼TeV. Secondly, the amount of background for a detector at the bottom of the atmosphere is enormous. To assure proper identification of unaccompanied primary protons, large hadronic calorimeters and additional veto arrays are needed. The expected attenuation of the proton flux after ∼10 interaction lengths is of the order of 10^{-5}. To reliably separate this small flux from other events, very efficient veto detectors and background estimates using Monte Carlo simulations are required. The attenuation length can then be calculated from

$$\lambda_{\text{prod}} = \frac{(X_{\text{bottom}} - X_{\text{top}})}{\ln(\Phi_{\text{top}}/\Phi_{\text{bottom}})}\,. \tag{2.10}$$

As it is impossible to veto all hadronic interactions along the cosmic ray passage through the atmosphere, this attenuation length can only be used to obtain a lower bound to the high energy particle production cross section $\sigma^{\text{prod}}_{\text{p-air}} = \langle m_{\text{air}} \rangle / \lambda_{\text{prod}}$. By design, the unaccompanied hadron flux is only sensitive to the particle production cross section, since interactions without particle production can not be separated from particles without any interaction.

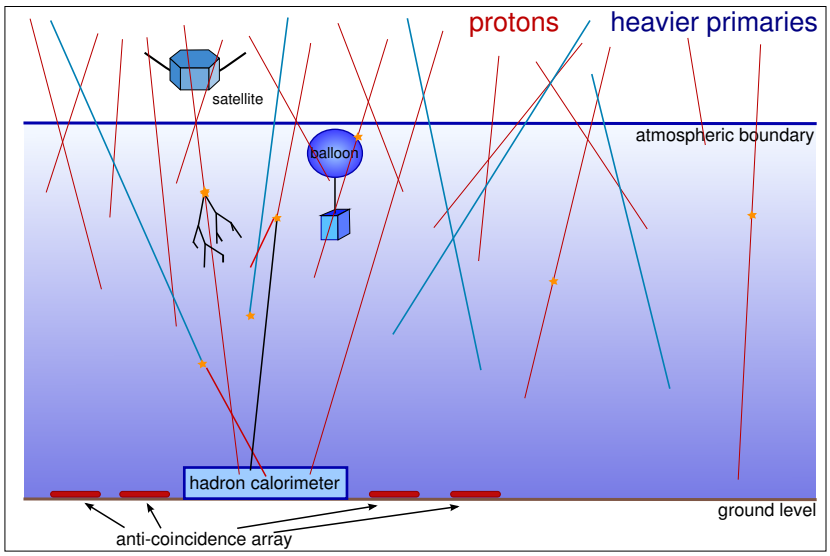


Figure 2.23: Primary cosmic ray proton flux as observed on satellites or balloons and in ground based calorimeters. The total cosmic ray flux generates a huge background for any detector within the atmosphere. This applies to ground based detectors as well as for balloon borne experiments.

2.6.2 Extensive air showers

In order to measure the cross section at even higher energies it is necessary to rely entirely on air shower data. The characteristics of the first few extremely high energy hadronic interactions during the startup of an EAS are paramount for the resulting air shower. Therefore, it must be possible to relate EAS observations like the shower maximum $X_{\max}$, or the total number of electrons $N_e(X)|_{X=X_{obs}} = N_e^{rec}$ and muons $N_\mu(X)|_{X=X_{obs}} = N_\mu^{rec}$ at a certain observation depth X_{obs} to the depth of the first interaction point, X_1, and the characteristics of the high energy hadronic interactions.

Air shower arrays. Figure 2.24 illustrates how a ground based air shower experiment can get a handle on the value of the cross section by measuring the frequency of air showers of the same energy at a given stage of their development. By selecting events of the same energy but different directions, the point of the first interaction has to vary with the angle, in order to observe the shower at the same development stage. The selection of showers of constant energy and stage depends on the particular detector setup, but the typical requirement is a constant (N_e^{rec}, N_μ^{rec}) at observation level. Examples of measurements of this type are [103, 104, 105, 107, 110]. Requiring an identical number of muons N_μ at the detector level does in good approximation select EAS of the same primary energy, because muons are stable as well as penetrating and therefore reflect the integrated shower development. On the other hand, showers with identical primary energy at the same stage of their shower development X_{stage} are assumed to yield the same number of electrons, N_e, since after the shower maximum the electromagnetic shower attenuation can be assumed to be universal. This is founded on the fact that at the shower maximum the total shower energy is almost completely transformed into a huge number of low energy electromagnetic particles.

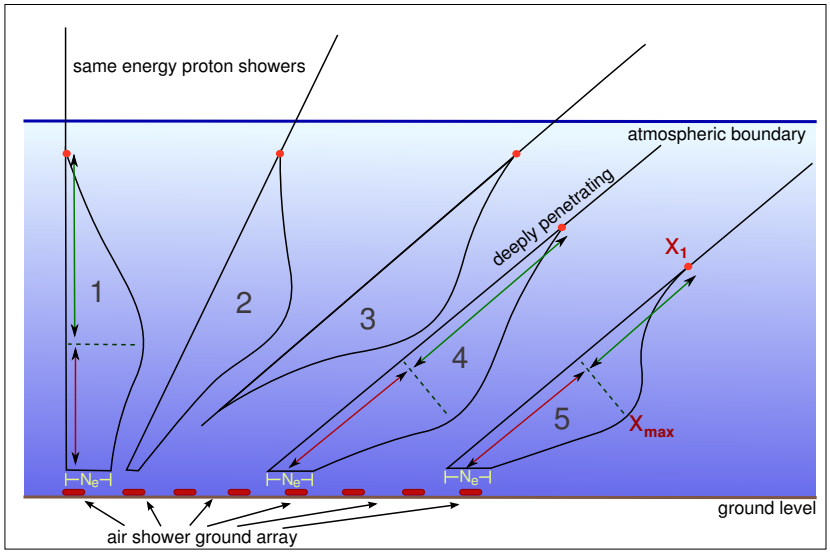


Figure 2.24: Equal energy proton induced air showers with an universal shower development before and after the shower maximum $X_{\max}$. The fourth shower (oblique) has the same number of electrons on the observation level N_e as the first shower (vertical). Showers with the same N_e and primary energy are assumed to have the same shower development stage, which corresponds to distance from $X_{\max}$. Therefore, the oblique shower needs a correspondingly deeper X_1 in order to reach the same N_e compared to the vertical event. The most righthand shower demonstrates an EAS with the same number of electrons in the detector, but a different development from X_1 to $X_{\max}$. Fluctuations of this kind are a major problem for ground based cross section measurements.

With the naming conventions given in Figure 2.25 the probability of observing a shower of a given energy E_0 and shower stage at the zenith angle θ can be written as

$$\frac{1}{N}\frac{\mathrm{d}N}{\mathrm{d}\cos\theta}\bigg|_{N_e^{\mathrm{rec}},\,N_\mu^{\mathrm{rec}}} = \int \mathrm{d}X_1 \int \mathrm{d}\Delta X_1 \int \mathrm{d}\Delta X_2 \;\; \frac{e^{-X_1/\lambda_{\mathrm{int}}}}{\lambda_{\mathrm{int}}} \cdot \; P_1(\Delta X_1)\cdot P_2(\Delta X_2) \cdot \; P_{\mathrm{res}}(X_{\mathrm{stage}}^{\mathrm{rec}},\, X_1+\Delta X_1+\Delta X_2)\,. \tag{2.11}$$

The additional factor of $\cos\theta$ from the projection of the array surface is not considered here. The distance X_{stage} between the first interaction point and the depth at which the shower reaches a given number of muons and electrons as defined by the selection criteria. The experimentally inferred shower stage at observation level $X_{\mathrm{stage}}^{\mathrm{rec}}$ does, in general, not coincide with the true stage due to the limited detector and shower reconstruction resolution. This effect is accounted for by the P_{res}. The functions P_1 and P_2 describe the shower-to-shower fluctuations. The probability of a shower having its maximum at $X_{\max} = X_1 + \Delta X_1$ is expressed by P_1, while P_2 denotes the probability of a given value of ΔX_2.

In cross section analyses Eq. (2.11) is often approximated by an exponential function of $\sec\theta$. Assuming that the integration of (2.11) over the distributions P_1, P_2, and P_{res} does not yield any generally non-exponential tail at large $\sec\theta$, it can be written as

$$\frac{1}{N}\frac{\mathrm{d}N}{\mathrm{d}\cos\theta}\bigg|_{N_e^{\mathrm{rec}},N_\mu^{\mathrm{rec}}} \propto e^{-X_{\mathrm{obs}}/\Lambda_{\mathrm{obs}}^{\mathrm{S}}} \propto e^{-\sec\theta/\Lambda_{\mathrm{obs}}^{\mathrm{S}}}\,. \tag{2.12}$$

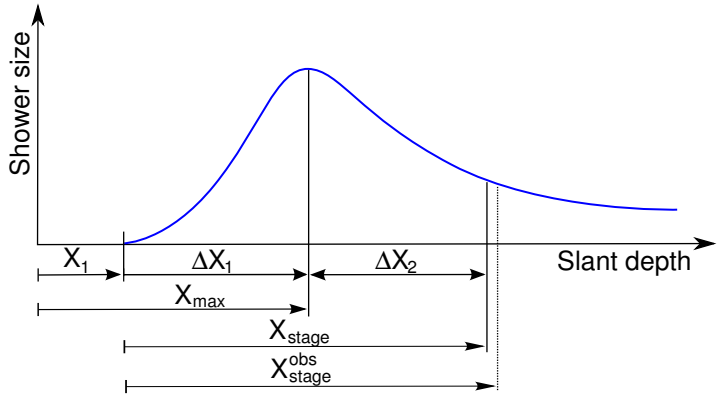


Figure 2.25: Definition of variables to characterize EAS longitudinal profiles. Zero slant depth is where the cosmic ray particle enters the atmosphere. The first interaction occurs at X_1, which initiates subsequent shower development. At $X_{\max}$ the shower reaches its maximum particle number, corresponding to a slant depth distance to X_1 of ΔX_1. The shower is attenuated over ΔX_2 before it reaches the observation level $X_{\text{obs}} = X_1 + X_{\text{stage}}$, which is located in a total slant depth distance of X_{stage} from X_1. However, due to uncertainties to measure the shower geometry as well as the shower size, the measured $X^{\text{obs}}_{\text{stage}}$ is generally not identical to the true shower development stage X_{stage}.

However, the slope parameter $\Lambda^{\text{S}}_{\text{obs}}$ does not coincide with the interaction length λ_{int} due to non-Gaussian fluctuations and a possible angle-dependent experimental resolution. The measured attenuation length can be written as

$$\Lambda^{\text{S}}_{\text{obs}} = \lambda_{\text{int}}\, k_{\Delta X_1}\, k_{\Delta X_2}\, k^{\text{S}}_{\text{res}} = \lambda_{\text{int}}\, k_{\text{S}} \qquad \text{with} \qquad k_{\text{S}} = k_{\Delta X_1}\, k_{\Delta X_2}\, k^{\text{S}}_{\text{res}}\,. \tag{2.13}$$

The k-factors $k_{\Delta X_1}$, $k_{\Delta X_2}$ and $k^{\text{S}}_{\text{res}}$ comprise the contributions to $\Lambda^{\text{S}}_{\text{obs}}$ from the corresponding integrations. However, these integrations are difficult to perform separately and the individual k-factors are not known in most analyses[3]. Finally we can see that the k_{S}-factor is related to shower fluctuations before and after $X_{\max}$ as well as the detector resolution. The importance of shower fluctuations on the $(N^{\text{rec}}_{\text{e}}, N^{\text{rec}}_{\mu}) = \text{const}$ method is discussed in [111].

Telescope detectors. With experiments being capable of directly observing the shower development in the atmosphere using fluorescence telescopes, the distribution of $X_{\max}$ could be used directly to obtain a handle on the value of the cross section at the highest energies [112, 113].

The direct observation of the position of the shower maximum allows us to simplify (2.11) by removing the term P_2 due to the shower development after the shower maximum. Furthermore the detector resolution P_{res} is much better under control for $X_{\max}$ and can be well approximated by a Gaussian distribution. The resulting distribution of reconstructed $X^{\text{rec}}_{\max}$ is

$$\frac{\mathrm{d}p(X^{\text{rec}}_{\max})}{\mathrm{d}X^{\text{rec}}_{\max}} = P(X^{\text{rec}}_{\max}) = \int \mathrm{d}X_1 \int \mathrm{d}\Delta X_1 \;\; \frac{e^{-X_1/\lambda_{\text{int}}}}{\lambda_{\text{int}}} \cdot\; P_1(\Delta X_1) \cdot\; P_{\text{res}}(X^{\text{rec}}_{\max} - X_{\max})\,, \tag{2.14}$$

with $X_1 + \Delta X_1 = X_{\max}$.

[3] For an exception see [106].

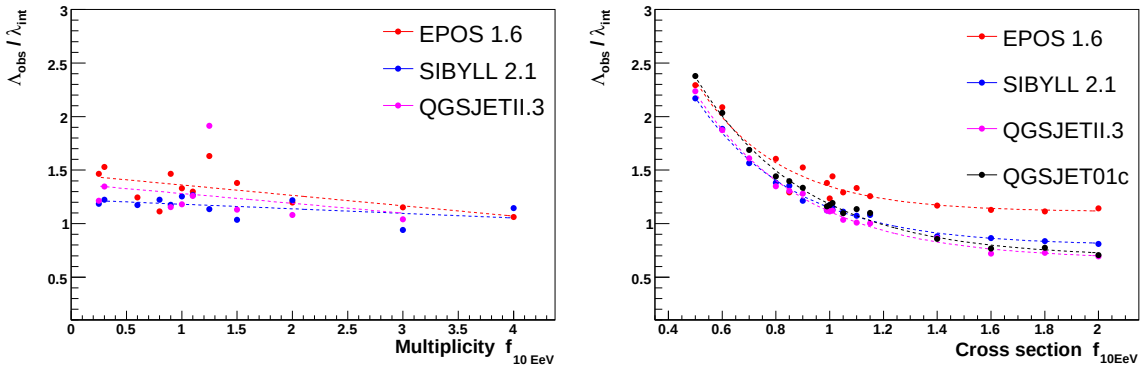


Figure 2.26: Results of a CONEX study with 10 EeV proton primaries, to infer the impact of changing interaction model characteristics on k_X-factors (in analogy to [114], see also Section 4.2.1 for details). The factors are calculated as $k_X = {}^{\Lambda_{obs}}/_{\lambda_{int}}$, where Λ_{obs} is the fitted exponential slope of the generated X_{max}-distribution after the peak of the distribution. The resulting impact is shown for of a changing secondary multiplicity $n^{mod}_{mult} = f_{10\,EeV} \cdot n^{org}_{mult}$ (left) and for a changing cross section $\sigma_{mod} = f_{10\,EeV} \cdot \sigma_{org}$ (right). The exact definition of the modification factor $f_{10\,EeV}$ is given in [114] and also in Section 4.2.1 (Equation (4.1) on page 74).

In analogy to (2.12) only the tail of $P(X^{rec}_{max})$ at large X^{rec}_{max} is approximated by an exponential distribution

$$P(X^{rec}_{max}) \propto e^{-X^{rec}_{max}/\Lambda^X_{obs}}, \tag{2.15}$$

whereas the exponential slope Λ_{obs} can be deduced from the convolution integral (2.14) as

$$\Lambda^X_{obs} = \lambda_{int}\, k_{\Delta X_1}\, k^X_{res} = \lambda_{int}\, k_X \qquad \text{with} \qquad k_X = k_{\Delta X_1}\, k^X_{res}\,. \tag{2.16}$$

Again $k_{\Delta X_1}$ and k^X_{res} are the contributions to Λ^X_{obs} from the corresponding integrations of (2.14).

"*k*-factor" technique. The approximation of an exponential attenuation of the frequency of air showers after the penetration of large amounts of atmosphere is the basis of the *k*-factor method [112, 113]. The exponential slope of the attenuation is typically denoted by Λ_{obs}, which is then related to the hadronic cross section by the *k*-factor

$$\Lambda_{obs} = k \cdot \lambda_{p-air}\,. \tag{2.17}$$

This method has been applied to data of air shower arrays (see Eq. (2.12)) as well as telescope detectors (see Eq. (2.15)). However, the resulting definition of the *k*-factor for a ground array, k_S, is (2.13) and for a telescope detector, k_X, is (2.16), which are not identical. The *k*-factors are obtained from massive air shower Monte Carlo simulations and are therefore depending on the models used in the simulations. Exemplary, the impact of changing characteristic features of high energy hadronic interaction models on the resulting k_X-factors is shown in Figure 2.26. This study was performed by modifying the high energy extrapolation of the cross section as well as of the secondary multiplicity of hadronic interactions in CONEX air shower simulations. The details on the performed modifications of the interaction characteristics are given in Section 4.2.1 or in [114]. The tail of the resulting X_{max}-distributions is then fitted by an exponential with the resulting slope Λ_{obs}. The k_X-factor is then calculated as

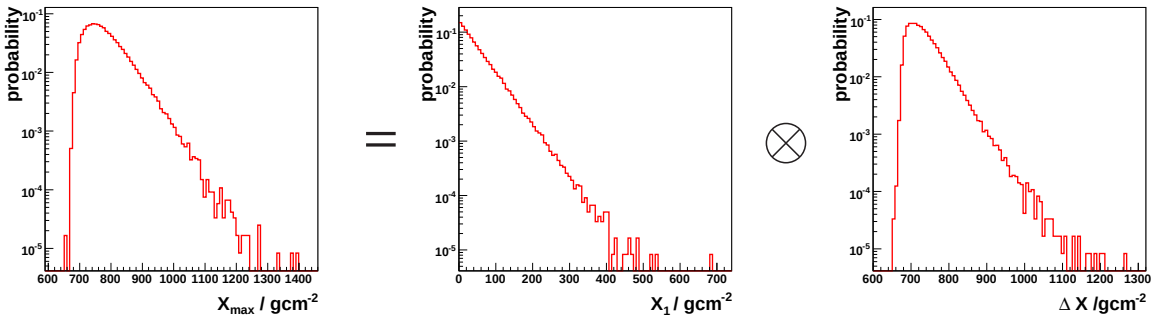


Figure 2.27: Principle of the $X_{\max}$-distribution decomposition method. The $X_{\max}$-distribution results from the convolution of the distributions of X_1 and ΔX_1, where $\Delta X_1 = X_{\max} - X_1$.

$k_X = {}^{\Lambda_{\text{obs}}}/_{\lambda_{\text{p-air}}}$. The found dependence of k_X on a changing multiplicity as well as cross section has never been taken into account by any air shower based cross section measurement.

$X_{\max}$-RMS method. For a short time it was believed that the proton-air cross section can be obtained just from the measurement of $X_{\max}$-fluctuations [115, 116]. In fact, the fluctuations are depending on the cross section, but nowadays it is well known that the RMS of the $X_{\max}$-distribution does mostly reflect the primary composition of cosmic rays. As a matter of fact, it is the best handle we currently have to learn about the primary mass composition. Only the extremely doubtful assumption of a pure proton cosmic ray composition may allow a measurement of the cross section this way.

Unfolding of the $X_{\max}$-distribution. A real improvement of the cross section measurement techniques was proposed by taking the air shower fluctuations more explicity into account [109]. This allows us to use not only the slope but more of the shape of the $X_{\max}$-distribution, by at the same time restricting the analysis to a range in $X_{\max}$, where the possible contribution from primaries other than protons is minimal. The ansatz unfolds the measured $X_{\max}$-distribution (2.14), by using a given ΔX_1-distribution to retrieve the original X_1-distribution (see Figure 2.27). The HiRes Collaboration claimed model independence of the used ΔX_1-distribution, leading to a model independent result for the cross section.

Indeed, this would have been a major step forward, since all the previous techniques are heavily depending on air shower Monte Carlo simulations and are therefore implicitly model dependent. Of course also the ΔX_1-distribution can not be accessed by observations, but has to be inferred entirely from simulations. Recently this triggered a discussion about the general shape and model dependence of the ΔX_1-distribution [117]. Ultimately this introduces a comparable amount of model dependence, as in the *k*-factor techniques (see Figure 2.28, left). This is a natural consequence of the fact that all air shower based analysis techniques are based on expression (2.14) in one or the other way.

Figure 2.28 (left) visualizes the dependence of the ΔX_1-distribution on hadronic interaction models. The ΔX_1-distribution, which mostly reflects the shower startup phase, is strongly depending on the parameters of the hadronic interaction models, like the cross

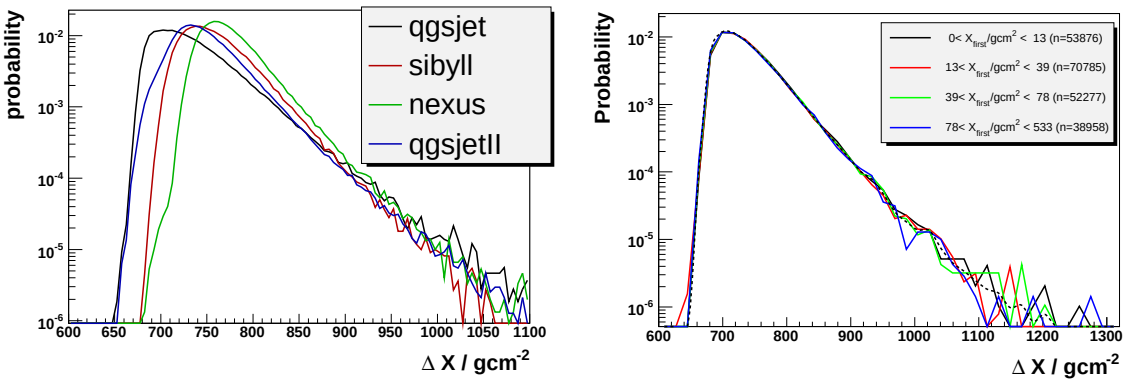


Figure 2.28: Model dependence of ΔX_1 (left) for 10 EeV proton simulations with CONEX. Independence of ΔX_1 of X_1 (right) for 10 EeV QGSJETII proton simulations with CONEX [117].

section itself but also the secondary multiplicity and others. This is intensively studied in Section 2.2.3.

To make the unfolding technique work in the first place, it has to be shown that ΔX_1 is independent from X_1 and thus $P_1(\Delta X_1|X_1) = P_1(\Delta X_1)$. As displayed in Figure 2.28 (right), the ΔX_1-distribution indeed does not depend on the depth of the first interaction [117]. The HiRes analysis [109] using the unfolding technique is lacking this demonstration.

2.6.3 Summary of measurement methods

It was shown that all methods of air shower based cross section measurements are indeed very similar and just variations of the same idea. They are all related to the single formulation given in (2.11). This is why they all suffer from the same limitations.

- The application of air shower Monte Carlo simulations introduces a model dependence for all measurement techniques based on the analysis of air shower observables.
- It can be shown that the $P_1(\Delta X_1)$-distribution is very sensitive to changes of the high energy hadronic interaction characteristics and thus makes $P_1(\Delta X_1) = f(\sigma, n_{\text{mult}}, ...)$ a function of σ , n_{mult} and other high energy model parameters [114]. Consequently this also makes the k-factors depending on the high energy interaction characteristics, as it is displayed in Figure 2.26. This effect certainly must be considered for any cross section analysis.
- Generally, the P_1 and P_2 distributions have a complex shape. The integrations of (2.11) and (2.14) to obtain the approximations (2.13) and (2.16) are leading to non-exponential contributions.
- Any non-exponential contribution creates a strong dependence of the fitted Λ_{obs} on the chosen fitting range [118]. A strong non-exponential contribution makes the k-factor analysis unusable.
- The values of all k-factors must be retrieved from massive Monte Carlo simulations. All analysis attempts so far have only calculated the combined factor of k_S, respectively k_X

(see [106] for a partial exception).

- k-factors depend on the resolution of the experiment and can therefore not be transferred simply to other experiments.
- k_X-factors are inherently different from k_S-factors and can therefore not be transferred from an X_{max}-tail analysis to that of ground based frequency attenuation or vice versa.
- It can not be disentangled whether a measurement of Λ_{obs} can be attributed to λ_{int} entirely or at least partly to changed fluctuations in ΔX_1 or ΔX_2.

2.7 Glauber theory and proton-proton cross section

Relating the proton-air cross section to that of proton-proton interactions makes it possible to compare cosmic ray measurements with accelerator data and corresponding models. Already in the 1950's it was realized that the cross section of a compound system is not the sum of the cross sections of its constituents [119]. Shadowing effects are leading to a much smaller cross section of the resulting compound system. This applies for hadrons, which are build out of partons as well as nuclei that are made out of nucleons.

In the following the scattering amplitude for proton-proton interactions is discussed. This will be used later for calculating the proton-air cross section.

2.7.1 Scattering amplitude and optical theorem

Figure 2.29 (left) illustrates the general case of scattering of the two initial state particles $|p_1(k_1)p_2(k_2)\rangle$ producing n_j particles of the final state j with $|p'_1(k'_1)...p'_{n_j}(k'_{n_j})\rangle$. The differential cross section for this process is

$$\mathrm{d}\sigma_j = \frac{1}{\mathcal{F}} \left| \langle p'_1(k'_1)...p'_{n_j}(k'_{n_j})|\mathcal{M}|p_1(k_1)p_2(k_2)\rangle \right|^2 \cdot \mathrm{d}\Phi_j \,, \tag{2.18}$$

where $\mathcal{F}$ denotes the flux factor and the n-body phase space reads

$$\mathrm{d}\Phi_j = (2\pi)^4 \prod_{i=1}^{n_j} \left[\frac{\mathrm{d}^3 k'_i}{(2\pi)^3 2E'_i} \right] \delta^4 \left(k_1 + k_2 - \sum_{i=1}^{n_j} k'_i \right) . \tag{2.19}$$

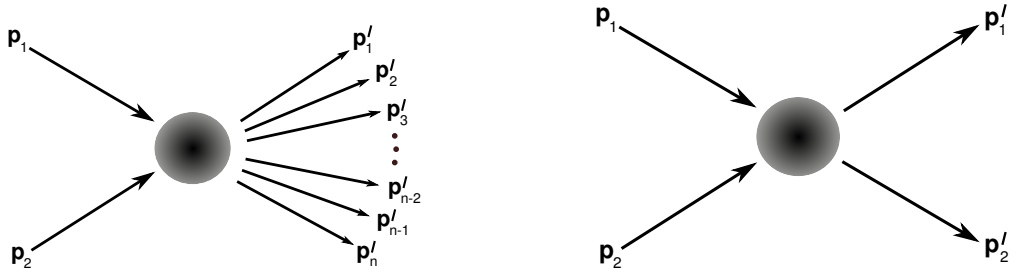


Figure 2.29: Left panel: Inelastic scattering with multiparticle production. Right panel: Two-body final state, which for the special case of $p'_1 = p_1$ and $p'_2 = p_2$ corresponds to the case of elastic scattering. Each particle p_i has an associated momentum k_i.

In the following for particle i the momentum k_i is not written explicitly any more, abbreviating $p_i(k_i)$ by just p_i, sometimes with the meaning of k_i. After summation over all possible final states j and integrating over the phase space $\mathrm{d}\Phi_j$, Equation (2.18) yields the total cross section

$$\sigma_{\text{tot}} = \sum_j \int \mathrm{d}\Phi_j \;\; \frac{1}{\mathcal{F}} \left| \langle p_1' ... p_{n_j}' | \mathcal{M} | p_1 p_2 \rangle \right|^2 . \tag{2.20}$$

The relation of the matrix element $\mathcal{M}$ to the $\mathcal{S}$-matrix is

$$\mathcal{S} = \mathbb{1} - \mathrm{i}(2\pi)^4 \delta^4 \left(p_1 + p_2 - \sum_{i=1}^{n_j} p_i' \right) \cdot \frac{\langle p_1' p_2' \dots p_{n_j}' | \mathcal{M} | p_1 p_2 \rangle}{(2E_1)(2E_2) \prod_i (2E_i')} . \tag{2.21}$$

Unitarity requires

$$\mathcal{S}^\dagger \mathcal{S} = \mathbb{1} . \tag{2.22}$$

Since the states are normalized to

$$\langle p(k') | p(k) \rangle = (2\pi)^3 \delta^3 (k - k') , \tag{2.23}$$

completeness is expressed by

$$\mathbb{1} = \sum_j \prod_{i=1}^{n_j} \left[\int \frac{\mathrm{d}^3 k_i'}{(2\pi)^3} \right] | p_1' p_2' \dots p_{n_j}' \rangle \langle p_1' p_2' \dots p_{n_j}' | . \tag{2.24}$$

Exploiting unitarity (2.22) for the general two-body final state $|p_1' p_2'\rangle$, as displayed in Figure 2.29 (right), yields

$$\begin{aligned} \mathcal{S}^\dagger \mathcal{S} = \mathbb{1} + \frac{(2\pi)^4 \delta^4 (p_1 + p_2 - p_1' + p_2')}{(2E_1)(2E_2)(2E_1')(2E_2')} \Big(& \langle p_1' p_2' | \mathrm{i}\mathcal{M} - \mathrm{i}\mathcal{M}^\dagger | p_1 p_2 \rangle + \\ & \frac{(2\pi)^4 \delta^4 (p_1 + p_2 - p_1' + p_2')}{(2E_1)(2E_2)(2E_1')(2E_2')} \langle p_1' p_2' | \mathcal{M}^\dagger \mathcal{M} | p_1 p_2 \rangle \Big) = \mathbb{1} \end{aligned} \tag{2.25}$$

and thus, by inserting the projector (2.24) and using expression (2.19), gives

$$\begin{aligned} & \langle p_1' p_2' | - \mathrm{i}\mathcal{M} + \mathrm{i}\mathcal{M}^\dagger | p_1 p_2 \rangle \\ & \qquad = \sum_j \int \mathrm{d}\Phi_j \, \langle p_1' p_2' | \mathcal{M}^\dagger | p_1' p_2' \dots p_{n_j}' \rangle \langle p_1' p_2' \dots p_{n_j}' | \mathcal{M} | p_1 p_2 \rangle . \end{aligned} \tag{2.26}$$

For the special case of elastic scattering into the forward region

$$p_1' \to p_1 \qquad p_2' \to p_2 \qquad t = (p_1 - p_1')^2 \to 0 \qquad s = (p_1 + p_2)^2 \to (2k_{\text{cms}})^2 \tag{2.27}$$

the differential cross section (2.18) can be formulated as [120]

$$\mathrm{d}\sigma_{\text{ela}} = \frac{|\langle p_1' p_2' | \mathcal{M} | p_1 p_2 \rangle|^2}{64\,\pi\, s\, k_{\text{cms}}} \, \mathrm{d}t = \frac{|A_{\text{ela}}(s,\, t)|^2}{64\,\pi\, s\, k_{\text{cms}}} \, \mathrm{d}t , \tag{2.28}$$

where A_{ela} is the elastic scattering amplitude. From (2.26) one can deduce

$$\langle p_1' p_2' | - \mathrm{i}\mathcal{M} + \mathrm{i}\mathcal{M}^\dagger | p_1 p_2 \rangle \overset{t \to 0}{=} \sum_j \int \mathrm{d}\Phi_j \cdot |\langle p_1 p_2 | \mathcal{M} | p_1 p_2 \rangle|^2 , \tag{2.29}$$

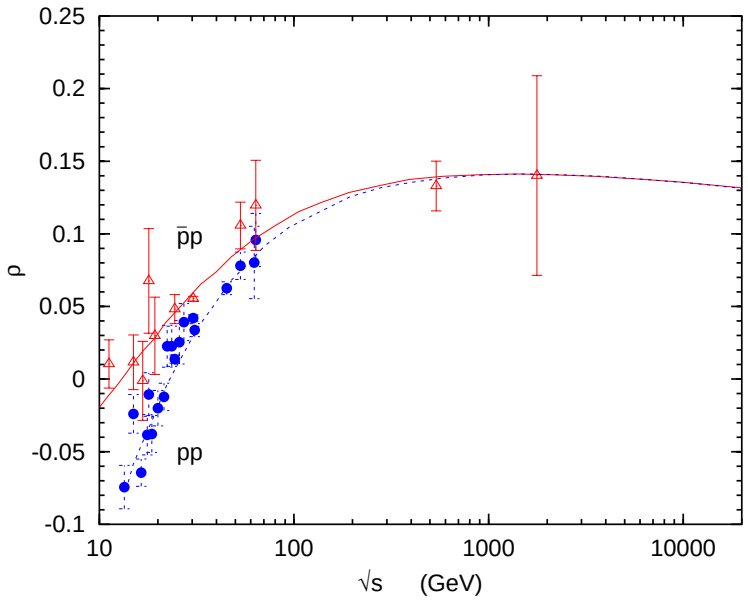


Figure 2.30: Measurements of $\rho(s)$ (see [121] and references therein). The data are shown together with the model by Donnachie and Landshoff [122].

while it is at the same time possible to directly obtain

$$\begin{aligned}\langle p_1' p_2'| - \mathrm{i}\mathcal{M} + \mathrm{i}\mathcal{M}^\dagger |p_1 p_2\rangle \\ &\overset{t\to 0}{=} \mathrm{i}\left(\langle p_1 p_2|\mathcal{M}|p_1 p_2\rangle^* - \langle p_1 p_2|\mathcal{M}|p_1 p_2\rangle\right) \\ &= 2\ \mathrm{Im}\,\langle p_1 p_2|\mathcal{M}|p_1 p_2\rangle = 2\ \mathrm{Im}\, A_{\mathrm{ela}}(s,\, t\to 0)\,. \end{aligned} \tag{2.30}$$

Combining (2.20) with (2.29) and (2.30) leads to the optical theorem

$$\sigma_{\mathrm{tot}} = \frac{1}{2k_{\mathrm{cms}}\sqrt{s}}\ \mathrm{Im}\ A_{\mathrm{ela}}(s,\, t\to 0)\,. \tag{2.31}$$

We find that

$$\sigma_{\mathrm{tot}} \propto \mathrm{Im}\ A_{\mathrm{ela}}(s,\, t\to 0) \qquad \text{and} \qquad \frac{\mathrm{d}\sigma_{\mathrm{ela}}}{\mathrm{d}t} \propto |A_{\mathrm{ela}}(s,\, t\to 0)|^2\,, \tag{2.32}$$

meaning that any process with $\sigma_{\mathrm{tot}} > 0$ necessarily has an elastic contribution with $\sigma_{\mathrm{ela}} > 0$. This is a direct consequence of unitarity. Experiments find a relation of

$$\rho(s) = \frac{\mathrm{Re}\ A_{\mathrm{ela}}(s,\, t\to 0)}{\mathrm{Im}\ A_{\mathrm{ela}}(s,\, t\to 0)} \approx 0.1\,, \tag{2.33}$$

which corresponds to a close to minimal contribution of elastic scattering according to (2.32). Figure 2.30 summarizes the available data. The Fourier transform of the amplitude $A(s,t)$ into the impact parameter space is

$$a(s,\, b) = \frac{1}{2\pi}\int \mathrm{d}^2\vec{\Delta}_\perp\ \ A_{\mathrm{ela}}(s,\, t)\cdot e^{-\vec{\Delta}_\perp\cdot\vec{b}}\,, \tag{2.34}$$

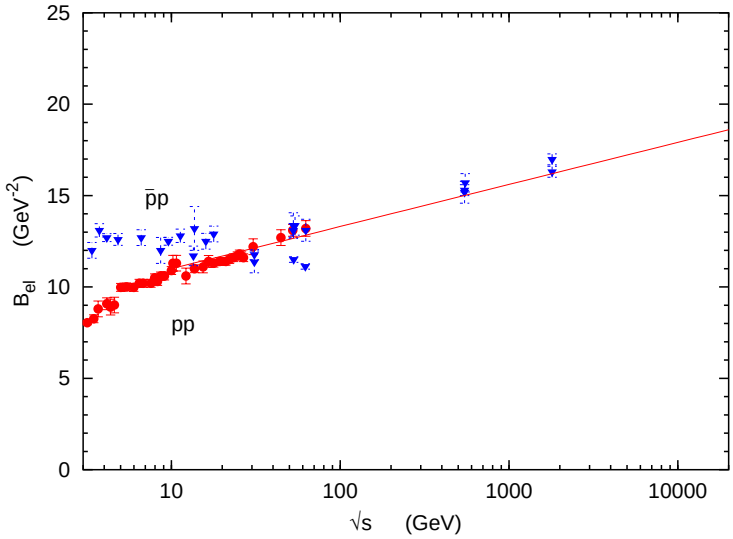


Figure 2.31: Measurements of the elastic scattering slope parameter $B_{\text{ela}}(s)$ (see [121] and references therein). The curve is a fit $B_{\text{ela}} \sim d^1 \log s$, with $d^1 = 0.25\,\text{GeV}^{-2}$.

where $\vec{\Delta}_\perp = \vec{k}_1' - \vec{k}_1$ is the momentum transfer within the plane perpendicular to the projectile motion. In the high energy limit $(\vec{\Delta}_\perp)^2 \approx |t|$. The phenomenological approximation

$$\frac{\mathrm{d}\sigma_{\text{ela}}}{\mathrm{d}t} = \frac{1}{64\,\pi\,s\,k_{\text{cms}}} \cdot |A_{\text{ela}}(s,\,t)|^2 \;\propto\; e^{-B_{\text{ela}}\,|t|} \tag{2.35}$$

is found to describe very well the available experimental data over a wide energy range. Figure 2.31 shows the available measurements of $B_{\text{ela}}(s)$. From (2.35) it follows that

$$A_{\text{ela}}(s,\,t) = A_{\text{ela}}(s,\,t=0) \cdot e^{-\frac{1}{2}B_{\text{ela}}(s)|t|}\,, \tag{2.36}$$

which allows us to determine the Fourier transformation (2.34), resulting into the Gaussian

$$a(s,b) \propto e^{-\frac{\vec{b}^2}{2\,B_{\text{ela}}(s)}}\,. \tag{2.37}$$

This describes the shape of the scattering region within the impact parameter space. It turns out to be circular shaped with a Gaussian radial profile having a width of $R_{\text{Gauss}} \propto \sqrt{B_{\text{ela}}}$. For center collisions (i.e. small impact parameters) the scattering is very efficiently (it is "*dark*"), while at the borders it only scatters weakly. In order to obtain the observed increase of the cross section with energy, several possibilities exist:

(a) The scattering region, and thus $B_{\text{ela}}(s)$, is getting larger.

(b) Scattering becomes more efficient and therefore $A_0(s)$ is increasing $\rightarrow$ the Gaussian shaped scattering region is becoming "*darker*" in the center. There is a natural limit to this, imposed by the conservation of the total probability (which means that the scattering region can never get "*darker*" than "*black*". This is called the *black disc limit*.

(c) For the total cross section, it might be that new final states are appearing at higher energies.

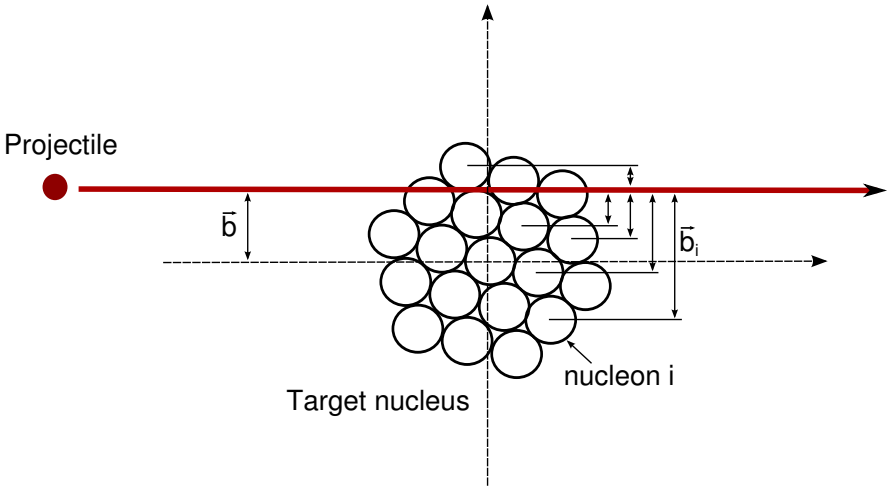


Figure 2.32: A projectile particle hitting a compound target nucleus made of n nucleons with an impact parameter $\vec{b}$. There is a separate relative impact parameter b_i for each of the nucleons i, which governs the scattering of the projectile from the individual nucleon i.

At ultra-high energies it is not expected to have a significant impact from (c). Also (b) will be strongly limited above a certain energy due to the *black disc limit*. The only relevant contribution in the high energy limit can come from possibility (a).

2.7.2 Glauber theory

It was realized by Glauber [119, 123] that the amplitude in the impact parameter space can be treated in analogy to diffraction in optics by using the phase shift relation

$$a(s,\, b) = 1 - e^{\mathrm{i}\chi(s,\, b)}\,, \tag{2.38}$$

with the Eikonal[4] function χ. Within this picture of the diffractive approximation, multiple scattering can be written as

$$\begin{aligned} a_{\mathrm{mult}}(s,\, b) &= 1 - e^{\mathrm{i}\chi_{\mathrm{mult}}(s,\, b)} \\ &= 1 - e^{\mathrm{i}\sum_j \chi_j(s,\, b_j)} = 1 - \prod_j e^{\,\mathrm{i}\chi_j(s,\, b_j)} = 1 - \prod_j \left(1 - a_j(s,\, b_j)\right)\,. \end{aligned} \tag{2.39}$$

It is postulated that the overall phase shift χ_{mult} corresponds to the sum of the individual phase shifts χ_j for each scattering process.

Figure 2.32 illustrates the interaction of a projectile particle with a nucleus build out of A nucleons. According to (2.39) the projectile scatters separately with each of the nucleons j of the nucleus. The resulting total cross section can then be written as

$$\sigma_{\mathrm{tot}} = \int \left[\prod_i \mathrm{d}^3\vec{r}_i\right] \int \mathrm{d}^2\vec{b} \left[\prod_i \Psi_i^*(\vec{r}_i)\right] \left[1 - \prod_i (1 - a(s,\, \vec{b}_i))\right] \left[\prod_i \Psi_i(\vec{r}_i)\right]\,, \tag{2.40}$$

[4]From the Greek word $\epsilon\iota\kappa\omega\nu$ with the meaning of image.

which needs to be integrated over all possible configurations of the nucleus, while $\Psi_i(\vec{r}_i)$ is the wave function of nucleon i and

$$\Psi_i^*(\vec{r}_i)\,\Psi_i(\vec{r}_i) = \rho_i(\vec{r}_i) \tag{2.41}$$

is the density distribution of the nucleon in the nucleus.

Knowing the nucleon-nucleon impact parameter amplitudes $a_i(s,\, b_i)$ one can calculate the multiple scattering amplitude and hence the total and elastic cross sections. From the definition of the amplitude (2.34) and the optical theorem (2.31) it is possible to derive the nucleon-nucleon amplitude. By using (2.33) it is straightforward to get

$$\begin{aligned} A_{\text{ela}}(s,\, t=0) &= \text{Re}A_{\text{ela}}(s,\, t=0) + \text{i}\,\text{Im}A_{\text{ela}}(s,\, t=0) && (2.42)\\ &= (1+\rho)\ \text{Im}A_{\text{ela}}(s,\, t=0)\,. && (2.43) \end{aligned}$$

This expression allows to rewrite (2.36) to obtain

$$\begin{aligned} A_{\text{ela}}(s,\, t) =& (1+\rho)\ \text{Im}A_{\text{ela}}(s,\, t=0)\cdot e^{-\frac{1}{2}B_{\text{ela}}(s)|t|}\\ =& (1+\rho)\,2\,k_{\text{cms}}\sqrt{s}\,\sigma_{\text{tot}}\cdot e^{-\frac{1}{2}B_{\text{ela}}(s)|t|}\,, \end{aligned} \tag{2.44}$$

while the last step corresponds to the application of the optical theorem (2.31). At this point we can insert (2.44) into (2.34) and calculate the nucleon-nucleon amplitude

$$a_i(s,\, b_i) = (1+\rho(s))\,\frac{\sigma_{\text{tot}}^{\text{pp}}(s)}{4\pi B_{\text{ela}}(s)}\,e^{-\frac{1}{2}\vec{b}_i^{\,2}/B_{\text{ela}}(s)}\,. \tag{2.45}$$

It is the combination of the expressions (2.40) and (2.45) that finally makes the connection between nucleon-nucleon and nucleon-nucleus scattering. With these formulations is possible to calculate the total cross section for nucleon-nucleus scattering just from the parameters $B_{\text{ela}}(s)$, $\rho(s)$, $\sigma_{\text{tot}}^{\text{pp}}$ and the possible configurations of the nucleus.

Figures 2.30 and 2.31 exemplify that due to the sparse data there are large uncertainties in the extrapolation of B_{ela} and ρ to cosmic ray energies. Since there are no theoretically well-founded functional forms of the used extrapolations this gives rise to large uncertainties for converting experimentally measured cosmic ray based proton-air cross sections into fundamental proton-proton cross sections.

Also the ultra-high energy hadronic interaction models have the Glauber formalism implemented in order to predict the hadron-air cross sections at cosmic ray energies. Differences in the application of the Glauber theory are mainly responsible for the non-vanishing offset between the predictions of the proton-air cross section of different interaction models even at lower energies (confer Figure 2.20).

3 Pierre Auger Observatory

In this section a brief overview of the Pierre Auger Observatory is given. The software simulation and reconstruction framework Offline is described, focusing on the extensions made for the purpose of this thesis. The relevant aspects of hybrid simulation and reconstruction are introduced. It was necessary to considerably improve the hybrid detector simulations in order to study detector properties like the resolution or the acceptance.

3.1 Experimental setup

Located in the vicinity of the small town of Malargüe in Argentina the Auger Collaboration is operating the presently largest air shower detector of the world (Figure 3.1) [13]. But it is not just size that makes it exceptional. Particularly the quality of data and analysis is expected to be much better compared to previous experiments. One of the most important features of the Pierre Auger Observatory is its hybrid-detector setup. There is the huge *surface detector* (SD) array [125], which is overlooked by the *fluorescence detector* (FD) telescopes [126]. Especially the redundancy of the hybrid-detector setup adds the possibility of cross checks as well as combined reconstruction with a very high reconstruction quality. This advantage is heavily exploited by the Auger Collaboration.

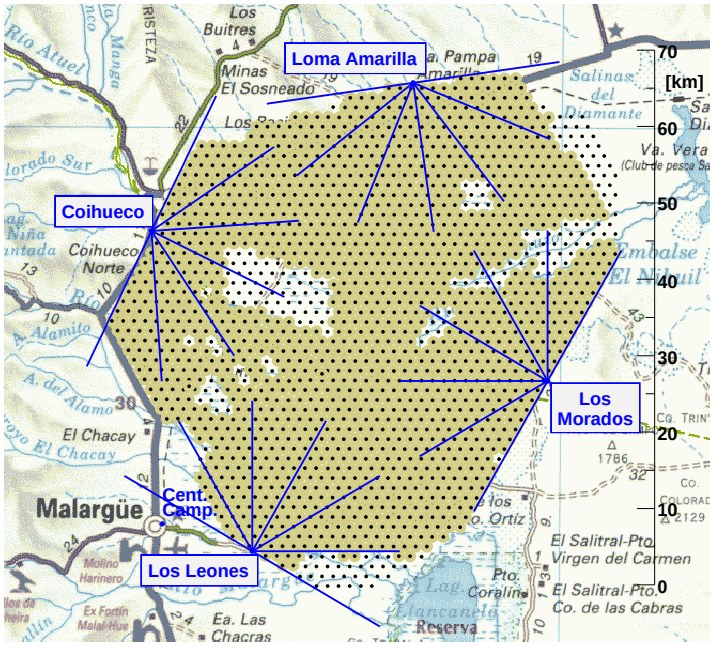


Figure 3.1: Status of the Pierre Auger Observatory as of October 2007 [124]. The four telescope buildings at Los Leones (LL), Los Morados (LM), Loma Amarilla (LA) and Coihueco (CO) host 6 telescopes each. The surface detector is close to completion (shaded area).

3.1.1 Surface detector

With a spacing of 1.5 km, the water Cherenkov tanks of the surface array are arranged on a hexagonal grid. An area of 3000 km^2 will be covered by 1600 of these tanks. Each tank is operating completely autonomously. Power is provided by solar panels with a battery buffer. All communications are done via WLAN. A set of communication towers provides the link to the *central data acquisition system* (CDAS), located in the town of Malargüe. All tanks are equipped with three photomultipliers (PMTs) looking from the top into 12 tons of water for Cherenkov radiation. Signals from all three PMTs are digitized with 40 MHz sampling rate. Data are taken continuously at low threshold to provide a high statistics integrated signal spectrum, which is used for the online calibration. The recorded ADC traces of each SD station are calibrated in relative units of the *vertical equivalent muon* (VEM) peak from this signal spectrum [127]. If the signal in a single ADC trace bin gets above the T1 threshold (1.8 VEM), the present ADC data are hold in a local memory for a short time to be later transferred to CDAS in case of a higher trigger request. If the signal reaches the T2 threshold (3.2 VEM), a T2 notice is sent to CDAS. A parallel trigger, working at a much lower threshold of 0.2 VEM, requires the signal to reach above the threshold for a certain duration (300 ns). These *Time-Over-Threshold* (ToT) triggers are also sent to CDAS. At CDAS, the central trigger algorithm collects all the incoming T2s and searches for clustering in time and space. If a possible coincidence is detected, all tanks of the array are asked to transfer their locally hold data [128].

3.1.2 Telescope detector

Each of the four telescope detector buildings (in the following called *eyes*) hosts 6 individual telescopes. The telescopes are built according to the Schmidt optics, consisting of a spherical mirror of radius 3.4 m, the aperture opening (radius 1.1 m) with a ring-shaped corrector lens on its outer 25 cm and a spherical pixel camera (radius 1.7 m). The center of the spherical mirror and camera are both placed at the center of the aperture, resulting in a fully concentric setup. Only the massive steel support of the camera, which is mounted on the floor between the mirror and the aperture, adds a small asymmetry.

All telescopes are taking data independently. Four layers of trigger algorithms have to be passed in order to qualify an event for readout [130]. The *first level trigger* (FLT) works on the level of individual pixels and requires the sum over 10 consecutive time bins of the ADC trace to be above the FLT threshold. The threshold is dynamically adjusted to produce a 100 Hz FLT trigger rate per pixel. A fast hardware-implemented *second level trigger* (SLT) collects all FLTs from one camera and searches for geometric patterns. The following *third level trigger* (TLT) is a software trigger still operating on telescope level. It also checks the time structure of an event. After the TLT all data from the telescopes of one eye are collected by the *eyePC*. On the eyePC, the combined mirror events have to pass the *eye level trigger* (T3), which performs a rudimentary event reconstruction to calculate the direction and time of impact on the ground. The information about the shower impact point is sent to CDAS in order to trigger data readout of the corresponding parts of the SD array.

Data from the telescopes are calibrated using the end-to-end calibration technique [131]. A uniform Lambertian light emitter (drum) is mounted directly in front of the aperture of

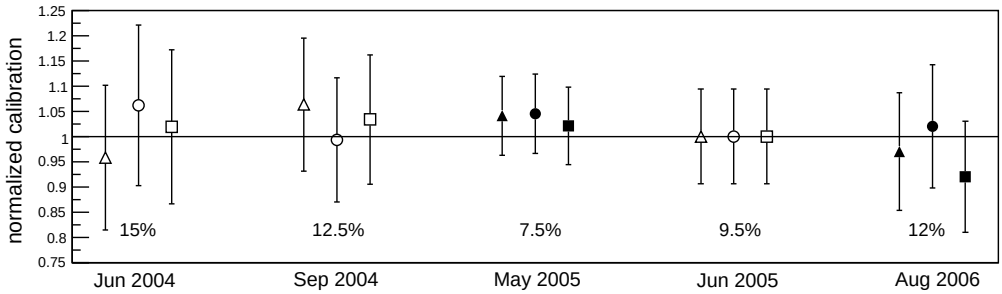


Figure 3.2: Absolute drum calibrations (open shapes) and the laser cross checks (solid shapes) for the past three years [129]. The different markers represent three different telescopes. The same telescopes are shown for each of the calibration campaigns. All points have been normalized to the June 2005 drum calibration by using the daily relative calibration data. The systematic uncertainty for each measurement campaign is stated below each group of points.

a telescope and the response of each pixel is measured. The calibration constant is then calculated as the ratio of the known number of photons entering the diaphragm, $n_{\mathrm{ph,dia}}^{\mathrm{drum}}$, and the total signal s_{PMT_i} recorded in pixel i

$$k_i^{\mathrm{cal}} = n_{\mathrm{ph,dia}}^{\mathrm{drum}} \;/\; s_{\mathrm{PMT}_i}. \tag{3.1}$$

Any signal recorded during shower observation is multiplied with this calibration constant to yield the number of photons entering the diaphragm without needing to know the details of the telescope optics or electronics. Figure 3.2 demonstrates the stability of the absolute drum calibration for the Auger telescope detectors.

3.1.3 Atmospheric monitoring

Since EAS are initiated by primary cosmic ray particles high above ground they propagate through the entire atmosphere. Thus it is essential to monitor the state of the atmosphere permanently. The total integrated amount of atmospheric matter as well as the detailed density profile have a noticeable effect on any air shower measurement [132, 133, 134]. For optical observations by telescopes it is necessary to constantly monitor scattering and attenuation of ultraviolet photons in the atmosphere in addition. Significant efforts are made to meet these requirements.

Radio sondes are launched frequently to measure vertical *temperature* and *density* profiles above Malargüe [135].

Cloud cameras installed on top of each telescope building are performing a full sky infrared imaging to detect *clouds*.

LIDAR (light detection and ranging) stations are placed next to all four FD buildings, to scan the sky for *clouds* and measure *aerosol backscattering* with high repetition lasers [136].

CLF and XLF (central laser facility and extra laser facility) facilities permanently generate a well known sequence of laser events in all telescope detectors to monitor *geometric alignment*, *timing* and *aerosol* scattering in the atmosphere [137].

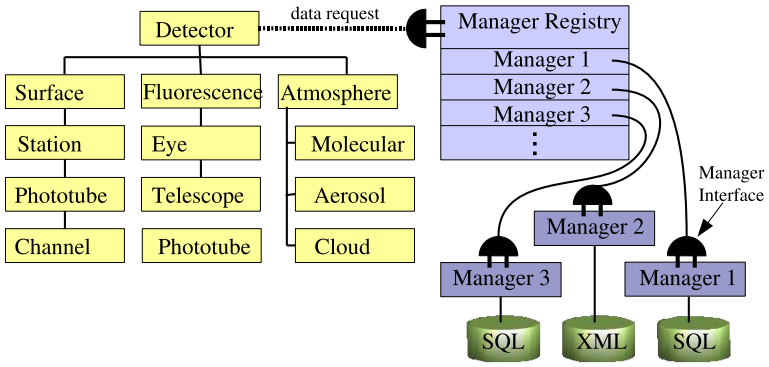


Figure 3.3: Time dependent detector description of the Offline framework.

HAM (horizontal attenuation measurement) lasers are operating below the FD field of view and shoot horizontally from one FD building to another to measure near ground *horizontal attenuation*.

APF (aerosol phase function) lasers shoot horizontally and in a short distance in front of the FD detectors to observe scattered laser light over a large scattering angle range. This is used to measure the Mie scattering *phase function* [138].

FRAM (fotometric robotic telescope for atmospheric monitoring) robot telescopes observe a large number of known bright ultraviolet stars in order to detect *clouds* and *attenuation* of light in the atmosphere [139].

The monitoring data are processed and stored in `SQL` databases to make them easily accessible during reconstruction and simulation tasks.

3.2 Offline software framework

Most of the offline computing tasks for the Pierre Auger Collaboration are performed in the Offline software framework [140], which is developed within the collaboration. The framework is especially designed to provide all functionality for processing data from a giant air shower detector.

An important feature of the Offline framework is the internal time dependent detector description (cf. Figure 3.3). All parts of the detector can be asked for their detailed status and configuration for all times the experiment was taking data. This includes configuration and size of the surface array, the status of each individual surface station, detailed configuration of the FD detector, uptime information for all relevant parts of the detector and last but not least all kind of atmospheric monitoring data. Even periods in time with mis-configured FD camera cabling are properly taken into account. The latter was introduced as a measure to improve the hybrid data quality during this thesis [141]. Figure 3.4 demonstrates how the mis-cabling of a telescope affects shower data. After mis-cabling is identified in any of the 11520 electronic channels of the FD detector, it can be easily taken into account during offline event reconstruction.

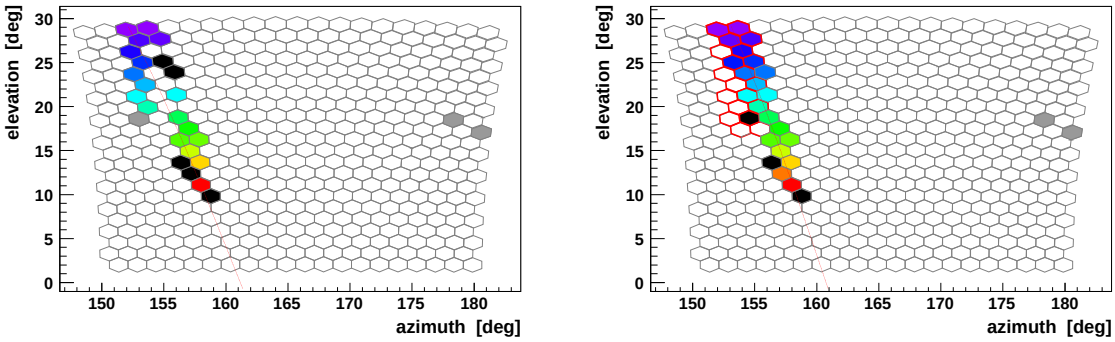


Figure 3.4: The `FdChannelMappingManager` in the Offline framework allows to correct air shower data with misconfigured camera electronics (left) to properly reconstruct the event (right). The corrected channels are surrounded in red. Pixels filled in black are rejected during reconstruction. The color code of the pixels indicates the timing (from blue to red). For the given event the reconstructed $X_{\max}$ changes from 242 to 600 $\mathrm{g\,cm^{-2}}$ [141].

The data needed to retrieve all this information are stored in `XML` files, if they are static and in `SQL` databases if they vary. The Offline framework provides an easy-to-use interface, completely hiding the detailed internal mechanisms of how and where the data are taken from, as illustrated in Figure 3.3.

The dimension of the experiment and the observed air shower events are too large to neglect the curvature of the Earth surface. The geometry package of the Offline framework allows a realistic representation of the curvature of the Earth. Cartesian coordinate systems can be defined in any place of the observatory, always considering proper tangential alignment with respect to the Earth ellipsoid (`WGS84`[5]). This feature of Offline greatly improves and simplifies all handling of Auger data.

3.2.1 Slant density profiles in a curved atmosphere

Obviously not only the Earth surface is curved but also the atmosphere with its differential density profile itself (see Figure 3.5). Since air showers develop via interactions with matter of the atmosphere, the total traversed matter thickness (slant depth) is primarily responsible for the number of interactions and hence for the shape of the resulting air shower profiles. In order to account for this, the functionality for numerical integration of atmospheric slant depth profiles was implemented in the Offline framework as part of this thesis [142].

Vertical density profiles of the atmosphere versus height above the Auger site near Malargüe, $\rho(h)$, are measured frequently using weather balloons. Combining the data of many balloon flights, it is possible to extract a set of monthly mean atmospheric density profiles, which can be used in the Offline framework. Detailed knowledge about the density of the atmosphere is paramount for analysis of air shower data [132, 133, 143]. To calculate the integrated matter thickness which is traversed for the trajectory from point $\vec{p}_1$ to $\vec{p}_2$ inside the curved atmosphere, it is necessary to integrate

[5] World Geodetic System, 1984.

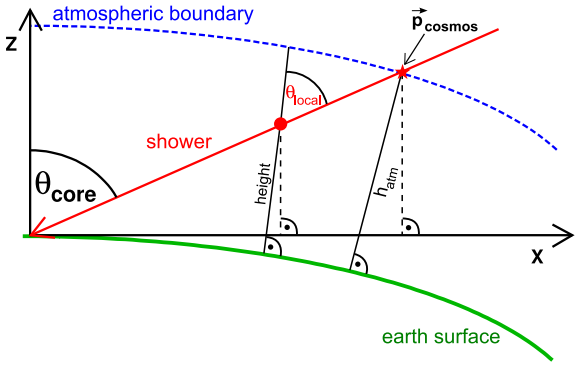


Figure 3.5: Difference between a flat and a curved geometry of the atmosphere. In the curved atmosphere, all heights and angles are calculated using verticals with respects to the elliptical Earth surface. Thus for all points along the shower axis, θ_{local} differs from θ_{core}.

$$\Delta X_{\text{slant}} = \int_{\vec{p}_1}^{\vec{p}_2} \mathrm{d}\vec{r}\, \rho(\vec{r}) \; . \tag{3.2}$$

Usually this is calculated in the approximation of a flat geometry as

$$\Delta X_{\text{slant}} = \Delta X_{\text{vert}} / \cos\theta_{\text{core}} \, , \tag{3.3}$$

where

$$\Delta X_{\text{vert}} = X_{\text{vert}}(h_2) - X_{\text{vert}}(h_1) = \int_{h_1}^{h_2} \mathrm{d}h\, \rho(h) \tag{3.4}$$

is the vertical matter thickness difference and θ_{core} the zenith angle of the straight trajectory. However, in Offline the integration of (3.2) is performed numerically with high precision by the `InclinedAtmosphereModel`. Figure 3.6 illustrates the importance of this numerical slant profile integration for inclined geometries. Properly considering the curved atmosphere makes a significant difference for reconstructed quantities like X_{max}, especially at large zenith angles (see Figure 3.7).

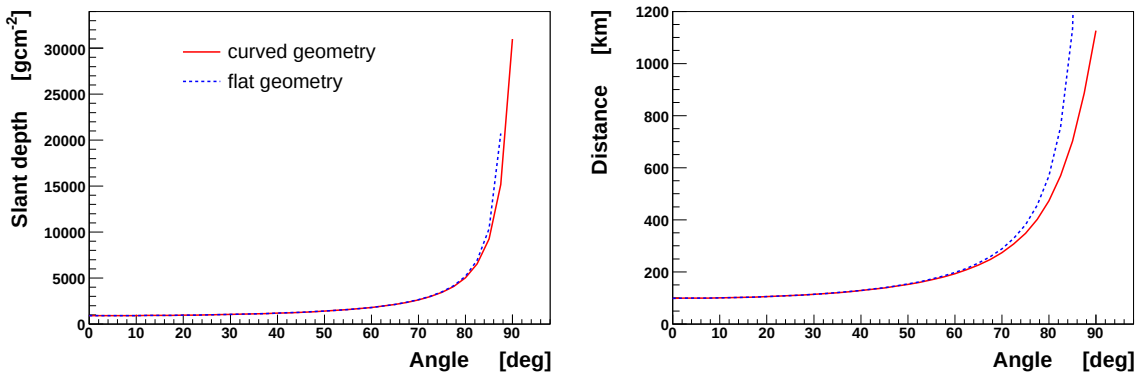
3.3 Hybrid event reconstruction

The hybrid reconstruction is based on FD data with additional timing information from SD, to improve the geometry fit. All algorithms are implemented within the Offline framework. In the following, all reconstruction steps are described for the example of the event shown in Figure 3.8.

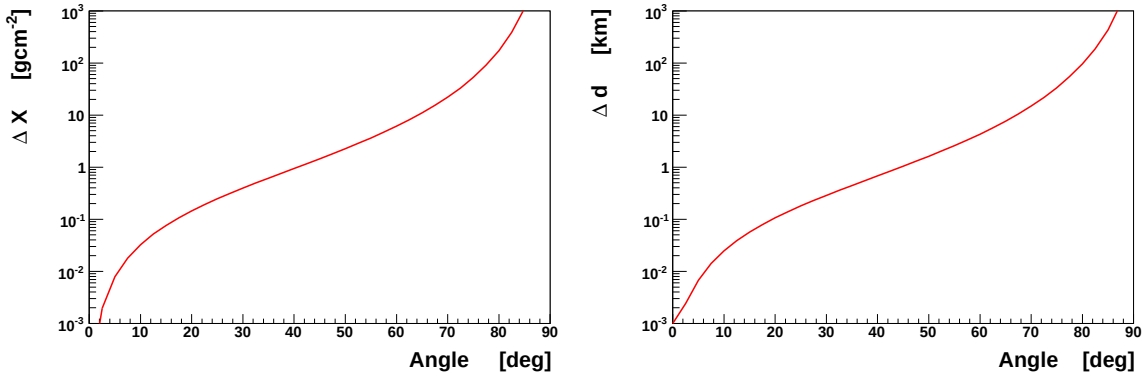
3.3.1 Pulse reconstruction

Each triggered FD pixel is searched for a clear shower signal in terms of a pulse with a required minimum signal to noise ratio. The minimum default ratio is set to 5. The signal to noise ratio for a pulse from bin k_{start} to k_{end} is

$$\frac{S}{N} = \frac{q_i}{(k_{\text{end}} - k_{\text{start}}) \cdot RMS_i} \, , \tag{3.5}$$



(a) Total slant depth (left) and geometric thickness of the atmosphere (right)



(b) Difference between flat and curved geometry for slant depth (left) and geometric thickness (right)

Figure 3.6: Comparison of the total slant depth integrated from the shower core along the shower axis and geometric distance from the core along the shower axis to the boundary of the atmosphere for curved and flat geometry. In the case of flat geometry there is a singularity at 90°. Even in total numbers a real divergence for large angles can be seen. This is better observed in the lower plots, where the difference between flat and curved model is plotted. Above 60° the difference in slant depth is becoming significant (>10 gcm^{-2}) reaching 100 gcm^{-2} around 80°.

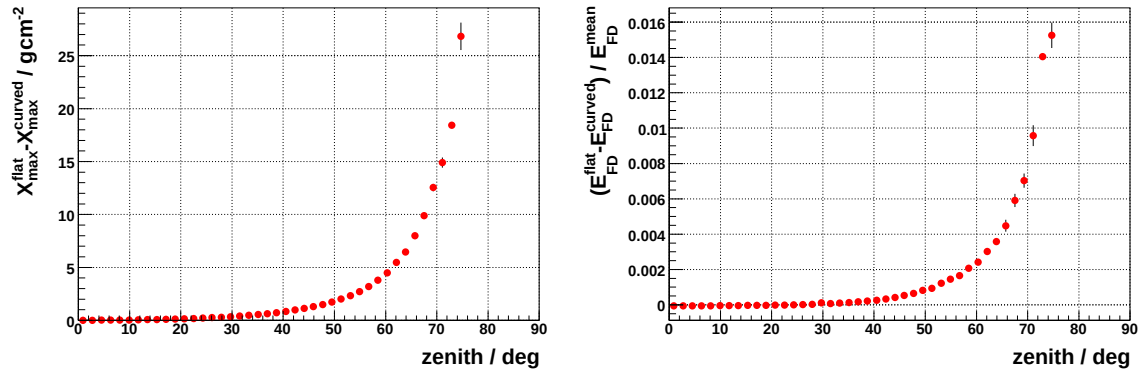


Figure 3.7: Impact of a curved atmosphere integration on the reconstructed shower properties shower maximum X_{max} (left) and total energy (right).

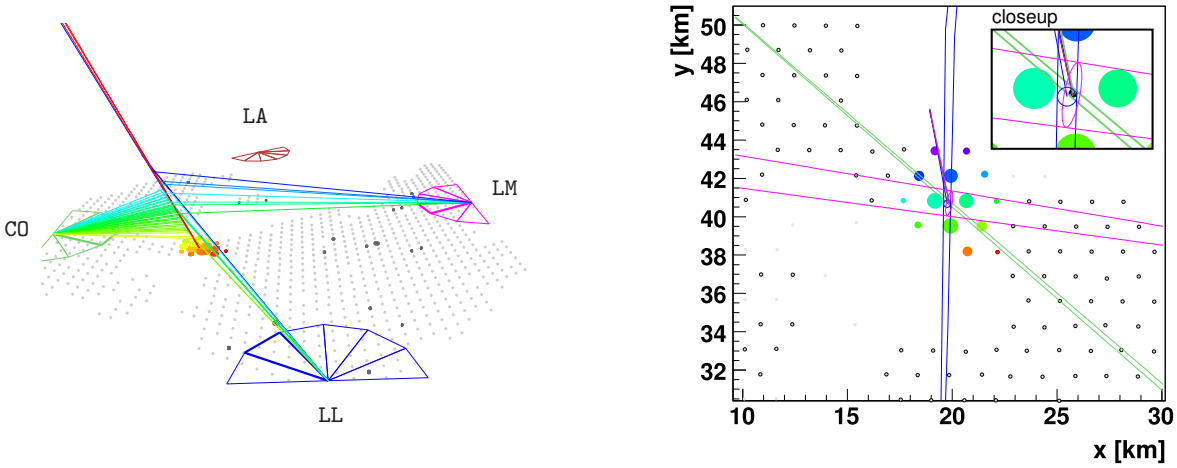


Figure 3.8: Auger event (`augerId`: 200608702501) observed by three telescope detectors and the surface array. The three dimensional picture on the left side shows timing along the shower axis (blue to red) of all detectors. On the right side the surface event is shown including timing (colors) as well as signal (size) of the stations and the projected shower detector planes (SDP) on ground including their uncertainties for all eyes (eye-color). The reconstructed shower core position including errors for all individual eyes (eye-colored ellipses within the SDP slices) and SD (black marker, at the intersection of all SDPs) and the reconstructed shower directions projected on the ground (colored lines starting from the core) are displayed. The independent fourfold reconstruction of this event (3×FD + SD) shows remarkable agreement. In the following part of this Section all steps of event reconstructions are pictured in detail on the example of this triple event.

where RMS_i is the baseline RMS of the trace, with the noise variance $V^i_{\text{noise}} = (RMS_i)^2$. The integrated signal over the bin range is

$$q_i = \sum_{k=k_{\text{start}}}^{k_{\text{end}}} s^i_k \,, \tag{3.6}$$

where s^i_k is the ADC count of pixel i at time bin k. The pulse timing is the centroid time ("signal weighted time") of all bins belonging to the pulse

$$t_i = \frac{1}{q_i} \sum_{k=k_{\text{start}}}^{k_{\text{end}}} t^i_k \, s^i_k \,. \tag{3.7}$$

To calculate the uncertainty of t_i, equation (3.6) is inserted into (3.7) and the uncertainties $\sigma_{t^i_k} = \sqrt{12}$ for the uniformly distributed timing of bin k as well as the Poissonian fluctuations of photoelectrons (PE) at the photocathode of the PMT i are propagated

$$\sigma_{s^i_k} = \sqrt{V^i_{\text{noise}} + s^i_k \, g} \,, \tag{3.8}$$

where $g = 7.9$ PE/photon is the factor to convert ADC counts into PEs. The two partial differentials of t_i are

$$\frac{\partial t_i}{\partial s^i_k} = \frac{1}{q_i} t^i_k - \frac{1}{q^2_i} \sum_{j=k_{\text{start}}}^{k_{\text{end}}} s^i_j \, t^i_j \qquad \text{and} \qquad \frac{\partial t_i}{\partial t^i_k} = \frac{1}{q_i} s^i_k \,, \tag{3.9}$$

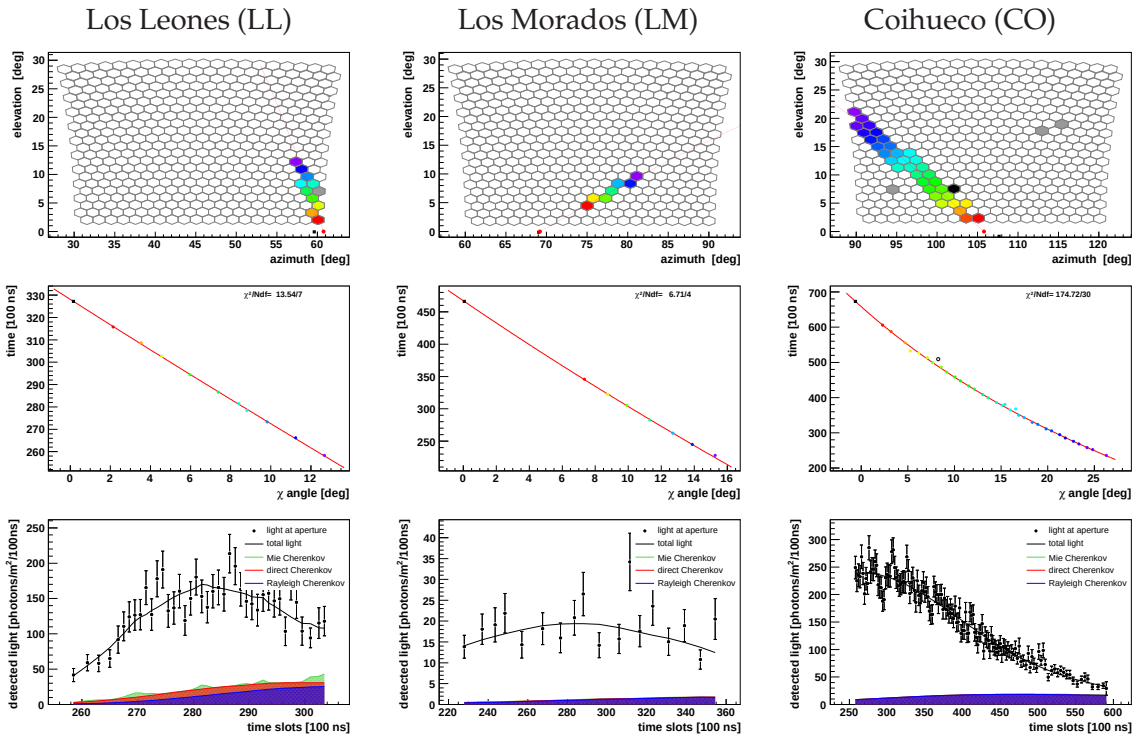


Figure 3.9: The triple FD event from Figure 3.8 as seen and reconstructed by the individual eyes. The upper pictures show the pixel camera view, including reconstructed timing of the pulses, the very few noise pixels (light grey), pulsed pixels rejected during the fits (black) and the resulting reconstructed SDP (red line). The plots in the middle show the hybrid time fits, including the SD stations (black squares) and all the pulsed pixels (colored circles), where the pixel colors are the same as in the plot above. The lower plots show the collected light at the aperture, including the reconstructed components of the arriving photon flux.

resulting in an uncertainty of t_i as

$$\begin{aligned}\sigma^2(t_i) &= \sum_{k=k_{\text{start}}}^{k_{\text{end}}} \left(\left(\frac{\partial t_i}{\partial s_k^i} \right)^2 \sigma_{s_k^i}^2 + \left(\frac{\partial t_i}{\partial t_k^i} \right)^2 \sigma_{t_k^i}^2 \right) \\ &= \sum_{k=k_{\text{start}}}^{k_{\text{end}}} \left(\left(t_k^i \, q_i - \sum_{j=k_{\text{start}}}^{k_{\text{end}}} s_j^i \, t_j^i \right)^2 q_i^{-4} \cdot \left(V_{\text{noise}} + s_k^i \, g \right) + \left(\frac{s_k^i}{q_i \sqrt{12}} \right)^2 \right) . \end{aligned} \tag{3.10}$$

3.3.2 Shower detector plane

The *shower detector plane* (SDP) is the plane comprising the shower axis and the telescope detector. Since the telescope detector is by definition part of the plane, the SDP can be displayed as a line on the camera. It can be reconstructed from the data of a telescope by minimizing the sum

$$\sum_i q_i \left(\vec{p}_i \cdot \vec{n}_\perp^{\,\text{SDP}} \right)^2 / \sigma_{\text{SDP}}^2 \tag{3.11}$$

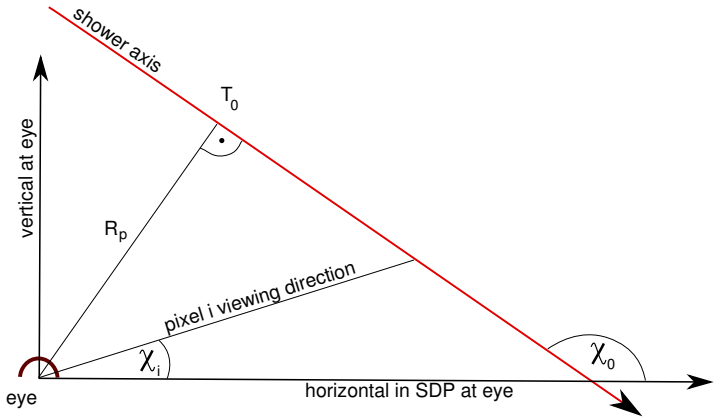


Figure 3.10: Geometry of an air shower within the shower detector plane. The geometry can be determined by the three parameters R_p (distance of closest approach), χ_0 (angle of the shower axis in the SDP) and T_0 (time of the shower axis at the point of the closest approach).

over all pulses i, with the two free parameters θ_{SDP} and ϕ_{SDP} to define the plane normal vector $\vec{n}_{\perp}^{\,\mathrm{SDP}}$ in spherical coordinates and the pixel pointing direction $\vec{p}_i$. The pointing uncertainty for the SDP fit, σ_{SDP}, was determined to be 0.35° by studying SDP fits of CLF laser shots with a known geometry. The top row of Figure 3.9 shows pixels on the FD camera with detected pulses indicated by their color-coded time t_i, together with the reconstructed SDP (red-line).

3.3.3 Hybrid time fit

From the perspective of a telescope, the projection of a shower on the camera evolves along the SDP. Each pulsed pixel on the camera can be associated with an angle χ_i within the plane of the SDP with respect to the horizontal axis at the camera (see Figure 3.10). The angular movement of the shower within the SDP in this representation is

$$t(\chi_i) = T_0 + \frac{R_p}{c} \tan\left(\frac{\chi_0 - \chi_i}{2}\right) . \tag{3.12}$$

To determine the three free parameters T_0, R_p and χ_0 the minimum of the function

$$\chi^2 = \sum_i \frac{(t_i - t(\chi_i))^2}{\sigma(t_i)^2} + \frac{(t_{\mathrm{SD}} - t(\chi_{\mathrm{SD}}))^2}{\sigma(t_{\mathrm{SD}})^2} \tag{3.13}$$

has to be found. The sum runs over all pulsed pixels i with the centroid pulse time t_i and the associated uncertainty $\sigma(t_i)$, adding the additional SD station time t_{SD} with the uncertainty $\sigma(t_{\mathrm{SD}})$, where χ_{SD} is the angle between the station position projection into the SDP and the local vertical. The middle row of Figure 3.9 displays the time fit (red line) that describes well the angular movement of the shower along the SDP. The improvement by using the single tank in the χ^2 of the hybrid time fit is displayed in Figure 3.11. While the mono cores are spread out along the line of the shower detector plane (SDP), the hybrid cores are well contained in a much smaller region with greatest extent in the direction perpendicular to the

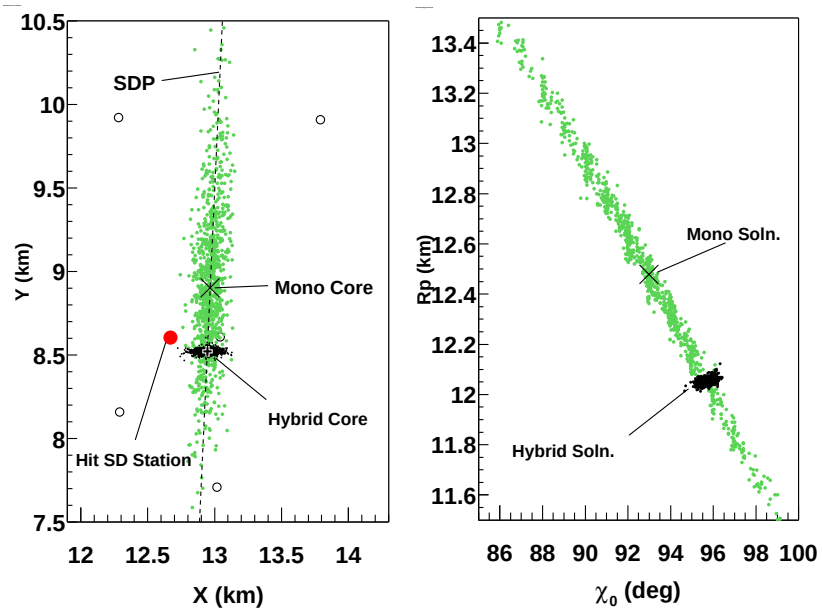


Figure 3.11: Bootstrap solutions for FD-only mono and hybrid fitting procedures for a selected hybrid event [144]. A bootstrapping [145] was performed by resampling the original data 1000 times. Left panel: core location solutions for mono and hybrid fits. Right panel: the corresponding (R_p, χ_0) solutions, which are the intrinsic geometry parameters of the hybrid geometry reconstruction.

line of the SDP. In mono reconstruction there is a strong correlation in the determination of R_p and χ_0.

3.3.4 Light collection

With the reconstructed geometry of the shower axis, photons can be collected in time bins using a circular collection area around the shower position at time t_{dia}. The radius ζ of the collection area is defined to be the angle between a pixel viewing direction and the position of the shower on the camera. It is chosen to give the largest possible signal to noise ratio. The collected light flux arriving at the detector in time bin i is

$$F_i = \frac{1}{A_{\text{dia}} N_{\text{pix}}} \sum_{n=1}^{N_{\text{pix}}} s_i^n \,, \tag{3.14}$$

where the sum runs over all pixels N_{pix} compatible with the ζ selection at time bin i and A_{dia} is the area of the diaphragm opening. Correspondingly the collected noise at time i is

$$N_i = \frac{1}{A_{\text{dia}} N_{\text{pix}}} \sqrt{\sum_{n=1}^{N_{\text{pix}}} \left(RMS_i^2 + q_i\, g\, (1 + V_g) \right)} \,, \tag{3.15}$$

where $V_g = 0.4$ is the gain variance [146]. In a simple iterative algorithm ζ is adopted to yield the maximal signal-to-noise

$$\frac{S}{N} = \frac{\sum_i F_i}{\sum_i N_i^2} \tag{3.16}$$

where the sum runs over all time bins i. The uncertainty of s_i^n is given in (3.8) and can be propagated to (3.14)

$$\sigma_{F_i} = \sqrt{\frac{1}{A_{\mathrm{dia}} N_{\mathrm{pix}}} \sum_n \left(\sigma_{s_i^n}\right)^2}. \tag{3.17}$$

The bottom row of Figure 3.9 shows the evolution of F_i (in units of photons per m^2 and per 100 ns time bin) versus time.

3.3.5 Profile reconstruction

Combining the reconstructed geometry, light flux at the diaphragm and the monitoring data of the atmosphere, it can be calculated how much light was lost due to absorption in the atmosphere from the shower axis to the diaphragm. At this point light production at the shower axis is known and can be attributed to fluorescence and Cherenkov light emission of the air shower [147, 148, 149]. It is a new feature of the Auger profile reconstruction to use *energy deposit* $\frac{\mathrm{d}E}{\mathrm{d}X}$ and not the *shower size* N^{e} as fundamental observable. The previously used "shower size" is lacking a strict definition, since it simply reflects the number of particles passing through a specific geometric layer. This number is strongly depending on the chosen lower energy cutoff. Also the angle of particles cannot be properly accounted for [150]. On the other hand, energy deposit is a well-defined quantity that can be calculated in Monte Carlo simulations and does not depend on the simulation threshold [151]. Moreover, energy deposit is the property of an air shower that can be most straightforwardly linked to fluorescence light production by

$$N_\gamma^{\mathrm{f}}(X_i) = Y_i^{\mathrm{f}}\, w_i\, \Delta X_i\,, \tag{3.18}$$

with the fluorescence yield Y_i^{f} [152, 153] at the slant depth X_i and the energy deposit $\omega_i = \mathrm{d}E/\mathrm{d}X$. Fluorescence emission is an isotropic process, whereas the photons from equation (3.18) are equally distributed on the sphere $4\pi r_i^2$ with r_i being the distance from X_i to the detector. Due to absorption of photons in the atmosphere T_i and the efficiency of photon detection ϵ the observed photon flux at the detector can be written as

$$y_i^{\mathrm{f}} = d_i\, Y_i^{\mathrm{f}}\, w_i\, \Delta X_i\,, \tag{3.19}$$

where the abbreviation $d_i = T_i \epsilon / 4\pi r_i^2$ is introduced.

The number of Cherenkov photons emitted at the shower axis is proportional to the number of electrons and positrons above the Cherenkov threshold energy. Other charged particles are completely outnumbered by electrons and positrons and do not contribute noteworthy to the light emission. Therefore the number of Cherenkov photons is

$$N_\gamma^{\mathrm{C}}(X_i) = Y^{\mathrm{C}}\, N_i^{\mathrm{e}}\, \Delta X_i, \tag{3.20}$$

where N_i^{e} denotes the number of electrons and positrons. Details of the Cherenkov light production are included in the Cherenkov yield factor Y_i^{C} [154, 155, 156, 80]. Although, Cherenkov photons are emitted in a narrow cone along the particle direction, they cover a considerable angular range with respect to the shower axis, because the charged particles are deflected from the primary particle direction due to multiple scattering. Given the fraction

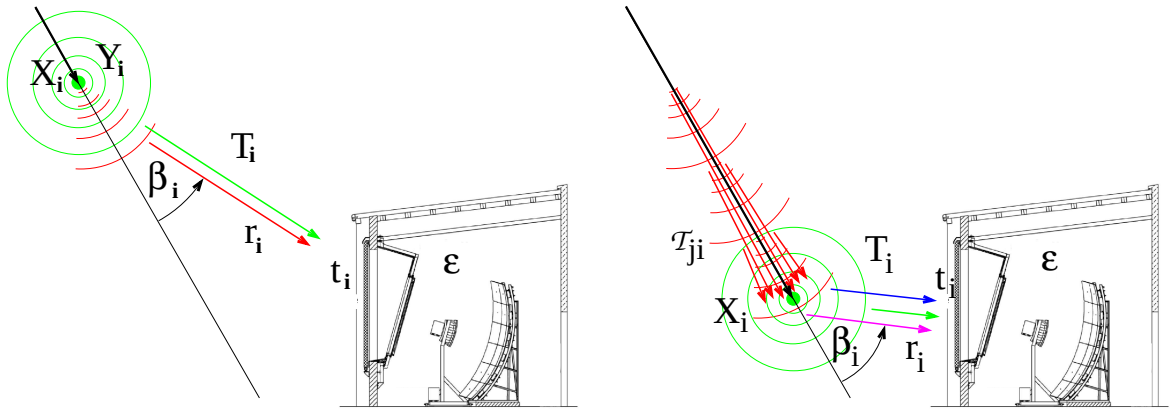


Figure 3.12: Shower light emission and detection at various shower stages. Left panel: Direct light is detected from isotropic fluorescence (green) and forward beamed Cherenkov (red) emission. Right panel: Scattered light (blue, magenta) out of the bright Cherenkov beam is detected simultaneously with fluorescence light [147, 149].

$f_C(\beta_i)$ of Cherenkov photons emitted at an angle β_i with respect to the shower axis [155, 80], the light flux at the detector aperture originating from direct Cherenkov light is

$$y_i^{\mathrm{Cd}} = d_i \, f_C(\beta_i) \, Y_i^{\mathrm{C}} \, \Delta X_i \, N_i^{\mathrm{e}}. \tag{3.21}$$

Due to the forward peaked nature of Cherenkov light production, an intense Cherenkov light beam builds up along the shower as it traverses the atmosphere (Figure 3.12). If a fraction $f_s(\beta_i)$ of the beam is scattered towards the observer it can contribute significantly to the total light received at the detector. In a simple one-dimensional model the number of photons in the beam at depth X_i is just the sum of Cherenkov light produced at all previous depths X_j attenuated on the way from X_j to X_i by $\mathcal{T}_{ji}$

$$N_\gamma^{\mathrm{beam}}(X_i) = \sum_{j=0}^{i} \mathcal{T}_{ji} \, Y_j^{\mathrm{C}} \, \Delta X_j \, N_j^{\mathrm{e}}. \tag{3.22}$$

In analogy to (3.21) the scattered Cherenkov light received at the detector is then

$$y_i^{\mathrm{Cs}} = d_i \, f_s(\beta_i) \sum_{j=0}^{i} \mathcal{T}_{ji} \, Y_j^{\mathrm{C}} \, \Delta X_j \, N_j^{\mathrm{e}}. \tag{3.23}$$

Finally, the total light received at the detector at the time t_i is obtained by adding the scattered and direct light contributions

$$y_i = y_i^{\mathrm{f}} + y_i^{\mathrm{Cd}} + y_i^{\mathrm{Cs}}. \tag{3.24}$$

It was realized that due to the universality of the energy spectra of the secondary electrons and positrons within an air shower, there exists a non-iterative solution for the reconstruction of a longitudinal shower profile from light detected by telescope detectors, while both fluorescence and Cherenkov light contributions are treated as signal [154]. The total

energy deposit of the air shower is just the sum of the energy loss of individual electrons

$$w_i = N^{\mathrm{e}}{}_i \int_0^\infty \mathrm{d}E \; f_{\mathrm{e}}(E, X_i) \; w_{\mathrm{e}}(E) = N_i^{\mathrm{e}} \; \alpha_i \,, \tag{3.25}$$

where $f_{\mathrm{e}}(E, X_i)$ denotes the normalized electron energy distribution, $w_{\mathrm{e}}(E)$ the energy loss per unit depth of a single electron with energy E and α_i is the average energy deposit per unit depth per electron at shower age $s_i = 3/(1 + 2X_{\max}/X_i)$. This is possible since the electron energy spectrum $f_{\mathrm{e}}(E, X_i)$ is universal in shower age s_i [80, 154, 156]. The relation (3.25) between N_i^{e} and w_i can be utilized to relate the total light at the detector (3.24) arriving at time t_i exclusively to the energy deposit w_i at slant depth X_i. For the solution of the problem, it is convenient to rewrite the relation between energy deposit and light at the detector in a matrix notation

$$\mathbf{y} = \mathbf{C}\mathbf{w}, \tag{3.26}$$

where

$$\mathbf{y} = (y_1, y_2, \ldots, y_n)^{\mathrm{T}} \tag{3.27}$$

is the n-component vector (histogram) of the measured photon flux at the aperture and

$$\mathbf{w} = (w_1, w_2, \ldots, w_n)^{\mathrm{T}} \tag{3.28}$$

is the energy deposit vector at the shower track. The elements of the *Cherenkov-fluorescence matrix* $\mathbf{C}$ can be found by a comparison with the coefficients in equations (3.19), (3.21) and (3.23)

$$C_{ij} = \begin{cases} 0, & i < j \\ c_i^{\mathrm{d}} + c_{ii}^{\mathrm{s}}, & i = j \\ c_{ij}^{\mathrm{s}}, & i > j \end{cases} \,, \tag{3.29}$$

where

$$c_i^{\mathrm{d}} = d_i \left(Y_i^{\mathrm{f}} + f_{\mathrm{C}}(\beta_i) \, Y_i^{\mathrm{C}} / \alpha_i \right) \, \Delta X_i \tag{3.30}$$

and

$$c_{ij}^{\mathrm{s}} = d_i \, f_{\mathrm{s}}(\beta_i) \, \mathcal{T}_{ji} \, Y_j^{\mathrm{C}} / \alpha_j \, \Delta X_j \,. \tag{3.31}$$

The solution of (3.26) can be obtained by inversion, leading to the energy deposit estimator

$$\widehat{\mathbf{w}} = \mathbf{C}^{-1}\mathbf{y} \,. \tag{3.32}$$

The statistical uncertainties of $\widehat{\mathbf{w}}$ can be obtained by error propagation

$$\mathbf{V}_{\mathbf{w}} = \mathbf{C}^{-1} \, \mathbf{V}_{\mathbf{y}} \left(\mathbf{C}^{\mathrm{T}} \right)^{-1} \,. \tag{3.33}$$

The bottom row of Figure 3.9 illustrates the resulting decomposition of the total light flux into fluorescence light, direct and scattered Cherenkov light.

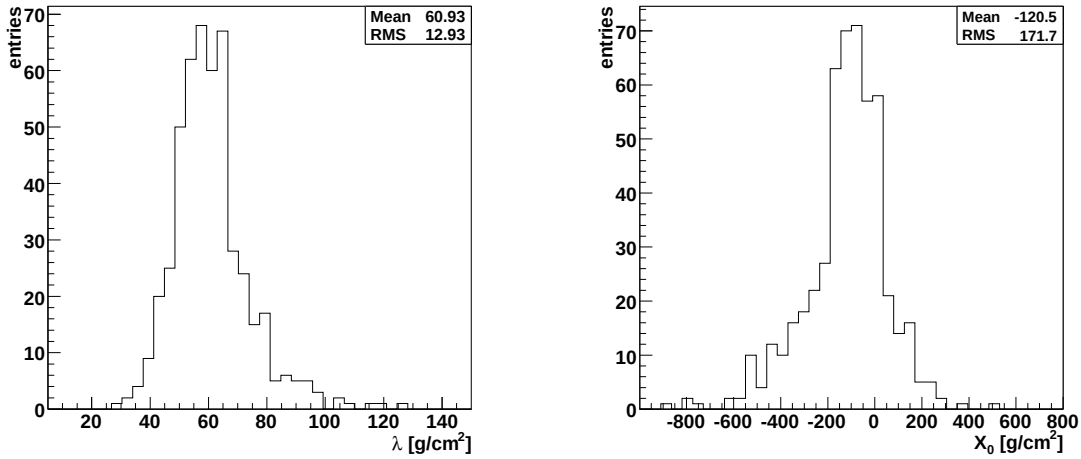


Figure 3.13: Gaisser-Hillas fit parameters Λ and X_0 as obtained from a set of highest quality hybrid data [147, 149].

3.3.6 Profile extrapolation and error propagation

For a determination of the total shower energy, it is necessary to integrate the energy deposit profile from the border of the atmosphere until the shower is completely absorbed. Therefore the observed $\mathrm{d}E/\mathrm{d}X$-profile has to be extrapolated beyond the telescope field of view boundaries using a Gaisser-Hillas (GH) function [157] in slant depth X

$$f_{\mathrm{GH}}(X) = \left(\frac{\mathrm{d}E}{\mathrm{d}X}\right)_{\mathrm{max}} \left(\frac{X - X_0}{X_{\mathrm{max}} - X_0}\right)^{\frac{X_{\mathrm{max}} - X_0}{\Lambda}} e^{\frac{X_{\mathrm{max}} - X}{\Lambda}}, \tag{3.34}$$

with the position of the shower maximum X_{max}, the maximum energy deposit $(\mathrm{d}E/\mathrm{d}X)_{\mathrm{max}}$ and the two shape parameters Λ and X_0. It is only possible to constrain all four GH-parameters for events where a large portion of the shower profile is observed before as well as after the shower maximum. For the general purpose, the χ^2 needs additional constraints on the two shape parameters

$$\chi^2_{\mathrm{GH}} = \sum_i \frac{\left(F_i - \sum_j c_{ij} f_{\mathrm{GH}}(X_i)\right)^2}{\sigma^2_{F_i}} + \left(\frac{\Lambda - \langle\Lambda\rangle}{V_\Lambda}\right)^2 + \left(\frac{X_0 - \langle X_0\rangle}{V_{X_0}}\right)^2, \tag{3.35}$$

where $\langle\Lambda\rangle$, $\langle X_0\rangle$ and V_Λ, V_{X_0} are the mean values and variances for Λ and X_0 taken from a dedicated study of a set of high quality events (see Figure 3.13). The integration of (3.34) yields the calorimetric energy

$$E_{\mathrm{cal}} = \int_0^\infty \mathrm{d}X \, f_{\mathrm{GH}}(X)\,. \tag{3.36}$$

The integration can be solved analytically by utilizing the substitutions

$$u = \frac{X - X_0}{\Lambda} \qquad \text{and} \qquad v = \frac{X_{\mathrm{max}} - X_0}{\Lambda} \tag{3.37}$$

in the Gaisser-Hillas function (3.34) to get

$$f_{\mathrm{GH}}(u) = \left(\frac{\mathrm{d}E}{\mathrm{d}X}\right)_{\max} \left(\frac{e}{v}\right)^{v} e^{-u}\, u^{v}\,, \tag{3.38}$$

which can be identified by the Γ-distribution. The above integral is then given by

$$E_{\mathrm{cal}} = \Lambda \left(\frac{\mathrm{d}E}{\mathrm{d}X}\right)_{\max} \left(\frac{e}{v}\right)^{v} \Gamma(v+1)\,, \tag{3.39}$$

which allows to rewrite (3.34) directly with E_{cal} as a parameter

$$f_{\mathrm{GH}}(u) = E_{\mathrm{cal}} \frac{e^{-u}\, u^{v}}{\Lambda\, \Gamma(v+1)}\,. \tag{3.40}$$

In this way the statistical uncertainty $\sigma_{\mathrm{flux}}(E_{\mathrm{cal}})$ can be directly obtained from the $\chi^2 + 1$ contour of (3.35).

Not all of the total energy of the primary cosmic ray particle is going into the electromagnetic component of the air shower. Neutrino production reduces the available energy as well as muons that need long trajectories to release their energy. To account for this invisible energy the calorimetric energy is multiplied by a correction factor

$$E_{\mathrm{tot}} = f_{\mathrm{inv}}\, E_{\mathrm{cal}}\,, \tag{3.41}$$

obtained from air shower simulations. Due to the stochastic nature of air showers, the correction factor is subject to shower-to-shower fluctuations. The statistical as well as systematic uncertainty of f_{inv} are determined in [158]. Since the meson decay probabilities, and thus the amount of neutrino and muon production decrease with energy, $f_{\mathrm{inv}} = f(E_{\mathrm{cal}})$ depends on the primary energy [159].

The uncertainties of the reconstructed parameters of the shower geometry need to be propagated into the profile parameters $X_{\max}$ and E_{cal}. The distances r_i from the telescope to the shower axis and therefore the transmission $T(r_i)$ as well as the geometry $1/(4\pi r_i^2)$ factors are affected. Furthermore, the shower direction is responsible for the resulting integrated atmospheric density profile and thus the conversion from position at the shower axis into slant depth. Finally, there is a strong dependence of the observed amount of direct and scattered Cherenkov light through the angles β_i.

The status of the atmosphere can be measured only with a limited precision. Whereas Rayleigh attenuation is a theoretically well understood process, the molecular density profiles and aerosol content of the atmosphere vary due to environmental influences and need to be well monitored in order to determine the slant depth and transmission coefficients needed for the profile reconstruction. Also the uncertainties in these measured atmospheric properties (see for instance [132, 160]) have to be propagated into the profile parameters.

Standard error propagation of the uncertainties of a set of parameters $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ into the quantity $q(\boldsymbol{\alpha})$ is

$$\sigma^2(q) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\mathrm{d}q}{\mathrm{d}\alpha_i} \frac{\mathrm{d}q}{\mathrm{d}\alpha_j} V_{ij}^{\alpha}\,, \tag{3.42}$$

where $\mathbf{V}^{\alpha}$ is the covariance matrix of the parameters $\boldsymbol{\alpha}$. In general, the derivatives $\mathrm{d}q/\mathrm{d}\alpha_i$ cannot be obtained analytically, since q can depend on the parameters $\boldsymbol{\alpha}$ in an arbitrarily

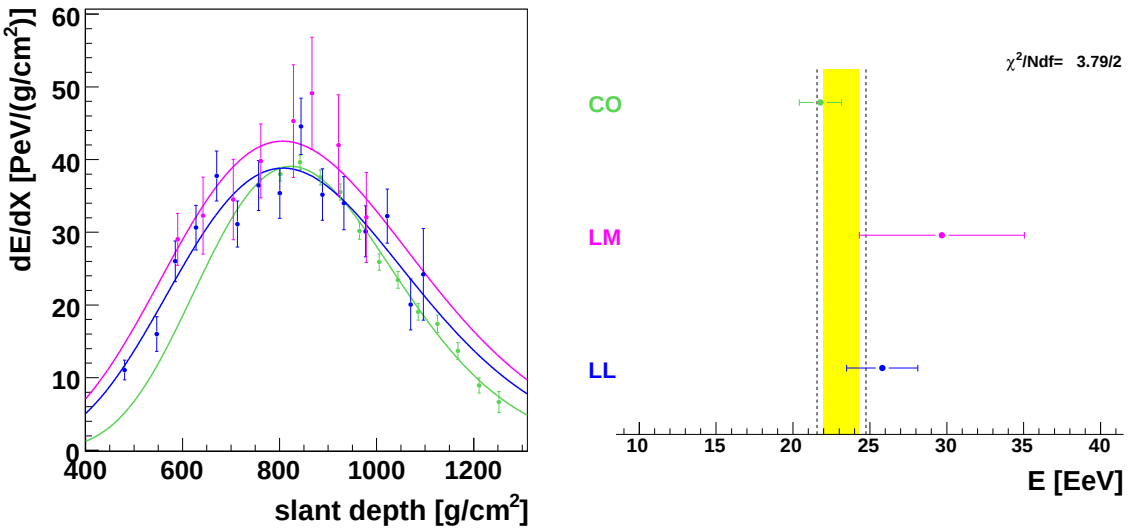


Figure 3.14: Reconstructed energy deposit profiles and profile parameters for the three individual eyes of the event from Figure 3.8. Left panel: reconstructed $\mathrm{d}E/\mathrm{d}X$-profiles for all three eyes (in eye-colors) with the propagated statistical uncertainties of the photon flux. Right panel: resulting energies including propagated geometric and atmospheric uncertainties for all eyes, together with the error-weighted mean energy (1 σ error band marked yellow).

complicated way. For a general application of (3.42) the derivatives have to be calculated numerically. With approximation

$$\Delta_i = \frac{\mathrm{d}q}{\mathrm{d}\alpha_i}\sigma_i \approx \frac{1}{2}\left[\, q(\alpha_i + \sigma_i) - q(\alpha_i - \sigma_i)\,\right] \tag{3.43}$$

it is possible to rewrite (3.42) as

$$\sigma^2_{\mathrm{geom}}(q) = \sum_{i=1}^{n}\sum_{j=1}^{n}\Delta_i\Delta_j\rho_{ij}\,, \tag{3.44}$$

where

$$\rho_{ij} = \frac{V^{\alpha}_{ij}}{\sqrt{V^{\alpha}_{ii}V^{\alpha}_{jj}}} \tag{3.45}$$

are the correlation coefficients of the parameters α_i and α_j.

Summarizing the above considerations and using (3.44) the statistical variance of the total energy (3.41) can be obtained as

$$\sigma^2_{\mathrm{stat}}(E_{\mathrm{tot}}) = E^2_{\mathrm{tot}}\,\sigma^2(f_{\mathrm{inv}}) + \left(\frac{\mathrm{d}f_{\mathrm{inv}}}{\mathrm{d}E_{\mathrm{cal}}}E_{\mathrm{cal}} + f_{\mathrm{inv}}\right)^2 \sum_i \sigma^2_i(E_{\mathrm{cal}})\,, \tag{3.46}$$

where i runs over the geometric, atmospheric and flux uncertainties. Since the invisible energy correction does not affect the depth of the shower maximum, its uncertainty is simply given by

$$\sigma_{\mathrm{stat}}(X_{\mathrm{max}}) = \sqrt{\sum_i \sigma^2_i(X_{\mathrm{max}})}\,. \tag{3.47}$$

Figure 3.15: Layout of the ADST data containers to store reconstructed as well as simulated event information.

Figure 3.14 summarizes the resulting energy deposit profiles and total energies on the example of the chosen triple event. It has to be noted that the errors in the profile plots are just the statistical uncertainties of the light flux conversion (3.33), while the errors on the total energies are including the propagated geometric and atmospheric uncertainties (3.46). This is the reason why the total energies are in good agreement with each other, while on the first look the profiles seem to be in disagreement. The discrepancy is caused by slightly differing geometric reconstructions of the three telescope events, leading to differing atmospheric corrections and as a consequence to an under- respectively over-estimation of the energy deposit profile at the shower axis. This illustrates the importance of a correct propagation of the uncertainties of the geometric reconstruction into the final profile parameters X_{max} and E_{tot}.

3.3.7 Data summary trees and event visualization

Participating in the development of a suited data format to work with reconstructed and simulated Auger data was also part of the work for this thesis. A general purpose ROOT-based [161] file format, to store the results from SD, FD as well as hybrid event reconstruction was implemented [162]. The format is called *advanced data summary trees* (ADST) and allows to store event data on a user-selectable level of detail (see Figure 3.15). This enables the development of standardized analysis algorithms within the large Auger Collaboration, which

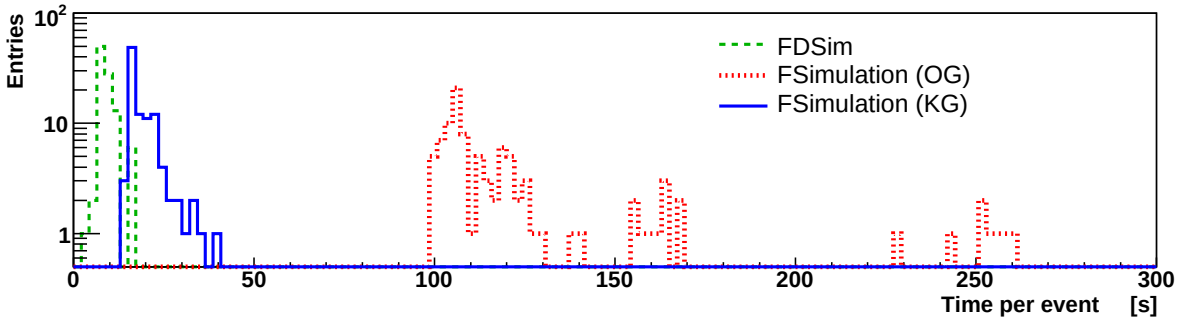


Figure 3.16: Computation time spent for one simulated event on an AMD64 2.1 GHz machine with 1 GB of memory. `FDSim` [164] is the original standalone C++ telescope detector simulation (no surface detector) and `FSimulation`(OG) the corresponding version translated into the $\overline{\text{Off}}\underline{\text{line}}$ framework, where OG stands for $\overline{\text{Off}}\underline{\text{line}}$-Group. `FSimulation`(KG), where KG stands for Karlsruhe-Group, is the software developed during this thesis, which is by now the officially accepted hybrid detector simulation for the Auger experiment.

can be easily exchanged between distant groups. The needed reconstructed event data are provided on a daily basis [163]. Additionally, we developed the powerful `EventBrowser` utility to visualize `ADST` data files. This is needed to get detailed insight in reconstructed as well as simulated events and is exploited to reach and assure a high level of quality of the algorithms used during event reconstruction as well as simulation. All of the plots from Figures 3.8, 3.9 and 3.14 have been created by the `EventBrowser`.

The `ADST` packaged allows all members of the Auger Collaboration to work with up-to-date reconstructed data, without requiring to install, configure and run the heavy $\overline{\text{Off}}\underline{\text{line}}$ framework. Only a valid ROOT installation is necessary.

Naturally, the analysis, with the aim to derive the proton-air cross section from Auger data presented here, relies on the reconstructed data in the `ADST` format.

3.4 Hybrid detector simulation

The simulation task is also embedded into the $\overline{\text{Off}}\underline{\text{line}}$ framework. Starting with the output from an EAS simulation program [49, 50, 66, 72] the signals induced in the SD and FD detectors have to be inferred as precise and fast as possible. As part of this thesis, the $\overline{\text{Off}}\underline{\text{line}}$ FD simulation software was significantly enhanced. The main improvements are

- Numerical integration of traversed matter in a curved atmosphere.
- Wavelength dependent ray-tracing.
- Simulated drum calibration.
- Support of upward going shower geometries.
- Use of the fully configurable $\overline{\text{Off}}\underline{\text{line}}$ detector description. Support for HEAT[6] electron-

[6]The High Elevation Auger Telescopes (HEAT) are currently built close to the Coihueco telescope site. First data are expected during 2008.

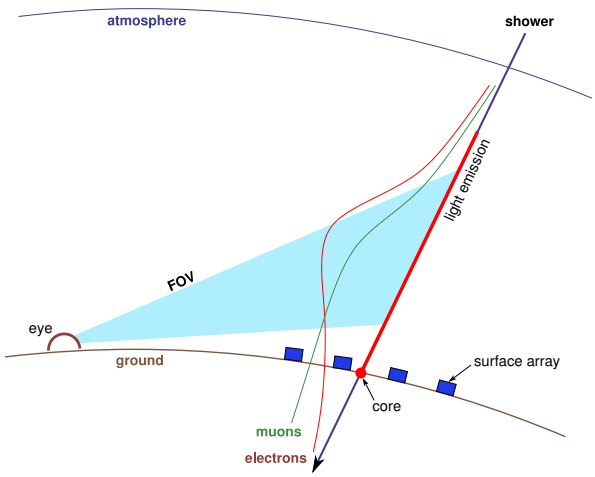


Figure 3.17: The hybrid detector simulation needs to calculate light emission, propagation in a curved atmosphere and the telescope response as well as the signal and timing induced by the air shower in the surface detectors. Detector simulations are based on generated longitudinal profiles of $\frac{dE}{dX}$, electron and muons.

ics and telescopes.

- Time dependent detector status.
- Per pixel noise taken from monitoring data.
- Fast surface detector simulation `SdSimpleSim`.

The full software bundle used for FD simulation in $\overline{\text{Off}}\underline{\text{line}}$ is called `FSimulation` (KG), where KG denotes the origin of the code and means Karlsruhe-Group. Thanks to extraordinary runtime stability, tested in thousands of hours of mass production runs in many institutions, the general high speed of the simulations (cf. Figure 3.16), the detailed implementation of all relevant physics and also the very user-friendly customization of options, the new `FSimulation` (KG) is the generally accepted and exclusively used simulation tool within the Auger Collaboration for FD and hybrid detector simulations.

3.4.1 Light generation

The calculation of fluorescence and Cherenkov photons produced by particles of the EAS, is the first step of the detector simulation. It is evident that the atmosphere indeed is a distinct part of the detector (see Figure 3.17). Thus, it is crucial to gather and use as much knowledge about the time dependent status of the atmosphere as possible.

Light production is calculated along the shower axis in steps of length dl that can be specified by the user (default value is c·100 ns). Only the part of the shower can be taken into account, where all needed atmospheric properties like aerosols, temperature, density and refractive index are defined by the used models. The distance l is defined along the shower axis starting from the shower core (usually intersection of the shower axis with ground level) and getting larger into the direction of shower motion.

Fluorescence photons are emitted by air molecules, mostly N_2, excited by the air shower cascade. Since the air shower simulation programs do provide direct information about the amount of energy deposited into the atmosphere, the number of fluorescence photons emitted per length along the shower axis, $\mathrm{d}l$, can be most accurately computed using this energy deposit $\mathrm{d}E/\mathrm{d}X$

$$\frac{\mathrm{d}n^{\mathrm{axis}}_{\mathrm{ph,fluo}}(l,\nu_{\mathrm{fluo}})}{\mathrm{d}l} = \frac{\mathrm{d}E}{\mathrm{d}X}(X_{\mathrm{slant}}) \cdot Y_{\mathrm{fluo}}(T,\rho,\nu_{\mathrm{fluo}}) \cdot \left(\frac{\mathrm{d}E}{\mathrm{d}X}\right)^{-1}_{0} . \tag{3.48}$$

Hereby $Y_{\mathrm{fluo}}(T,\rho,\nu_{\mathrm{fluo}})$ is the measured fluorescence yield, including temperature and density dependence, for the ultraviolet fluorescence emission bands ν_{fluo} of interest. The yield Y_{fluo} is given in units of photons/m for electrons of a fixed energy E_0, and the normalization constant $\left(\frac{\mathrm{d}E}{\mathrm{d}X}\right)_0$ is the energy deposit of electrons at this energy.

Particles of the EAS with an energy above the Cherenkov emission threshold E_{Ckov} emit Cherenkov photons. The total number of generated Cherenkov photons per length along the shower axis is proportional to the number of particles above E_{Ckov}. Since the shower is totally dominated by electromagnetic particles, this is in good approximation the number of electrons above the energy threshold $N^{\mathrm{e}}_{\mathrm{Ckov}}$. The number of Cherenkov photons emitted is then

$$\frac{\mathrm{d}n^{\mathrm{axis}}_{\mathrm{ph,Ckov}}(l,\nu_{\mathrm{Ckov}})}{\mathrm{d}l} = Y_{\mathrm{Ckov}}(s,\nu_{\mathrm{Ckov}}) \cdot N^{\mathrm{e}}_{\mathrm{Ckov}}(X_{\mathrm{slant}}) , \tag{3.49}$$

where Y_{Ckov} is the parameterized total Cherenkov yield for all EAS electrons above E_{Ckov} [154, 155, 156, 80]. The calculation of Cherenkov light is always integrated over a wavelength range around ν_{Ckov} in order to assure seamless coverage of the full wavelength interval. Moreover, the emission of Cherenkov photons is strongly forward directed, resulting in the formation of a bright Cherenkov beam along the shower axis:

$$n^{\mathrm{axis}}_{\mathrm{ph,\,beam}}(l=0,\,\nu_{\mathrm{Ckov}}) = 0 \tag{3.50}$$

$$\begin{aligned} n^{\mathrm{axis}}_{\mathrm{ph,\,beam}}(l,\,\nu_{\mathrm{Ckov}}) = &\left(\mathrm{d}l\,\frac{\mathrm{d}n^{\mathrm{axis}}_{\mathrm{ph,\,Ckov}}(l,\,\nu_{\mathrm{Ckov}})}{\mathrm{d}l} + n^{\mathrm{axis}}_{\mathrm{ph,beam}}(l-\mathrm{d}l,\,\nu_{\mathrm{Ckov}})\right) \\ &\times \quad T_{\mathrm{Mie}}\,T_{\mathrm{Rayleigh}} . \end{aligned} \tag{3.51}$$

The beam itself is attenuated while moving a distance $\mathrm{d}l$ by the transmissivity of the atmosphere due to Mie (T_{Mie}) and Rayleigh (T_{Rayleigh}) scattering. Of course the transmissivities T are always a function of wavelength and geometric distance (Mie), respectively grammage distance (Rayleigh).

3.4.2 Light propagation in the atmosphere

To assure a constant quality of the simulation along the shower axis, the light flux reaching the telescopes is calculated with a time resolution $\mathrm{d}t$, which is a fixed ratio of the final ADC sampling time. The default ratio is 10 and the sampling time for standard telescopes 100 ns while it is 25 ns in the case of HEAT telescopes.

The first problem to solve is the purely geometric question where and when the shower exactly enters respectively leave the field of view (FOV) of a telescope (cf. Figure 3.18). For

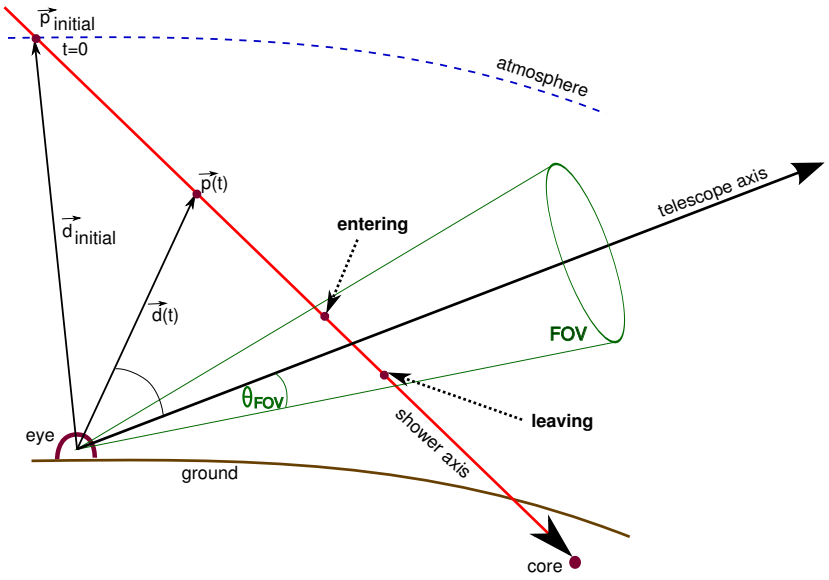


Figure 3.18: Geometry of a shower entering and leaving the field of view (FOV) of a telescope.

this purpose the FOV of a telescope is approximated by a viewing cone along the telescope axis $\vec{n}_{\text{tel}}$ with an opening angle of θ_{FOV}. For any point $\vec{p}(t)$ on the shower axis

$$\vec{p}(t) = \vec{p}_{\text{initial}} + t \cdot c \cdot \vec{n} \tag{3.52}$$

the angular distance of the vector connecting the position of the eye $\vec{p}_{\text{eye}}$ and $\vec{p}(t)$

$$\vec{d}(t) = \vec{p}(t) - \vec{p}_{\text{eye}} \tag{3.53}$$

and $\vec{n}_{\text{tel}}$ can be calculated as

$$\cos\theta = \frac{\vec{n}_{\text{tel}} \cdot \vec{d}(t)}{|\vec{n}_{\text{tel}}| \cdot |\vec{d}(t)|} \,. \tag{3.54}$$

To find the times t_1 and t_2, where the shower enters and leaves the FOV (3.54) needs to be solved for the condition $\cos\theta = \cos\theta_{\text{FOV}}$. This requires taking the square of (3.54) after combining with (3.53) and (3.52)

$$\cos^2\theta_{\text{FOV}} = \frac{(\vec{n}_{\text{tel}} \cdot \vec{d}_{\text{initial}})^2 + 2ct(\vec{n}_{\text{tel}} \cdot \vec{d}_{\text{initial}})(\vec{n}_{\text{tel}} \cdot \vec{n}) + (ct)^2(\vec{n}_{\text{tel}} \cdot \vec{n})^2}{\vec{d}^{\,2}_{\text{initial}} + 2(\vec{d}_{\text{initial}} \cdot \vec{n}) \cdot ct + (ct)^2 \cdot \vec{n}^{\,2}} \tag{3.55}$$

and solving the resulting quadratic equation for t:

$$\begin{aligned} 0 = {} & t^2 \cdot c^2 \left(\cos^2\theta_{\text{FOV}} - (\vec{n}_{\text{tel}} \cdot \vec{n})^2\right) \\ & + \; t \cdot 2c \left(\cos^2\theta_{\text{FOV}} \cdot (\vec{d}_{\text{initial}} \cdot \vec{n}) - (\vec{n}_{\text{tel}} \cdot \vec{d}_{\text{initial}}) \cdot (\vec{n}_{\text{tel}} \cdot \vec{n})\right) \\ & + \; \cos^2\theta_{\text{FOV}} \cdot \vec{d}^{\,2}_{\text{initial}} - (\vec{n}_{\text{tel}} \cdot \vec{d}_{\text{initial}})^2 \,. \end{aligned} \tag{3.56}$$

Since the FOV cone as defined in Figure 3.18 after quadration extends mathematically also to the "negative" side (cf. Figure 3.19), the solutions t_1 and t_2 need to be carefully checked

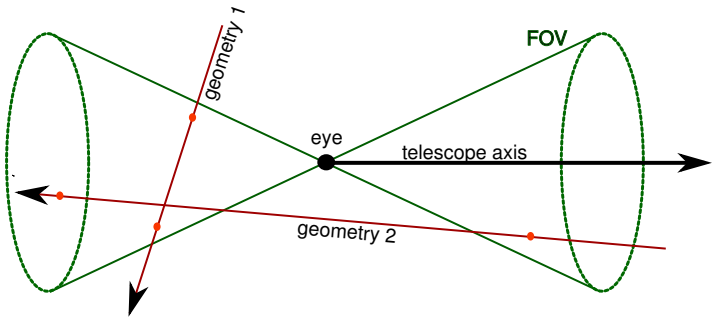


Figure 3.19: Example of two shower geometries having invalid intersection points (circles) with the FOV cone.

for physical correctness. It also has to be checked that the shower is well confined inside the atmosphere, meaning it cannot start before entering the atmosphere and cannot proceed after hitting ground. Finally, two times t_{enter} (the smaller value of t_1 and t_2) and t_{leave} (the larger value of t_1 and t_2) are obtained for each of the telescopes, where the shower geometry intersects the FOV. These times are easily transformed into times at the diaphragm of the telescope by

$$t_{\text{dia}} = t + |\vec{p}(t) - \vec{p}_{\text{eye}}| \cdot c^{-1} \,. \tag{3.57}$$

Since the FOV cone is slightly larger than the real sensitive FOV of a telescope, using the times t_{enter} and t_{leave} also enables scattered light from just outside the real FOV to be correctly simulated. At this point the light flux of fluorescence and Cherenkov photons at the diaphragm of the telescope can be calculated in time bins of dt over the time interval from $t^{\text{dia}}_{\text{enter}}$ to $t^{\text{dia}}_{\text{leave}}$. The back-transformation from time at the diaphragm t_{dia} into time t and thus position $\vec{p}(t)$ along the shower axis is best calculated within the known SDP (see Figure 3.10) as

$$t(t_{\text{dia}}) = \frac{t_0^2 + R_p^2/c^2 - t_{\text{dia}}^2}{2t_0 - 2t_{\text{dia}}} \,, \tag{3.58}$$

where t_0 is the time from $\vec{p}_{\text{initial}}$ to the point, where R_p is located on the shower axis. The transformation into l is just $l = l_{\text{initial}} + t \cdot c$, with l_{initial} is the distance along the shower axis from the core $\vec{p}_{\text{core}}$ to $\vec{p}_{\text{initial}}$.

For the isotropically emitted fluorescence light the number of photons arriving at the diaphragm per area in a given time interval dt can then be written as

$$\frac{\mathrm{d}n^{\text{dia}}_{\text{ph,fluo}}(t_{\text{dia}}, \nu_{\text{fluo}})}{\mathrm{d}A} = \frac{\mathrm{d}n^{\text{axis}}_{\text{ph,fluo}}(l, \nu_{\text{fluo}})}{\mathrm{d}l} \cdot \mathrm{d}t \cdot c \cdot \frac{1}{4\pi r^2} \cdot T_{\text{Mie}} \cdot T_{\text{Rayleigh}} \,. \tag{3.59}$$

For Mie or Rayleigh scattered photons out of the Cherenkov beam it is

$$\frac{\mathrm{d}n^{\text{dia}}_{\text{ph,Ckov-scattered}}(t_{\text{dia}}, \nu_{\text{Ckov}})}{\mathrm{d}A} = n_{\text{ph, beam}}(l, \nu_{\text{Ckov}}) \cdot P_{\text{scattered}} \cdot T_{\text{Mie}} \cdot T_{\text{Rayleigh}} \,, \tag{3.60}$$

where $P_{\text{scattered}}$ is the geometric scattering amplitude, taking care of the scattering angle as well as the distance to the observer. And for Cherenkov light directly emitted into the tele-

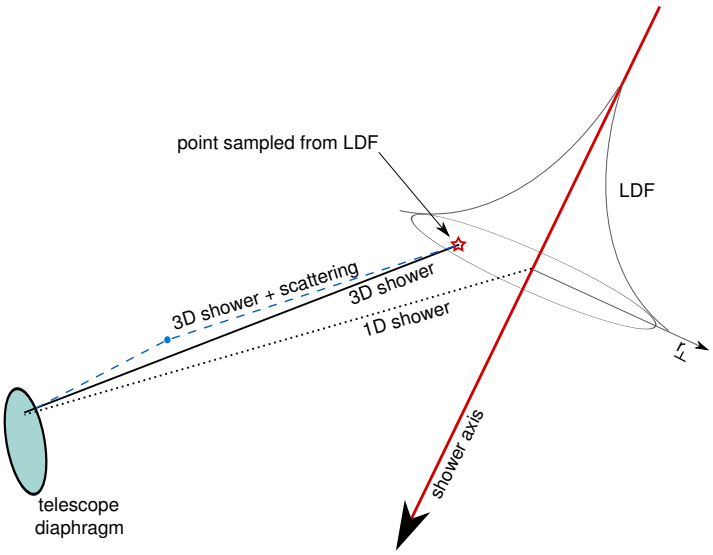


Figure 3.20: The three dimensional structure of an air shower leads to a broadening of the angular and timing distribution of arriving photons.

scope the number of photons is

$$\frac{\mathrm{d}n^{\mathrm{dia}}_{\mathrm{ph,Ckov-direct}}(t_{\mathrm{dia}}, \nu_{\mathrm{fluo}})}{\mathrm{d}A} = \frac{\mathrm{d}n_{\mathrm{ph,Ckov}}(l, \nu_{\mathrm{Ckov}})}{\mathrm{d}l} \cdot \mathrm{d}t \cdot c \cdot P_{\mathrm{direct}} \cdot T_{\mathrm{Mie}} \cdot T_{\mathrm{Rayleigh}}, \tag{3.61}$$

with P_{direct} being the geometric probability of direct Cherenkov detection, taking into account the parameterized angular distribution of electrons at the shower axis, as well as the geometry of the Cherenkov light cone.

3.4.3 Shower structure

Up to this point the shower was handled as a purely one-dimensional object. This is of course an oversimplification, which we have to correct for. The lateral distribution of $\frac{\mathrm{d}E}{\mathrm{d}X}$ [165] can be used to blur the shower structure. For each time interval $\mathrm{d}t$ the total light flux at the diaphragm summed over all wavelengths

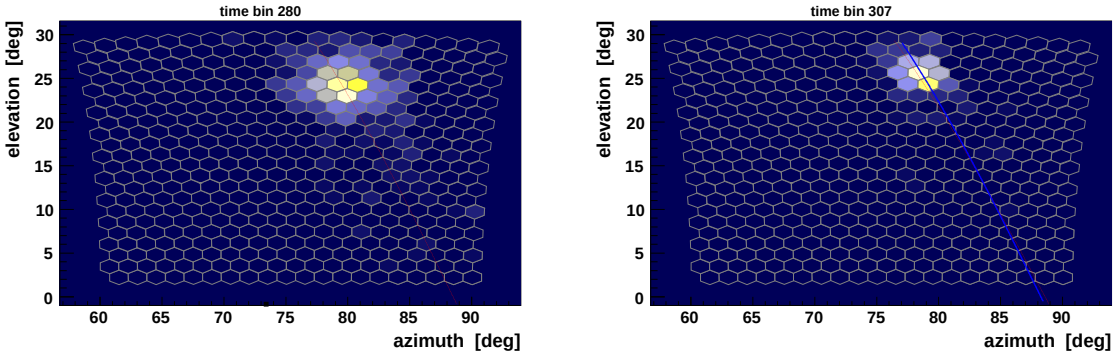
$$\frac{\mathrm{d}n^{\mathrm{dia}}_{\mathrm{ph}}(t_{\mathrm{dia}})}{\mathrm{d}A} = \sum_i \frac{\mathrm{d}n^{\mathrm{dia}}_{\mathrm{ph,all}}(t_{\mathrm{dia}}, \nu_i)}{\mathrm{d}A} \tag{3.62}$$

is converted into a number of photons

$$n_{\mathrm{ph,dia}}(t_{\mathrm{dia}}) = \frac{\mathrm{d}n^{\mathrm{dia}}_{\mathrm{ph}}(t_{\mathrm{dia}})}{\mathrm{d}A} \cdot A_{\mathrm{dia}} \cdot \cos\theta^{\mathrm{dia}}_{\mathrm{1D}}. \tag{3.63}$$

To limit the needed computing resources an upper limit n_{limit} can be defined by the user to limit the number of photons per time bin generated for ray-tracing

$$n_{\mathrm{raytrace}}(t_{\mathrm{dia}}) = \begin{cases} n_{\mathrm{ph,dia}}(t_{\mathrm{dia}}) & \text{for} \quad n_{\mathrm{ph,dia}}(t_{\mathrm{dia}}) \leq n_{\mathrm{limit}} \\ n_{\mathrm{limit}} & \text{for} \quad n_{\mathrm{ph,dia}}(t_{\mathrm{dia}}) > n_{\mathrm{limit}} \end{cases}. \tag{3.64}$$



(a) data (b) simulation

Figure 3.21: Left panel: Image of a reconstructed shower passing in a distance of (0.99±0.02) km in front of the Los Leones eye. Lighter colors indicate a larger signal. Shown is the shower image on the camera arriving within a single 100 ns time bin. The resulting circular shape is a consequence of the combination of lateral and time spread of light emission at the shower axis. Right panel: Monte Carlo simulation, performed with the reconstructed parameters of the real event. It shows the same circular structure. Probably due to a mis-reconstruction of the real event or less scattering of light in the atmosphere, the lateral extent seems to be smaller compared to the measured event.

By default n_{limit} is set to 40, which assures sufficient sampling even for a three dimensional shower structure without generating artificial fluctuations. Photons are generated with a weight

$$\omega_i(t_{\text{dia}}) = \frac{n_{\text{ph,dia}}(t_{\text{dia}})}{n_{\text{raytrace}}} \cdot A_{\text{dia}} \cdot \cos\theta_i^{\text{dia}}\,, \tag{3.65}$$

which conserves the total light flux at the diaphragm

$$\sum_i \omega_i(t_{\text{dia}}) = \frac{\mathrm{d}n_{\text{ph}}^{\text{dia}}(t_{\text{dia}})}{\mathrm{d}A} \cdot A_{\text{dia}} \cdot \cos\theta_{1D}^{\text{dia}}\,. \tag{3.66}$$

The wavelength of each photon is randomly selected from the probability distribution of wavelengths at the diaphragm. Photons are generated uniformly on the surface of the diaphragm, and with the injection time t_{dia} at which the photon flux at the diaphragm was calculated for. With (3.58) the time and thus position at the shower axis, where the photons were produced can be calculated. Using the properties of the atmosphere at this point (temperature, density) and the age s of the shower, the lateral extend of electrons can be calculated. Figure 3.20 illustrates how this information can be used to resample the photon emission point within the local shower coordinate system, by sampling a uniform azimuth angle ϕ and a radius $r_\perp$ according to the lateral $\frac{\mathrm{d}E}{\mathrm{d}X}$ distribution. This newly sampled point is then used to calculate the arrival direction of the photon i on the diaphragm (θ_i^{dia} and ϕ_i^{dia}) and to correct the arrival time $t_{\text{dia}}^{\text{3D}}$. Strictly, the lateral distribution model for $\frac{\mathrm{d}E}{\mathrm{d}X}$ is only valid for fluorescence light production, whereas it is currently also used for the Cherenkov light emission. Only for large shower ages $s > 1.5$ there is some sign of disagreement which can be attributed to an underestimated shower lateral extent (compare Figure 3.20). One of the future improvements of the simulations will be the implementation of a lateral Cherenkov

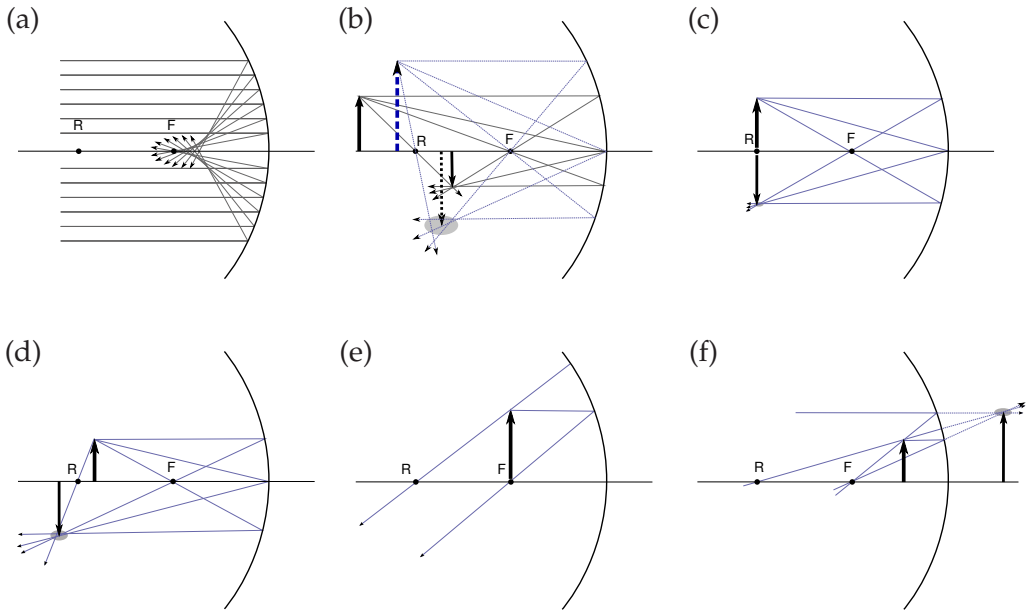


Figure 3.22: Image construction for a spherical optics with radius R and focal surface at $F = R/2$. Focal caustic for parallel incoming light (a). Objects in front of the radius point R are seen scaled down and upside-down between R and F (b). The effect of spherical abberations (shaded ellipses) is more significant, if larger parts of the mirror are used. This strongly limits correct imaging of the optics to a small central part of the mirror. Objects located on the radius point are not scaled, but seen upside-down (c). Between the radius and the focal point F objects are magnified in front of R (d). On top of the focal point no image is produced (e). Between the focal surface and the mirror objects are seen magnified behind the mirror and correctly aligned (f).

distribution model [166, 167, 168], which is not yet available, but worked on within the Auger Collaboration [169]. Another thing not yet implemented is the blurring effect by Mie and Rayleigh scattering in the atmosphere. Figure 3.21 illustrates how the currently implemented shower structure reproduces real events.

3.4.4 Ray-tracing in the telescope

The generated photons on the diaphragm are then traced through the geometry of the telescope optics. Since the geometry is a simple non-imaging spherical Schmidt optics, the ray-tracing can easily be done using a custom-written code. The radius of the focusing mirror is $R = 3.4\,\mathrm{m}$ and all components of the optics are arranged with respect to the center of the mirror. Spherical optics allows to observe a large field of view by focusing light from a wide range of directions on the focal plane, which itself is a sphere with radius $R/2$ located at $R/2$. However, spherical optics produce large spherical abberations because they do not focus light on a single point but on a caustic surface (see Figure 3.22 (a)), which in fact leads to a blurred spot on the focal plane. The radius of this spot is rapidly increasing if a larger

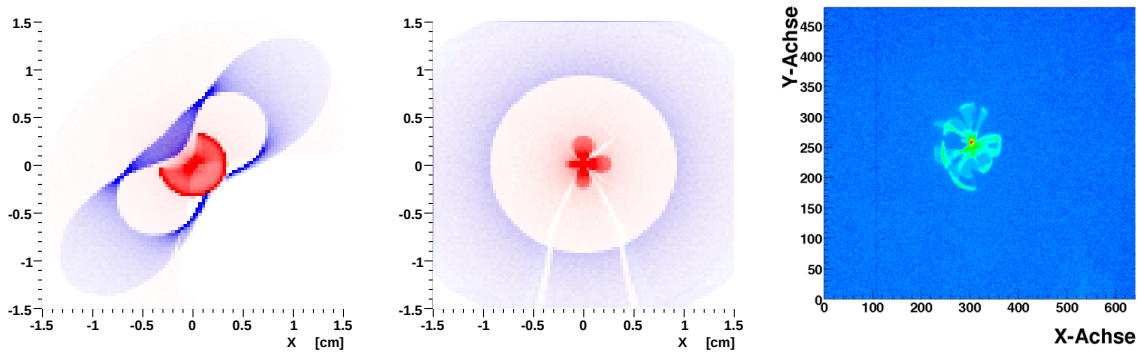


Figure 3.23: Simulated light spot on top of pixel 1 at the corner of the camera (left), on top of pixel 210 in the center of the camera (middle), and measured light spot produced by the ultraviolet light of the star Vega on the center of the camera (right). For the simulated spots the intensity of blue shows light, which passed through the corrector lens, and red is light, which was not refracted by the lens. In the measured spot one can see a fine structuring caused by the segmentation of the mirror.

part of the mirror surface is used. Figure 3.22 (b) shows how the spherical abberations are getting more significant if a larger part of the mirror is illuminated. For an imaging optics the spherical abberations have to be corrected for by an additional corrector lens. For the Auger telescopes the abberations only need to be limited to produce spots on the focal surface well smaller than the size of the used PMT pixels. This is why a diaphragm opening window of radius 0.85 m is used in order to limit the available mirror area, keeping the spot diameter within about 0.6° while the pixel field of view is 1.5°. In order to increase the sensitivity of the Auger telescopes the light collection area was almost doubled by increasing the diaphragm opening from 0.85 m to 1.1 m and installing a corrector ring on the outer 25 cm to compensate the increased spherical abberations, still keeping the spot size within 0.6° (see Figure 3.23).

Ray-tracing of photons in the telescope optics takes care of reflection and refraction at the corrector lens, mirror, light collectors and PMT window (c.f. Figure 3.24). Transmission and reflectivity of all materials is properly considered. Photons reaching the photocathode of a PMT are added to the corresponding photon signal $n_{\mathrm{ph}}^{\mathrm{PMT}_i}(t)$, which is binned with the ADC sampling time of 100 ns (25 ns for HEAT). A correction factor $\omega_{\mathrm{binning}} = \mathrm{d}t_{\mathrm{ADC}}/\mathrm{d}t_{\mathrm{dia}}$ is used due to the different time binning of photons at the diaphragm $\mathrm{d}t_{\mathrm{dia}}$, which is typically 10 ns, and the time binning on pixel level $\mathrm{d}t_{\mathrm{ADC}}$, determined by the 100 ns ADC sampling time. Figure 3.25 shows a comparison of the simulated angular efficiency of photon detection by the PMTs mounted in the camera with a dedicated measurement [170].

To take into account the measured wavelength dependent end-to-end telescope efficiency relative to $\nu_{\mathrm{norm}} = 380\,\mathrm{nm}$

$$\epsilon_{\mathrm{eye}}^{\mathrm{rel}}(\nu) = \frac{\epsilon_{\mathrm{eye}}(\nu)}{\epsilon_{\mathrm{eye}}(\nu_{\mathrm{norm}})}\,, \tag{3.67}$$

the photon weight after ray-tracing is multiplied by a correction factor

$$\omega_{\mathrm{relative}}(\nu) = \frac{\epsilon_{\mathrm{eye}}^{\mathrm{rel}}(\nu)}{\epsilon_{\mathrm{tel}}^{\mathrm{rel}}(\nu)}\,, \tag{3.68}$$

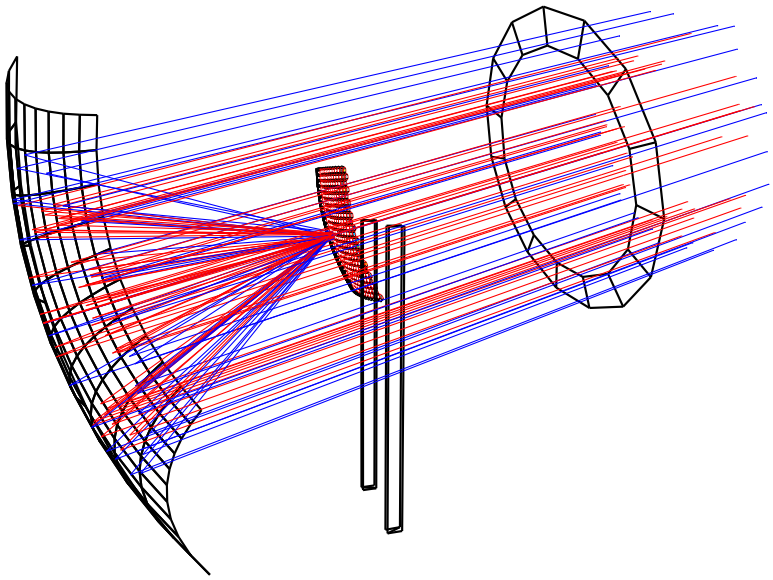


Figure 3.24: Ray-traced photons through the geometry of the telescope. Red photons are passing directly through the diaphragm opening. Blue photons are refracted by the corrector lens.

where $\epsilon_{\text{tel}}(\nu)$ is the combined efficiency of all optical components used during ray-tracing simulations

$$\epsilon_{\text{tel}}(\nu) = \epsilon_{\text{filter}}(\nu)\ \epsilon_{\text{lens}}(\nu)\ \epsilon_{\text{mirror}}(\nu)\ Q_{\text{eff}}(\nu)\,. \tag{3.69}$$

By using this telescope response (relative to $\nu_{\text{norm}} = 380\,\text{nm}$) the simulated events can be consistently reconstructed with the same settings as used for real data. The last correction which is applied for each pixel i is the simulated drum calibration constant

$$\omega_i^{\text{cal}} = \frac{k_i^{\text{sim,drum}}}{k_i^{\text{cal}}}\,, \tag{3.70}$$

where k_i^{cal} is the calibration constant going to be applied during reconstruction and $k_i^{\text{sim,drum}}$ is the response of the pixel i to a uniform light flux of a simulated drum in front of the diaphragm.

3.4.5 Background light

All telescopes permanently record background light from bright stars, the milky way, the moon, the atmosphere and other sources [171]. This background light has no relevant time structure on the scale of an FD event (1000 μs) and is responsible for most of the fluctuations seen in the ADC traces. Obviously this background light flux $\Phi_{\text{bg}}^{\text{dia}}$ is slowly changing with time (moon cycle, atmospheric condition, stars in the FOV) and also dependent on the viewing direction (zenith angle dependence,...). The background photon flux can be calculated from the recorded variance of ADC traces by subtracting the measured variance caused by electronics noise $V_{\text{noise}} = 3.5\,(\text{ADCcounts})^2$

$$V_{\text{bg}} = V_{\text{trace}} - V_{\text{noise}}\,. \tag{3.71}$$

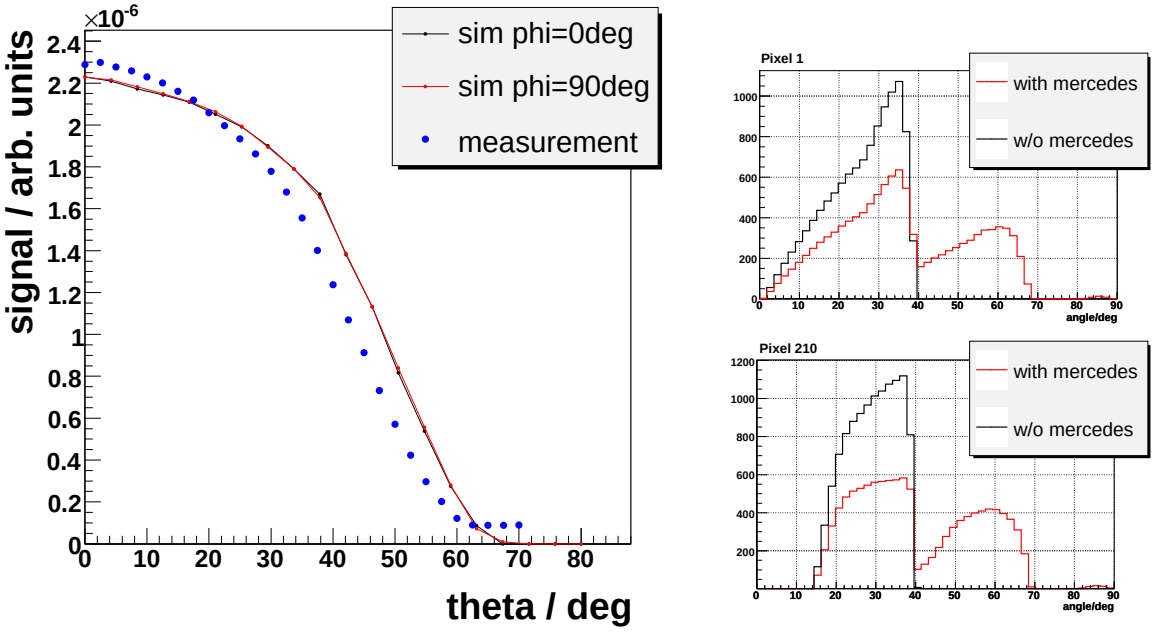


Figure 3.25: Left panel: Efficiency of photon detection for a PMT of the FD camera for different photon impact angles [170]. Right panel: Simulated distribution of photon impact angles on the corner pixel 1 and the central pixel 210. The strong impact of the light collectors (called *mercedes*) can be clearly identified.

The fluctuations after the PMT can be expressed as convolution of the Poissonian fluctuations of photo electrons at the cathode $\sigma^2_{\mathrm{PE}} = \langle n_{\mathrm{PE}} \rangle$ convoluted with Gaussian broadening due to the dynode chain

$$\sigma^2_{\mathrm{PE,\,amp}} = \sigma^2_{\mathrm{PE}} + \sigma^2_{\mathrm{dynodes}} = \langle n_{\mathrm{PE}} \rangle \cdot (1 + V_{\mathrm{g}}) \tag{3.72}$$

with $\sigma^2_{\mathrm{dynodes}} = \langle n_{\mathrm{PE}} \rangle \cdot V_{\mathrm{g}}$ and the measured gain variance $V_{\mathrm{g}} = 0.41$. Multiplying (3.72) with the gain g and the quantum efficiency Q_{eff} transforms the units from number of PEs to ADC counts and can then be directly compared to (3.71), which gives the mean number of background photons $\langle n_{\mathrm{bg}} \rangle$ at each pixel i per ADC sampling time bin $\mathrm{d}t_{\mathrm{ADC}}$

$$\left\langle n^{\mathrm{PMT}_i}_{\mathrm{ph,bg}} \right\rangle = \frac{V_{\mathrm{bg}}}{g^2 (1 + V_{\mathrm{g}}) Q_{\mathrm{eff}}} \,. \tag{3.73}$$

3.4.6 Electronics and sampling

Given the shower photon signal and the background photon flux at each PMT the number of generated PEs at the photocathode at time t_{ADC} can be calculated for each bin of the ADC traces as

$$n^{\mathrm{PMT}_i}_{\mathrm{PE}}(t_{\mathrm{PMT}}) = \mathrm{Poisson}\left[(n^{\mathrm{PMT}_i}_{\mathrm{ph,bg}} + n^{\mathrm{PMT_i}}_{\mathrm{ph,shower}}) \cdot Q_{\mathrm{eff}} \right] \,. \tag{3.74}$$

The PEs can then be converted into an ADC signal and associated fluctuations by

$$s_{\mathrm{ADC}}^{\mathrm{PMT}_i}(t_{\mathrm{ADC}}) = \left(n_{\mathrm{PE}}^{\mathrm{PMT}_i}(t_{\mathrm{PMT}}) + \mathrm{Gauss}\left[0,\, \sigma_{\mathrm{dynodes}}\right] - \left\langle n_{\mathrm{ph,bg}}^{\mathrm{PMT}_i} \right\rangle \cdot Q_{\mathrm{eff}}\right) \cdot g \tag{3.75a}$$

$$+ \quad \mathrm{Gauss}\left[0,\, \sigma_{noise}\right] \tag{3.75b}$$

$$+ \quad \mathrm{baseline} \tag{3.75c}$$

with the Gaussian fluctuations by the dynode chain $\sigma^2_{\mathrm{dynodes}} = V_{\mathrm{g}} \cdot n_{\mathrm{PE}}^{\mathrm{PMT}_i}(t_{\mathrm{PMT}})$ and the Gaussian fluctuations added by the electronics σ_{noise}. In line (3.75a) the Poissonian fluctuations are combined with the Gaussian fluctuations and then reduced by the mean number of background light PEs. The latter subtraction is performed, because the ADC baseline is added separately later, and the background photons are used only to produce the right fluctuations. The amplified PEs are then multiplied with the gain g to convert the result into ADC counts. The electronics adds another Gaussian noise term with $\sigma^2_{\mathrm{noise}} = 3.5\,(\mathrm{ADCcounts})^2$ (line (3.75b)). Finally the baseline is added in line (3.75c). Now the electronics filter characteristics has to be taken into account by convoluting $s_{\mathrm{ADC}}^{\mathrm{PMT}_i}(t_{\mathrm{ADC}})$ with a filter response function

$$s_{\mathrm{ADC}}^{\mathrm{PMT}_i}(t_{\mathrm{ADC}}) = \int \mathrm{d}t\; s_{\mathrm{ADC}}^{\mathrm{PMT}_i}(t)\; \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}(t-t_{\mathrm{ADC}})^2/\tau^2}\,, \tag{3.76}$$

where $\tau = \sqrt{\frac{\log(2)}{2\pi f_{\mathrm{cutoff}}}}$ and f_{cutoff} =3.1 MHz is the Gaussian filter 3 dB cutoff frequency [172].

3.4.7 Trigger

With the simulated ADC trace for each pixel the FD trigger can be simulated. The trigger system for the Auger telescope detectors is organized in three layers [130]. For each PMT the first level trigger (FLT) condition is a running boxcar sum over 10 ADC bins to be larger than the FLT threshold. In the real DAQ the threshold is dynamically adjusted to yield a 100 Hz FLT trigger rate per PMT. For the purpose of simulations the threshold has to be set to a level, which corresponds to the 100 Hz trigger rate.

With all the FLT pixels of a camera the second level trigger (SLT) algorithm can search for the geometric SLT patterns. In real DAQ the SLT is implemented on the hardware level. It reads two camera columns every 100 ns, resulting in a total time of 1 μs to read all 20 columns of a camera. Thus the timing resolution of the SLT trigger is 1 μs. The first SLT trigger on a camera triggers readout of the electronics and defines the time window of an event: the event starts 300 time bins before the SLT and lasts for 1000 ADC bins. It is therefore necessary to simulate a time span starting at least 300 ADC time bins before the shower signal arrives at the camera and 700 ADC time bins after the shower has left the camera. Since the FLT and SLT triggers are hardware implemented, they are reprogrammed in C++ for the purpose of simulations.

For each SLT trigger the third level trigger (TLT) is evaluated. The TLT is the last trigger step working on camera level. It combines timing and geometric information in order to separate direct muon hits and lightning from real shower events.

The last trigger step is the eye trigger (T3) which combines all camera events. The TLT and T3 algorithms are directly taken from the FD-DAQ software, to assure exactly the same processing of simulated and real data.

3.4.8 SD signal generation

Since no ground particles are generated by CONEX, a very simple and fast approach was followed to generate the needed SD detector response. The utility was implemented in the `SdSimpleSim` module of the Offline framework and can therefore easily be used within any detector simulation run. For a detailed description see Appendix A, here only the functionality as it was used to generate the detector simulations for this work is described.

During hybrid event reconstruction the SD station information is used to provide timing at a well known location. The station signal itself is only needed for station selection, in combination with the geometry of the station with respect to the reconstructed FD-only shower axis. This means it is needed to generate a realistic tank timing close to the shower core, though a flat shower front can be assumed in good approximation. Shower front fluctuations as well as detector timing resolution are taken into account. According to [173] the timing uncertainty is

$$\sigma_{t_i}^2 = \sigma_{t0}^2 + (k\,\rho_i^2\cos\theta)^2 \tag{3.77}$$

with ρ_i the shower plane distance of station i, $\sigma_{t0} = 26.84\,\mathrm{ns}$ the detector time resolution and $k = 5.368 \cdot 10^{-5}\,\mathrm{ns/m^2}$.

The integrated tank signal itself does not need to be very precise, meaning we can use a simple model for the lateral particle distributions and normalize it to the longitudinal profile at the shower axis given by CONEX.

Realistic noise tanks in the SD array are important, since mis-selection of the hybrid tank is a frequent source of mis-reconstruction and therefore an important factor for the overall reconstruction efficiency. The single tank trigger cannot be simulated in detail, since no ADC

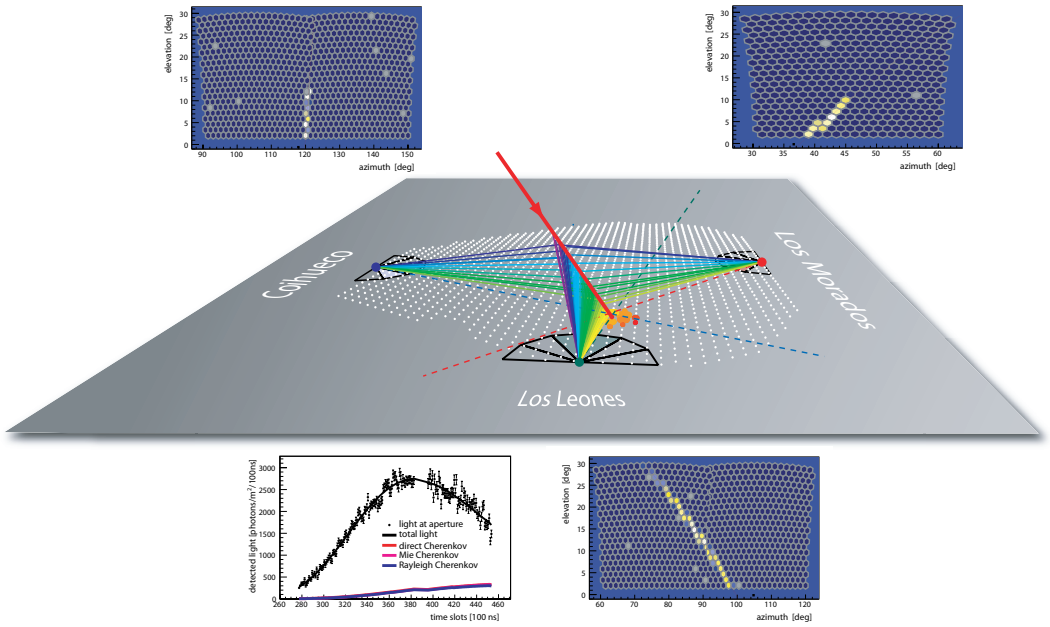


Figure 3.26: Resulting event after full detector simulation and reconstruction. The incident primary particle is a proton with 100 EeV.

signal trace is generated. Thus, lateral trigger probability (LTP) parameterizations [174, 175] are used to evaluate the individual tank trigger probabilities.

Only SD stations flagged as alive by the SD monitoring data [176] are considered during simulation. This automatically reflects the growth of the array with time as well as all technical problems which occurred during SD data-taking.

In Figure 3.26 the final product of a hybrid detector simulation after the standard event reconstruction procedure is illustrated. The visualization is done with the 3D-mode of the `EventBrowser`.

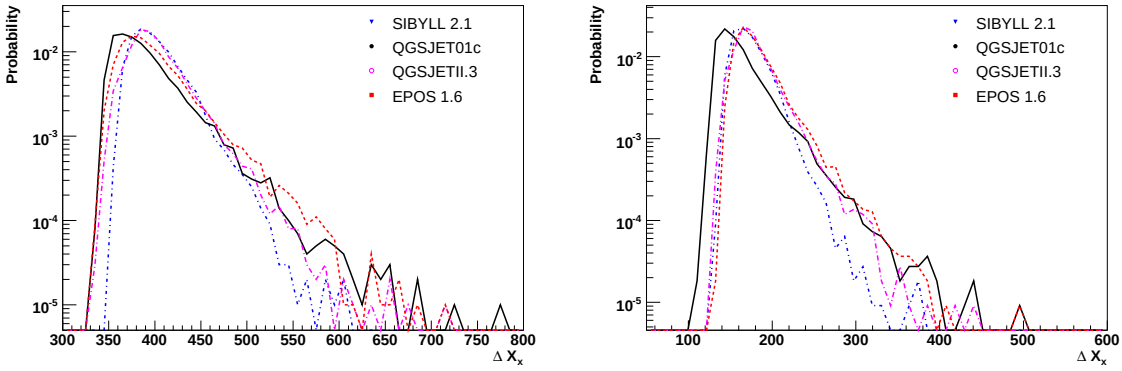
4 A novel method to derive the proton-air cross section

To derive the proton-air cross section based on the hybrid dataset of the Pierre Auger Observatory, a novel method is developed. The fundamental formulation is necessarily identical to (2.14), as it is described in Section 2.6.2. For this reason, there is large similarity to previous analysis approaches. However, two new aspects are added which are demonstrated to be important for the cross section measurement. Since the method is an indirect measurement of the first interaction point X_1 based on the observation of air showers, the impact of a changing cross section on the resulting shower development is fully incorporated within the approach. Additionally, the impact of air shower selection due to the acceptance of the detector is directly accounted for. Finally, the definition of an *invisible cross section* is discussed. However, it is demonstrated to have little impact on the resulting cross sections.

4.1 Motivation of X_{max} as observable

The depth of the maximum of the shower development X_{max} is a very well established observable. It can be directly observed by fluorescence telescopes and Section 3.3.6 describes in detail how X_{max} is reconstructed and how geometric and atmospheric uncertainties are properly propagated into the uncertainty of X_{max}. Nevertheless, already in 1982 it was proposed by [112] to use the depth of the quarter maximum $X_{1/4}$ instead of X_{max} in order to improve the sensitivity to the first interaction. At this time no detector was available to measure the shower profile accurately enough for a reliable estimate of $X_{1/4}$. However, the detector of the Pierre Auger Observatory can measure profiles very precisely down to even smaller intensities as $X_{1/4}$.

The principle idea in choosing an observable X_x located at closer distance to the first interaction X_1 is to enhance the correlation from X_x to X_1. But in fact the shape of the electromagnetic bulk of the air shower, as it is observed with a telescope detector is determined by the few ultra-high energy interactions during the shower startup phase (cf. Figure 2.10). After most of the shower energy is transferred into electromagnetic particles, the shower profile at any development stage, from the shower maximum X_{max} to much earlier development stages X_x, is already predetermined. Therefore no extra information is contained in X_x compared to X_{max}. Figure 4.1 depicts ΔX_x-distributions for $x = 1/4$ and $1/10$. The model differences are very similar, while the general width of the ΔX_x-distributions is getting narrower for smaller x. This might be an interesting feature, since the fluctuations of the correlation function ΔX_x cover the fluctuations originating from X_1, which is leading to a loss of sensitivity if the fluctuations of X_1 are getting too small compared to ΔX_x. Depending on the magnitude of the proton-air cross section, the expected RMS of X_1 is about $\sim$45 gcm^{-2}, while the width of ΔX_x approximately corresponds to the β parameter, as shown in Figure 4.2 (middle), which changes from $\sim$21 gcm^{-2} to $\sim$12 gcm^{-2} for decreasing x. However, the experimental difficulties in measuring X_x are getting larger for a smaller x, mostly because of the requirement of a simultaneous and good observation of X_{max} and X_x. Also atmospheric corrections will get much more important, since generally X_x is much further away from the detector than X_{max}. In total this will lead to a deterioration of the achievable resolution, compared to the $\sim$20 gm^{-2} for X_{max}, which at least compensates the gains by the reduced fluctuations of ΔX_x.



(a) $\Delta X_x = X_{1/4} - X_1$ (b) $\Delta X_x = X_{1/10} - X_1$

Figure 4.1: Shower development from the first interaction point X_1 up to a specific fraction of the shower maximum $(\mathrm{d}E/\mathrm{d}X)_{\max}$. The principle model differences do not vanish while moving to earlier shower stages.

On the other hand, it has to be investigated if the inherent model dependence of $\Delta X_1 = X_{\max} - X_1$, as displayed in Figure 2.28 (left), may be reduced by alternatively using the correlation $\Delta X_x = X_x - X_1$. In Figures 4.1 and 4.2 it is demonstrated that there is no benefit in doing so. This is again due to the fact that the model dependence is produced mostly in the hadronic interactions at ultra-high energy during the shower startup phase, where the subsequent electromagnetic shower cascade is initiated. Thus, there is no reason to move from the well defined observable $X_{\max}$ to anything less strict, like the position of shower development at a certain fraction of the maximum.

4.2 Impact of mass composition and hadronic interaction features on air showers

4.2.1 Dependence of air shower observables on features of the hadronic interactions

In order to explore the impact of uncertainties of the present high energy hadronic interaction models on the interpretation of air shower observables, the CONEX program is modified to change some of the interaction characteristics during EAS simulation. To achieve this, individual interaction characteristics are altered by the energy dependent factor

$$f(E) = \begin{cases} 1 & E \leq 1\,\mathrm{PeV} \\ 1 + (f_{10\,\mathrm{EeV}} - 1) \cdot \log_{10}(E/1\,\mathrm{PeV}) / \log_{10}(10\,\mathrm{EeV}/1\,\mathrm{PeV}) & E > 1\,\mathrm{PeV} \end{cases} . \quad (4.1)$$

This factor gets 1 below 1 PeV, where the models are constraint by accelerator data. Above 1 PeV the factor grows linearly with the logarithm of the energy, reaching the normalization value of $f_{10\,\mathrm{EeV}}$ at 10 EeV. This reflects the increasing uncertainty of the extrapolations with energy. The factor $f(E)$ is then used to modify specific characteristic properties of the secondary particle distributions created in hadronic interactions. Note that with this approach the strict idea of correlations as shown in Figure 2.15 is violated, since only individual properties are changed, thus leaving the parameter space provided by the original model. Nevertheless one can get a clear impression of how the resulting EAS properties are depending

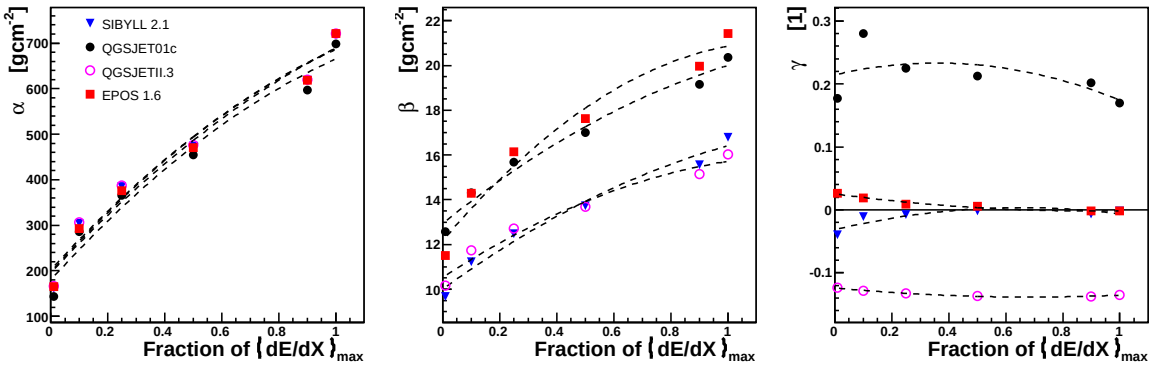


Figure 4.2: Dependence of the parameters of the parameterization of the ΔX_x-distribution on the chosen fraction of the shower maximum $(\mathrm{d}E/\mathrm{d}X)_{\max}$ (see Section 4.4 for a description of the parameterization). Even for very small fractions (down to 0.01) the model dependence does not improve significantly. The lines are polynomial fits of 2$^{\mathrm{nd}}$ degree to guide the eye.

on the specific interaction characteristics. Obviously, for large deviations from the original model, these results are getting less reliable and have to be treated with some caution.

For these studies a modified CONEX version is implemented that evaluates $f(E)$ after each hadronic interaction and enters a secondary particle *resampling* algorithm[7] for the case of $f(E) \neq 1$. The resampling algorithm is written in order to specifically change individual properties of the secondary particle distributions, while conserving all of the important features like total energy, charge, particle types and energy fractions in different particle types as far as possible. In the following the resampling algorithm is utilized to study the impact of the secondary multiplicity on the resulting air shower development. This is motivated by the enormous model uncertainty of the predictions for the secondary particle multiplicity, discussed in Section 2.2.3 (cf. Figure 2.14 on page 18).

Multiplicity. In order to change the multiplicity of secondaries, first of all, the leading particle needs to be excluded to prevent a change of the elasticity at the same time. The remaining particles are grouped together with respect to their type: nucleons, pions, kaons, gammas, electrons and muons. To yield a different multiplicity, particles are removed or duplicated at random from these groups depending on an increasing ($f(E) > 1$) or decreasing ($f(E) < 1$) multiplicity. After the particle number has been adapted to be $n_{\mathrm{mult}}^{\mathrm{prior}} \cdot f(E)$ the *kinetic* energy per particle group is scaled to yield exactly the same *total* energy as prior to the resampling. This assures energy conservation as well as a constant energy ratio between the particle type groups. To conserve the charge, particles are changed to their antipartners within particle type groups, until the total charge balance is restored up to a maximum charge offset of ± 1. Finally, all particle momenta are re-computed. Since CONEX is a 1D-EAS simulation code, no special efforts are made to conserve p_{t}.

Although, it is recognized that this approach to determine the impact of a changing multiplicity on the resulting shower development may not be the ultimate one, it certainly is an important investigation with relevance for the interpretation of air shower data. The devel-

[7] Available at `http://www-ik.fzk.de/~rulrich`.

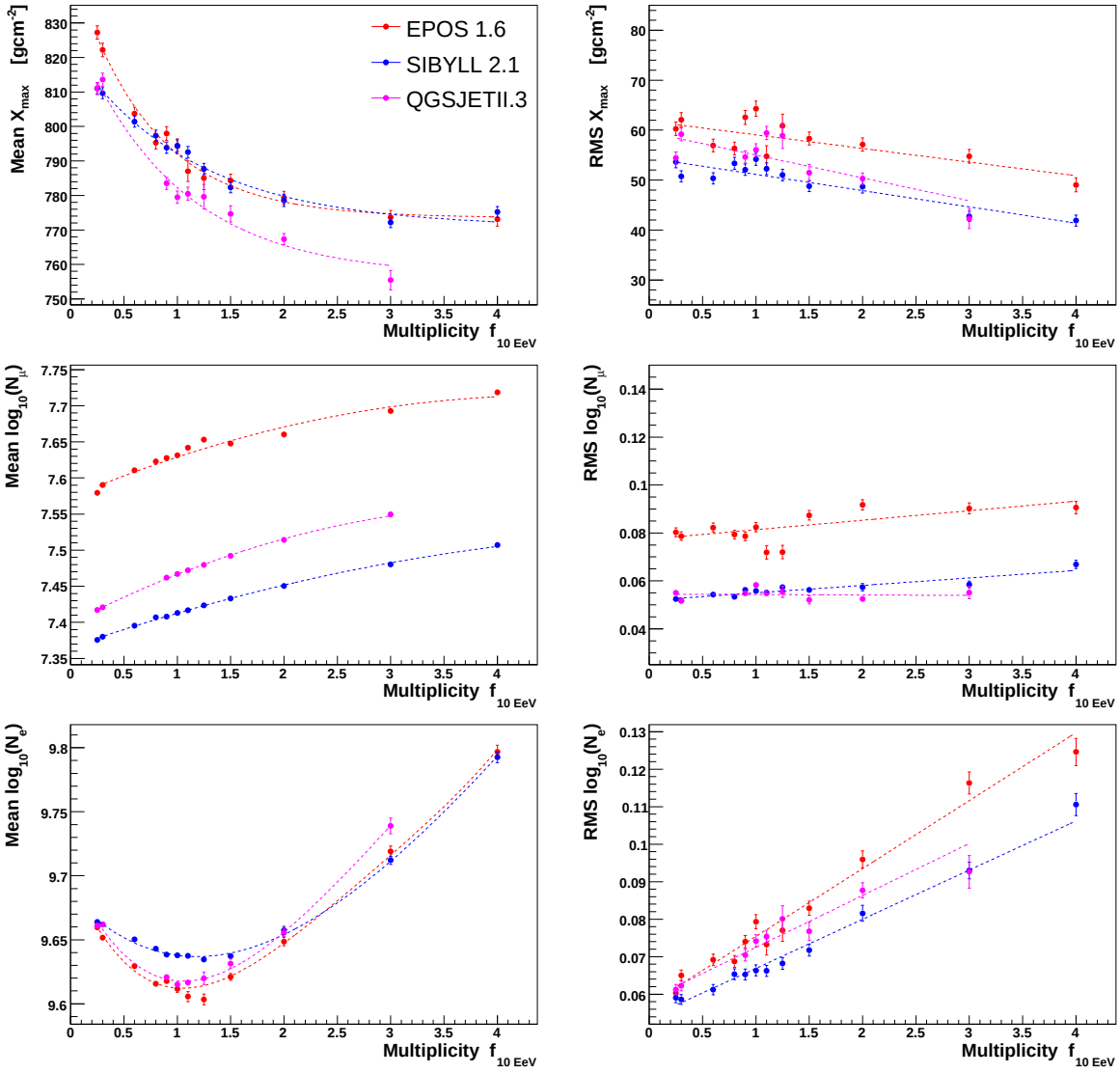


Figure 4.3: Impact of a modified secondary multiplicity on the mean and RMS of the observables X_{max}, $N_{\mathrm{e}}(X_{\mathrm{obs}} = 1000\,\mathrm{gcm}^{-2})$ and $N_{\mu}(X_{\mathrm{obs}} = 1000\,\mathrm{gcm}^{-2})$ for proton initiated air showers at 10 EeV. The dashed lines are just to guide the eye.

oped method allows us to resample secondary particles of hadronic interactions in order to increase or decrease the multiplicity with as few assumptions as possible. No particle energy spectra or any other model is utilized. Of course artificial fluctuations are introduced since only existing particles can be deleted or duplicated. But the qualitative shape of the particle energy spectra, the energy ratio between particle groups, the charge balance and the total energy as well as the leading particle are preserved as far as possible.

The effect of a changed multiplicity on EAS observables is summarized in Figures 4.3 and 4.4. The simulations are done for proton primaries at 10 EeV with CONEX and the interaction models EPOS, SIBYLL and QGSJETII. The CONEX tracking energy threshold is set to 1 PeV, to require full interaction Monte Carlo simulation for all interactions above

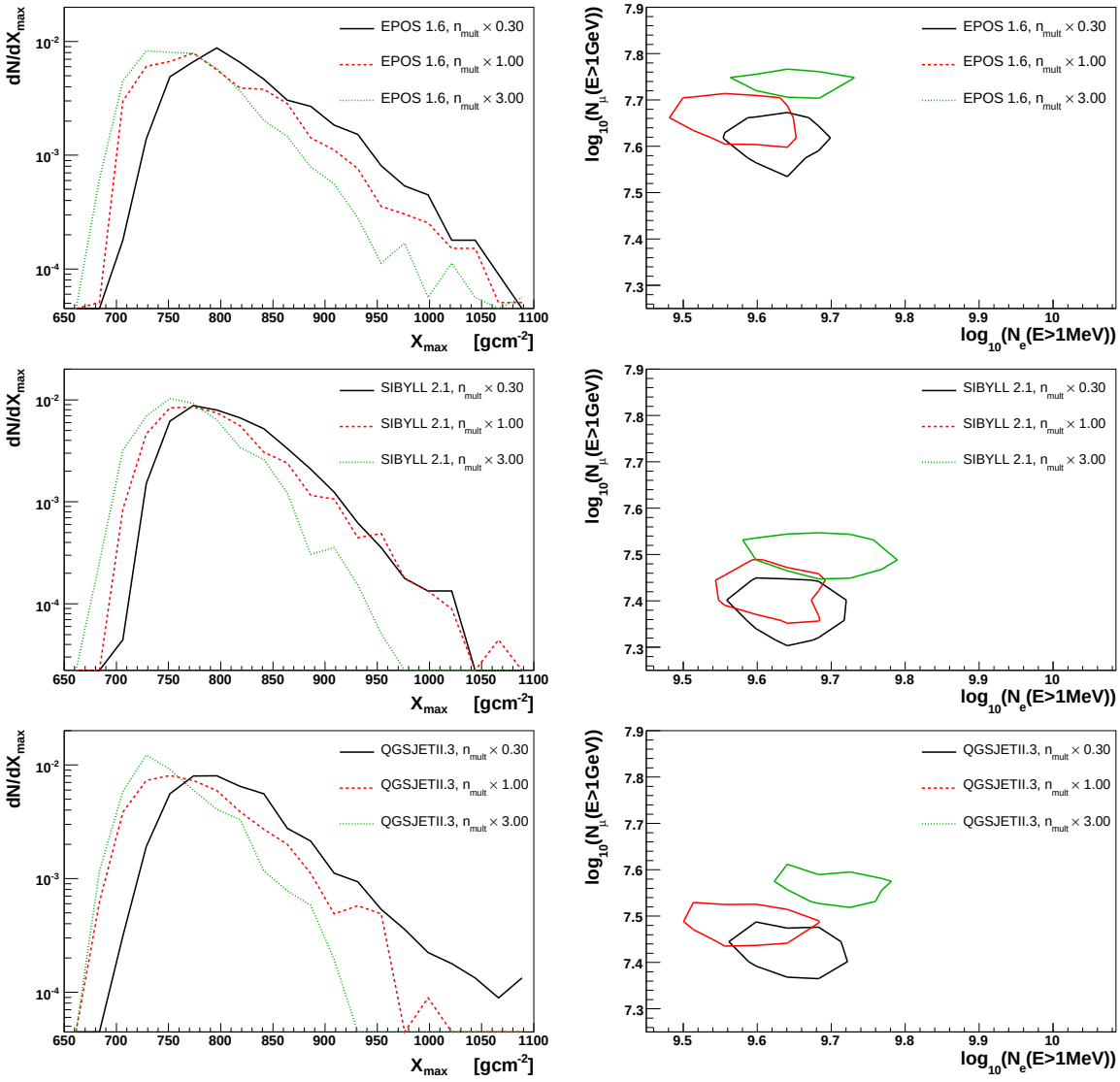


Figure 4.4: Distributions of the observables $X_{\max}$ (left) and $N_\mu(X_{\text{obs}} = 1000\,\text{gcm}^{-2})$ versus $N_\text{e}(X_{\text{obs}} = 1000\,\text{gcm}^{-2})$ (right) for modified secondary multiplicities with $f_{10\,\text{EeV}}$ values 0.3, 1 and 3 simulated for the interaction models EPOS 1.61, SIBYLL 2.1 and QGSJETII.3. The FWHM contour of the distributions are displayed in the right column. The simulations are preformed for proton initiated air showers at 10 EeV.

1 PeV. For each value of $f_{10\,\text{EeV}}$, 1000 showers are simulated. In Figure 4.3 the mean and RMS values for the resulting $X_{\max}$, N_μ and N_e distributions are shown for all the simulated multiplicity factors $f_{10\,\text{EeV}}$.

With an increasing multiplicity n_{mult} the $X_{\max}$-distribution shifts to shallower depth, as expected from the Heitler model. This is a consequence of distributing the available energy onto a growing number of particles. The resulting lower energy electromagnetic subshowers reach their maximum earlier. The impact on the RMS of the $X_{\max}$-distribution is small, but there is a significant trend to less fluctuations for an increasing number of secondaries. Interestingly, the number of muons N_μ itself is affected only slightly, besides the changing

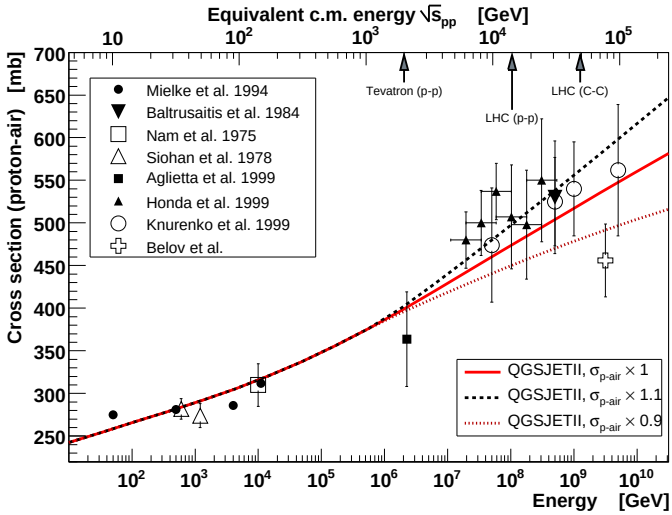


Figure 4.5: Example of a modified cross section for QGSJETII.3 for a 10% increase and decrease of $f_{10\,\mathrm{EeV}}$. See Figure 2.20 for the references.

number of produced secondaries and thus muons per interaction. This is due to the fact that muons are mostly produced by decaying mesons. While, for the case of an increasing multiplicity, the number of secondaries and thus mesons is growing, the energy per particle is reduced correspondingly. This is leading to an earlier dropping of the energy per meson below the critical energy, where they then stop to interact and hence no longer participate in multiplicative particle production, but just decay. This is partly compensating the effect of increasing particle number by the increased multiplicity, and takes care of the muon number after 1000 gcm^{-2} of shower development being relative stable with respect to a changing multiplicity. The impact on the electron number at the observation level X_{obs} =1000 gcm^{-2} shows a pronounced minimum close to $f_{10\,\mathrm{EeV}} = 1$ for all interaction models. The rising trend in the direction of smaller n_{mult} can be explained by the increase of X_{max} and therefore the shower maximum comes closer to the observation level. On the other hand the rising trend in the direction of larger n_{mult} is again just the consequence of a generally growing number of secondary particles. It is an interesting finding that all models are built in order to predict a close to minimal number of electrons N_{e}. In contrary to muons, the RMS is significantly changing while n_{mult} gets larger. This can be explained by comparing to Figure 2.9, where it is shown that N_{e}-fluctuations are strongly depending on the distance to the shower maximum.

In Figure 4.4 some of the relevant distributions for three selected values of $f_{10\,\mathrm{EeV}}$ for all interaction models are shown. The left column demonstrates how the X_{max}-distributions are shifted to deeper respectively shallower depth by changing the multiplicity. The right column illustrates how the N_{e}/N_{μ}-distribution moves slightly to the lower left for a decreasing multiplicity and to the upper right for an increasing multiplicity.

Cross section. Modifying the cross section can be implemented straightforwardly. No secondary particle resampling is needed. Only the extrapolated value of all hadronic cross sections has to be multiplied by $f(E)$ (see Figure 4.5). This is not only done for the primary

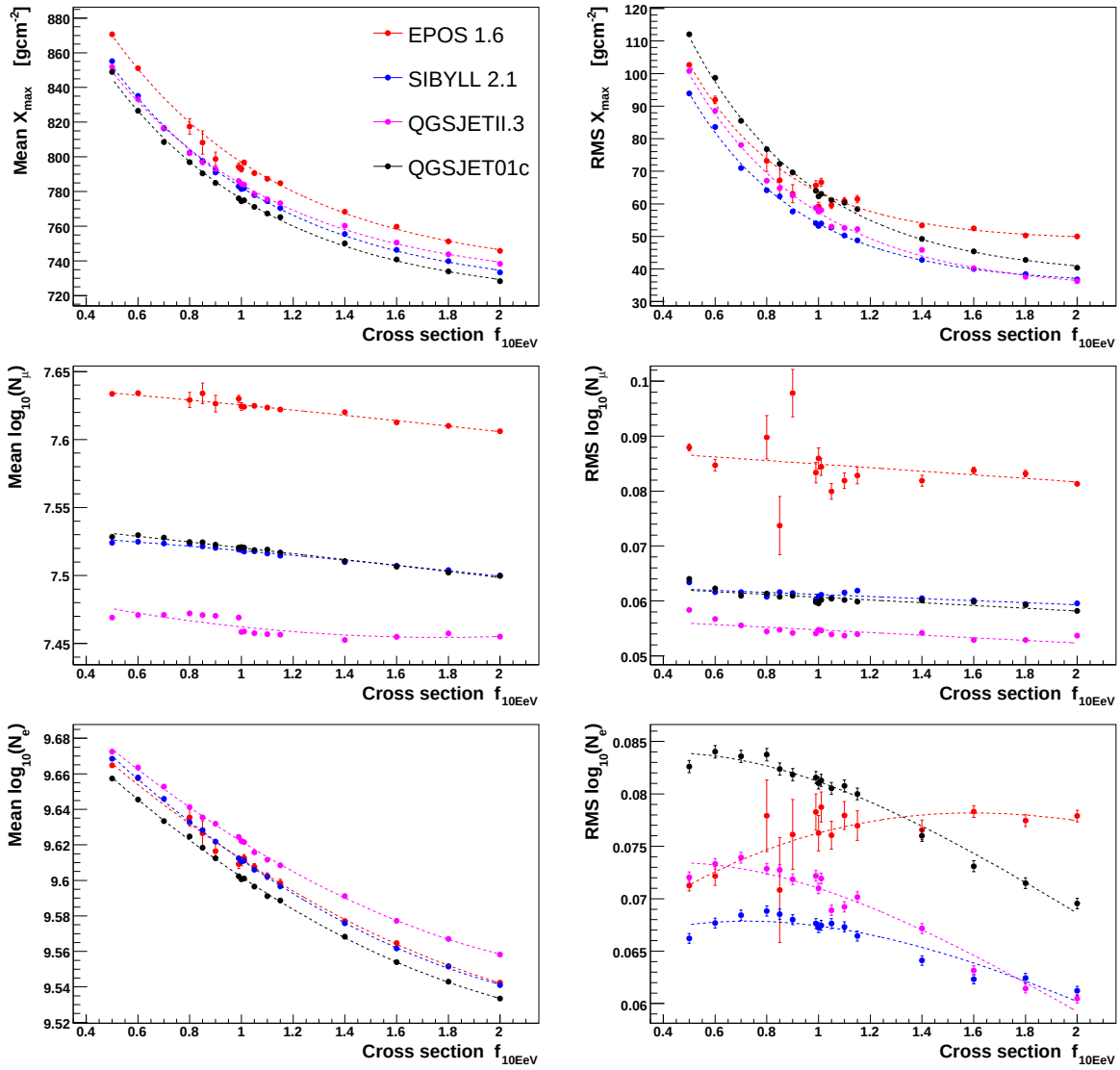


Figure 4.6: Impact of a modified cross section on the mean and RMS of the observables $X_{\max}$, $N_e(X_{\text{obs}} = 1000\,\text{gcm}^{-2})$ and $N_\mu(X_{\text{obs}} = 1000\,\text{gcm}^{-2})$ for proton initiated air showers at 10 EeV. The dashed lines are just to guide the eye.

particle and hence $\sigma_{\text{p-air}}$, but for all corresponding hadronic interaction cross sections in the EAS above the chosen transition energy of 1 PeV. The impact on the distribution of $X_{\max}$, N_μ and N_e is shown in Figures 4.6 and 4.7.

The mean as well as the RMS of the $X_{\max}$-distribution are dropping with a larger cross section. The effect is very pronounced, since the depth of the first interaction X_1 is affected as well as the shower startup phase. Both effects are pointing into the same direction. This makes $X_{\max}$ a sensitive observable for a cross section measurement. The impact on the muon number N_μ is not very large. Since the shower maximum moves away from the observation level with increasing cross section, we just see the slow decrease of the muon number at late shower development stages, while the fluctuation of N_μ stays basically constant (compare

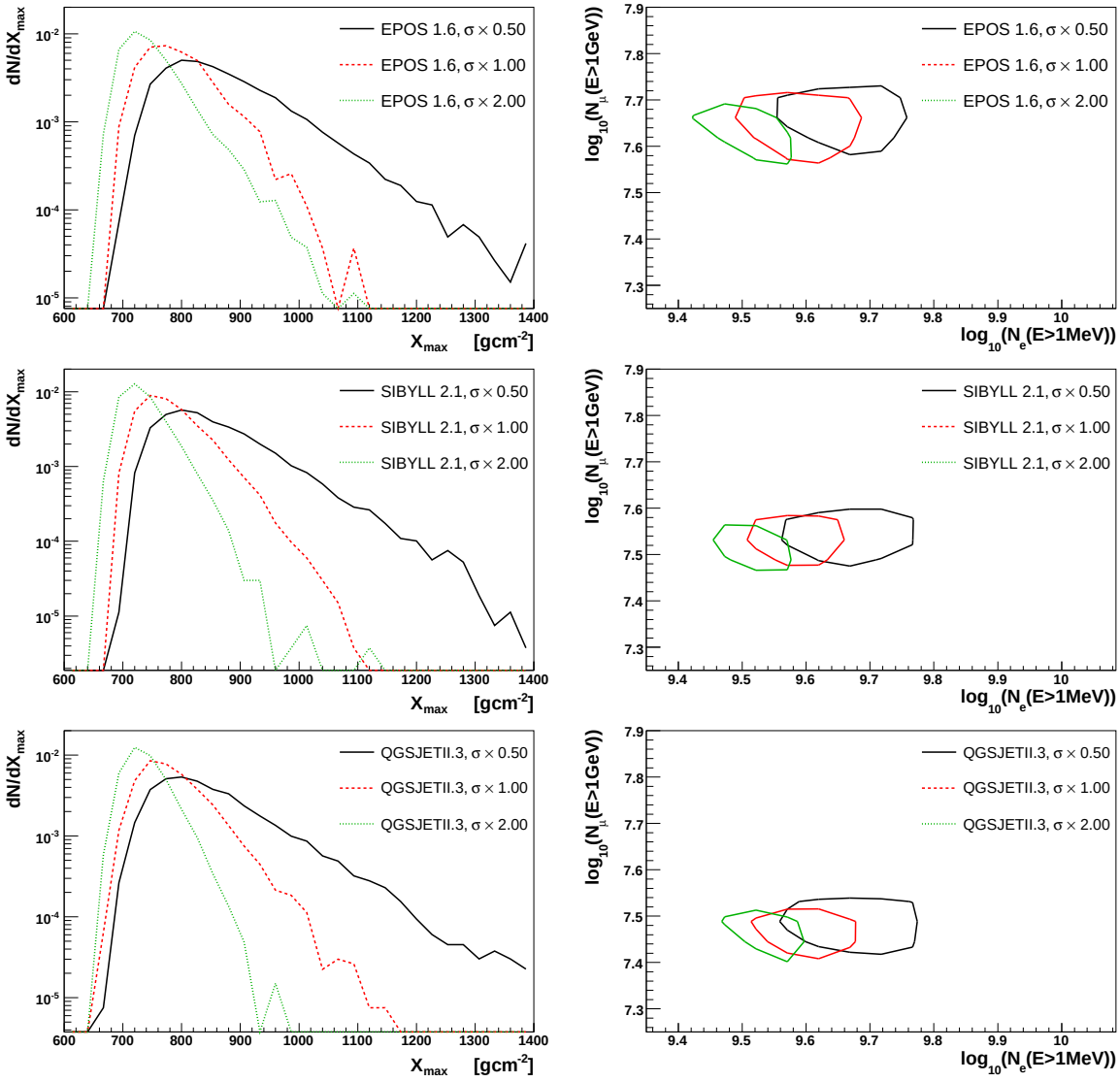


Figure 4.7: Distributions of the observables $X_{\max}$ (left) and $N_\mu(X_{\text{obs}} = 1000\,\text{gcm}^{-2})$ versus $N_\text{e}(X_{\text{obs}} = 1000\,\text{gcm}^{-2})$ (right) for modified cross sections with $f_{10\,\text{EeV}}$ values 0.5, 1 and 2 simulated for the interaction models EPOS 1.61, SIBYLL 2.1 and QGSJETII.3. The FWHM contour of the distributions are displayed in the right column. The simulations are preformed for proton initiated air showers at 10 EeV.

Figure 2.9). The generally large number of muons in EPOS is a consequence of an increased importance of baryons as secondaries of hadronic interactions [96]. This leads to a larger number of hadronic interactions, and subsequently to the production of more muons. The mean electron number as well as its fluctuations depend strongly on the distance of $X_{\max}$ from the observation level (also compare Figure 2.9). Combined with the influence of the modified cross section on $X_{\max}$ this explains well the strong decrease of the mean N_e as well as the RMS with increasing cross section. For very low cross sections, the shower maximum comes closer to the observation level, which can be observed as a decrease of the fluctuations in N_e. In contrary to all other models, EPOS shows a slight increase in the fluctuations of N_e,

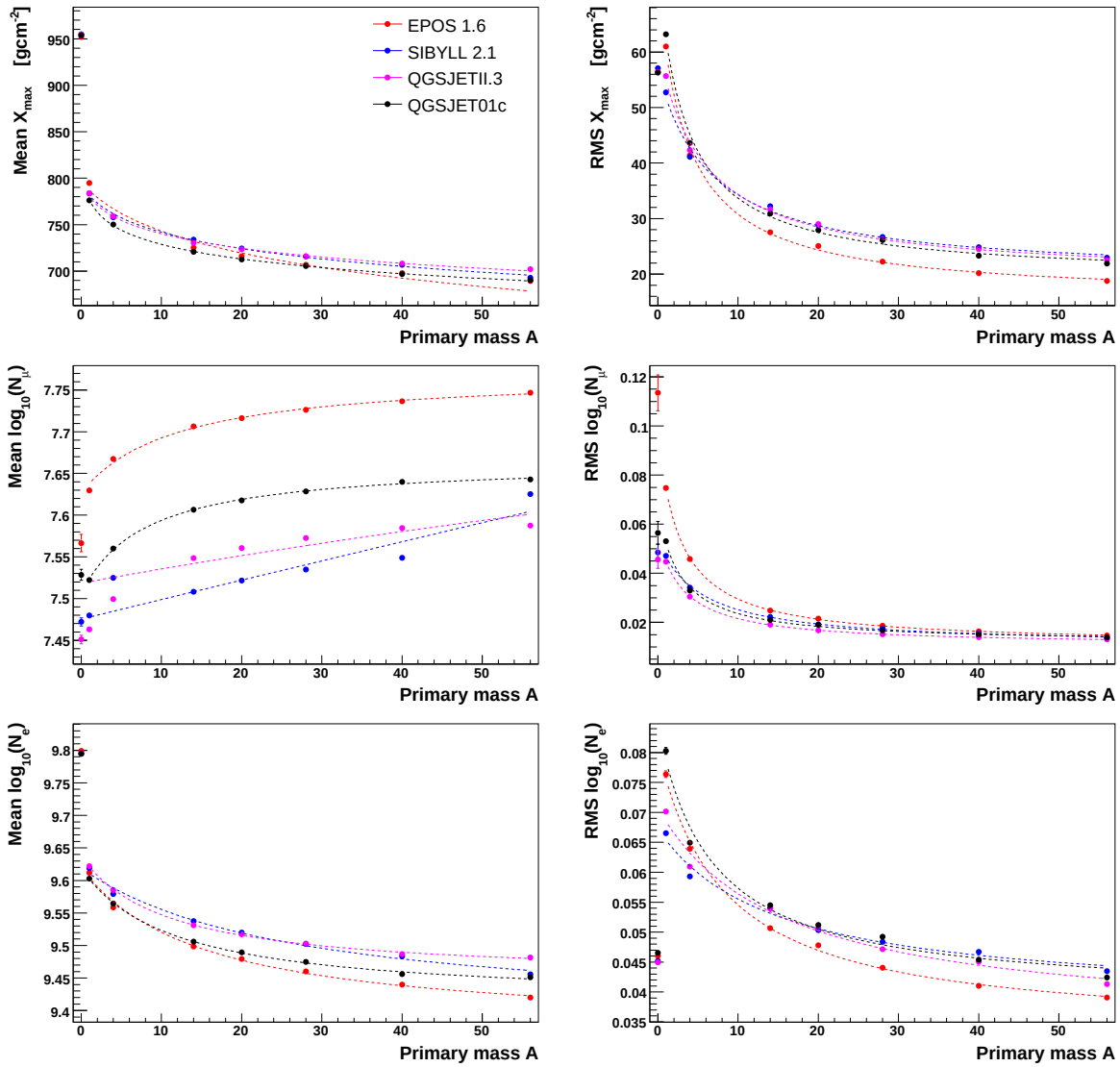


Figure 4.8: Impact of different primary cosmic ray particles on the mean and RMS of the observables X_{max}, $N_{\mathrm{e}}(X_{\mathrm{obs}} = 1000\,\mathrm{gcm}^{-2})$ and $N_{\mu}(X_{\mathrm{obs}} = 1000\,\mathrm{gcm}^{-2})$ for photon ($A = 0$), proton ($^{1}_{1}$H), helium ($^{4}_{2}$He), nitrogen ($^{14}_{7}$N), neon ($^{20}_{10}$Ne), silicon ($^{28}_{14}$Si), calcium ($^{40}_{20}$Ca) and iron ($^{56}_{26}$Fe) initiated air showers at 10 EeV. The dashed lines are just to guide the eye. They exclude the photon primaries, since they behave fundamentally different than hadronic primaries.

which is probably due to the production of π^0 mesons with a hard energy spectra.

The relevant distributions for three selected values of $f_{10\,\mathrm{EeV}}$ are shown in Figure 4.7. The left column demonstrates the dramatic change of the shape accompanied by a small shift of the X_{max}-distributions. The right column illustrates a clear shifting of the N_{e}/N_{μ}-distributions along a slightly inclined line.

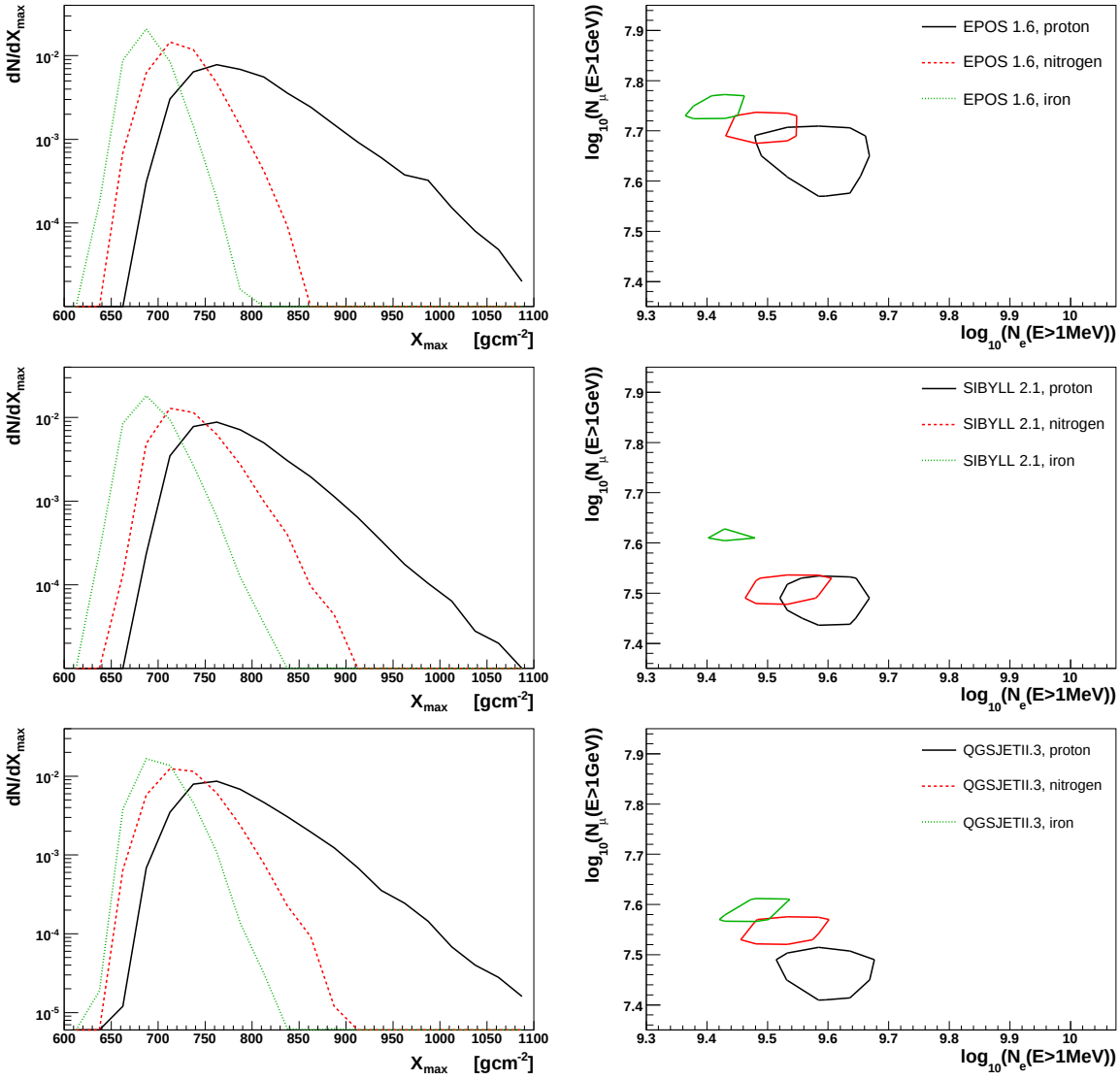


Figure 4.9: Distributions of the observables $X_{\max}$ (left) and $N_\mu(X_{\rm obs} = 1000\,\mathrm{gcm}^{-2})$ versus $N_{\rm e}(X_{\rm obs} = 1000\,\mathrm{gcm}^{-2})$ (right) for primary cosmic rays with mass $A = 1$ (proton), 7 (nitrogen) and 26 (iron) simulated for the interaction models EPOS 1.61, SIBYLL 2.1 and QGSJETII.3. The FWHM contour of the distributions are displayed in the right column. The simulations are preformed for proton initiated air showers at 10 EeV.

4.2.2 Impact of primary mass composition on air shower observables

It is phenomenologically well understood how nuclei of different mass A produce shower maxima at different depth $X_{\max}(A)$ and how shower-to-shower fluctuations decrease with A. A very successful model to relate mass A to proton primaries is the *semi-superposition* model [94]. It describes the resulting fluctuations from nucleus-nucleus interactions very well. The model assumes that only the N out of the A nucleons of the incident cosmic ray nucleus participate in the interaction that are located in the geometric overlap region of the projectile and the target nucleus [177]. The mean number of interacting nucleons is related

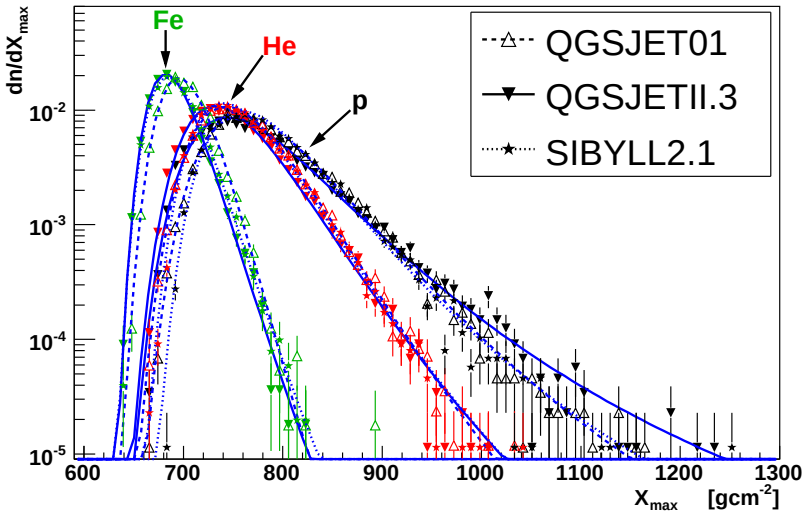


Figure 4.10: Impact of primary particle mass on $X_{\max}$ at 10 EeV for several models.

to the cross section

$$\langle N \rangle = \frac{A\, B\, \sigma_{\mathrm{pp}}}{\sigma_{\mathrm{AB}}}\,, \tag{4.2}$$

with the mass number A of the projectile and B of the target. Since $\langle N \rangle \leq A$ the predicted fluctuations of the semi-superposition model are in between the proton and the *superposition* model. The superposition model assumes that the A nucleons of the projectile are completely independent with an energy of E_0/A. Since the nucleons are not completely independent, but weakly bound (~ 8 MeV) in the nucleus, the predicted fluctuations are too small.

The relative change of the $X_{\max}$-distribution from a pure proton to a pure mass A primary composition can be evaluated using CONEX simulations. The impact of the primary mass A on the mean and RMS of the $X_{\max}$, N_{e} and N_{μ}-distributions is summarized in Figure 4.8. Since photon primaries are not hadronic, they are discussed separately. The model differences for photon induced showers are much smaller, since the simulation of electromagnetic interactions is performed with EGS4 [178], regardless of the hadronic interaction model. Also the extrapolation of electromagnetic interactions to air shower energies is not expected to produce large uncertainties. Only photo-hadronic interactions result in a small dependence on hadronic interaction models. This is why the mean and RMS of the N_{μ}-distributions for photon induced air showers are the only ones strongly affected by model dependence. Especially the mean and RMS of the $X_{\max}$-distribution and also the mean and the RMS of the N_{e}-distribution exhibit only a negligible model dependence for photon induced air showers.

The mean of $X_{\max}$ can be described by a $\propto \log(A^{-1})$ behavior. This is a consequence of the distribution of the total primary energy of the nucleus E_0 onto the A nucleons $E = {}^{E_0}/_{A}$. Only the EPOS model shows a slight deviation from $\propto \log(A^{-1})$. Furthermore, the RMS of $X_{\max}$ shows a very strong dependence, which makes it one of the prime composition-sensitive observables [26, 52, 27, 28, 179].

The mean muon number N_μ is growing with A, because of the increase in the number of hadronic interactions and therefore secondary mesons that subsequently decay into muons. The decreasing RMS with A is a consequence of the overall increased particle number.

Mostly as a consequence of the decreasing X_{max} and thus the shower maximum moving away from the surface detector, the mean number of electrons is decreasing with A. At the same time the RMS of N_{e} is getting smaller, due to the decreasing RMS of X_{max}.

In Figure 4.9 several of the resulting distributions of X_{max} and N_{e} versus N_μ are shown. The left row demonstrates the strong impact of the primary particle type on the resulting X_{max}-distribution. On the right side it is shown how the surface array observables N_{e} and N_μ can be utilized to analyze the primary mass composition. In Figure 4.10 it is illustrated how the X_{max}-distributions for proton, helium and iron primaries are positioned relative to each other for several models.

4.3 Parameterization of the X_{max}-distribution

Combining the results of the work so far, it is now possible to construct a model of the observable X_{max}. The fundamental description of the distribution of observed X_{max} in terms of $\sigma_{\mathrm{p-air}}$ and other hadronic interaction properties is based on (2.14) as described in Section 2.6.2, which is repeated here

$$\frac{\mathrm{d}P(X^{\mathrm{rec}}_{\mathrm{max}})}{\mathrm{d}X^{\mathrm{rec}}_{\mathrm{max}}} = P(X^{\mathrm{rec}}_{\mathrm{max}}) = \int \mathrm{d}X_1 \int \mathrm{d}\Delta X_1 \;\; \frac{e^{-X_1/\lambda_{\mathrm{p-air}}}}{\lambda_{\mathrm{p-air}}} \cdot \; P_1(\Delta X_1) \cdot \; P_{\mathrm{res}}(X^{\mathrm{rec}}_{\mathrm{max}} - X_{\mathrm{max}}) \,,$$

with $X_1 + \Delta X_1 = X_{\mathrm{max}}$. It is stated that the $X^{\mathrm{rec}}_{\mathrm{max}}$-distribution can be deduced from double folding the exponential distribution of X_1 with the shower development ΔX_1 and the detector resolution. In this form it does only apply for detectors with an ideal acceptance, which is absolutely not the case for telescope detectors (see Figure 4.11). Therefore, the additional folding of the detector acceptance into the final $X^{\mathrm{rec}}_{\mathrm{max}}$-distribution needs to be introduced. The acceptance function is defined as

$$P_{\mathrm{acc}}(X^{\mathrm{rec}}_{\mathrm{max}},\, X^{i}_{\mathrm{low}},\, X^{i}_{\mathrm{up}}) = \begin{cases} 1 & \text{for } \; X^{i}_{\mathrm{low}} < X^{\mathrm{rec}}_{\mathrm{max}} < X^{i}_{\mathrm{up}} \\ 0 & \text{otherwise} \end{cases} \,. \tag{4.3}$$

Because of the complexity of the resulting model, a direct unfolding of the measured $X^{\mathrm{rec}}_{\mathrm{max}}$-distribution to obtain the original X_1-distribution, as it was proposed by the HiRes Collaboration [109], is not feasible any more. Thus, the developed forward folded model is used as input for a log-likelihood minimization. This is important, since it allows us to extend the fit to regions of the X_{max}-distribution with very few or even zero entries by assuming correct Poissonian fluctuations. Additionally, by substituting $\sigma_{\mathrm{p-air}} = \lambda_{\mathrm{p-air}} / \langle m_{\mathrm{air}} \rangle$ during the application of the X_{max}-model, the uncertainty on the cross section can be derived directly from the $\chi^2_{\mathrm{logP}} + 1$-contour line. The only remaining free parameters of the model are the cross section $\sigma_{\mathrm{p-air}}$ and the additionally introduced parameter X_{shift}

$$P(X^{\mathrm{rec}}_{\mathrm{max}}) = f\left(\sigma_{\mathrm{p-air}},\, X_{\mathrm{shift}}\right) \,, \tag{4.4}$$

where X_{shift} allows us to shift the P_1-distribution to $\Delta X_1 \rightarrow \Delta X_1 + X_{\mathrm{shift}}$. The explicit introduction of the X_{shift} parameter is necessary in order to reduce the model dependence

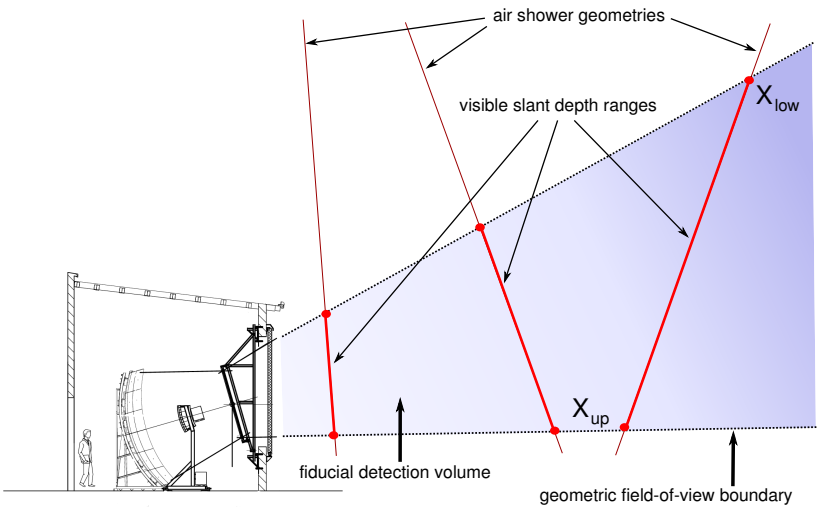


Figure 4.11: Fiducial detection phase space. For any given shower geometry there exist the points X_{low} and X_{up}, where the air shower enters respectively leaves the field-of-view of a telescope. Since X_{low} and X_{up} are most conveniently expressed in integrated traversed atmospheric matter (slant depth), they also depend on the changing atmospheric density profile.

of the reconstruction. This can be easily understood by looking at the P_1-distributions in Figure 2.28 (left) (on page 34), which are shifted with respect to each other. It is found that the X_{shift} parameter is highly model dependent and comprises many additional effects of the characteristics of high energy hadronic interactions on the X_{max}-distribution. In Section 4.2.1 it is demonstrated how the secondary multiplicity may act as the expected prime cause for such an additional shift (see Figure 4.3, top row and Figure 4.4, left column). But also other less important effects like the elasticity may contribute to X_{shift}. The only critical assumption related to the existence of X_{shift} for the cross section reconstruction is that any additional and unknown changing characteristics of hadronic interactions at extreme energies contributes only to a global shift of X_{max} and thus ΔX_1, while leaving the shape of the distributions unchanged. More details on X_{shift} are given in Section 4.4.2. The full model of $X^{\text{rec}}_{\text{max}}$ can now be written as

$$\frac{\mathrm{d}N(X^{\text{rec}}_{\text{max}})}{\mathrm{d}X^{\text{rec}}_{\text{max}}} = \left(\int \frac{\mathrm{d}N_{\text{acc}}(X^{\text{rec}}_{\text{max}})}{\mathrm{d}X^{\text{rec}}_{\text{max}}} \cdot \mathrm{d}X^{\text{rec}}_{\text{max}} \right)^{-1} \cdot \frac{\mathrm{d}N_{\text{acc}}(X^{\text{rec}}_{\text{max}})}{\mathrm{d}X^{\text{rec}}_{\text{max}}}, \tag{4.5}$$

with the distribution

$$\begin{aligned} \frac{\mathrm{d}N_{\text{acc}}(X^{\text{rec}}_{\text{max}})}{\mathrm{d}X^{\text{rec}}_{\text{max}}} = \sum_{i=1}^{n_{\text{events}}} P_{\text{acc}}(X^{\text{rec}}_{\text{max}}, X^i_{\text{low}}, X^i_{\text{up}}) \int_{X_1^{\text{min}}}^{X_1^{\text{max}}} \mathrm{d}X_1 \int_{\Delta X_1^{\text{min}}}^{\Delta X_1^{\text{max}}} \mathrm{d}\Delta X_1 \; \frac{e^{-X_1/\lambda_{\text{p-air}}}}{\lambda_{\text{p-air}}} \\ \cdot \; P_1(\Delta X_1 + X_{\text{shift}}, \lambda_{\text{p-air}}) \\ \cdot \; P_{\text{res}}(X^{\text{rec}}_{\text{max}} - X_{\text{max}}, \sigma_{X_{\text{max}}}(E^i)) \,. \end{aligned} \tag{4.6}$$

The normalization (4.5) is required in order to preserve the overall probability, since the detector acceptance P_{acc} is cutting into the tails of the distributions for every shower geometry i and thus destroys the conservation of probability. The sum in (4.6) runs over a set of n_{events} reconstructed shower events with $(E^i,\, X^i_{\mathrm{low}},\, X^i_{\mathrm{up}})$ in order to sample from realistic distributions for the energy and the detector acceptance.

The numerical evaluation of (4.5) is a computational demanding task, since the triple folding corresponds to an algorithmic expense of $N \cdot N \cdot N$, which means for a typical N of 1000 a single evaluation results in $\sim 10^9$ computation steps. During a minimization with `MINUIT` [180] the function (4.5) is called several thousand times. Therefore, some effort is spent in order to keep the computational demands low. Firstly, Romberg integration is used, which allows us to achieve a maximum precision with the least number of function evaluations. Secondly, the integrations in (4.6) are carried out only over the intervals in X_1 and ΔX_1, where a non-vanishing contribution to $X^{\mathrm{rec}}_{\mathrm{max}}$ is expected. This is from

$$X_1^{\mathrm{min}} = \max\left(0,\, X^{\mathrm{rec}}_{\mathrm{max}} + \Delta^{\mathrm{start}}_{X_{\mathrm{max}}} - \Delta X_1^{\mathrm{stop}}\right)$$

to

$$X_1^{\mathrm{max}} = \max\left(0,\, X^{\mathrm{rec}}_{\mathrm{max}} + \Delta^{\mathrm{stop}}_{X_{\mathrm{max}}} - \Delta X_1^{\mathrm{start}}\right)$$

respectively from

$$\Delta X_1^{\mathrm{min}} = \min\left(\max\left(\Delta X_1^{\mathrm{start}},\, X^{\mathrm{rec}}_{\mathrm{max}} + \Delta^{\mathrm{start}}_{X_{\mathrm{max}}} - X_1\right),\, \Delta X_1^{\mathrm{stop}}\right)$$

to

$$\Delta X_1^{\mathrm{max}} = \max\left(\min\left(\Delta X_1^{\mathrm{stop}},\, X^{\mathrm{rec}}_{\mathrm{max}} + \Delta^{\mathrm{stop}}_{X_{\mathrm{max}}} - X_1\right),\, \Delta X_1^{\mathrm{start}}\right),$$

where the parameterized P_1 and P_{res}-distributions are assumed to be non-zero over the range from $\Delta^{\mathrm{start}}_{X_{\mathrm{max}}}$ to $\Delta^{\mathrm{stop}}_{X_{\mathrm{max}}}$ and from $\Delta X_1^{\mathrm{start}}$ to $\Delta X_1^{\mathrm{stop}}$. The corresponding values depend on the chosen parameterizations and are given in Sections 4.4 and 5.2.

For the remaining parts of this section, where the performance of the X_{max}-model is thoroughly tested, only air shower profiles on the CONEX level plus additional Gaussian detector resolution are utilized. Thus it is not required to correct for the detector acceptance and therefore $P_{\mathrm{acc}} = 1$. The Gaussian detector resolution is chosen to be $20\,\mathrm{g cm^{-2}}$. And for any application of the log-likelihood fit an X_{max}-distribution with 2000 entries and a binning of $10\,\mathrm{g cm^{-2}}$ is generated.

4.4 Shower development and the ΔX_1-distribution

It is an important feature of the method presented here that the dependence of the ΔX_1-distribution on the characteristics of hadronic interactions at ultra-high energy is taken into account. In all previous EAS-based cross section measurements only the X_1-distribution is considered to change with $\lambda_{\mathrm{p-air}}$. But if the cross section changes at the primary cosmic ray energy, which is responsible for the fluctuation of the first interaction points X_1, it must also be different at lower energies and therefore affect the other very high energy interactions during the shower startup phase of the EAS. The resulting continuous variation of the cross section is illustrated in Figure 4.5. Also the ΔX_1-distribution must be affected by $\lambda_{\mathrm{p-air}}$. This

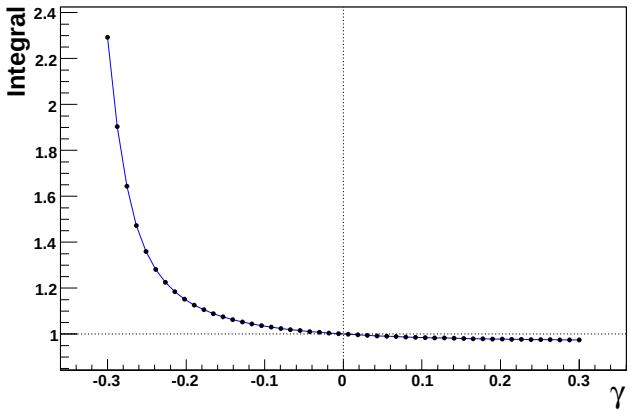


Figure 4.12: Numerically integrated normalisation of the ΔX_1-parameterization, as a function of the introduced γ parameter. For $\gamma = 0$ the parameterization gets identical to the Moyal function.

resembles well the fact that the first interaction can not be studied isolated from the resulting EAS development. It is only the superposition of the first interaction and the following ultra-high energy interactions during the shower startup phase (cf. Figure 2.10), which is responsible for the resulting air shower, as it is then observed by air shower detectors.

In order to build the $X_{\max}^{\text{rec}}$-model as given in (4.4), a good parameterization of ΔX_1 is needed. To achieve this, a modified version of the Moyal distribution is introduced

$$P_1(\Delta X_1) = \frac{N}{\beta \cdot \sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(z + e^{-z\cdot\left(z^2\right)^{\gamma}}\right)} \qquad \text{with} \qquad z = \frac{\Delta X_1 - \alpha}{\beta}\,. \tag{4.7}$$

The additional third degree of freedom γ makes the parameterization more flexible, allowing us to improve the description around the peak of the ΔX_1-distributions. The normalization value N is not considered a free parameter, but is always numerically computed to correctly normalize the distribution. Since the Moyal function with the two parameters α and β itself is normalized, we only have to correct for the γ parameter

$$N = \frac{1}{f(\gamma)}\,, \tag{4.8}$$

where the functional form of $f(\gamma)$ is displayed in Figure 4.12. The extended Moyal function (4.7) has a pronounced peak around the α parameter and vanishes rapidly for decreasing as well as increasing ΔX_1. There is hardly any contribution to the integration in (4.4) outside of the interval

$$\Delta X_1^{\text{start}} = \alpha - 10\,\beta \qquad \text{and} \qquad \Delta X_1^{\text{stop}} = \alpha + 30\,\beta\,,$$

which is therefore adopted to be the range of ΔX_1 to which any evaluation of (4.7) can be restricted to.

It will be shown in Section 4.5 that it is beneficial to require a minimum inelasticity of 0.01 for the first interaction to qualify as the starting point of an air shower. The quality of the resulting parameterization of P_1 can be enhanced by this requirement, while the resulting model dependence is slightly reduced. Thus, in the following, if not explicitly stated

Table 4.1: Resulting parameters of the energy dependence of ΔX_1-distributions.

Model	Index i	α_i^e [gcm^{-2}]	β_i^e [gcm^{-2}]	γ_i^e [1]
QGSJETII.3	0	-1322.33±75.12	253.02±32.51	4.66±1.05
	1	164.98±8.19	-25.52±3.53	-0.49±0.11
	2	-3.01±0.22	0.68±0.10	0.01±0.00
QGSJET01c	0	-771.55±69.57	102.77±36.97	-16.27±3.07
	1	101.99±7.58	-9.53±4.02	1.76±0.33
	2	-1.29±0.21	0.27±0.11	-0.05±0.01
SIBYLL 2.1	0	-602.08±65.62	58.31±31.18	-2.19±1.73
	1	82.27±7.10	-3.71±3.37	0.23±0.19
	2	-0.65±0.19	0.08±0.09	-0.01±0.01

otherwise, the minimum inelasticity for all analyses is always $k_{\text{inel}}^{\text{min}} = 0.01$. See Section 4.5 for the details.

4.4.1 Energy dependence of P_1

In order to determine the energy dependence of the parameters α, β and γ, the parameterization (4.7) is fitted to ΔX_1-distributions generated by CONEX at several primary energies. The energy dependence can then be interpolated by a polynomial function of 2$^{\text{nd}}$ degree

$$\begin{aligned}
\alpha^e(E) &= \alpha_0^e + \alpha_1^e \cdot \log_{10}(E/\text{eV}) + \alpha_2^e \cdot \log_{10}^2(E/\text{eV}) \\
\beta^e(E) &= \beta_0^e + \beta_1^e \cdot \log_{10}(E/\text{eV}) + \beta_2^e \cdot \log_{10}^2(E/\text{eV}) \\
\gamma^e(E) &= \gamma_0^e + \gamma_1^e \cdot \log_{10}(E/\text{eV}) + \gamma_2^e \cdot \log_{10}^2(E/\text{eV}) \,.
\end{aligned} \tag{4.9}$$

The chosen polynomial interpolation is well suited to reproduce the found dependences over the considered energy range. The results are illustrated in Figure 4.13 and the parameters are listed in Table 4.1. See Appendix B.1 for a complete listing of the fits and a tabular overview. A first result is that even the energy dependence of ΔX_1 is depending strongly on the high energy interaction model used during air shower simulations.

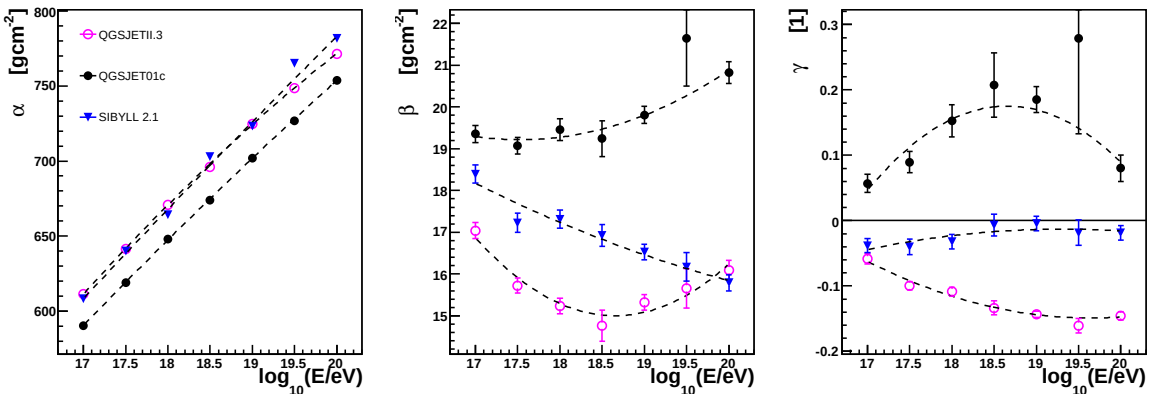


Figure 4.13: Parameterization of the energy dependence of the ΔX_1-distributions. The dashed lines denote the polynomial interpolations given in the text.

Table 4.2: Resulting parameters of the cross section dependence of ΔX_1-distributions.

Model	Index i	α_i^e [gcm^{-2}]	β_i^e [gcm^{-2}]	γ_i^e [1]
QGSJETII.3	0	758.14±0.99	8.60±0.35	-0.06±0.01
	1	-40.26±1.62	7.22±0.69	-0.14±0.02
	2	6.73±0.61	-0.04±0.04	0.05±0.01
QGSJET01c	0	735.30±0.81	10.36±0.46	0.44±0.04
	1	-41.83±1.36	11.46±1.05	-0.47±0.06
	2	8.27±0.52	-0.16±0.05	0.14±0.02
SIBYLL 2.1	0	771.91±0.83	10.13±0.28	0.16±0.02
	1	-62.61±1.38	6.28±0.49	-0.30±0.04
	2	12.55±0.52	0.04±0.03	0.10±0.01

4.4.2 Dependence of P_1 on the hadronic interaction characteristics

It is a fact that ΔX_1-distributions also depend on the characteristics of the hadronic interactions in the startup of air showers. So far this has been neglected by any air shower based cross section measurement. It is one of the main accomplishments of this thesis to work out the dependence of ΔX_1 from a changing cross section. In fact, it seems to be hopeless trying to measure the proton-air cross section without taking this into account. As argued previously, any other feature of high energy interactions is assumed to get compensated by the X_{shift} parameter, which corresponds to a global shift of the P_1-distribution as $\Delta X_1 \rightarrow \Delta X_1 + X_{\text{shift}}$. This assumption has to be kept in mind, while interpreting the resulting proton-air cross sections. Again, it must be stressed that any previous measurement approach not only ignored the changed shower development by a modified cross section, but also completely overlooked the possibility of any other unknown change of the characteristics of hadronic interactions contributing to the found signal in the analyzed distribution. Besides, there is clear evidence that the assumption of a cosmic ray proton beam, which is underlying any cross-section analysis, is indeed requiring a shifting of the X_{max}-distribution. This is can be seen in Figure 2.19 (right). Section 4.2.1 explains how the secondary multiplicity could act as the main source of such a shift. However, while it is possible to shift the X_{max}-distribution, the resulting impact on the shape of the X_{max}-distribution can not be determined explicitly. At least it demonstrated to be small (cf. Figure 4.3, upper-right).

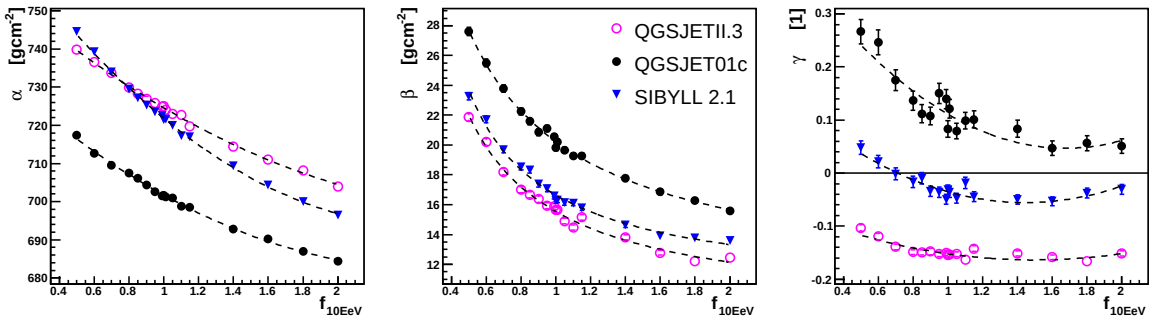
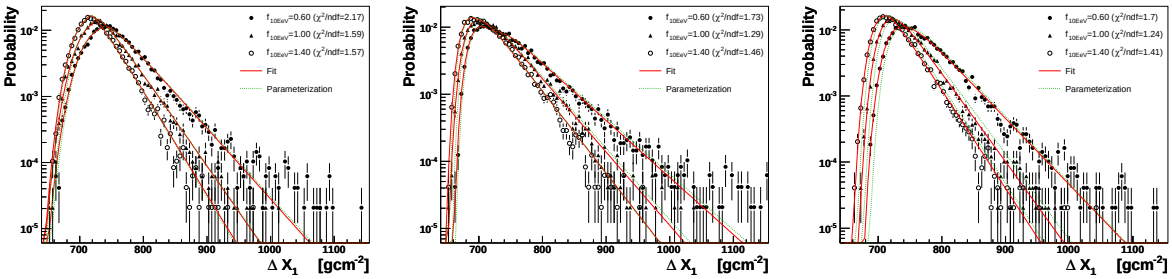


Figure 4.14: Parameterization of the $f_{10\,\text{EeV}}$-dependence of the ΔX_1-parameterizations with respect to a changing cross section. The dashed lines denote the interpolations as discussed in the text.



(a) QGSJETII.3 (b) QGSJET01c (c) SIBYLL 2.1

Figure 4.15: Examples of fitted and parameterized ΔX_1-distributions using the extended Moyal distribution (4.7).

In order to infer the dependence of P_1 on $\lambda_{\mathrm{p-air}}$ the modified CONEX version, as described in Section 4.2.1, was used to generate ΔX_1-distributions for several values of $f_{10\,\mathrm{EeV}}$. The modification factor $f(E)$ assumes that the model uncertainties rise proportional to the logarithm of the energy, while there are no model uncertainties at energies below 1 PeV, where accelerator data are available. These assumptions certainly are reasonable and conservative. Of course there are implications on the interpretation of the results of this thesis. A measured proton-air cross section at ultra-high energy does include the modified high energy interactions occurring during the air shower development at lower energies, assuming an interpolation of the measured cross section in $\log E$ down to the predictions of the hadronic model at 1 PeV. In other words, the result of the analysis is not just a cross section measurement at a fixed energy E_{obs}, but a line from the interaction model prediction of the cross section at 1 PeV up to the measured cross section at E_{obs}, like it is illustrated in Figure 4.5.

The resulting dependence of ΔX_1 on a changing cross section for air showers at 10 EeV is shown in Figure 4.14. See Appendix B.2 for a complete list of fits and a tabular overview. The dependence of the α, β and γ parameters of the modified Moyal distribution on $f_{10\,\mathrm{EeV}}$ is parameterized as follows

$$\begin{aligned}
\alpha^{\mathrm{x}}(f_{10\,\mathrm{EeV}}) &= \alpha_0^{\mathrm{x}} + \alpha_1^{\mathrm{x}} \cdot f_{10\,\mathrm{EeV}} + \alpha_2^{\mathrm{x}} \cdot f_{10\,\mathrm{EeV}}^2 \\
\beta^{\mathrm{x}}(f_{10\,\mathrm{EeV}}) &= \beta_0^{\mathrm{x}} + \beta_1^{\mathrm{x}} / (f_{10\,\mathrm{EeV}} - \beta_2^{\mathrm{x}}) \\
\gamma^{\mathrm{x}}(f_{10\,\mathrm{EeV}}) &= \gamma_0^{\mathrm{x}} + \gamma_1^{\mathrm{x}} \cdot f_{10\,\mathrm{EeV}} + \gamma_2^{\mathrm{x}} \cdot f_{10\,\mathrm{EeV}}^2 \,.
\end{aligned} \tag{4.10}$$

The resulting parameters of the interpolations are listed in Table 4.2. Furthermore, it is convenient to introduce the relative change as

$$\begin{aligned}
\Delta\alpha^{\mathrm{x}}(f_{10\,\mathrm{EeV}}) &= \alpha^{\mathrm{x}}(f_{10\,\mathrm{EeV}}) - \alpha^{\mathrm{x}}(1) \\
\Delta\beta^{\mathrm{x}}(f_{10\,\mathrm{EeV}}) &= \beta^{\mathrm{x}}(f_{10\,\mathrm{EeV}}) - \beta^{\mathrm{x}}(1) \\
\Delta\gamma^{\mathrm{x}}(f_{10\,\mathrm{EeV}}) &= \gamma^{\mathrm{x}}(f_{10\,\mathrm{EeV}}) - \gamma^{\mathrm{x}}(1) \,.
\end{aligned} \tag{4.11}$$

4.4.3 Parameterization of the ΔX_1-distribution

Combining the found dependence of ΔX_1-distributions on the energy E and the cross section modifier $f_{10\,\mathrm{EeV}}$, it is now possible to construct a parameterization of P_1 in terms of E and

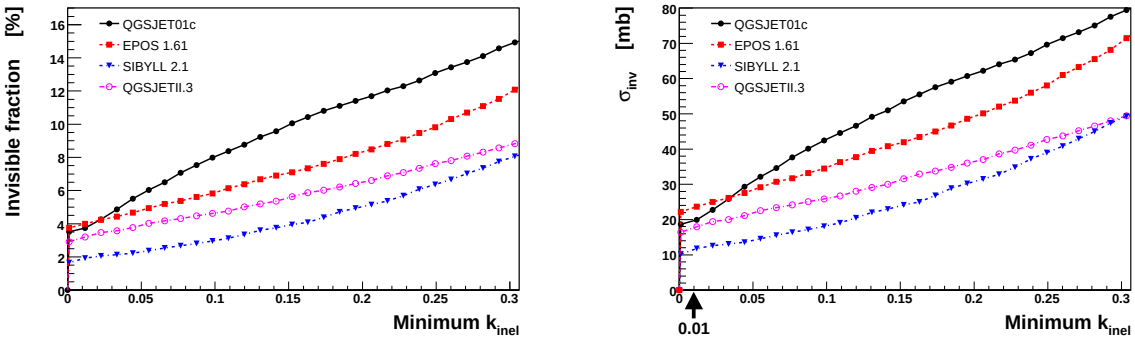


Figure 4.16: Model predictions for the fraction (left) and cross section (right) for interactions at 10 EeV with $k_{\text{inel}} \leq k_{\text{inel}}^{\text{min}}$. Naturally all models start with $\sigma_{\text{inv}} = 0$ for $k_{\text{inel}}^{\text{min}} = 0$. Due to the diffractive peak, σ_{inv} jumps to ∼20 mb at $k_{\text{inel}}^{\text{min}}$ just above 0. For increasing minimal inelasticities the model predictions are getting larger.

$\sigma(E)$. Firstly, the principle energy dependence of P_1 is evaluated by calculating $\alpha^{\text{e}}(E)$, $\beta^{\text{e}}(E)$ and $\gamma^{\text{e}}(E)$ using (4.9) and the results from Table 4.1. Secondly, the effect of the changed cross section is added. To achieve this, the first step is to calculate the corresponding deviation of the cross section $\sigma(E)$ extrapolated to 10 EeV assuming (4.1), which is

$$f_{10\,\text{EeV}} = 1 + \left(\frac{\sigma(E)}{\sigma_{\text{model}}(E)} - 1 \right) \frac{\log_{10}(10\,\text{EeV}/1\,\text{PeV})}{\log_{10}(E/1\,\text{PeV})}, \tag{4.12}$$

where $\sigma_{\text{model}}(E)$ is the original interaction model prediction of the cross section at energy E. The resulting factor $f_{10\,\text{EeV}}$ is then used to evaluate $\Delta\alpha^{\text{x}}(f_{10\,\text{EeV}})$, $\Delta\beta^{\text{x}}(f_{10\,\text{EeV}})$ and $\Delta\gamma^{\text{x}}(f_{10\,\text{EeV}})$ using (4.11) and the results listed in Table 4.2. In accordance to (4.1) the energy dependence is assumed to be logarithmic in energy and vanishing below 1 PeV

$$F(E) = \begin{cases} 0 & E \leq 1\,\text{PeV} \\ \log_{10}(E/1\,\text{PeV}) / \log_{10}(10\,\text{EeV}/1\,\text{PeV}) & E > 1\,\text{PeV} \end{cases}. \tag{4.13}$$

This yields the final set of parameters

$$\begin{aligned} \alpha(E, \sigma(E)) &= \alpha^{\text{e}}(E) + F(E) \cdot \Delta\alpha^{\text{x}}(f_{10\,\text{EeV}}) \\ \beta(E, \sigma(E)) &= \beta^{\text{e}}(E) + F(E) \cdot \Delta\beta^{\text{x}}(f_{10\,\text{EeV}}) \\ \gamma(E, \sigma(E)) &= \gamma^{\text{e}}(E) + F(E) \cdot \Delta\gamma^{\text{x}}(f_{10\,\text{EeV}}). \end{aligned} \tag{4.14}$$

In Figure 4.15 the resulting parameterization is plotted together with the fits and the CONEX simulations for a few example distributions. A complete list of all generated ΔX_1-distributions together with the fits and parameterization is given in Appendix B.

4.5 Invisible cross section

Elastic and diffractive interactions of primary cosmic ray particles in the atmosphere are not contributing significantly to the air shower development. For a cross section measurement which relies on the interpretation of air shower observables there can not be any sensitivity to these interactions. An effective way to classify interactions is to consider the *inelasticity*

$k_{\mathrm{inel}} = 1 - E_{\mathrm{max}}/E_{\mathrm{tot}}$, where E_{max} is the energy of the leading particle and E_{tot} the total energy. It is a measure of how much of the total energy was used for new particle production. Elastic interactions have $k_{\mathrm{inel}} = 0$ and diffractive interactions k_{inel} very close to 0. It is possible to require a minimum inelasticity $k_{\mathrm{inel}}^{\mathrm{min}}$ to qualify an interaction for being relevant for air shower development. With respect to an air shower based cross section analysis it is possible to define an *invisible* fraction of the total cross section

$$\sigma_{\mathrm{inv}} = \sigma(k_{\mathrm{inel}} < k_{\mathrm{inel}}^{\mathrm{min}}) \,, \tag{4.15}$$

with a vanishing importance for air shower development. The model predictions for σ_{inv} at 10 EeV are shown in Figure 4.16. It is also found that the elasticity of interaction models itself is a source of uncertainty.

If interactions below a minimal inelasticity are omitted from the cross section analysis, the resulting production cross section needs to be increased by the model dependent correction σ_{inv} to get the production cross section (for the definition of the components of the cross section see Figure 2.22 on page 28). This can be achieved by considering only interactions with $k_{\mathrm{inel}} \geq k_{\mathrm{inel}}^{\mathrm{min}}$ for the construction of the ΔX_1-distribution, which is then used for the cross section analysis. Since standard air shower simulations are by default properly considering every type of interaction, their application does usually implicitly correct for the invisible cross section.

In Figure 4.18 the resulting ability to reconstruct the input cross section for several choices of $k_{\mathrm{inel}}^{\mathrm{min}}$ is illustrated. Apparently the only significant impact on the result is for the case of $k_{\mathrm{inel}}^{\mathrm{min}} = 0.01$, corresponding just to the omission of the diffractive peak. Larger choices of the minimal inelasticity do not help to further constrict the resulting model dependence. For this thesis $k_{\mathrm{inel}}^{\mathrm{min}} = 0.01$ is generally adopted for all analyses, if not stated otherwise explicity.

4.6 Fitting range and stability

The interval of the X_{max}-distribution used to fit the model is important for the resulting statistical as well as systematical uncertainties. The former is more obvious, since a reduction of the dataset is clearly leading to a reduced statistical power of the reconstruction. The latter is because of the contamination of the X_{max}-distribution by cosmic ray primaries other than

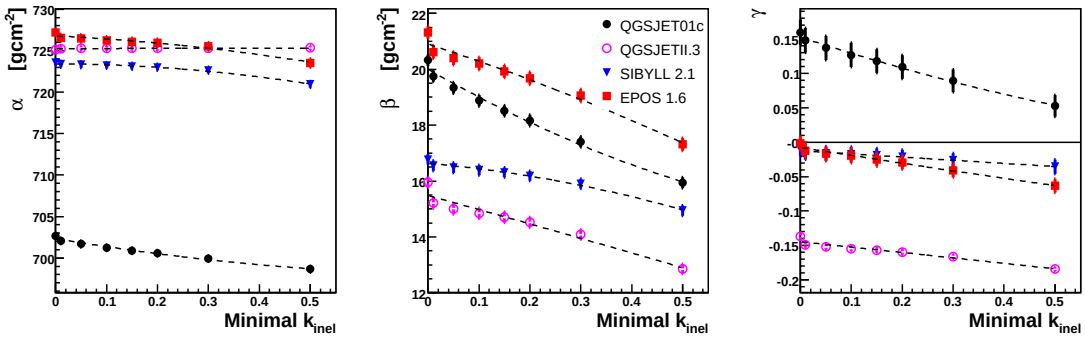


Figure 4.17: Impact of a minimal required inelasticity $k_{\mathrm{inel}}^{\mathrm{min}}$ on the parameterization of ΔX_1 at 10 EeV. The dashed lines are 2$^{\mathrm{nd}}$-order polynomial fits to the simulations, plotted in order to guide the eye.

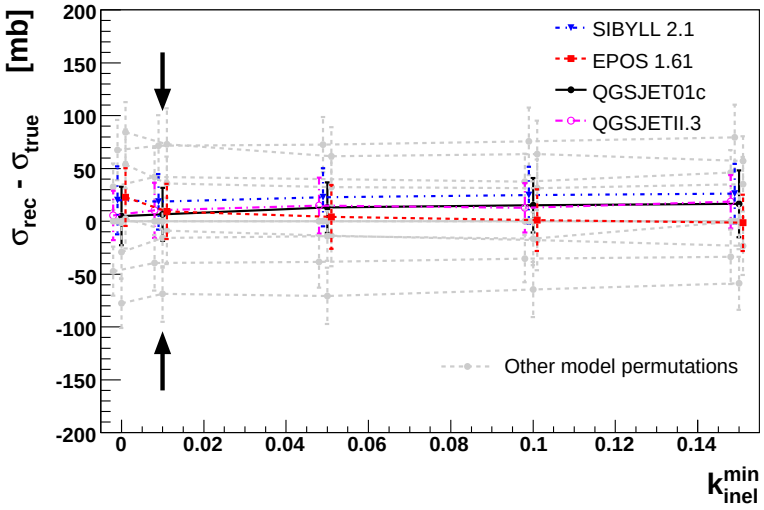


Figure 4.18: Impact of the chosen $k_{\text{inel}}^{\text{min}}$ on the model dependence of the cross section analysis at 10 EeV. Each point is the resulting mean of 100 reconstructions and the error bars denote the corresponding RMS. The adopted X_{max}-resolution is a Gaussian with a width of 20 gcm^{-2}. All generated X_{max}-distributions have 2000 entries. The graphs explicitly stating a model are denoting the cases, where the X_{max} as well as ΔX_1-distribution are generated by the given model.

protons. All primary nuclei are heavier than protons and are producing shallower X_{max} compared to protons. Primary photons, on the other hand, are deeply penetrating and have a larger X_{max} than protons. A restricted fitting range in X_{max} can thus be used to enrich the considered fraction of protons and reduce a possible contamination of other cosmic ray primaries.

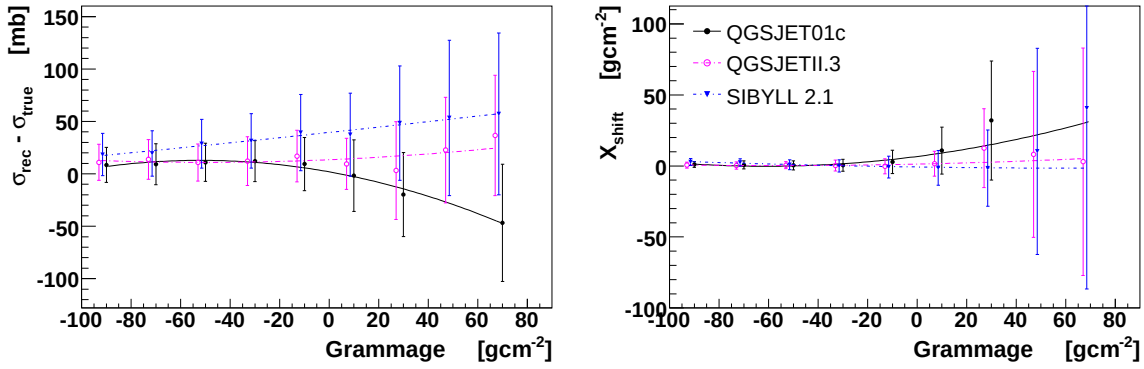
The resulting impact of the chosen fitting range on the performance of the reconstruction is shown in Figure 4.19. The position of the peak of the X_{max}-distribution, X_{peak}, is used as a reference to define the fitting range. The starting point as well as the ending point of the fit are expressed only relative to X_{peak}. This is what is plotted in Figure 4.19. The peak of the X_{max}-distribution is determined by a log-likelihood fit of

$$\frac{\mathrm{d}N}{\mathrm{d}X_{\text{max}}} = N \cdot e^{-2\,\sigma^{-2}\,(X_{\text{max}}-X_{\text{peak}})^2\,(X_{\text{max}}-X_{\text{peak}}+3\,\delta)^{-2}} \tag{4.16}$$

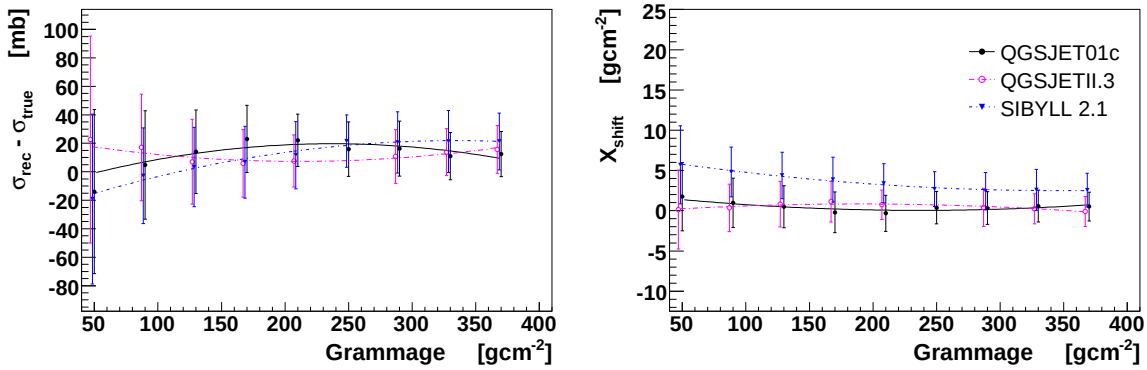
to the X_{max}-distribution [181]. Thus, the pre-fit of (4.16) needs to performed prior to any cross section analysis, in order to define the fitting range.

Evidently it is beneficial to chose the fitting range as large as possible, in order to get the smallest resulting statistical uncertainties. However, it is possible to shift the beginning of the fitting range relatively close to X_{peak} without loosing too much statistical power (cf. Figure 4.19 (a)). It is found that within the systematic uncertainties resulting from the parameterization of P_1 and the statistical uncertainties of the reconstruction, the beginning of the fitting range can be set to 50 gcm^{-2} in front of X_{peak}. In the following this is the adopted default choice, if not stated otherwise.

Correspondingly the choice of the end of the fitting range has a similar impact on the reconstruction. In Figure 4.19 (b) it is shown how the reconstruction is degrading, while



(a) Start of the fitting range, $X_{\mathrm{max}}^{\mathrm{start}}$ (relative to the peak of the X_{max}-distribution)



(b) End of the fitting range, $X_{\mathrm{max}}^{\mathrm{stop}}$ (relative to the peak of the X_{max}-distribution)

Figure 4.19: Impact of the chosen fitting range in $X_{\mathrm{max}}^{\mathrm{rec}}$ on the resulting cross section (left) and X_{shift} (right) at 10 EeV. Each point corresponds to the mean of 100 reconstructions and the error bars denote the resulting RMS. The adopted X_{max}-resolution is a Gaussian with a width of 20 gcm^{-2}. Both, the reconstructed X_{max} as well as the ΔX_1-distribution are produced with the specified interaction model. Each X_{max}-distribution contains 2000 entries.

choosing a shorter fitting range. Since the photon fraction of ultra-high energy cosmic ray primaries is already strongly constrained [182] there is no special need to restrict the upper end of the fitting range. In the following the end of the fitting range for the log-likelihood fit of the X_{max}-distribution is therefore set to the value of the maximum X_{max} plus 40 gcm^{-2}.

In order to analyze the ability of the reconstruction method to recover a changing input cross section, air shower simulations with a modified cross section (cf. Section 4.2.1) have to be performed. The resulting X_{max}-distributions are then reconstructed and the found cross sections are compared to the modified input cross sections $\sigma_{\mathrm{mod}} = f_{10\,\mathrm{EeV}} \cdot \sigma_{\mathrm{org}}$. The result is given in Figure 4.20. Within a given interaction model, it is possible to retrieve a changing cross section over a wide range. On the other hand, the resulting maximum model systematics is getting large with increasing cross section. The very different behavior to reconstruct smaller compared to larger cross section is resulting from the increasing importance of the fluctuations of X_1 for the final X_{max}-distribution. This facilitates the measurement of a small cross section, while for large cross sections X_{max}-distributions are mostly shaped by fluctuations of the shower development P_1 and the detector resolution, making a measurement of

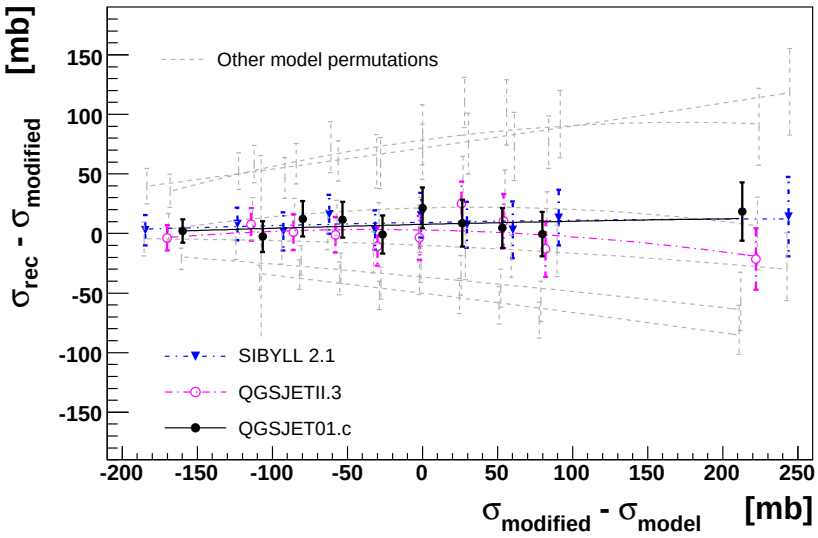


Figure 4.20: Ability to reconstruct a changing input cross section at 10 EeV. Each point is the mean of 100 reconstructions and the error bars denote the RMS. The adopted X_{max}-resolution is a Gaussian with a width of 20 gcm^{-2} and all X_{max}-distributions contain 2000 entries. The graphs explicitly stating a model are denoting the cases, where the X_{max} as well as ΔX_1-distribution are generated by the given model.

the cross section very difficult.

It must be stated that the magnitude of the found model dependence is not a property of the chosen cross section reconstruction ansatz. The model predictions for X_{max}-distributions are diverging too much by themselves. Every cross section analysis based on air shower data has to deal with a comparable model dependence. This also applies for the *k*-factor methods and especially for ground array based analyses, which acquire additional model dependence by the air shower development from X_{max} to the observation level.

4.7 Comparison to previous measurements techniques

Previous reconstruction techniques assume a static ΔX_1-distribution or *k*-factor, which are independent of the modified cross section. As the reconstructed cross section deviates from the predicted value of the hadronic interaction model used to generate the ΔX_1-distribution or *k*-factor, a discrepancy of the data with the air shower simulations is inevitable. Obviously the data can not be described by the models used for the air shower simulations and it is certainly not enough to adopt just the cross section of the first interaction point, while keeping the rest of the air shower simulations, which are shown to be wrong. The reconstruction ansatz presented in this thesis is considering a changing air shower development for a changing cross section by explicitly parameterize the dependence of P_1 on the cross section. Compare Figure 4.21 to the previously shown Figure 4.20 to get an impression on the improvements made to cross section analyses. By ignoring the impact of a changing cross section on the resulting air shower development, the reconstruction of a cross section differing from the original model prediction is not feasible. From Figure 4.21 it can be deduced

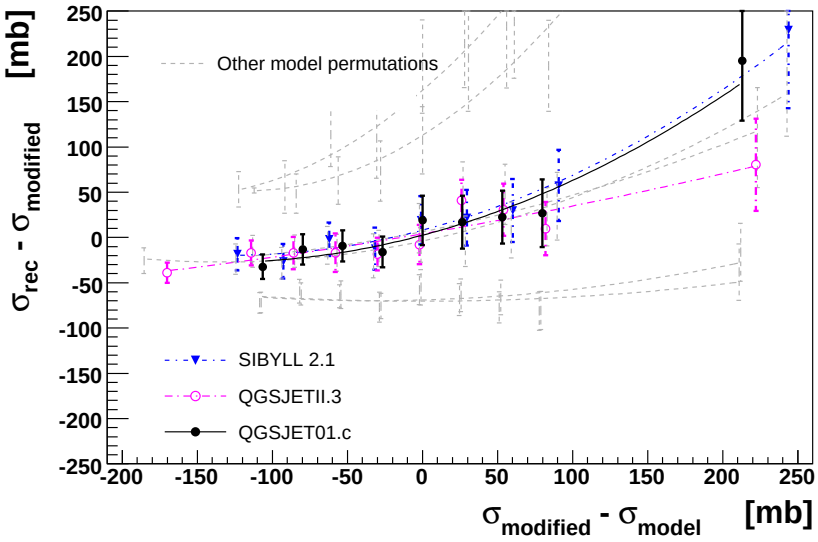


Figure 4.21: The resulting systematics of previous reconstruction techniques, which assumed a constant ΔX_1-distribution or k-factor. Each point is the mean of 100 reconstructions and the error bars denote the RMS. The adopted $X_{\max}$-resolution is a Gaussian with a width of $20\,\mathrm{gcm}^{-2}$. For each $X_{\max}$-distribution 2000 events are generated.

that any cross section measurement which is smaller than the prediction of the used model, in fact must be an *underestimation* of the true value. This analysis technique is limited to the verification respectively falsification of the predictions of the interaction model used for the air shower simulations, while any deviation can not be easily interpreted as a cross section measurement. In addition the resulting overall model dependence is becoming even more pronounced.

4.8 Primary composition

There are large uncertainties on the primary mass composition of cosmic rays at ultra-high energies. Therefore it is important to study the impact of a mixed composition on the results of the cross section analysis. It is well known how primaries of different type contribute to the $X_{\max}$-distribution at different places. Photon primaries produce shower maxima at large depths and thus contribute at $X_{\max}$ well above the peak of a purely proton induced $X_{\max}$-distribution. All nuclei heavier than protons have a larger cross section compared to protons, and are therefore producing shower maxima at shallower depths compared to protons. This will contribute to the $X_{\max}$-distribution at depths smaller than the peak of a purely proton induced $X_{\max}$-distribution. In the following the study of heavy primaries is restricted to helium, nitrogen and iron. Iron primaries are the heaviest nuclei that can be produced during stellar fusion. Elements beyond iron are generally rare, because they are only produced in limited quantities during supernova explosions. Therefore iron is usually believed to be the heaviest element that possibly contributes to the total cosmic rays flux. Helium, as the second most abundant element in the universe, is a very probable cosmic ray primary. There are no other neighboring elements as clearly separated as proton and

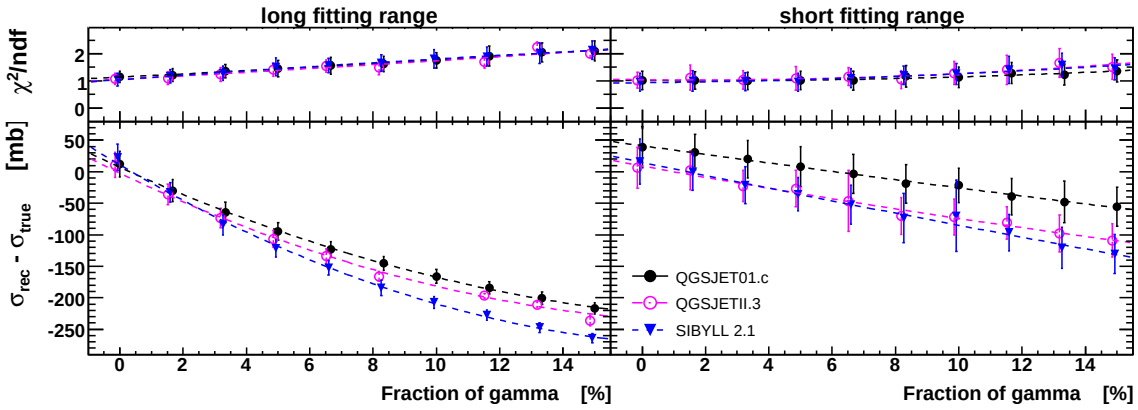


Figure 4.22: Impact of a mixed primary composition on the results of the cross section analysis. Each data point contains 2000 entries, blurred with an assumed Gaussian-like resolution of 20 gcm^{-2}. Shown is the mean (marker) and the RMS (error bars) for 100 independent samples. The lines are 2^{nd}-order polynomial fits in order to guide the eye.

helium ($A_{\mathrm{He}} = 4\,A_{\mathrm{H}}$ and $Z_{\mathrm{He}} = 2\,Z_{\mathrm{H}}$). However, currently there is little hope to separate helium from proton induced air showers, mainly because of the deficiencies of the hadronic interaction models. The light nuclei of lithium, beryllium and boron have only marginal abundances in the universe, and are therefore believed to have no importance for cosmic rays. On the other hand, the following three nuclei carbon, nitrogen and oxygen (CNO) are overabundant in the universe because of the stellar CNO fusion cycle. Thus they may well constitute a significant fraction of the cosmic ray flux.

The impact of a cosmic ray beam with a mixed composition on the results of the cross section reconstruction is exemplarily studied for photon, helium, nitrogen (CNO) and iron admixtures. In Section 4.6 it is described how the fitted range of the X_{max}-distribution can be adopted to possibly reduce the impact of non-proton cosmic ray primaries. This is demonstrated here by showing the results for two choices of fitting ranges:

Long fitting range: Starting from 50 gcm^{-2} in front of X_{peak} up to the largest X_{max} of the selected events plus 40 gcm^{-2}.

Short fitting range: For photon primaries the upper bound of the fitting range is restricted to $X_{\mathrm{peak}} + 150\,\mathrm{gcm}^{-2}$, while for hadronic primaries the beginning of the fitting range is set to X_{peak}.

The impact of photon primaries is shown in Figure 4.22. Even a few percent of photon primaries have a devastating impact on the cross section analysis. An underestimation by 50 mb is evoked by only 2% of photons. This is approximately what can be achieved with the current limits on the photon fraction [182]. It is interesting to note that the mean χ^2 rises from ~ 1 for no photon contribution to ~ 2.5 for 15% photon content. This feature is almost model independent. Photon primaries significantly distort the X_{max}-distribution at large X_{max}, where only few proton events are located. The resulting total X_{max}-distribution rapidly gets incompatible with a pure proton distribution for a rising photon fraction and causes the increase of the χ^2. There is no indication for such behavior found in the data of

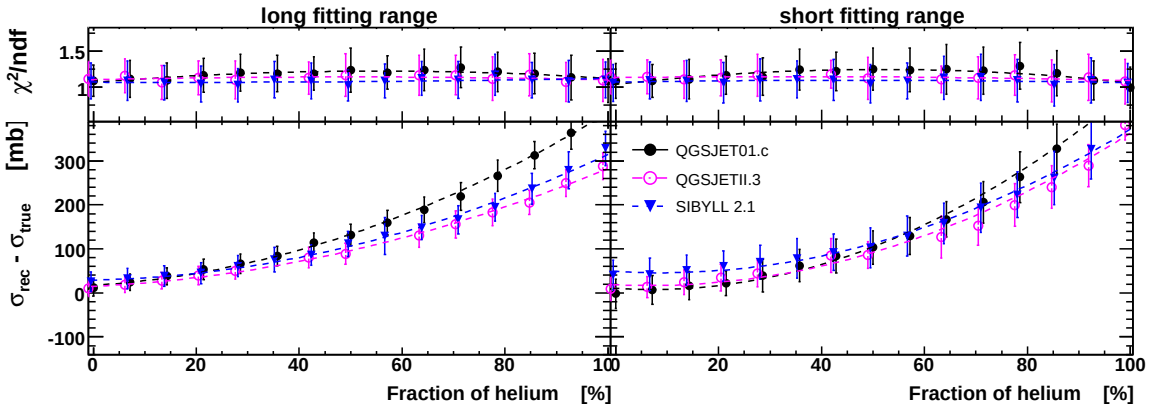


Figure 4.23: Impact of a mixed proton/helium composition on the results of the cross section analysis. Each data point contains 2000 entries, blurred with an assumed Gaussian-like resolution of 20 gcm^{-2}. Shown is the mean (marker) and the RMS (error bars) for 100 independent samples. The lines are 2^{nd}-order polynomial fits in order to guide the eye.

the Pierre Auger Observatory, which is supporting the 2% limit on the photon fraction. The theoretical lower limit on the photon fraction above 10 EeV is as low as $10^{-4} - 10^{-3}$ [183]. It is resulting from considerations about the pion production in ultra-high energy proton interactions with microwave background photons (GZK effect). Thus, in the future the photon-induced systematic uncertainty on the cross section analysis may vanish with the availability of better experimental results on the photon limit. Therefore no special efforts to restrict the fitting range in the high X_{max}-tail of the X_{max}-distribution are made during the analysis of hybrid data. A shorter fitting range is also having a considerable negative impact on the quality of the reconstruction.

The situation for hadronic contaminations of the proton beam is much less pleasant. There are no equally strong theoretical and experimental arguments to restrict the fraction of heavy cosmic ray primaries at ultra-high energy compared to the photon-limit. However, re-

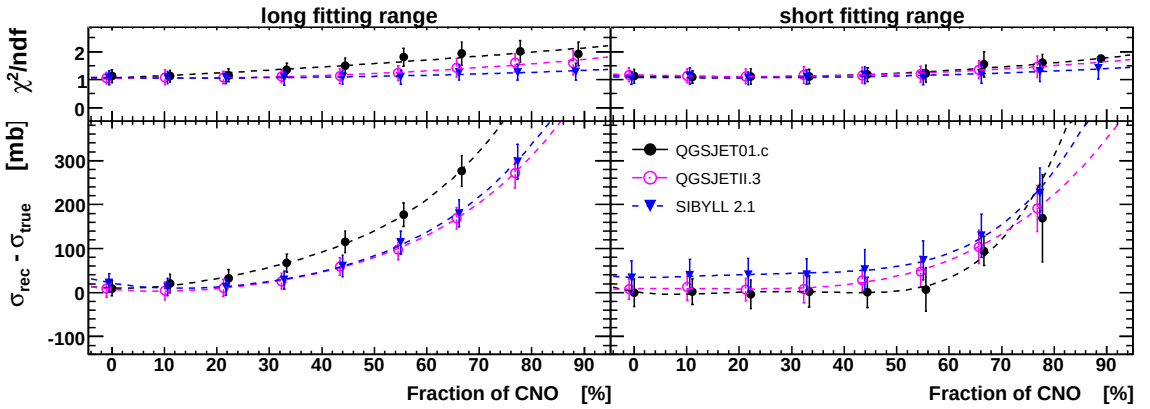


Figure 4.24: Impact of a mixed proton/CNO composition on the results of the cross section analysis. Each data point contains 2000 entries, blurred with an assumed Gaussian-like resolution of 20 gcm^{-2}. Shown is the mean (marker) and the RMS (error bars) for 100 independent samples. The lines are polynomial fits in order to guide the eye.

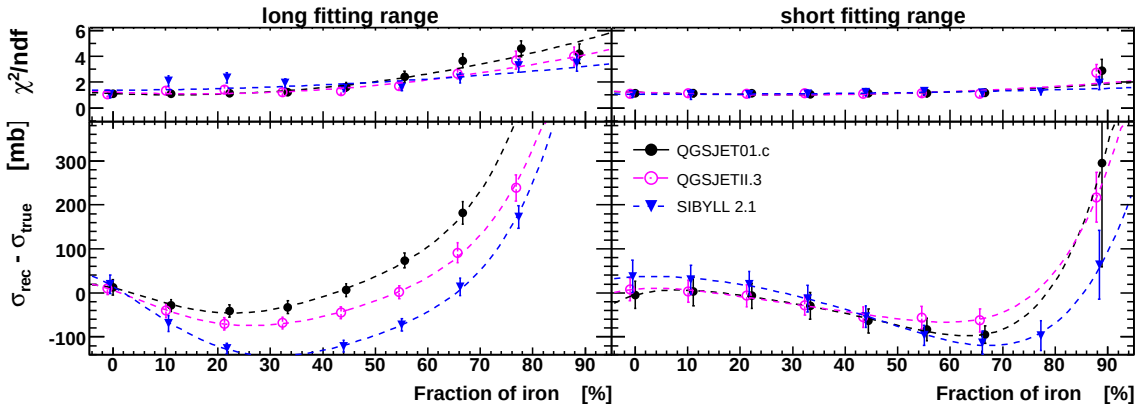


Figure 4.25: Impact of a mixed proton/iron composition on the results of the cross section analysis. Each data point contains 2000 entries, blurred with an assumed Gaussian-like resolution of 20 gcm^{-2}. Shown is the mean (marker) and the RMS (error bars) for 100 independent samples. The lines are polynomial fits in order to guide the eye.

cently a series of observations has been published, which can be exploited in supporting the proton-dominance in cosmic rays at ultra-high energy. Namely the observation of a GZK-like feature in the cosmic ray flux by the HiRes [39] as well as Auger [21] Collaborations and most importantly the correlation of the highest energy events with close-by astrophysical objects [42]. In Section 5.4 this is discussed in more detail and exploited in order to support a proton dominated cosmic ray beam at ultra-high energies.

Figures 4.23 to 4.25 summarize the impact of heavier nuclei on the cross section analysis. It is most interesting to note that any amount of helium admixture from 0 up to 100% is fitted with no degeneration of the resulting χ^2. There is a smooth transition to reconstruct the proton cross section for 100% protons and increasingly overestimate the cross section at larger helium fractions. The "short fitting range" examples demonstrate that even the restriction of the analysis to the tail of the X_{max}-distribution does not significantly help to restrict the impact of helium contaminations. The impact of CNO and iron admixtures has a different characteristics. Firstly, there is some impact on the resulting χ^2, since these simple-minded two-component scenarios are producing artificial features in the X_{max}-distributions, like double-peaked distributions or breaks in the exponential tail of the distributions. Also the impact on the resulting cross sections are manifold. Up to about $\sim$ 30% of CNO, respectively $\sim$ 60% for the short fitting range, the impact on the cross section analysis is small. For fractions up to $\sim$ 50% of iron the cross section is even underestimated by up to 100 mb. This behavior can be shifted to higher fractions of iron by choosing a shorter fitting range. Along with the strong impact on the cross section goes a significant increase of the χ^2. This is a feature of the superposition of two shifted distributions $X^{\mathrm{A}}_{\mathrm{max}}$ and $X^{\mathrm{B}}_{\mathrm{max}}$. The expectation value of the resulting distribution

$$\mathrm{E}[\epsilon\, X^{\mathrm{A}}_{\mathrm{max}} + (1-\epsilon)\, X^{\mathrm{B}}_{\mathrm{max}}] = \epsilon\, \mathrm{E}[X^{\mathrm{A}}_{\mathrm{max}}] + (1-\epsilon)\, \mathrm{E}[X^{\mathrm{B}}_{\mathrm{max}}] \tag{4.17}$$

shifts proportional to the fraction ϵ, while the variance

$$\begin{aligned}\mathrm{Var}[\epsilon\, X^{\mathrm{A}}_{\mathrm{max}} + (1-\epsilon)\, X^{\mathrm{B}}_{\mathrm{max}}] \\ = \epsilon\, \mathrm{Var}[X^{\mathrm{A}}_{\mathrm{max}}] + (1-\epsilon)\, \mathrm{Var}[X^{\mathrm{B}}_{\mathrm{max}}] + \epsilon(1-\epsilon)\, \left(\mathrm{E}[X^{\mathrm{A}}_{\mathrm{max}}] - \mathrm{E}[X^{\mathrm{B}}_{\mathrm{max}}]\right)^2 \end{aligned} \quad (4.18)$$

is additionally increased by an additional term for $\epsilon \neq 0$ and $\epsilon \neq 1$ [184]. Thus, the shape of the distribution is strongly changing due to admixtures of heavier elements, leading to the observed behavior. For very large fractions of heavy nuclei ($> 50\%$) the cross section is rapidly getting overestimated.

5 Data analysis

For the following analysis the reconstructed hybrid data from the beginning of 2004 until the end of October 2007 was used. The dataset is taken from the *Auger Observer* [163], which is based on the reconstruction of Offline version v2r3p1-ADST and stored in the `ADST` data format.

5.1 Hybrid data selection and quality cuts

Only a small fraction of the 753036 triggers recorded in the considered time period is suited for the proton-air cross section measurement. The quality of a reconstructed event strongly depends on the amount of light collected by the telescopes and thus primarily on the distance of the event as well as the energy of the cosmic ray primary. In order to select a dataset with high quality event reconstruction, a set of quality criteria, as listed in Table 5.1, needs to be applied. From the study of full detector Monte Carlo simulations as well as stereo data, it is found that these cuts are resulting in a resolution of X_{max} better than $20\,\mathrm{gcm}^{-2}$ above an energy of 1 EeV (cf. Section 5.2). This corresponds to a ≈5% effect for a typical observed RMS of an X_{max}-distribution of the order of $60\,\mathrm{gcm}^{-2}$. The quality selection can be classified as follows:

Calibration. The availability of valid calibration constants for the FD detector limits the dataset to the time after the beginning of December 2004 [129]. Since the SD array was small prior to this date and just two FD buildings were in operation, not much data is lost by this requirement. In addition this cut removes all data from Loma Amarilla, since so far no calibration constants are available for this building.

Hybrid geometry. Requiring a minimum distance of less than 2 km from the hybrid tank to the shower core rejects FD-mono events with a random SD trigger. The more restrictive fiducial core-tank distance cut takes into account the primary mass dependent single tank trigger probability. It is well known that the tank trigger efficiency depends on the mass of the primary particle and may therefore induce a bias already on the hybrid trigger level. As it is demonstrated to be possible to use simulated lateral tank trigger probabilities for proton and iron [174, 175, 185] to develop energy dependent cuts, that minimize any mass dependent trigger bias [186]. The tank trigger probability decreases with zenith angle and increases with primary energy. This can be exploited to select events while keeping the trigger probability close to unity [186]

$$d_{\mathrm{tank}} \leq \begin{cases} 750\,\mathrm{m} & \log_{10}(E/\mathrm{EeV}) < 0 \\ 750\,\mathrm{m} + 833\,\mathrm{m}\ \log_{10}(E/\mathrm{EeV}) & 0 \leq \log_{10}(E/\mathrm{EeV}) < 1.5 \\ 2000\,\mathrm{m} & \log_{10}(E/\mathrm{EeV}) \geq 1.5 \end{cases} \tag{5.1}$$

$$\cos(\theta) \geq \begin{cases} 0.55 & \log_{10}(E/\mathrm{EeV}) < 0 \\ 0.55 - 0.25\ \log_{10}(E/\mathrm{EeV}) & 0 \leq \log_{10}(E/\mathrm{EeV}) < 1.5 \\ 0.175 & \log_{10}(E/\mathrm{EeV}) \geq 1.5 \end{cases} . \tag{5.2}$$

The algorithms used to reconstruct the longitudinal shower profiles work for events with large contributions of Cherenkov light [147]. However, for shower geometries

pointing towards the telescope within small angles, the geometric uncertainties are getting large, due to the compressed arrival time of photons at the detector and also due to the resulting short projection of the shower on the camera. Since the intensity of direct Cherenkov light depends exponentially on the viewing angle with respect to the shower axis, events with a minimum viewing angle smaller than 20° are rejected.

Shower profile. In order to yield a resolution in $X_{\max}$ of $\sim 20\,\mathrm{gcm}^{-2}$, it is necessary to select events with an accurate profile reconstruction. First of all, this requires the shower maximum $X_{\max}$ being located within the observed slant depth range of the telescope. Since the uncertainties of the measured light flux as well as of the geometry are propagated into the parameters of the shower profile, the estimated uncertainties of the energy and $X_{\max}$ are well suited to reject poorly reconstructed events. Events hampered by atmospheric effects like clouds or fog can be identified by a large reduced χ^2 of the Gaisser-Hillas (GH) fit, since parts of the profiles are obscured due to absorption, or enhanced via the reflection of the bright Cherenkov beam.

The significance of a reconstructed shower profile can be judged by comparing it to a straight line fit. Reconstructions of the GH-parameters with a difference in χ^2 of less than 4 compared to a straight line are considered to be "insignificant" and are rejected.

Events involving more than a single mirror are affected by a flaw in the current implementation of the light collection. Light close to the boundaries of the telescopes is lost. Shower profiles with a resulting gap of more than 20% of their total length are discarded.

After applying this quality selection 16696 events remain for further analysis (cf. Table 5.1). The impact of the cuts on hybrid data as well as on simulated events is displayed in Figure 5.1. Shown are the distributions relevant for quality selection with all quality cuts

Table 5.1: Impact of the applied quality criteria on the hybrid dataset of the Pierre Auger Observatory from 01/2004 up to 10/2007. The cuts are applied in the sequence as listed in this table.

Cut	Excluded events [%]	Remaining events
Total hybrids	-	753036
Calibration:		
Deselect Loma Amarilla	6.1	707240
Calibration available	8.9	643999
Geometry reconstruction:		
Hybrid tank closer than 2 km	65.7	221058
Minimal viewing angle < 20 deg	27.1	161178
Fiducial core tank distance	12.3	141324
Profile reconstruction:		
$X_{\max}$ observed	58.7	58302
$\sigma(X_{\max}) < 40\,\mathrm{gcm}^{-2}$	48.0	30305
$\sigma(E_{\mathrm{tot}})/E_{\mathrm{tot}} < 0.2$	0.5	30142
$\chi^2_{\mathrm{GH}}/\mathrm{ndf} < 2.5$	2.0	29530
$\chi^2_{\mathrm{GH}} - \chi^2_{\mathrm{line}} < 4$	42.9	16850
Profile gap, $(\Delta X)_{\mathrm{hole}} < 20\,\mathrm{gcm}^{-2}$	0.9	16696

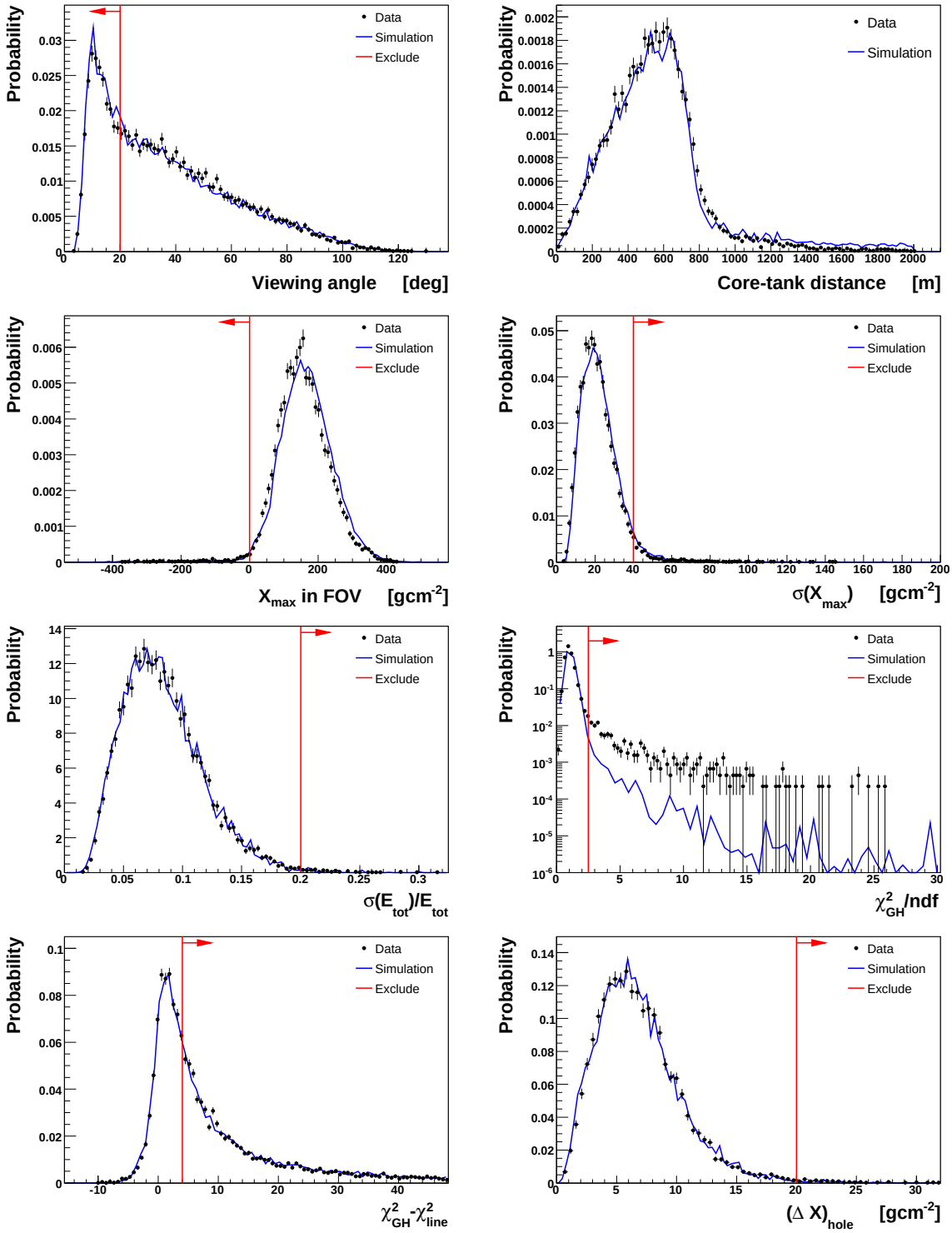


Figure 5.1: Distribution of the quantities, which are used for quality selection. For each of the distributions all quality cuts, except the one concerning the shown quantity, are applied ($\rightarrow$ *N-1*-cut). The markers show the hybrid data, while the lines denote simulated events after full hybrid detector Monte Carlo simulations. The simulations are based on CONEX with primary protons from 100 PeV up to 1 ZeV for a flux $\propto E^{-3}$. The vertical line with the arrow indicates the region which is rejected. The plot in the upper-right corner shows the fiducial core-tank distance cut, which has a running cut depending on the reconstructed energy and zenith angle, so it can not be displayed as a vertical line.

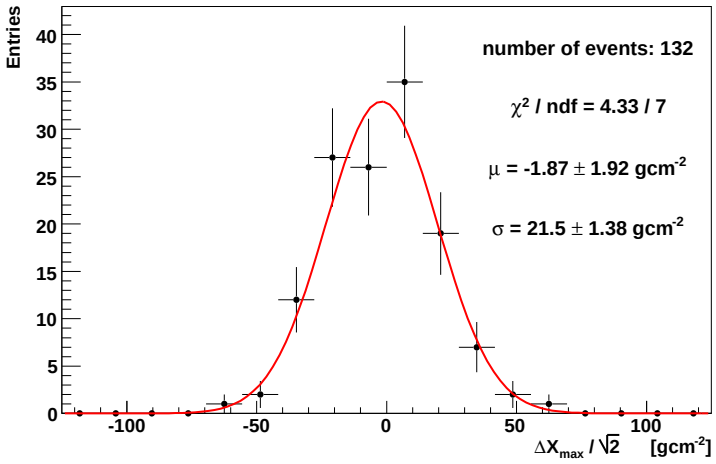


Figure 5.2: Distribution of $(X^1_{\max} - X^2_{\max})/\sqrt{2}$ of the 132 stereo events remaining after quality selection. The line is a Gaussian fit to the data, whose parameters are listed. The RMS of the distribution is $21.36 \pm 1.3\,\mathrm{gcm}^{-2}$ and the mean -1.94 $\pm$ 1.9 gcm^{-2}, which corresponds well to the Gaussian fit. The distribution of $\log_{10}(E/\mathrm{eV})$ for the stereo events has a mean of 19.11 and an RMS of 0.24.

used, except the one going to be applied to the displayed quantity. This type of plots are therefore called (N-1)-plots.

The general agreement between the proton simulations and the hybrid data is remarkably good. Only the (N-1)-distribution of the reduced χ^2_{GH} shows a large deviation between hybrid data and simulations. This is expected, since small-scale atmospheric effects, as clouds, layers of fog or other atmospheric fluctuations, are not included in the detector Monte Carlo simulations, because they are not monitored reliably. Thus, the cut on χ^2_{GH} is placed at the position, where the data significantly start to deviate from simulations.

The distribution of minimum viewing angles shows a peak at small angles. This is caused by the enhanced trigger probability of events pointing into the telescope due to the bright forward-peaked Cherenkov beam. Because of the large systematic uncertainties introduced by these geometries this peak has to be rejected.

There is one distribution where the cut seems to be placed too strong. The $\chi^2_{\mathrm{GH}} - \chi^2_{\mathrm{line}}$-distribution is rejecting almost 43% of the events, by directly cutting into the distribution. These events are either faint, distant showers or very low energy events of poor quality.

5.2 Hybrid $X_{\max}$-resolution

It is possible to study the attainable $X_{\max}$-resolution after reconstruction and quality selection by generating Monte Carlo events including the full detector simulation. An important feature of the Auger Observatory is that a fraction of events is observed by more than one single telescope building and can therefore be reconstructed independently. This was exemplified with a triple event in Section 3.3. The independent observation of the same shower event can be used to cross-check the performance and general understanding of the reconstruction. During the considered time period there are 132 pairs of simultaneous observations fulfilling the quality criteria. The resulting distribution of $(X^1_{\max} - X^2_{\max})/\sqrt{2}$, which

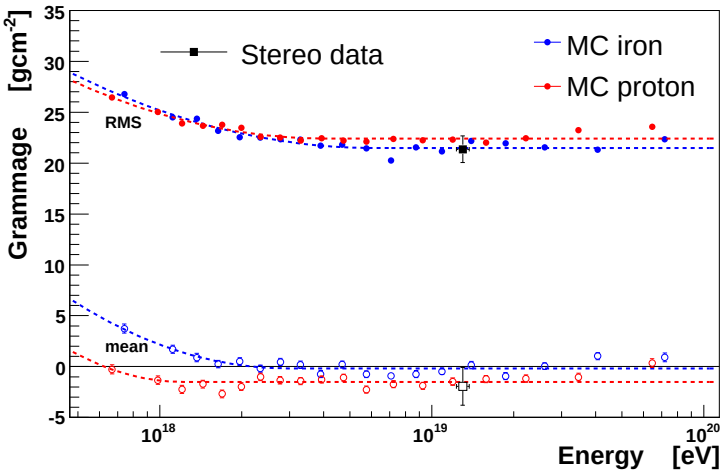


Figure 5.3: Mean and RMS of the $\Delta X_{\max}$-distribution obtained after reconstruction and quality selection of Monte Carlo events after full detector simulation. At energies above 2 EeV the RMS is almost constant and in very good agreement to the stereo data. The stereo data point shows the corresponding values for the distribution of the 132 stereo events after the same quality selection. For the purpose of this plot, atmospheric fluctuations that are not part of the Monte Carlo simulation have been quadratically added to the simulations (see text and Table 5.2).

can be interpreted as the measured $X_{\max}$-resolution, is shown in Figure 5.2. However, this is not sufficient for a full understanding of the $X_{\max}$-resolution over a wide energy range. But it can be used in order to verify the results obtained from the simulations at a single point in energy. In Figure 5.3 the mean and RMS of the distributions of $\Delta X_{\max} = X_{\max}^{\text{rec}} - X_{\max}^{\text{true}}$ resulting from the full simulation dataset is illustrated. Estimates on how atmospheric fluctuations and non-uniformities are contributing additional fluctuations to the measurement of $X_{\max}$ are available [187]. In the simulations the atmosphere is implemented perfectly uniformly over large distances, because there is no reliable monitoring data available for small scale weather effects like clouds or fog. Thus, it is required to explicitly add the fluctuations as listed in Table 5.2 to the simulations before comparing to the stereo data. If these uncertainties are quadratically added to the mean RMS obtained from the simulation of proton and iron of 18.4 gcm^{-2}, the resulting total RMS gets 22.3 gcm^{-2}. This is in good agreement

Table 5.2: Atmosphere-induced fluctuations of the distribution of reconstructed $X_{\max}$. Values adopted from [187].

Effect	**Impact on $X_{\max}$**
Molecular atmosphere:	
Horizontal uniformity	1 gcm^{-2}
Variations in air density profile	6 gcm^{-2}
Aerosols:	
Horizontal uniformity	9 gcm^{-2}
Vertical aerosol optical depth	6 gcm^{-2}
Wavelength dependence	1 gcm^{-2}
Differential scattering cross section	2 gcm^{-2}

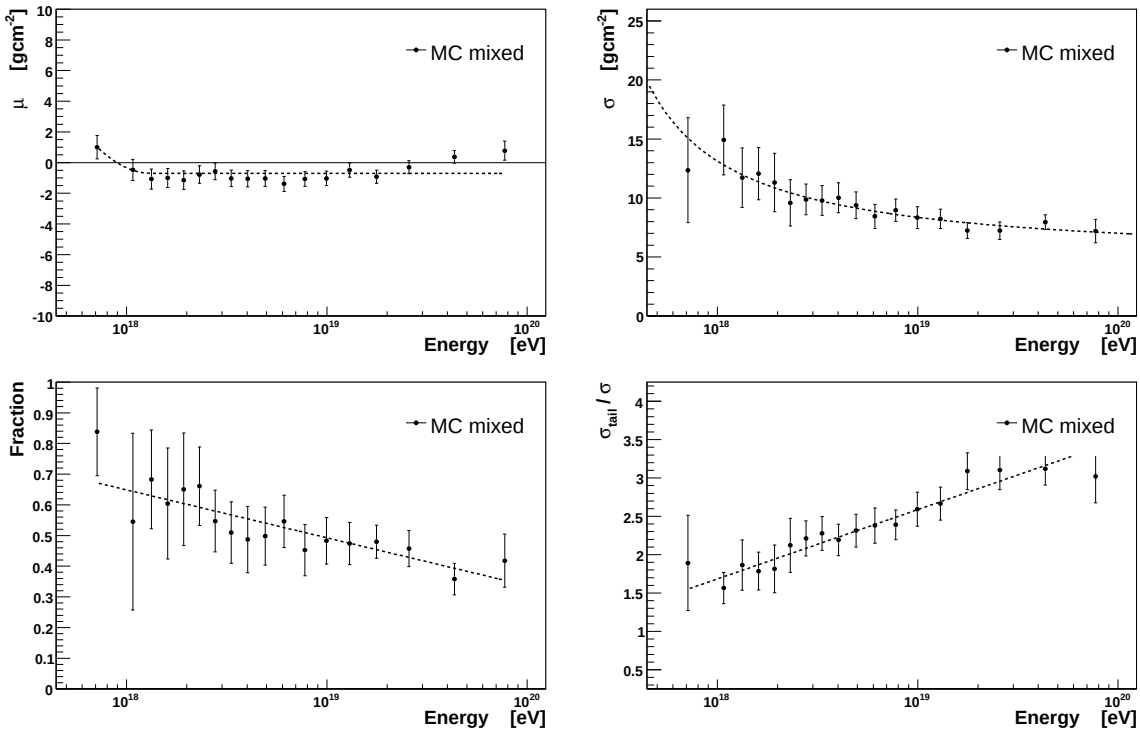


Figure 5.4: Parameterization of the X_{max}-resolution using a double-Gauss model. Shown is the dependence of the parameters on the energy. The markers denote the results from simulations with a mixed composition of 50% protons and 50% iron.

with the stereo data, which is found to have an RMS of $21.36 \pm 1.3\,\mathrm{gcm}^{-2}$.

The differences of the RMS between proton and iron induced air showers are small. This is important, since it shows that the detector resolution does not depend on the primary particle type, but only on the energy. Furthermore given the fact that the X_{max}-distribution of proton and iron have a considerably different width, this result indicates that no dependence of the X_{max}-resolution on the cross section is expected. For the resulting mean of the ΔX_{max}-distribution the differences between proton and iron simulations are larger. However, for the purpose of the cross section analysis a small bias in the absolute X_{max}-scale is irrelevant, because it can be absorbed by the free parameter X_{shift}.

For the high statistics simulations used to study the X_{max}-resolution over the full energy range it is found that a double-Gauss model

$$P_{\mathrm{res}}(\Delta X_{\mathrm{max}}) = \frac{1-\epsilon_{\mathrm{res}}}{\sigma_{X_{\mathrm{max}}}\sqrt{2\pi}}\, e^{-0.5(\Delta X_{\mathrm{max}}-\mu_{\mathrm{res}})^2/\sigma^2_{X_{\mathrm{max}}}} + \frac{\epsilon_{\mathrm{res}}}{\sigma_{X_{\mathrm{max}}}\delta_{\mathrm{res}}\sqrt{2\pi}}\, e^{-0.5(\Delta X_{\mathrm{max}}-\mu_{\mathrm{res}})^2/(\sigma_{X_{\mathrm{max}}}\,\delta_{\mathrm{res}})^2} \tag{5.3}$$

needs to be used in order to parameterize the resolution. The two Gaussians are centered around the same mean value μ_{res}. The narrow Gaussian has a width of $\sigma_{X_{\mathrm{max}}}$, the broader one of $\sigma_{X_{\mathrm{max}}}\,\delta_{\mathrm{res}}$. The relative importance of the two Gaussians is regulated by the fraction ϵ_{res} of the integrated area of the broader Gaussian to the total. The resulting dependence of

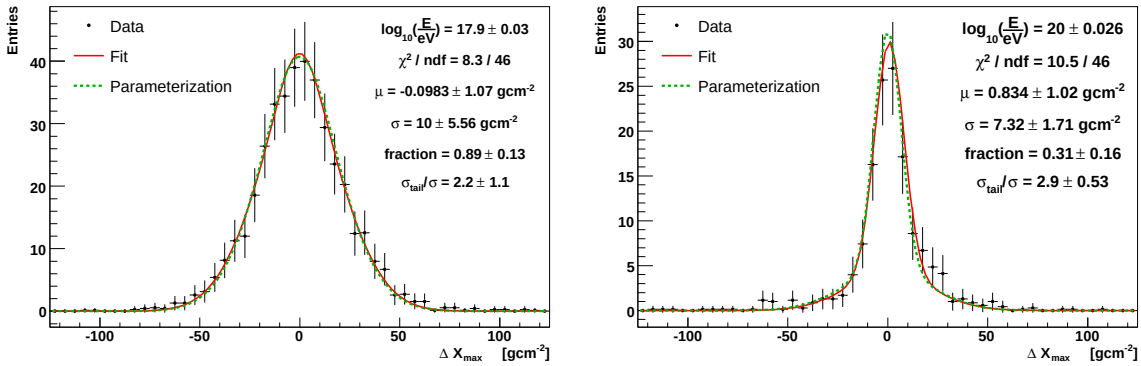


Figure 5.5: $X_{\max}$-resolution as obtained from Monte Carlo simulations for two energy intervals. The fit (solid line) and parameterization (dotted line) of the distribution by the double-Gauss (5.3) are shown.

the parameters on the energy is presented in Figure 5.4. It can be parameterized as follows:

$$\begin{aligned} \mu_{\text{res}}(E) &= 0 \\ \sigma_{X_{\max}}(E) &= 4.84 + 5.04/(\log_{10}(E/\text{eV}) - 17.32) \\ \delta_{\text{res}}(E) &= 5.00 - \log_{10}(E/\text{eV}) \cdot 0.23 \\ \epsilon_{\text{res}}(E) &= -12.45 + \log_{10}(E/\text{eV}) \cdot 0.79\,. \end{aligned} \tag{5.4}$$

In Figure 5.5 the $\Delta X_{\max}$-distribution of simulated events at two of the energy intervals is illustrates. The double-Gauss model (5.3) with the interpolated parameters of (5.4) is shown together with the initial fits to the distribution.

The interval on which the double-Gauss model gives non-vanishing probabilities can be restricted to

$$\Delta^{\text{start}}_{X_{\max}} = -8\,\sigma_{X_{\max}}(E^i) \quad \text{and} \quad \Delta^{\text{stop}}_{X_{\max}} = +8\,\sigma_{X_{\max}}(E^i)\,. \tag{5.5}$$

It can be concluded that the $X_{\max}$-resolution for hybrid reconstruction is well understood and can be efficiently parameterized by a double-Gaussian. Since atmospheric effects are inducing additional fluctuations, it is necessary to increase the RMS determined by simulations by quadratically adding the sources of uncertainties listed in Table 5.2. This can be done on the level of the final parameterization. The total variance of the double-Gauss is

$$\text{Var}_{\text{res}} = (1-\epsilon_{\text{res}})^2\,\sigma^2_{\text{res}} + \epsilon^2_{\text{res}}\,(\sigma_{\text{res}}\,\delta_{\text{res}})^2\,. \tag{5.6}$$

Quadratically adding of the total atmospheric fluctuations $\text{Var}_{\text{atm}} = 159\,\text{g}^2\text{cm}^{-4}$ from Table 5.2 can be done by adopting the new effective width $\sigma^{\text{atm}}_{X_{\max}}$ for the parameterization (5.3)

$$\sigma^{\text{atm}}_{X_{\max}} = \sqrt{\frac{(1-\epsilon_{\text{res}})^2\,\sigma^2_{\text{res}} + \epsilon^2_{\text{res}}\,(\sigma_{\text{res}}\,\delta_{\text{res}})^2 + \text{Var}_{\text{atm}}}{(1-\epsilon_{\text{res}})^2 + \epsilon^2_{\text{res}}\,\delta^2_{\text{res}}}}\,. \tag{5.7}$$

5.3 Detector acceptance

One of the requirements for a good quality reconstruction demands that the shower maximum is located within the field of view of the telescope. The field of view is an upward

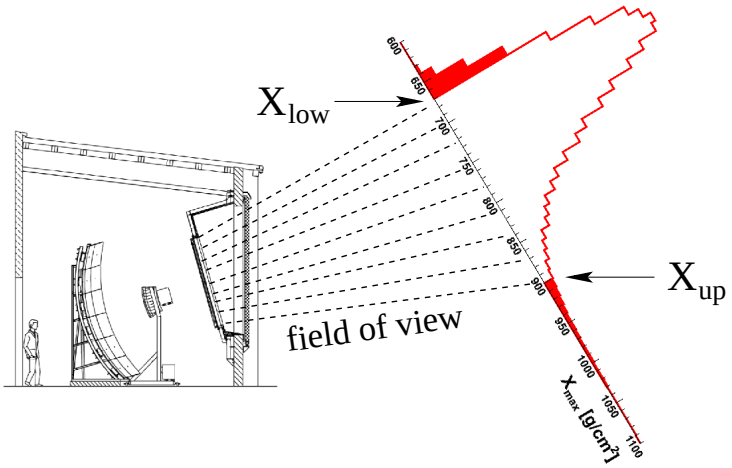


Figure 5.6: Selection of events due to the limited acceptance of a telescope detector [186].

looking cone with an opening angle of 15°. In Figure 5.6 it is illustrated, how this imposes a direct geometry dependent selection of events in terms of their X_{max}. For this reason, it follows that the acceptance, in terms of X_{max}, depends on the primary composition as well as on features of hadronic interaction at ultra-high energy, since both are known to have significant impact on the depth of the shower maximum.

Furthermore, the acceptance of a telescope detector depends strongly on atmospheric conditions. Obviously the telescope detector efficiency depends on absorption of UV light in the atmosphere, distance from the detector to the shower axis and the number of charged particles in the field of view.

In order to obtain the shape of the real underlying X_{max}-distribution of cosmic rays, it is mandatory to understand the acceptance of the detector and the efficiency of the reconstruction. This has to be properly considered during any analysis of X_{max}-data.

The detector acceptance function P_{acc} is incorporated in the X_{max}-model (4.5) and (4.6) for the proton-air cross section measurement. For each shower geometry i the model only accounts for the resulting $X^{\mathrm{rec}}_{\mathrm{max}}$-distribution in the slant depth range from X^{i}_{low} up to X^{i}_{up}. There is no contribution by the shower geometry i to the total $X^{\mathrm{rec}}_{\mathrm{max}}$-distribution from the tails of the distribution that are located outside of the visible slant depth range.

To calculate the values of X^{i}_{low} and X^{i}_{up} for a reconstructed shower event i, it is not sufficient to consider only the telescope, in which the shower was detected. The situation of an inclined event triggering a telescope with two neighboring telescopes, of which one was actively taking data and the other was turned off at this moment, is illustrated in Figure 5.7. The corresponding range in slant depth, which needs to be considered for the $X^{\mathrm{rec}}_{\mathrm{max}}$-model ranges from the upper border of the active telescope to the boundary of the triggered telescope. Obviously this requires information on the status of all telescopes at any time. This kind of up-time can be calculated on the basis of minimum bias files, background data-taking and monitoring data stored in the raw data themselves [188]. However, it has to be computed offline, after all relevant data were collected and copied from Malargüe. Currently, the up-time information is available until the end of July 2007 [189], which is cutting down the available dataset from 16696 events to 15074 events.

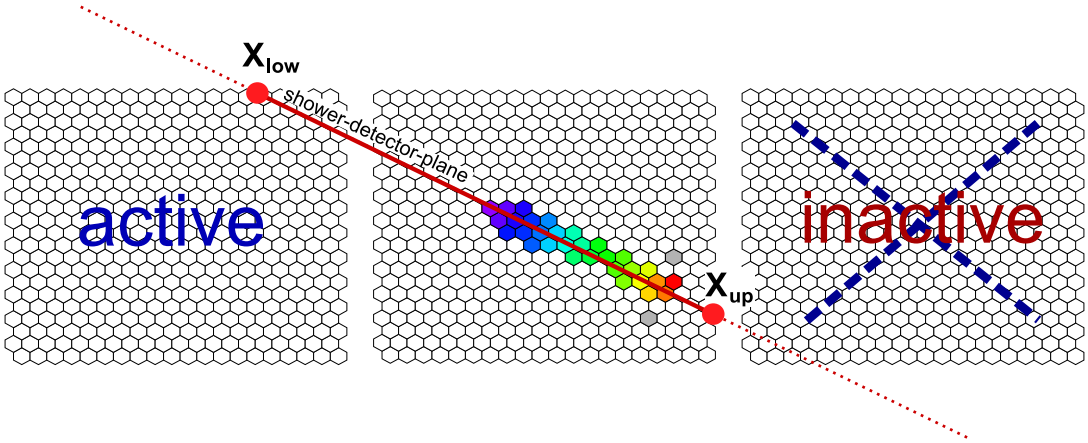


Figure 5.7: Definition of the visible slant depth range corresponding to a reconstructed shower geometry. In this example a shower is recorded by the central telescope, while the left telescope is active but the right telescope is not taking data.

The resulting distributions of X_{max}, X_{low} and X_{up} of the selected dataset are shown in Figure 5.8 (left). Clearly, both sides of the X_{max}-distribution must be affected by the limitations of the field of view.

Furthermore it is not sufficient to consider just the geometry of the field of view for the telescope acceptance. The optical transmissivity of the atmosphere is generating a much more complex behavior. Absorption of light leads to a drop of the detection efficiency for large distances from the telescope to the detector [191]. Since the atmospheric aerosol density is much more dense close to the ground compared to higher altitudes, this distant dependent drop of the efficiency is also a function of the altitude. Generally it is possible to see much further looking upward through the less dense layers of the atmosphere compared to the much more opaque atmosphere close to the ground. Therefore the fiducial field of view of a telescope detector starts to significantly deviate from the purely geometric definition for

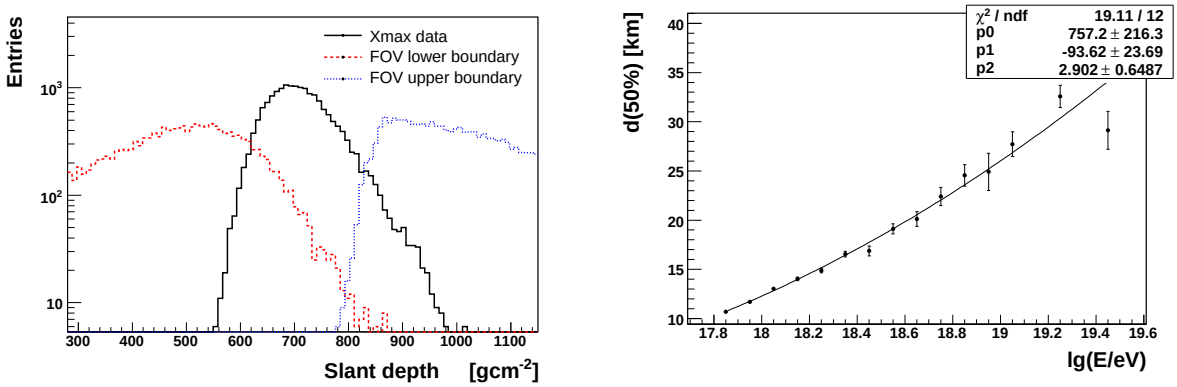


Figure 5.8: Left panel: Distribution of X_{max}, X_{low} and X_{up} of the selected dataset. Right panel: The distance, where the observed frequency of events is dropping below 50% of $N \propto R\mathrm{d}R$ that is expected from a purely geometric acceptance [190]. Points denote the estimates obtained from data, the solid line indicates a parameterization with a polynomial of 2$^{\mathrm{nd}}$-degree.

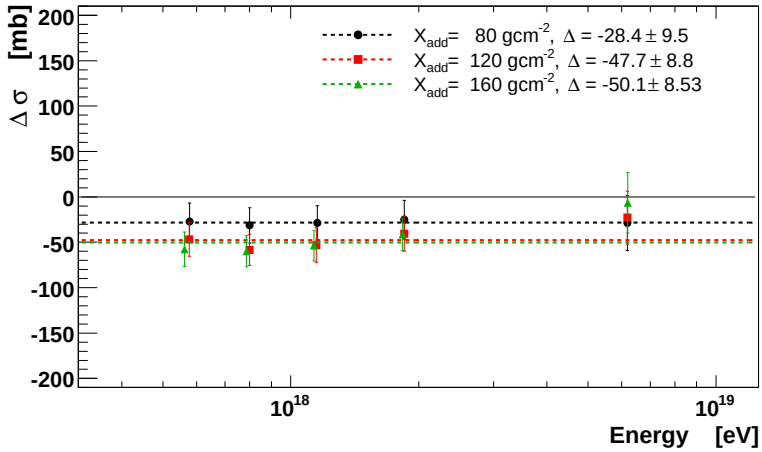


Figure 5.9: The reduction of the fiducial slant depth range by the additional margin X_{add} is used to remove boundary effects from the geometrical field of view definition. The reference line at zero is for $X_{\text{add}} = 40\,\text{gcm}^{-2}$.

large distances. This is especially dangerous because the detection efficiency is first dropping close to the ground and thus for large X_{max}. The tail of the resulting X_{max}-distribution gets suppressed mimicking smaller fluctuations and a larger cross section. In order to define the distance, where the detection efficiency starts to significantly deviate from the expectation of the geometry of the field of view, we search for the distance where the observed frequency of events drops below 50% of the geometric scaling $N \propto R\mathrm{d}R$. This is done on the basis of the quality selected hybrid data. In Figure 5.8 (right) the dependence of the resulting distances on the energy is illustrated. The remaining dataset shrinks to 12805 events after the application of this fiducial distance cut.

Before the application of the developed X_{max}-model (4.5) to data distributions, one more cross check of the implemented acceptance model is performed. The calculated field of view interval in slant depth from X_{low} up to X_{up} can be further narrowed by the additional margin X_{add} to remove eventual boundary effects ($X_{\text{up}} \rightarrow X_{\text{up}} - X_{\text{add}}$ and $X_{\text{low}} \rightarrow X_{\text{low}} + X_{\text{add}}$). This can be performed on the basis of the hybrid data. The cross section analysis is performed for several values of X_{add} and the resulting impact on the cross section is shown in Figure 5.9. The reference reconstruction (zero line) is for $X_{\text{add}} = 40\,\text{gcm}^{-2}$. There is a global shift to smaller cross sections, which saturates at $\approx -45\,\text{mb}$ for $X_{\text{add}} \geq 120\,\text{gcm}^{-2}$. For all further analysis the value of $X_{\text{add}} = 120\,\text{gcm}^{-2}$ is therefore adopted to further reduce the fiducial slant depth range. Of course this also impacts the event selection, since some events may have an X_{max} that is incompatible with this restricted slant depth range. This is causing the loss of just 23 events.

5.4 Proton hypothesis

It is the fundamental assumption of any proton-air cross section analysis with cosmic rays that the air showers used for the measurement are indeed initiated by primary protons. In general this is very difficult to assure. It was demonstrated in Section 4.8 how the contamination with other primaries distorts the analysis. Thus, it is an important step to argue in

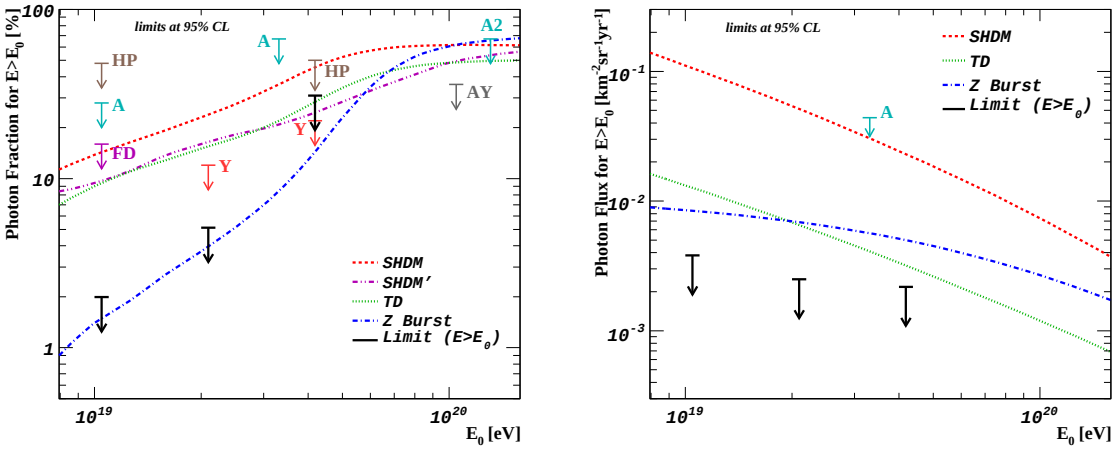


Figure 5.10: Limit on the primary photon fraction (left) and flux (right) from the Pierre Auger Collaboration and previous experiments compared to predictions by SHDM (super heavy dark matter), TD (topological defects) and Z-Burst models [182]. The labels A, HP and Y refer to the experiments AGASA, Haverah Park and Yakutsk array, while FD stands for the Auger hybrid limit. For the references to data and models see [192].

favor of a proton dominated cosmic ray composition. It is very fortunate that recent results from the Pierre Auger Collaboration can be exploited in terms of this argument.

In Figure 5.10 the latest results on the photon limit are summarized [182, 192]. At the energy relevant for this work, the limits on the photon fraction are on the 2% level. By comparing to Figure 4.22 the associated maximum systematic impact that is expected from photon cosmic ray primaries on the cross section analysis can be estimated to be $< +50$ mb.

Even more conclusive is the instrumentalisation of the recently published correlation of ultra-high energy cosmic ray arrival directions with the location of close-by active galactic nuclei objects [42]. It can be interpreted as the first observational evidence for a proton dominated composition of extragalactic cosmic rays at ultra-high energies $E > 57$ EeV. The typical energy of events available for the proton-air cross section analysis is somewhat smaller. The highest energy bin with 500 events has a mean energy of $\approx$ 13 EeV. However, no abrupt change of the cosmic ray composition is expected on such a short energy interval as from 13 to 57 EeV. Therefore the following argumentation, to require the events measured at the Pierre Auger Observatory at energies exceeding 57 EeV to be most probably protons, can be directly used as a claim for a proton beam for the cross section analysis.

The Larmor-radius of charged cosmic ray particles in the magnetic field of our galaxy is

$$r_{\text{larmor}} \approx 1.1\,\text{kpc} \cdot Z^{-1} \left(\frac{E_{\text{CR}}}{\text{EeV}}\right) \left(\frac{B}{\mu\text{G}}\right)^{-1} . \tag{5.8}$$

For ultra-high energy cosmic ray particles the Larmor-radius in the galactic magnetic field $B_{\text{gal}} = 3\,\mu$G is rapidly becoming larger than the scale of our galaxy, which is of the order of 1 to 10 kpc. Since $r_{\text{larmor}} > r_{\text{gal}}$, these particles can not be confined nor isotropized in the magnetic fields of our galaxy. Therefore, the sources of these particles must be distributed relatively homogeneous around us in order to produce the observed level of isotropy. For

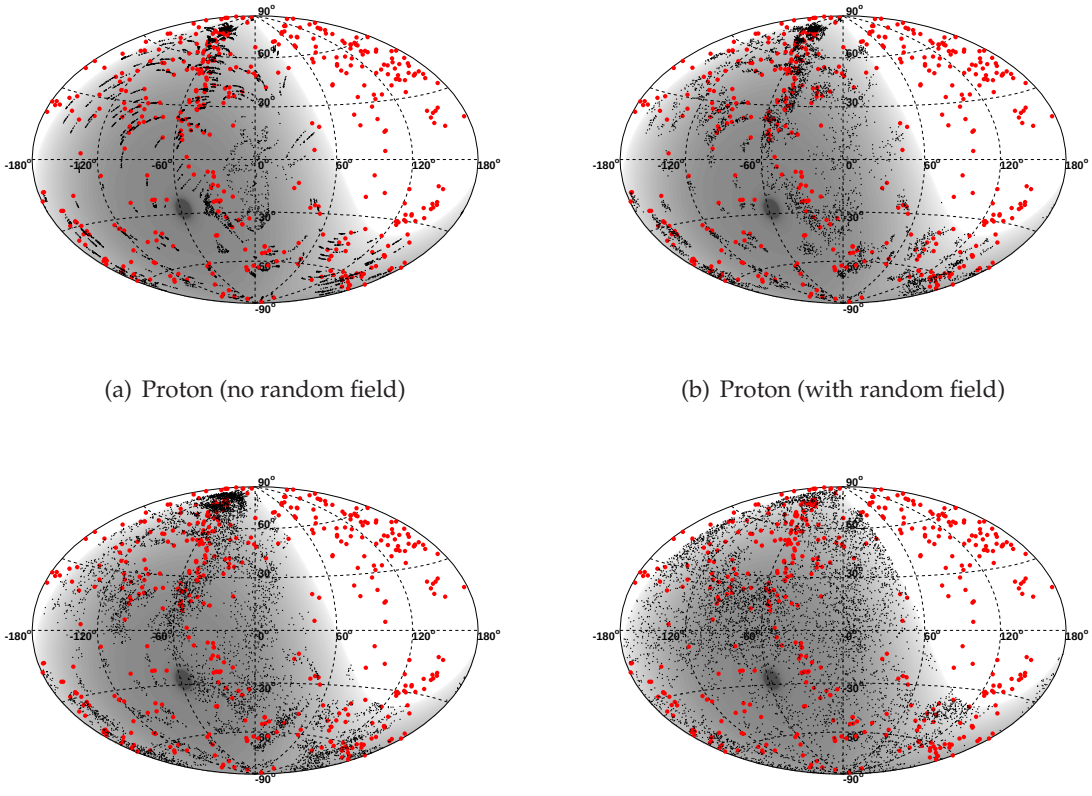
any constituent of our own galaxy this can not be achieved. Seen from the location of the Earth our galaxy is an extremely inhomogeneous object. This is a strong argument for these ultra-high energy cosmic ray particles being of extragalactic origin. However, it could not be ruled out for a long time that these ultra-high energy cosmic rays are not produced relatively close to the Earth by the decay of very heavy and exotic particles. With the new results of the Auger Observatory this is now strongly disfavored for two reasons. Firstly, decays are always associated with the accompanied production of ultra-high energy photons, which are not observed. In Figure 5.10 existing limits on the fraction and flux of ultra-high energy photons are summarized. Especially the limit of the Pierre Auger Collaboration [182, 193] is strongly restricting existing models. Secondly, the clustering of the highest energy cosmic rays observed by the Pierre Auger Observatory is not compatible with the assumed much more homogenous distribution of possible exotic particles. Furthermore, the found correlation of cosmic ray arrival directions with the location of close-by AGN is strongly favoring the scenario of cosmic ray acceleration at astrophysical sources.

The correlation observed by the Auger Collaboration on an angular scale of $\sim 3.1^\circ$ can be used to restrict astrophysical scenarios. Cosmic rays are deflected by extragalactic magnetic fields [194]. For the case of diffusive propagation, the expected width of the space angle deflection distribution after traveling a distance r through the field B with a coherence length l_c is given by [195]

$$\alpha_{\mathrm{rms}} = \frac{2ZeB}{\pi E}\sqrt{r l_c} = 1.1^\circ\, Z \left(\frac{E}{100\,\mathrm{EeV}}\right)^{-1} \left(\frac{r}{10\,\mathrm{Mpc}}\right)^{1/2} \left(\frac{l_c}{1\,\mathrm{Mpc}}\right)^{1/2} \left(\frac{B}{1\,\mathrm{nG}}\right) . \qquad (5.9)$$

The strength and spacial structure of extra galactic magnetic fields are highly uncertain, but typically adopted values are $B \sim$ nG and $l_c \sim$ Mpc. For the AGNs correlating to the data of the Pierre Auger Observatory one can obtain $r \approx 75\,\mathrm{Mpc}$ for cosmic ray energies of $E \approx 57\,\mathrm{EeV}$ the deflection angle gets $\alpha_{\mathrm{rms}} \approx 6^\circ$. This is already larger than the observed angular scale of the correlation. However, it is believed that (5.9) can only provide an upper limit, since the real structure of extragalactic magnetic fields is much more complex. There are large volumes in extragalactic space, which are empty from both matter and magnetic fields (*voids*). More sophisticated Monte Carlo simulations of extragalactic propagation, where the local magnetic field strength is coupled to the (dark) matter distribution yield deflection angles of $\approx 1^\circ$ for cosmic rays at 40 EeV, strongly depending on the traversed matter configurations and thus the direction of the particles [196]. Deflections of this size do not influence the correlation signal.

However, the magnetic fields of our galaxy are at least 3 orders of magnitude larger compared to the extragalactic fields. While the fields can be constrained by Faraday-rotation observation of distant pulsars [197], there are still large uncertainties on their strength and structure. Nevertheless it is possible to estimate the impact of the galactic magnetic fields on the spread of cosmic ray arrival directions. For this purpose a magnetic field configuration according to the *bisymmetric spiral structure* (BSS) model was adopted for a galactic Monte Carlo propagation model [198, 199]. The AGNs that were selected by the Auger Collaboration for the observed cosmic ray-AGN correlation are assumed to emit an isotropic flux of cosmic rays. A random magnetic field component with a strength of 100 % of the regular component and a coherence length of $l_{\mathrm{ran}} = 100\,\mathrm{pc}$ is assumed.



(a) Proton (no random field)

(b) Proton (with random field)

(c) Helium (with random field)

(d) Carbon (with random field)

Figure 5.11: Arrival direction sky map in galactic coordinates showing the deflection by galactic magnetic fields from the selected AGNs towards Earth for different primary particles. Red large circles denote the selected AGNs used by the Auger Collaboration to identify the correlation signal, small black dots indicate the arrival direction of cosmic rays on Earth ($E \geq 57\,\text{EeV}$). The SD acceptance is illustrated by the shaded area [179].

No extragalactic propagation effects like energy losses or deflection are considered. This means particles are not tracked from the position of the AGN, but starting from a sphere around our galaxy with a radius of 0.3 Mpc and an energy spectrum as it is measured by the Auger Observatory above 57 EeV of $\Phi(E) \propto E^{-4.1}$ [200]. The rigidity dependence of the correlation is studied by generating proton ($Z = 1$), helium ($Z = 2$) and carbon ($Z = 6$) primary particles. In Figure 5.11 the resulting distribution of cosmic ray arrival directions on Earth is shown. As can be seen, the original source structures get washed out by the random magnetic field component and with decreasing rigidity of the particle trajectories. The fraction of events within an angle α around the original source increases with rigidity, as illustrated in the left panel of Figure 5.12. Even for protons and when shifting the energies to the upper bound of our current systematic energy uncertainty, $E' = 1.25 \cdot E$, only less than 30% of the cosmic rays are expected to point back to their original source within $\alpha < 3.1^\circ$. However, in the data the original sources of particles are not known, and thus the minimum

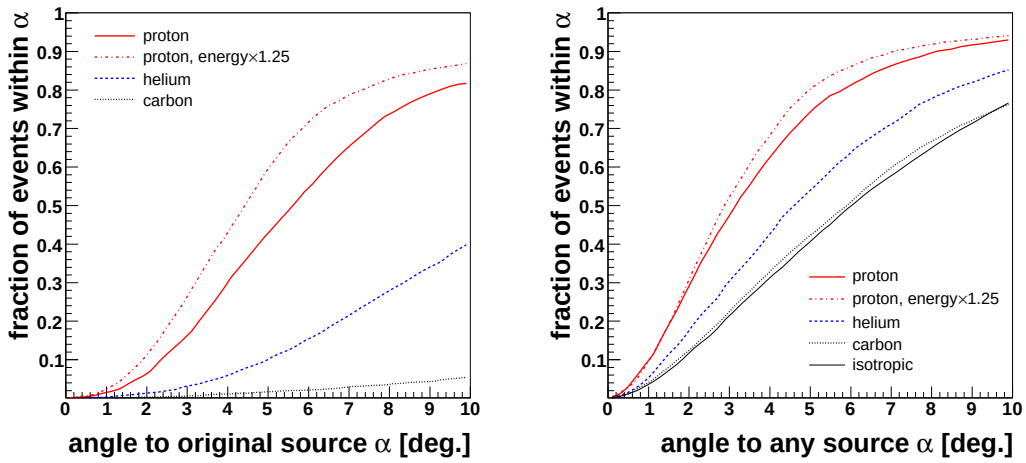


Figure 5.12: Results from an extragalactic charged particle propagation Monte Carlo study. Shown is the angular deflection for cosmic rays originating from a given set of AGNs. Left panel: angle to original source, right panel: angle to closest source. The correlation signal found by the Auger Collaboration corresponds to a fraction of ∼0.7 events within an angular radius of 3.3°. This can only be achieved with proton primaries [179].

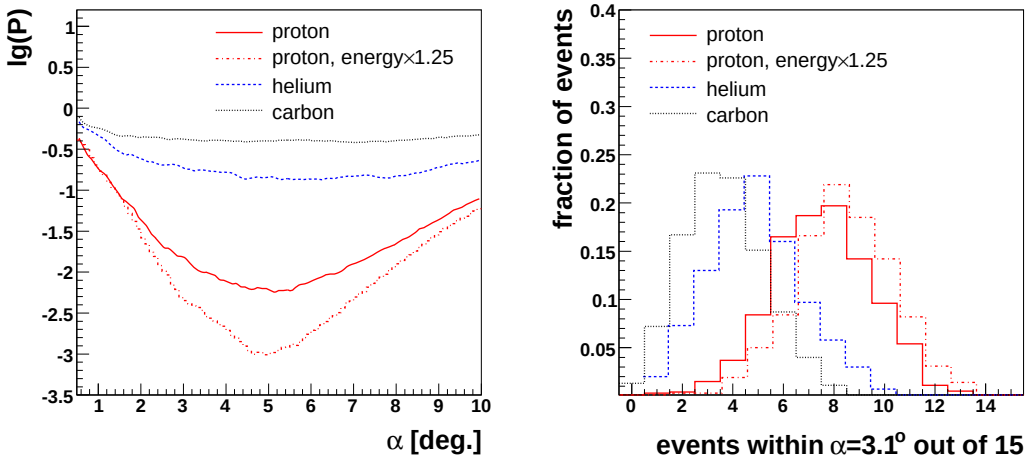


Figure 5.13: Expected correlations for the BSS galactic magnetic field model given 15 observed events [179].

angle to *any* source is shown in Figure 5.12 (right). In this case, the correlated event fractions are at the 50 % level for protons, whereas for carbon primaries there is hardly any difference to an isotropic distribution of the arrival directions.

A significance scan in analogy to the analysis in [42] is performed by generating 1000 Monte Carlo data sets. Each with $N_{\mathrm{sel}} = 15$ particles sampled to originate from AGNs with an energy above the prescription threshold of $E \geq 57\,\mathrm{EeV}$ and propagated to Earth. For all configurations an angular scan is performed, i.e. for each search window α the number of correlating events $N_{\mathrm{cor}}^{\mathrm{MC}}(\alpha)$ is derived and the probability P to observe by chance $N_{\mathrm{cor}}^{\mathrm{MC}}$ or more events from an isotropic distribution of arrival directions following the SD sky cov-

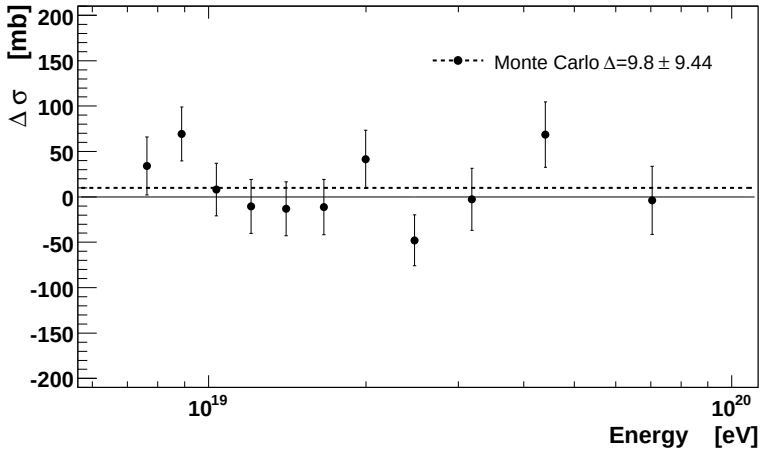


Figure 5.14: Comparison of the derived cross section from a dataset generated by full detector Monte Carlo simulations with the input cross section of the used high energy hadronic interaction model QGSJET01.

erage [201] is calculated. In Figure 5.13 (left) the average probability to obtain the same correlation from an isotropic distribution is shown. The maximum significance of the correlation (minimum chance probability) of the proton simulations is for an angle of $\sim 5^\circ$. The fraction of Monte Carlo data sets with at least 12 correlating events is 2% (5%) for protons with default (shifted) energy scale (cf. Figure 5.13, right). Since none of the 1000 Monte Carlo experiments for helium and carbon primaries resulted in more than 11 correlated events, the corresponding probabilities must be at the 0.1% level or smaller and thus it can be concluded that, for a galactic magnetic field similar to the one used for this study, only proton primaries can explain the observed AGN correlations. By taking into account extragalactic magnetic fields and energy losses this conclusion can only get stronger.

5.5 Proton-air cross section

A final cross check of the analysis method can be performed on the base of a dataset generated by full detector Monte Carlo simulations. Air shower profiles are generated with the CONEX air shower generator and the hadronic interaction model QGSJET01 for proton primaries with an energy spectrum $\propto E^{-3}$ (cf. Figure 5.1). The cross section analysis with the P_1-parameterization based on the QGSJET01 interaction model, can reliably retrieve the true input cross section over a wide energy range. This is displayed in Figure 5.14.

Now we are in the position to present the results obtained from the analysis of air shower data. Two examples of the fits to quality selected hybrid data are given in Figure 5.15. In the energy intervals at higher energies, the X_{max}-distribution outside of the fit range is described well by the proton model. At lower energies there is an increasing overestimation of the measured X_{max}-distribution at depth prior to X_{peak}. Most probably this finding needs interpretation in terms of the primary mass composition. This is beyond the scope of this thesis and also our current understanding of the data is not yet sufficient. But it can be speculated that this behavior can only be achieved by the transition from a very heavy to a very light composition. The resulting fits generally have a high quality with a reduced χ^2 close

to one. Appendix C gives a complete overview of the fits to the hybrid data. A complete overview over all free parameters of the reconstruction as well as the resulting χ^2 is given in Figure 5.16. As a compromise between bin width and statistics within a bin, the energy ranges are chosen such as the two highest energy bins contain 500, the third bin 1000 and all the other lower energy bins 2000 events. This choice allows to make the energy-dependence better visible for very high energies. The X_{shift}-parameter shows an almost linear behavior in $\log_{10}(E/\text{eV})$.

In the lower energy bins the measured cross section is comparable to the predictions of the high energy hadronic interaction models, but the X_{shift}-parameter is indicating significant deviations. To explain a shift of this magnitude just by a changing multiplicity of the interactions, might not suffice (cf. Figure 4.3). At the highest energies, on the other hand, X_{shift} gets small, but above $10^{18.4}$ eV the cross section tends to increase rapidly. In total it seems the interaction models need adjustment over the hole energy range.

5.6 Systematic uncertainties

Several sources of systematic uncertainties have been already identified in previous sections of this work. At this place a complete list is given and the overall impact is quantified.

There are several uncertainties due to the chosen reconstruction ansatz and the choice and quality of the parameterizations. Furthermore, there are more general sources of uncertainties, on the level of the physics distributions that will equally apply to other cross section analysis. First of all this is the primary mass composition of cosmic rays, but also the chosen fitting range. Atmospheric fluctuations of the density profile are of importance both for ground based air shower arrays as well as telescope detectors, while fluctuations of the optical properties of the atmosphere (aerosols) are relevant only for the latter.

Parameterization of the X_{max}-resolution. The X_{max}-resolution is parameterized on the basis of events generated by a full detector Monte Carlo simulation over a wide energy range. The found resolution is cross checked with the available 132 stereo events showing a very good agreement. However, the uncertainty of the stereo measurement of

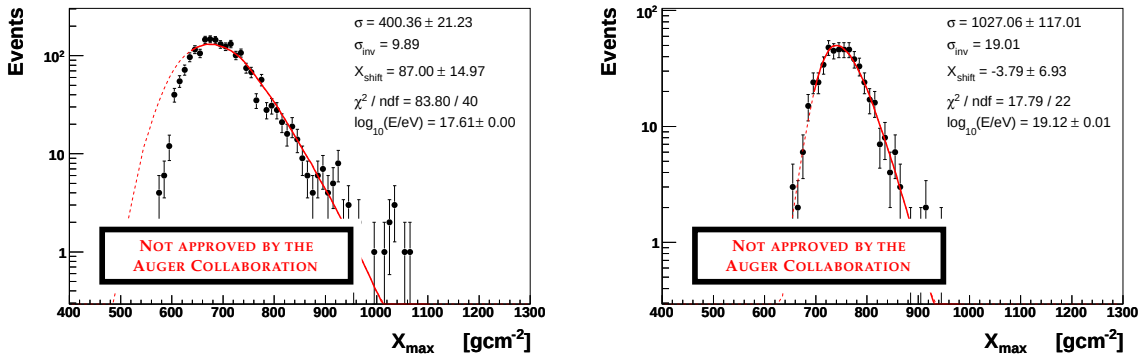


Figure 5.15: Two example fits of the X_{max}-model to hybrid data with the P_1-distribution generated with SIBYLL. The left panel shows the 2000 lowest energy events and the right panel the 500 highest energy events of the considered dataset. The fitted X_{max}-model to the data is shown as a solid red line over the interval where the fit is performed and it is prolongated as dashed line.

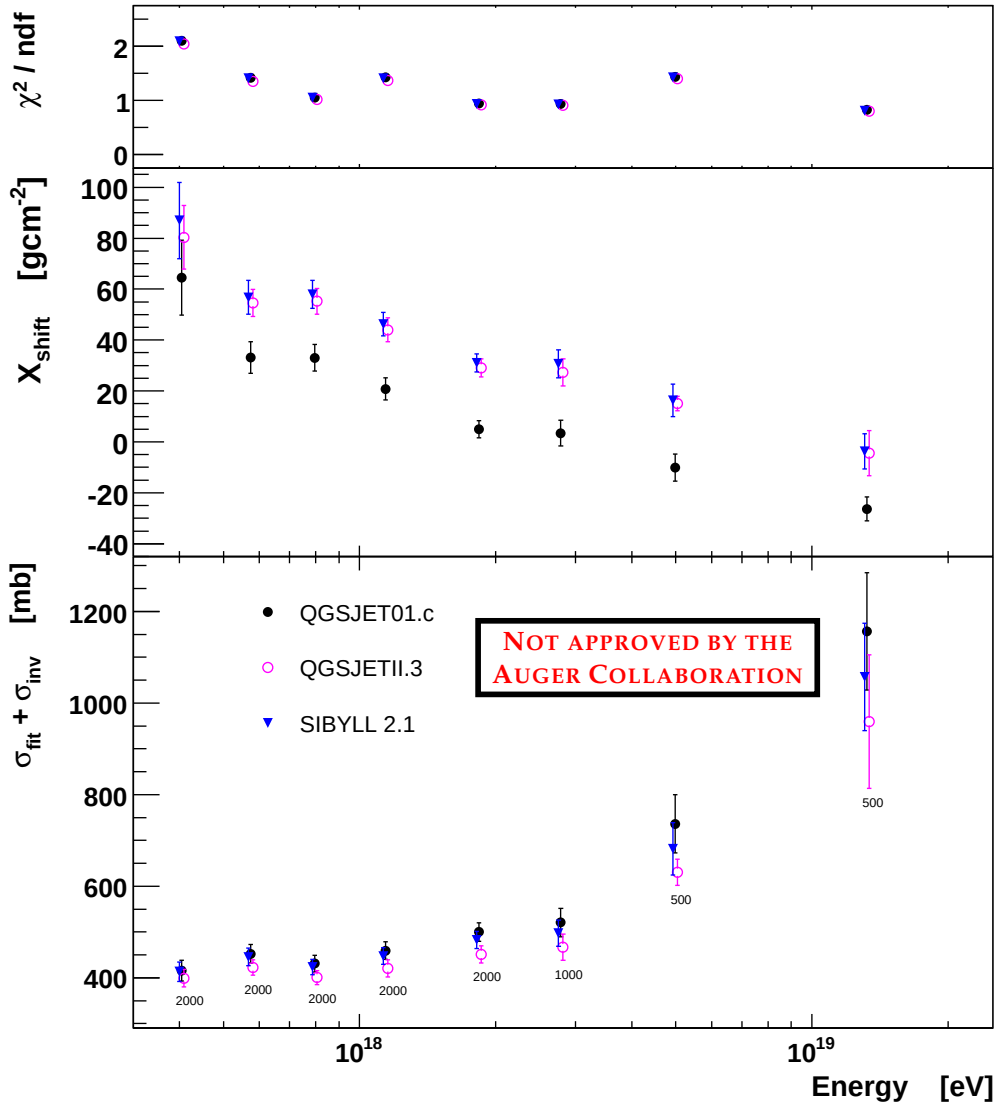


Figure 5.16: Complete overview of the results from the cross section analysis of hybrid data. At each energy bin the number of events used for the analysis is indicated. The error bars are only statistical.

$\pm 1.3\,\mathrm{gcm}^{-2}$ can be propagated into the parameterization and thus into the result of the cross section analysis. The result of doing this are shown in Figure 5.17. An energy dependent deviation from the original resolution model is identified, which can be parameterized with a polynomial of 2$^{\mathrm{nd}}$-degree

$$\begin{aligned}\Delta^{-}_{\mathrm{res}}(E) &= -1698.2\,\mathrm{mb} + 194.6\,\mathrm{mb}\,\log_{10}(E/\mathrm{eV}) - 5.6\,\mathrm{mb}\,\log^2_{10}(E/\mathrm{eV})\\ \Delta^{+}_{\mathrm{res}}(E) &= +2092.1\,\mathrm{mb} - 239.2\,\mathrm{mb}\,\log_{10}(E/\mathrm{eV}) + 6.8\,\mathrm{mb}\,\log^2_{10}(E/\mathrm{eV})\,. \end{aligned} \tag{5.10}$$

Reconstruction ansatz. The performance of the reconstruction is thoroughly studied in Section 4. Three independent sets of parameterizations of P_1 are available, one for each of the models QGSJET01, QGSJETII and SIBYLL. A large variety of Monte Carlo generated X_{max}-distributions is used to check the quality of the reconstruction method.

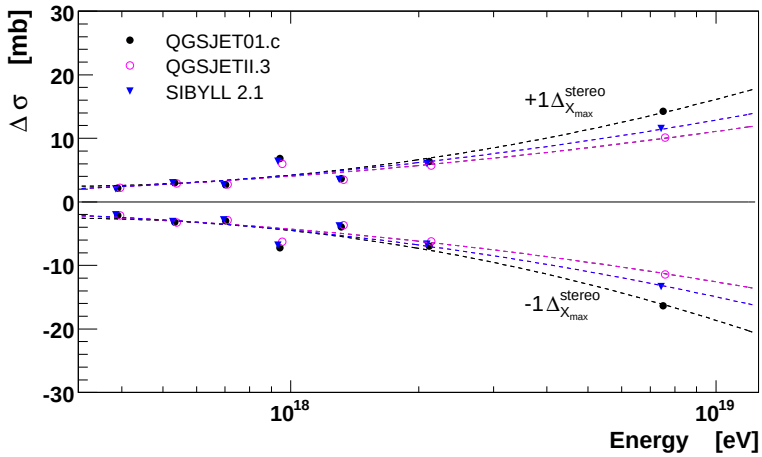


Figure 5.17: Difference in the resulting cross section for the X_{max}-resolution model with $\pm 1\Delta^{\mathrm{stereo}}_{X_{\mathrm{max}}} = 1.3\,\mathrm{gcm}^{-2}$, where the latter is the uncertainty of the X_{max}-resolution as derived from the stereo data. The reference (at zero) is the reconstruction with the X_{max}-resolution model as described in the text. The lines are 2$^{\mathrm{nd}}$-order polynomial fits. The fit to the QGSJET01 points is used as an estimator of the corresponding systematic uncertainty.

Two parameters are identified that are mainly responsible for the resulting performance of the cross section reconstruction. These are the quality of the parameterization of ΔX_1 and the chosen fitting range. In Figure 4.18 it is shown how the parameterization can be optimized by requiring a minimal inelasticity of 0.01. The fitting range is also studied intensively and Figure 4.19 is summarizing the results. The chosen fitting range in X_{max} starting from 50 gcm^{-2} in front of X_{peak} until 40 gcm^{-2} beyond the largest X_{max} was done in order to prevent any significant additional systematic uncertainty.

Especially in Figure 4.20 the high quality that is finally achieved is demonstrated by the ability of the reconstruction ansatz to retrieve the true input cross section over a wide range. While the reconstruction of smaller input cross sections compared to the original model cross section works very well and without any additional uncertainty, for an increased cross section the reconstruction becomes more difficult. But only for very large cross sections a significant underestimation seems to occur. The corresponding systematic uncertainty is small and assumed to be zero and independent of the cross section.

Based on the results from Section 4, the total uncertainty associated to the reconstruction ansatz is estimated to be approximately -25 mb and +5 mb at 10 EeV.

Invisible cross section. The correction for the invisible cross section for the chosen $k^{\mathrm{min}}_{\mathrm{inel}}$ is resulting in an uncertainty on the level of 1%, slightly depending on the energy. This is just the maximal span between the predictions of the different models divided by two. The result is illustrated in Figure 5.18.

Atmosphere. The properties of the atmosphere fluctuate with time and are only measured with a limited precision. This induces an additional broadening of the observed physics

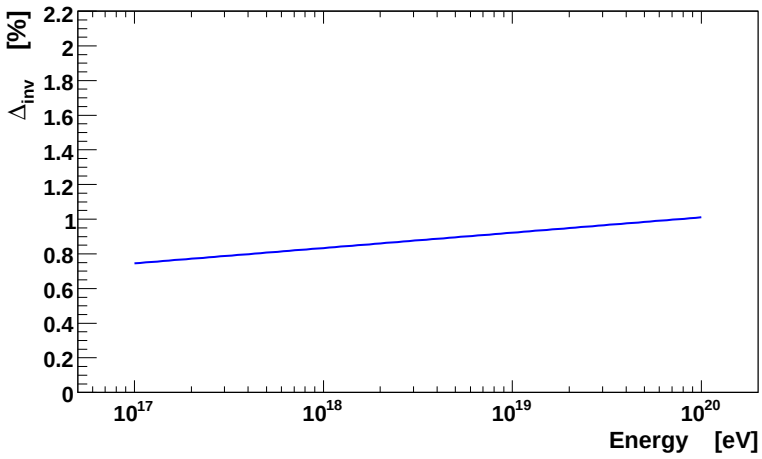


Figure 5.18: Uncertainty associated to the invisible cross section correction.

distributions like X_{max}. It is important to assure that this is not mixed up with the signal from a changing cross section. The atmosphere above the Pierre Auger Observatory is permanently monitored as an integral part of the standard data taking procedure. However, only ≈60% of all hybrid events have an associated measurement of the aerosol profile of the atmosphere. For the remaining 40% an average atmosphere is used during the reconstruction. In Figure 5.19 the results of the cross section analysis for these two different data samples are compared. Fortunately the parameterization of the mean aerosol profile used for the analysis of the data with no aerosol measurement, produces cross sections that are in good agreement with the data that have an associated aerosol measurement. Therefore no systematic uncertainty on the cross section needs to be attributed to fluctuations of the aerosol content in the atmosphere.

Acceptance. Any deviation from the acceptance model as it is used in the X_{max}-model dis-

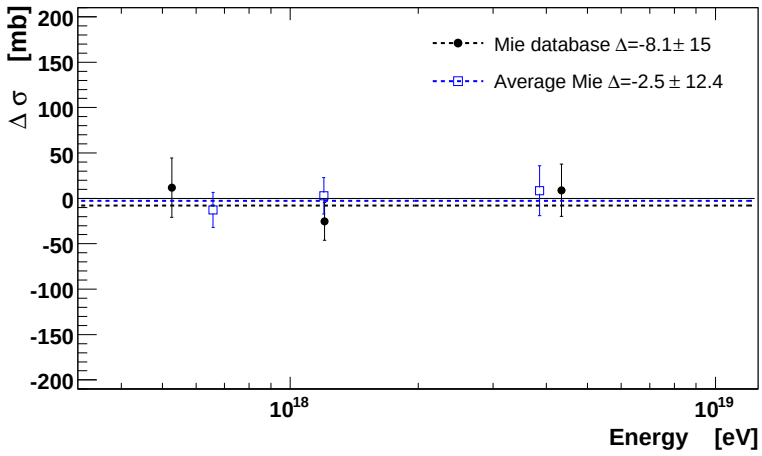


Figure 5.19: Comparison of the resulting cross sections of the subset of events with an accompanied valid atmospheric aerosol measurement (Mie database, about 60% of the events) and with the average aerosol model (average Mie) to the reconstruction of all events.

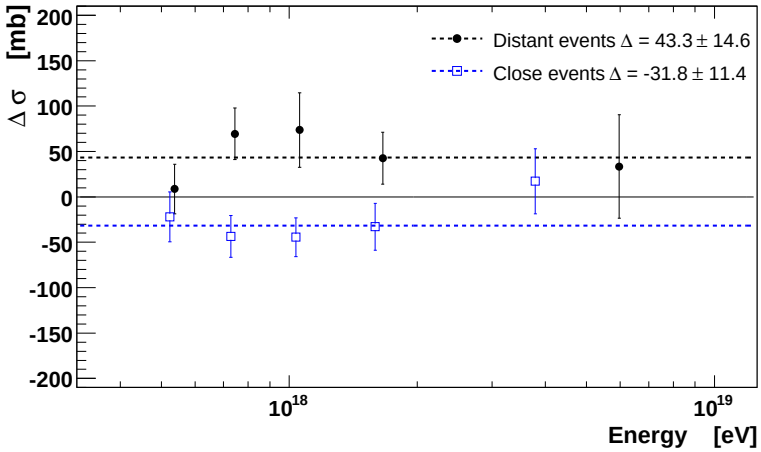


Figure 5.20: Independent analysis of events located close or far from the telescope. The reference line at zero is the result of the analysis of the total dataset.

torts the resulting X_{max}-distribution. Checking the acceptance is possible by analyzing subsamples of the dataset that have an inherently different acceptance. Suited for this purpose are the classification of events with respect to their distance from the telescope as well as their reconstructed zenith angle. Because of the geometry of the field of view of a telescope, events closer to the telescope have a more restricted fiducial slant depth range than events at larger distances. Inclined shower events traverse much more atmospheric matter compared to more vertical events. Therefore inclined events allow to observe more of the tail of the X_{max}-distribution. In fact the limited extent of the atmosphere is strongly cutting into the X_{max}-distribution for vertical events.

In Figures 5.20 and 5.21 the resulting impact on the cross section analysis are demonstrated. For each energy bin the data is divided in two independent sub-samples: close and far events as well as vertical and inclined events. As can be seen, the corresponding cross sections are measured to be significantly different for these sub-samples. Cor-

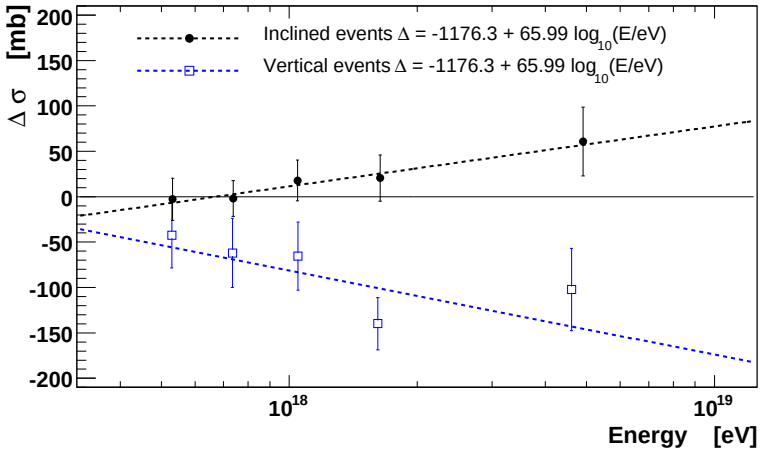


Figure 5.21: Independent analysis of data subsets according to their zenith angle. The reference line at zero is the result of the analysis of the total dataset.

Table 5.3: By ignoring the cross section dependence of the correlation function P_1, the model dependence of the analysis becomes much more severe [117]. Systematic biases of the order of several 100 mb are observed during the reconstruction of the cross section for 10 EeV proton simulations.

data	Δ X$_1$ model used for reconstruction							
	QGSJET		SIBYLL		NEXUS		QGSJETII	
data model	σ_{rec}	X_{shift}	σ_{rec}	X_{shift}	σ_{rec}	X_{shift}	σ_{rec}	X_{shift}
QGSJET01	**537.1±5.5**	1.7±0.8	465.3±4.1	-29.8±0.8	447.9±3.2	-46.9±0.8	467.6±3.7	-19.6±0.7
SIBYLL	802.8±12	36.1±0.7	**609.2±5.7**	0.7±0.5	572.8±4.5	-17.4±0.5	613.3±5.7	10.6±0.5
NEXUS	749.9±10	55.1±1.2	569.1±3.9	19.5±0.9	**543.1±4.0**	1.5±0.5	576.5±4.6	29.5±0.8
QGSJETII	697.1±8.2	27.0±0.8	553.2±4.3	-7.5±0.6	521.6±3.7	-26.1±0.3	**562.3±4.4**	2.1±0.4

respondingly, acceptance effects have a large impact on the cross section analysis. As a matter of fact, due to the very much reduced theoretical systematic uncertainty caused by the model dependence, the observed differences dominate now the systematic uncertainties. Clearly, a better understanding of these effects is needed to improve the experimental precision of the proton-air cross section analysis.

It is an additional complication that in the case of a mixed composition cosmic ray beam, for different choices of event subsets we might introduce selection effects in term of the primary mass. For example for vertical events, the X_{max} for proton events at ultra-high energy is generally already below the ground level and therefore an effective depletion of the selected proton fraction occurs. Thus, it is far from trivial to disentangle the impact caused by the acceptance, the available fitting range in slant depth and a possible changing mass composition on the resulting cross section measurement.

In order to quantify the bias associated with the acceptance, the maximum deviation from the distance (Figure 5.20) and zenith (Figure 5.21) study is taken as an estimates of the systematic uncertainty. At the lower energies this amounts to around 50 mb growing to up to 150 mb at the highest energies. The effect is not symmetric and the uncertainty for an underestimation of the cross sections is larger.

Composition. As demonstrated in Section 4.2.2 the primary composition of cosmic rays has a major impact on the X_{max}-distribution and thus on the cross section analysis (cf. Section 4.8).

Because of the available photon limits the influence from photon contamination can be estimated and we associate a maximum of +50 mb systematic shift of the cross sections due to an admixture of photons.

For the heavy nuclei contamination of the cosmic ray proton beam no equivalently strong conclusion can be drawn. Since the mass composition is very uncertain, in principle a large variety of systematic shifts of the cross section to lower values are possible. However, the fits of the proton model to the X_{max}-data, generally describes the distribution well even outside of the fitted range. This can be read as indication of a pure-proton or at least single-component composition. Only at the two lowest energy bins of the presented cross section analysis there seems to be a significant overestimation of the X_{max}-distribution outside of the fit interval (at X_{max} prior to the peak of the distribution). Interestingly also the energy dependence of the X_{shift}-parameter seems

to show a feature at low energies.

Model dependence. The model dependence can be deduced directly from the different reconstruction of the data by the P_1-distributions generated with different models. For the finally specified cross section measurement the mean value as well as the mean statistical error of the individual reconstructions are calculated. The associated model dependent systematics is then the maximum deviation between the models divided by two. In the lower energy bins this amounts to ± 5 mb while in the upper energy bins it becomes a significant source of uncertainty with up to ± 90 mb.

The improvements with respect to previous analysis techniques can be clearly seen by looking at the differences of Figures 4.20 and 4.21. Furthermore, the model dependence of the analysis technique without consideration of the cross section dependence of the correlation function P_1 is quantified to be ≈ 200 mb at 10 EeV [117] (see Table 5.3).

The investigated sources of systematic uncertainty are listed in Table 5.4 together with an estimate of their impact on the resulting cross section. The energy dependence of these systematic uncertainties in millibarn, as they apply to the final cross section measurement, is shown in Figure 5.22. The dominant source of uncertainty is coming from the detector acceptance. Below $10^{18.3}$ eV all other uncertainties are smaller than 20 mb. Only at higher energies the model dependence is rapidly growing, eventually becoming a comparable strong source of uncertainty than the acceptance.

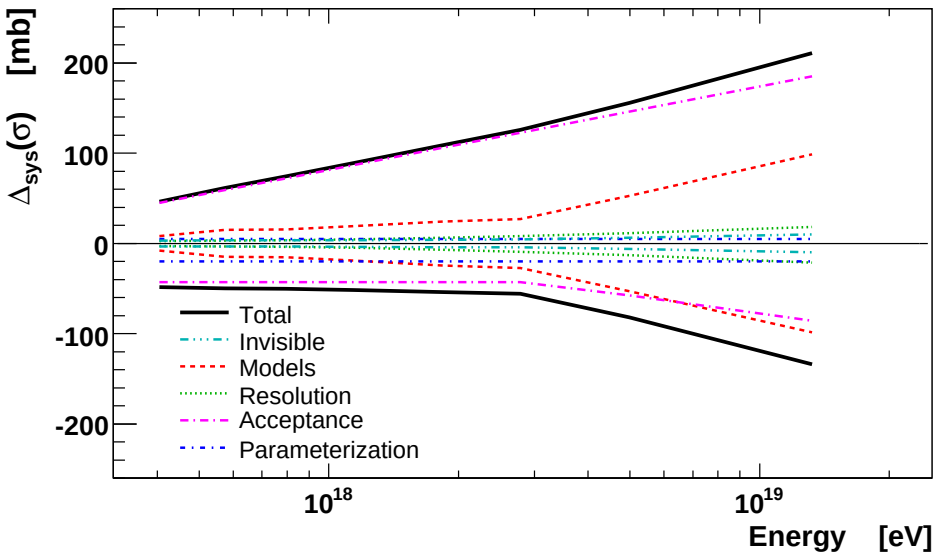


Figure 5.22: Decomposition of the total systematic uncertainty versus energy. It is only at the highest energies that the model dependent uncertainties get significant. The uncertainties caused by the unknown primary composition are not shown.

Table 5.4: Overview of the study of systematics uncertainties. The impact of the primary composition is discussed separately and is not mixed up with the other points.

Source	Impact
Acceptance	50 to 150 mb
Model dependence	5 to 90 mb
Parameterization of P_1	+5 and -25 mb
Parameterization of P_{res}	5 to 20 mb
Invisible cross section	$\sim 1\%$
Atmosphere (aerosols)	vanishing
Primary composition (photons)	up to +10%
Primary composition (heavy nuclei)	up to -100%

5.7 Discussion of the results

5.7.1 Proton-air cross section and comparison to model predictions

The combined results of the cross section analysis based on the hybrid data are presented in Figure 5.23. The mean value of the reconstruction with the three P_1-models is given together with the mean statistical uncertainty. The estimation of the systematic uncertainties as derived in the previous section is quadratically added to the statistical error bars to give an impression of the total uncertainty. The results are summarized in Table 5.5. There is an additional uncertainty, which has not been noted up to now. The energy scale of the hybrid reconstruction is uncertain on the 25% level. This is not propagated into the final results, but the implications are minor. The results of the proton-air as well of the proton-proton analysis can be shifted by up to 25% in energy. Within the total uncertainty, the predictions of QGSJET01, QGSJETII and NEXUS are in good agreement with the data below $10^{18.4}$ eV. Due to the systematic uncertainty of this measurement, the somewhat larger cross section of SIBYLL and EPOS can not be excluded.

Above $10^{18.4}$ eV a strong increase of the reconstructed cross section is observed. This is a consequence of the measured X_{max}-distribution getting much narrower than any proton simulation would allow (cf. Figure 5.24). It is worthwhile noting that the systematic uncertainty of this analysis are correlated with respect to the energy, i.e. whereas the overall scale of $\sigma_{\mathrm{p-air}}$ is uncertain at the 100 mb level, the measured shape is much less uncertain.

5.7.2 Mass composition of primary cosmic rays

It is too early to draw reliable conclusions about the mass composition at ultra-high energy from air shower data. More data need to be acquired and the understanding of the detectors and of the interaction physics have to be improvents. However, a strong impact of a mixed composition on the results of the cross section analysis was demonstrated in Section 4.8. Therefore we can review the result obtained in this work also in the context of a possible mixed composition scenario.

The primary assumption for the presented analysis is the ability to attain a cosmic ray proton beam, even if the primary mass composition is mixed. For the presented work this was done for example by considering only the deeply penetrating part of the X_{max}-

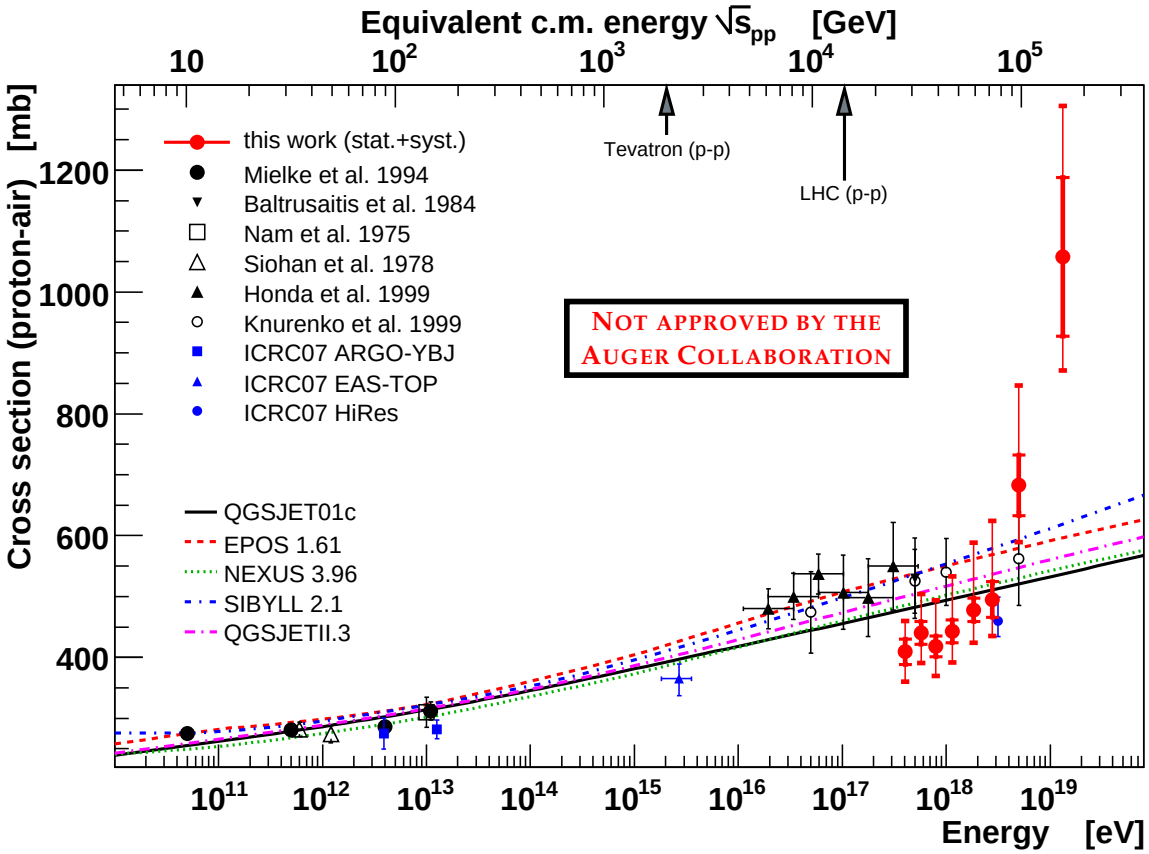


Figure 5.23: Final cross section results plotted together with other data and model predictions. All measurement are well above the LHC energy. The inner error bars indicate statistical (thick line) and the outer error bars the total of systematic and statistical uncertainties. The uncertainties caused by the unknown mass composition are not included.

Table 5.5: The proton-air cross section with its statistical and systematical uncertainties. The uncertainties caused by the unknown mass composition are not included.

Energy $\log_{10}(E/\mathrm{eV})$	Events	σ_{p-air} [mb]	Δ_{stat} [mb]	Δ^{-}_{syst} [mb]	Δ^{+}_{syst} [mb]
17.6	2000	409	20.9	44.0	45.9
17.8	2000	440	18.8	45.8	61.1
17.9	2000	418	17.0	46.0	74.0
18.1	2000	443	18.7	47.5	89.0
18.3	2000	478	19.2	50.1	109
18.4	1000	495	29.4	51.8	126
18.7	500	683	49.5	79.5	156
19.1	500	1060	130	133	211

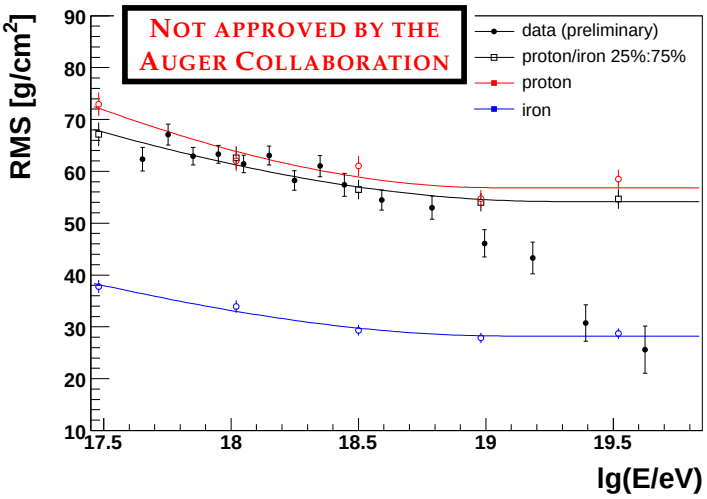


Figure 5.24: The fluctuations of the X_{max}-distribution plotted together with model predictions by QGSJETII [190]. Above an energy of $10^{18.5}$ eV the data considerably start to deviate from the model expectations for a proton or even mixed (25% proton, 75% iron) composition.

distribution. Based on the found impact of the chosen fitting range on the analysis for the case of a mixed primary composition we can perform the data analysis for the "short fitting range" scenario as it is introduced in Section 4.8. In this spirit the data is reanalysed for two modified fitting ranges

Late start. The fit starts at X_{peak}. This proved to be useful to suppress the impact of heavy admixtures.

Early stop. The fit ends at $X_{\mathrm{peak}} + 150\,\mathrm{gcm}^{-2}$. This helps to suppress the impact of a photon contribution.

The results of this study are shown in Figure 5.25. The "early stop" fitting range is indeed leading to systematic lower reconstruction of the cross section, as it would be expected in the presence of photon primaries. The offset of -39 gcm^{-2} points to maximum photon fraction on the level of a few percent, which is compatible with the photon limit of 2%. On the other hand the "late start" fitting range is leading to a systematic shift by an offset of +58 gcm^{-2}. This could be caused by a contribution of heavy nuclei.

In any case on the present level the results cannot be easily transformed into quantitative estimates on the mass composition. Systematic effect caused just by the choice of the fitting range itself must be considered. Also the resulting quality of the short fits is much worse than for the standard fitting range.

However, the offsets identified in Figure 5.25 can be taken to justify the assumptions made earlier, concerning the estimated maximum impact of systematic uncertainty caused by the primary composition.

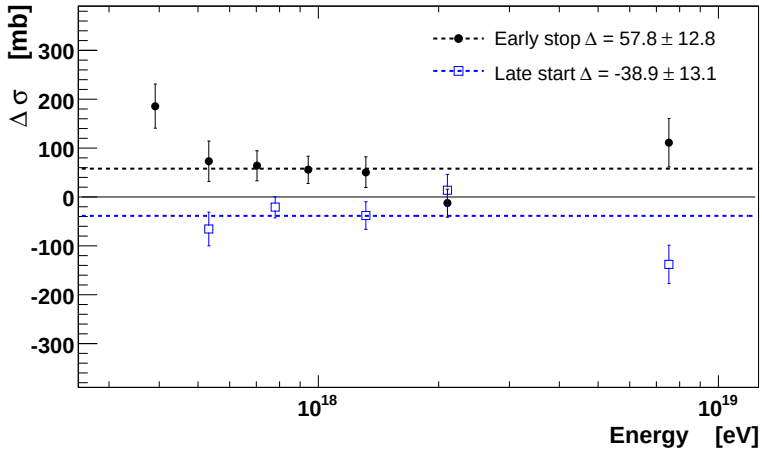


Figure 5.25: Investigation of an alternative reconstruction of the hybrid data with different fitting ranges.

5.7.3 Proton-proton cross section and the Glauber theory

In a final analysis step, the derived proton-air cross section can be converted into proton-proton cross section by application of the Glauber theory (cf. Section 2.7.2). There are additional uncertainties entering in this conversion process, which is mainly the extrapolation of the slope parameter B_{ela} to the needed ultra-high energies.

For a measured value of a proton-air cross section at a specific energy the Glauber theory needs a value for B_{ela} in order to convert to the proton-proton cross section. The extrapolation of B_{ela} to cosmic ray energies is very uncertain, therefore it is helpful to plot the trajectory of possible B_{ela} as a function of $\sigma_{\text{tot}}^{\text{pp}}$ for each measurement of a proton-air cross section. For each of the eight energy bins of the proton-air cross section analysis this is done in Figure 5.26. The trajectory in the plane of B_{ela} and $\sigma_{\text{tot}}^{\text{pp}}$ is shown for the proton-air value, the $\pm 1\sigma_{\text{stat}}$ contour as well as the total statistical plus systematic uncertainty. Up to center of mass energies of $\log_{10}(\sqrt{s}\,/\,\text{GeV}) = 4.9$ the conversion can be done by using a mean Donnachie-Landshoff (DL) parameterization. It has to be stressed again, that the DL predictions are very uncertain at these energies and a large systematic uncertainty is present due to the enormous extrapolation from the few available data points. This additional systematics is not included in the results of this analysis. However, the plots in Figure 5.26 should be taken as a reference. These plots give a good impression on the possible results of the conversion. Furthermore, with the availability of new data on the slope parameter or an improved parameterization, the value of the proton-proton cross section can be obtained directly from these diagrams. For the two highest energy measurements the model of DL clearly fails to describe the data. For proton-proton cross sections larger than $\approx 225\,\text{mb}$ the DL model hits the black disc limit, which requires a minimum value of the slope parameter. For very high cross section the black disc limit might be a good approximation, thus the two highest energy data points are converted by the assumption of this limit. In Figure 5.27 the proton-proton cross sections including their statistical and systematic uncertainty are given. The systematics is just the one arising from the proton-air analysis.

The plots from Figure 5.26 at ultra high energy can also be exploited in order to give

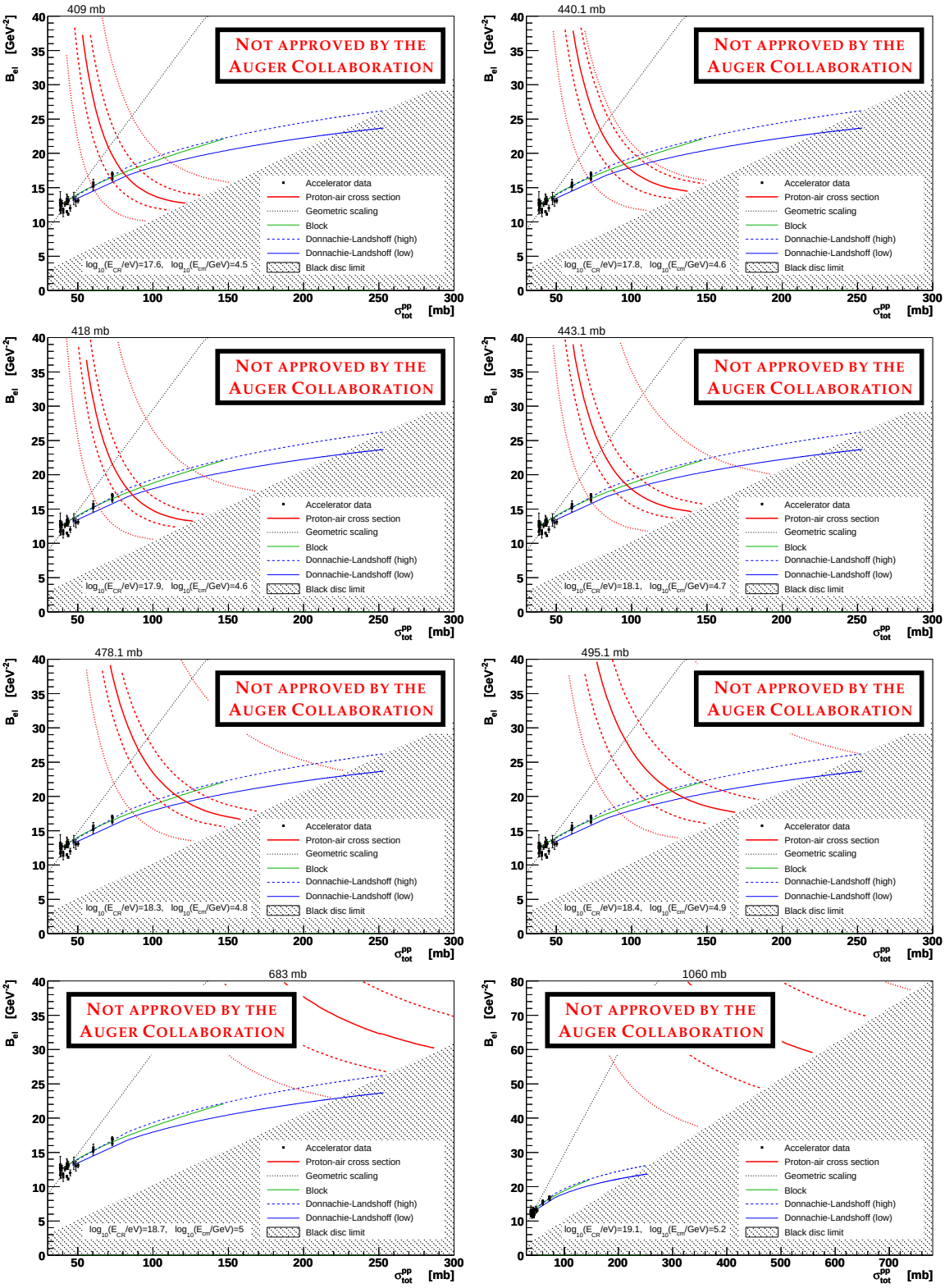


Figure 5.26: Each proton-air cross section measurement defines a curve in the $(B_{\mathrm{ela}} - \sigma_{\mathrm{tot}}^{\mathrm{pp}})$-plane. The solid red line corresponds to the mean value of the proton-air cross section, while the dashed lines are the $\pm 1\sigma_{\mathrm{stat}}$ values and the dotted lines the total uncertainty including the systematics. The highest energy interval (bottom right) shows a larger section of the plane. For the references see [121].

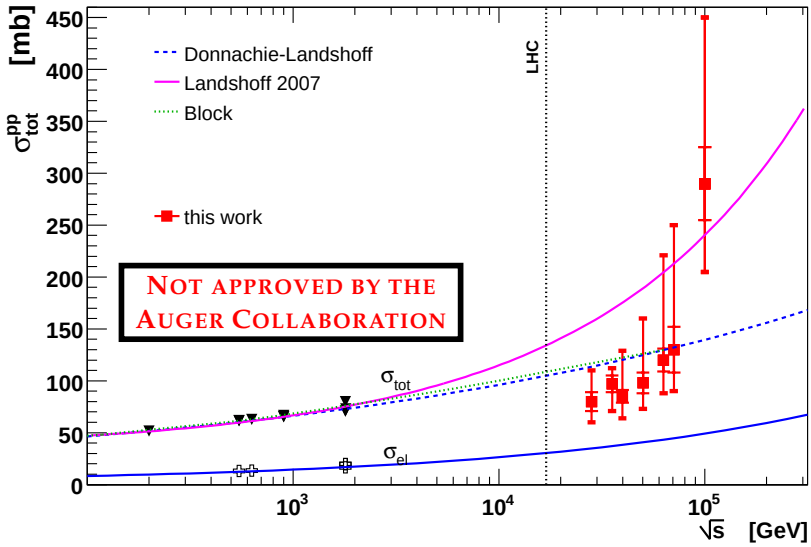


Figure 5.27: Resulting proton-proton cross section after application of the Glauber formalism. The inner error bars denote statistical and the outer error bars systematical uncertainties. The highest energy point at $\log_{10}(\sqrt{s}\,/\,\mathrm{GeV}) = 5.2$ with $560 \pm 100_{\mathrm{stat}} - 210_{\mathrm{syst}} + 180_{\mathrm{syst}}$ mb is not plotted.

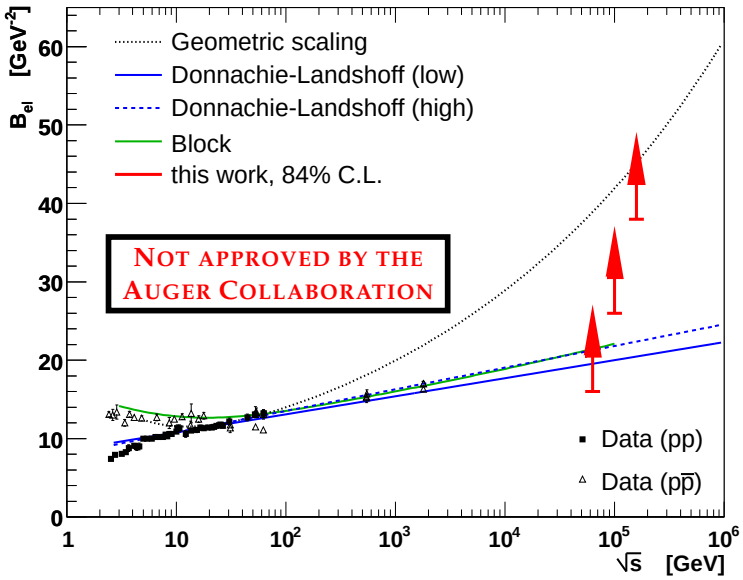


Figure 5.28: Uncertainty of the extrapolation of the slope parameter B_{ela} to cosmic ray energies (for the references see [121]). The analysis presented here can be used to put limits on B_{ela} at ultra-high energy.

a lower limit on the slope parameter at cosmic ray energies. If we account for the total uncertainties (stat.+syst.) we can derive limits at the three highest energy intervals. These limits are illustrated in Figure 5.28.

The upcoming data from the LHC will help to considerably improve the understanding

of cosmic ray data. If the cosmic rays we observe at the highest energies is indeed a proton beam, we have to expect a drastic change in the interaction physics close to $\log_{10}(\sqrt{s}/\mathrm{GeV}) \approx 5$. This is well above the LHC energy. At the LHC the possible predictions spanned by DL and Landshoff 2007 are both possible, with a slight preference of a lower cross section at LHC energies.

Cosmic rays currently provide the only possibility for studying interactions at center of mass energies of $\log_{10}(\sqrt{s}/\mathrm{GeV}) > 5$. By further developing the analysis technique, and improvements in understanding of the atmosphere, the detector acceptance and the nature of ultra-high energy cosmic rays it might be possible to study ultra-high energy interaction physics based on cosmic rays with good precision in the near future.

6 Summary

The aim of this work has been a first analysis of hybrid data of the Pierre Auger Observatory to measure the proton-air cross section.

There are different methods that have been applied in the past for measuring the proton-air cross section. These methods use either ground-based (particle detector) observables or telescope data (fluorescence or Cherenkov light signal). In preparation of performing a cross section analysis of hybrid data, these existing analysis methods have been reviewed and shortcomings have been pointed out.

- All cross section measurement techniques depend on hadronic interaction models because the relation between the first interaction point and the used shower observable has to be obtained from simulations.
- The previous methods treat the proton-air cross section in an inconsistent way. The measured cross section is attributed only to the distribution of the first interaction point, but does not enter the shower evolution. The shower evolution is always simulated with the default cross section of the interaction model used in the analysis.
- The correlation between the first interaction point and the depth of the shower maximum X_{max} is more direct than that with ground-based particle observables.
- The k-factor method to account for the fluctuations between the first interaction point and the observed shower stage depends very sensitively on the chosen range for fitting.
- Experimental effects are important for all of the measurement approaches. From this it follows that k-factors cannot be transferred between different experiments.
- Because of the different stages of the observed shower development, k-factors from ground arrays are inherently different from k-factors of detectors that can directly observe the shower maximum.

Based on these results, a new method for deriving the high energy proton-air cross section using X_{max}-data has been developed and verified by simulation studies. This method has a much smaller dependence on the hadronic interaction model applied in the analysis than previous methods. The method developed in this thesis is characterized by

- Interpolation of the cross section starting from model predictions at an energy at which the cross sections are known from accelerator data up to the air shower energy. The correlation between the first interaction point and the depth of the shower maximum is parameterized as a function of the interpolated cross section and primary energy. This way the cross section derived from data is not only determining the distribution of the first interaction point, but also influences the shower evolution and, hence, the correlation between the first interaction point and X_{max}.
- Empirical determination of the invisible cross section as

$$\sigma_{\mathrm{inv}} = \sigma\left(k_{\mathrm{inel}} \geq 0.01\right) , \tag{6.1}$$

  to which any air shower measurement is not sensitive to.

- Due to the inclusion of the acceptance in the developed X_{max}-model, there is no need to strongly constrain the dataset to the phase-space that guarantees the undistorted observation of the complete X_{max}-distribution. A definition of the fiducial slant depth range based on the geometry of the field of view of a telescope has been developed. It has been shown that the presented analysis method has the potential to use almost the full hybrid statistics.
- Inclusion of the detector resolution in the analysis. In this work, the X_{max}-resolution is parameterized and verified with stereo data.

The simulation studies needed for the cross section analysis required the extension of existing software packages and the development of new simulation algorithms. The hybrid simulation as well as reconstruction algorithms have been significantly improved. These contributions are now integrated in the official releases of the $\overline{\mathrm{Off}}\underline{\mathrm{line}}$ framework and include

- Fast hybrid simulations with `SdSimpleSim`.
- Numerical slant depth integration of the traversed matter in a curved atmosphere.
- Extensive revision and re-implementation of the telescope simulation algorithms with a focus on the relevant physics, computing speed, runtime stability and user customization.
- Numerical error propagation for arbitrary calculations.
- Development of the `ADST` data format and the `EventBrowser` visualization tool.

The hybrid data of the southern Pierre Auger Observatory from 2004 to October 2007 have been chosen for the analysis of this work. After quality selection, more than 12000 events are available for the cross section measurement. Applying the developed analysis method to this data set and assuming a proton dominated composition, the following results have been obtained

- Below $10^{18.4}$ eV, the cross section is compatible with QGSJET01, QGSJETII and NEXUS. EPOS and SIBYLL are less favored, but are not excluded.
- Above $10^{18.4}$ eV, the cross section is increasing much stronger than expected by standard extrapolations.
- The uncertainty induced by the model dependence is small.
- The mean depth of the shower maximum derived from data and predicted by models is different. The difference is steadily decreasing $\propto \lg(E/\mathrm{eV})$ for all models.

The systematic uncertainties of the measurement have been studied in detail. The following aspects have been investigated using simulations as well as data.

- Impact of a mixed hadronic composition on the cross section analysis.
- Effect of photon primaries.

- Stability of the analysis method to reconstruct simulated data generated by any one interaction model and reconstructed with another model.
- Quality of the used parameterizations of the X_{max}-resolution and the corresponding correlation function.
- Aerosol content of the atmosphere.
- Identification of systematic effects that are attributed to the understanding of the detector acceptance, by the reconstruction of subsets of the total event selection.

After the significant reduction of the systematic uncertainties due to hadronic interaction models achieved in this work, the remaining uncertainties are dominated by the understanding of the detector acceptance.

The AGN-correlation recently reported by the Auger Collaboration strongly favors protons as primary particles. The correlation is found for events above 57 EeV. The data analyzed in this work cover the range from 0.4 to 79 EeV. Above 57 EeV only very few events are available. However, it is at the highest energies where the presented analysis indicates a very rapid increase of the cross section. On the other hand, a rapid increase of the measured cross section could be naturally explained by a change of composition to heavier elements. There are indeed indications of such a change, if Pierre Auger Observatory data are interpreted with existing models. However, it is known that these models are not providing a consistent description of the data.

The Glauber formalism, which is one of the foundations of all cosmic ray interaction models, is used to convert the measured proton-air into proton-proton cross sections. There are considerable systematic uncertainties entering this conversion. Again, within the proton hypothesis, the impact on the characteristics of ultra-high energy proton-air interactions is tremendous. The observed increase of the cross section is difficult to explain within the framework of standard Glauber theory. In this case, it must be interpreted as evidence for new interaction physics above a center of mass energy of $6 \cdot 10^4$ GeV.

It is the challenge of the near future to bring these results in accordance to each other. Probably it is the higher statistics data of the northern Auger Observatory that will provide a much deeper insight in the interaction physics at ultra-high energy.

References

[1] C.D. ANDERSON. Phys. Rev., **43**, 491-494, 1933.

[2] J.C. STREET AND E.C. STEVENSON. Phys. Rev., **52**, 1003-1004, 1937.

[3] C.M.G. LATTES, G.P.S. OCCHIALINI, AND C.F. POWELL. Nature, **160**, 486-492, 1947.

[4] C.M.G. LATTES, G.P.S. OCCHIALINI, AND C.F. POWELL. Nature, **160**, 453-456, 1947.

[5] C.M.G. LATTES, H. MUIRHEAD, G.P.S. OCCHIALINI, AND C.F. POWELL. Nature, **159**, 694-697, 1947.

[6] G. ALEXANDER *et al.* [LEP Collaboration]. Phys. Lett., **B276**, 247-253, 1992.

[7] G. QUAST. Eur. Phys. J., **C33**, S641-S643, 2004.

[8] Y. FUKUDA *et al.* [Super-Kamiokande Collaboration]. hep-ex/9807003 and Phys. Rev. Lett., **81**, 1562-1567, 1998.

[9] Q.R. AHMAD *et al.* [SNO Collaboration]. nucl-ex/0204008 and Phys. Rev. Lett., **89**, 011301, 2002.

[10] S.M. FABER AND J.S. GALLAGHER. Ann. Rev. Astron. Astrophys., **17**, 135-183, 1979.

[11] S.R. COLEMAN AND S.L. GLASHOW. hep-ph/9812418 and Phys. Rev., **D59**, 116008, 1999.

[12] F.R. KLINKHAMER AND M. RISSE. arXiv:0709.2502 [hep-ph] and to appear in PRD, 2007.

[13] J. ABRAHAM *et al.* [Pierre Auger Collaboration]. Nucl. Instrum. Meth., **A523**, 50, 2004.

[14] J. BLÜMER [Pierre Auger Collaboration]. J. Phys., **G29**, 867-879, 2003.

[15] J.P. WEFEL. private communication, 2007.

[16] N.L. GRIGOROV, V.E. NESTEROV, I.D. RAPOPORT, I.A. SAVENKO, AND G.A. SKURIDIN. Yad. Fiz., **11**, 1058-1069, 1970.

[17] T. ABU-ZAYYAD *et al.* [HiRes Collaboration]. astro-ph/0208301 and Astropart. Phys., **23**, 157-174, 2005.

[18] V.A. DERBINA *et al.* [RUNJOB Collaboration]. Astrophys. J., **628**, L41-L44, 2005.

[19] T. ANTONI *et al.* [KASCADE Collaboration]. astro-ph/0505413 and Astropart. Phys., **24**, 1-25, 2005.

[20] D.J. BIRD *et al.* [HiRes Collaboration]. Astrophys. J., **424**, 491-502, 1994.

[21] T.T. YAMAMOTO [Pierre Auger Collaboration]. arXiv:0707.2638 [astro-ph] and Proc. of 30^{th} Int. Cosmic Ray Conf., Merida, Mexico, 2007.

[22] J BLÜMER, R. ENGEL, AND J.R. HÖRANDEL. *Cosmic rays from the knee to the highest energies*, to be published.

[23] G.V. KULIKOV AND G.B. KHRISTIANSEN. Sov. Phys. JETP, **35**, 441, 1959.

[24] H. ULRICH *et al.* Nucl. Phys. (Proc. Suppl.), **122**, 218-221, 2003.

[25] A.D. ERLYKIN AND A.W. WOLFENDALE. astro-ph/0112553 and Astropart. Phys., **18**, 151-164, 2002.

[26] D.J. BIRD *et al.* [HiRes Collaboration]. Phys. Rev. Lett., **71**, 3401-3404, 1993.

[27] T. ABU-ZAYYAD *et al.* [HiRes-MIA Collaboration]. astro-ph/0010652 and Astrophys. J., **557**, 686-699, 2001.

[28] R.U. ABBASI *et al.* [HiRes Collaboration]. astro-ph/0407622 and Astrophys. J., **622**, 910-926, 2005.

[29] A.M. HILLAS. J. Phys., **G31**, R95-R131, 2005.

[30] D. ALLARD, A.V. OLINTO, AND E. PARIZOT. astro-ph/0703633 and submitted to Astron. Astrophys., 2007.

[31] P.L. BIERMANN AND J.P. CASSINELLI. astro-ph/9305003 and Astron. Astrophys., **277**, 691, 1993.

[32] P.L. BIERMANN, T.K. GAISSER, AND T. STANEV. astro-ph/9501001 and Phys. Rev., **D51**, 3450-3454, 1995.

[33] V. BEREZINSKY, A.Z. GAZIZOV, AND S.I. GRIGORIEVA. astro-ph/0502550 and Phys. Lett., **B612**, 147-153, 2005.

[34] A.M. HILLAS. astro-ph/0607109 and Conf. on Cosmology, Galaxy Formation and Astro-Particle Physics on the Pathway to the SKA, Oxford, England, 2006.

[35] V. BEREZINSKY. arXiv:0710.2750 [astro-ph] and Proc. of 30$^{\text{th}}$ Int. Cosmic Ray Conf., Merida, Mexico, 2007.

[36] G.T. ZATSEPIN AND V.A. KUZMIN. JETP Lett., **4**, 78-80, 1966.

[37] K. GREISEN. Phys. Rev. Lett., **16**, 748-750, 1966.

[38] M. TAKEDA *et al.* astro-ph/9807193 and Phys. Rev. Lett., **81**, 1163-1166, 1998.

[39] R. ABBASI *et al.* [HiRes Collaboration]. astro-ph/0703099 and submitted to Phys. Rev. Lett., 2007.

[40] T. STANEV, R. ENGEL, A. MÜCKE, R.J. PROTHEROE, AND J.P. RACHEN. astro-ph/0003484 and Phys. Rev., **D62**, 093005, 2000.

[41] V.S. BEREZINSKY, S.I. GRIGOREVA, AND B.I. HNATYK. Nucl. Phys. (Proc. Suppl.), **151**, 497-500, 2006.

[42] AUGER COLLABORATION. arXiv:0711.2256v1 [astro-ph] and Science, **318**, 939-943, 2007.

[43] R.J. PROTHEROE AND A.P. SZABO. Phys. Rev. Lett., **69**, 2885-2888, 1992.

[44] J.P. RACHEN AND P.L. BIERMANN. astro-ph/9301010 and Astron. Astrophys., **272**, 161-175, 1993.

[45] G. BENFORD AND R.J. PROTHEROE. arXiv:0706.4419 [astro-ph], 2007.

[46] J. BRAU (ED.) *et al.* ILC-REPORT-2007-001, 2007.

[47] W. KOLHÖRSTER *et al.* Naturw., **26**, 567, 1938.

[48] P. AUGER, R. MAZE, AND T. GRIVET-MAYER. Compt. Rend. Acad. Sci. (Ser. II), **206**, 1721-1722, 1938.

[49] D. HECK, J. KNAPP, J.N. CAPDEVIELLE, G. SCHATZ, AND T. THOUW. FZKA-6019, 1998.

[50] T. BERGMANN *et al.* astro-ph/0606564 and Astropart. Phys., **26**, 420-432, 2007.

[51] W. HEITLER. Rev. Mod. Phys., **21**, 113–121, 1949.

[52] M. UNGER [Pierre Auger Collaboration]. Proc. of 30$^{\text{th}}$ Int. Cosmic Ray Conf., Merida, Mexico, 2007.

[53] N.N. KALMYKOV, S.S. OSTAPCHENKO, AND A.I. PAVLOV. Nucl. Phys. (Proc. Suppl.), **B52**, 17-28, 1997.

[54] S. OSTAPCHENKO. hep-ph/0505259 and Phys. Rev., **D74**, 014026, 2006.

[55] R. ENGEL, T.K. GAISSER, T. STANEV, AND P. LIPARI. Proc. of 26th Int. Cosmic Ray Conf., Salt Lake City, Utah, **1**, 415, 1999.

[56] R.S. FLETCHER, T.K. GAISSER, P. LIPARI, AND T. STANEV. Phys. Rev., **D50**, 5710-5731, 1994.

[57] K. WERNER AND T. PIEROG. arXiv:0707.3330 [astro-ph] and AIP Conf. Proc., **928**, 111-117, 2007.

[58] S. OSTAPCHENKO. hep-ph/0412332 and Nucl. Phys. Proc. Suppl., **151**, 143-146, 2006.

[59] L.D. LANDAU AND I. POMERANCHUK. Dokl. Akad. Nauk Ser. Fiz., **92**, 535-536, 1953.

[60] A.B. MIGDAL. Phys. Rev., **103**, 1811-1820, 1956.

[61] B. ROSSI AND K. GREISEN. Rev. Mod. Phys., **13**, 240-309, 1941.

[62] J. MATTHEWS. Astropart. Phys., **22**, 387-397, 2005.

[63] J. ALVAREZ-MUNIZ, R. ENGEL, T.K. GAISSER, J.A. ORTIZ, AND T. STANEV. astro-ph/0205302 and Phys. Rev., **D66**, 033011, 2002.

[64] S.Z. TILAV. `http://proquest.umi.com/pqdlink?did=746945291&Fmt=7&clientId=79356&RQT=309&VName=PQD`. PhD thesis, UMI-92-23696, 1992.

[65] A.M. HILLAS. Nucl. Phys. (Proc. Suppl.), **52B**, 29-42, 1997.

[66] S.J. SCIUTTO. astro-ph/9911331 and `http://www.fisica.unlp.edu.ar/auger/aires`, 1999.

[67] M. KOBAL [Pierre Auger Collaboration]. Astropart. Phys., **15**, 259-273, 2001.

[68] P. BILLOIR. `http://www-ik.fzk.de/corsika/corsika-school/talks/5_eas-phys/billoir_universality.pdf`. VIHKOS Corsika School, 2005.

[69] P. BILLOIR. *Natural and Artificial Fluctuations in the Auger Surface Detector*, GAP2001-005, 2001.

[70] R. ENGEL. private communication, 2007.

[71] T.K. GAISSER. *Cosmic Ray and Particle Physics*. Cambridge University Press, 1990.

[72] H.J. DRESCHER AND G.R. FARRAR. astro-ph/0212018 and Phys. Rev., **D67**, 116001, 2003.

[73] H.H. MIELKE, M. FOELLER, J. ENGLER, AND J. KNAPP. J. Phys., **G20**, 637-649, 1994.

[74] T. ANTONI *et al.* [KASCADE Collaboration]. astro-ph/0102443 and Astropart. Phys., **16**, 245-263, 2002.

[75] T. ANTONI *et al.* [KASCADE Collaboration]. astro-ph/0106494 and J. Phys., **G27**, 1785-1798, 2001.

[76] R. ENGEL [Pierre Auger Collaboration]. arXiv:0706.1921 [astro-ph] and Proc. of 30th Int. Cosmic Ray Conf., Merida, Mexico, 2007.

[77] W.D. APEL *et al.* [KASCADE Collaboration]. J. Phys., **G34**, 2581-2593, 2007.

[78] J. MILKE *et al.* Acta Phys. Polon., **B35**, 341-349, 2004.

[79] J. KNAPP, D. HECK, S.J. SCIUTTO, M.T. DOVA, AND M. RISSE. astro-ph/0206414 and Astropart. Phys., **19**, 77-99, 2003.

[80] F. NERLING, J. BLÜMER, R. ENGEL, AND M. RISSE. Astropart. Phys., **24**, 421-437, 2006.

[81] M. GILLER, A. KACPERCZYK, J. MALINOWSKI, W. TKACZYK, AND G. WIECZOREK. J. Phys., **G31**, 947-958, 2005.

[82] R. ENGEL. astro-ph/0504358 and Nucl. Phys. Proc. Suppl., **151**, 437-461, 2006.

[83] J. KNAPP, D. HECK, AND G. SCHATZ. Comparison of hadronic interaction models used in air shower simulations and of their influence on shower development and obsevables. in Wissenschaftliche Berichte FZKA 5828, Forschungszentrum Karlsruhe, 1996.

[84] T. PIEROG, R. ENGEL, AND D. HECK. astro-ph/0602190 and Czech. J. Phys., **56**, A161-A172, 2006.

[85] V.N. GRIBOV. Sov. Phys. JETP, **26**, 414, 1968.

[86] T.K. GAISSER AND F. HALZEN. Phys. Rev. Lett., **54**, 1754, 1985.

[87] L. DURAND AND P. HONG. Phys. Rev. Lett., **58**, 303, 1987.

[88] P.D.B. COLLINS. *An Introduction to Regge Theorie & High Energy Physics*. Cambridge University Press, 1977.

[89] N.N. KALMYKOV AND S.S. OSTAPCHENKO. Sov. J. Nucl. Phys., **50**, 315-318, 1989.

[90] N.N. KALMYKOV AND S.S. OSTAPCHENKO. Phys. Atom. Nucl., **56**, 346-353, 1993.

[91] A.B. KAIDALOV AND K.A. TER-MARTIROSIAN. Sov. J. Nucl. Phys., **39**, 979, 1984.

[92] A.B. KAIDALOV. Surveys High Energ. Phys., **13**, 265-330, 1999.

[93] S. OSTAPCHENKO. hep-ph/0602139 and Phys. Lett., **B636**, 40-45, 2006.

[94] J. ENGEL, T.K. GAISSER, T. STANEV, AND P. LIPARI. Phys. Rev., **D46**, 5013-5025, 1992.

[95] K. WERNER, F.M. LIU, AND T. PIEROG. hep-ph/0506232 and Phys. Rev., **C74**, 044902, 2006.

[96] T. PIEROG AND K. WERNER. astro-ph/0611311 and Proc. of 30$^{\text{th}}$ Int. Cosmic Ray Conf., Merida, Mexico, 2007.

[97] H.J. DRESCHER, M. HLADIK, S. OSTAPCHENKO, T. PIEROG, AND K. WERNER. hep-ph/0007198 and Phys. Rept., **350**, 93-289, 2001.

[98] M. HLADIK, H.J. DRESCHER, S. OSTAPCHENKO, T. PIEROG, AND K. WERNER. hep-ph/0102194 and Phys. Rev. Lett., **86**, 3506-3509, 2001.

[99] N.L. GRIGOROV *et al.* Proc. of 9$^{\text{th}}$ Int. Cosmic Ray Conf, London, **1**, 860, 1965.

[100] G.B. YODH, Y. PAL, AND J.S. TREFIL. Phys. Rev. Lett., **28**, 1005-1008, 1972.

[101] R.A. NAM, S.I. NIKOLSKY, V.P. PAVLUCHENKO, A.P. CHUBENKO, AND V.I. YAKOVLEV. Proc. of 14$^{\text{th}}$ Int. Cosmic Ray Conf, Munich, page 2258, 1975.

[102] F. SIOHAN *et al.* J. Phys., **G4**, 1169-1186, 1978.

[103] M. HONDA *et al.* Phys. Rev. Lett., **70**, 525-528, 1993.

[104] M. AGLIETTA *et al.* Proc. of 26$^{\text{th}}$ Int. Cosmic Ray Conf., Salt Lake City, Utah, **1**, 143, 1999.

[105] T. HARA *et al.* Phys. Rev. Lett., **50**, 2058-2061, 1983.

[106] S.P. KNURENKO, V.R. SLEPTSOVA, I.E. SLEPTSOV, N.N. KALMYKOV, AND S.S. OSTAPCHENKO. Proc. of 26$^{\text{th}}$ Int. Cosmic Ray Conf., Salt Lake City, Utah, **1**, 372, 1999.

[107] ARGO COLLABORATION. Proc. of 30$^{\text{th}}$ Int. Cosmic Ray Conf., Merida, Mexico, 2007.

[108] R.M. BALTRUSAITIS *et al.* Phys. Rev. Lett., **52**, 1380-1383, 1984.

[109] K. BELOV. Nucl. Phys. (Proc. Suppl.), **151**, 197-204, 2006.

[110] N.N. NIKOLAEV. hep-ph/9304283 and Phys. Rev., **D48**, 1904-1906, 1993.

[111] J. ALVAREZ-MUNIZ, R. ENGEL, T.K. GAISSER, J.A. ORTIZ, AND T. STANEV. astro-ph/0209117 and Phys. Rev., **D66**, 123004, 2002.

[112] R.W. ELLSWORTH, T.K. GAISSER, T. STANEV, AND G.B. YODH. Phys. Rev., **D26**, 336, 1982.

[113] R.M. BALTRUSAITIS *et al.* [Fly's Eye Collaboration]. Nucl. Instrum. Meth., **A240**, 410, 1985.

[114] R. ULRICH, J. BLÜMER, R. ENGEL, F. SCHÜSSLER, AND M. UNGER. arXiv:0709.1392 [astro-ph] and Proc. of 30$^{\text{th}}$ Int. Cosmic Ray Conf., Merida, Mexico, 2007.

[115] J. LINSLEY. Nuovo Cim. Lett., **42**, 403, 1985.

[116] J. LINSLEY. Proc. of 19$^{\text{th}}$ Int. Cosmic Ray Conf., **6**, 1, 1985.

[117] R. ULRICH, J. BLÜMER, R. ENGEL, F. SCHÜSSLER, AND M. UNGER. astro-ph/0612205 and Proc. of XIV ISVHECRI, Weihai, China, 2006.

[118] J. ALVAREZ-MUNIZ, R. ENGEL, T.K. GAISSER, J.A. ORTIZ, AND T. STANEV. astro-ph/0402092 and Phys. Rev., **D69**, 103003, 2004.

[119] R.J. GLAUBER. Phys. Rev., **100**, 242-248, 1955.

[120] M.M. BLOCK AND R.N. CAHN. Rev. Mod. Phys., **57**, 563, 1985.

[121] R. ENGEL. Nucl. Phys. (Proc. Suppl.), **B82**, 221-231, 2000.

[122] A. DONNACHIE AND P.V. LANDSHOFF. hep-ph/9209205 and Phys. Lett., **B296**, 227-232, 1992.

[123] R.J. GLAUBER AND G. MATTHIAE. Nucl. Phys., **B21**, 135-157, 1970.

[124] J. OEHLSCHLÄGER. private communication, 2007.

[125] J. RIDKY [Pierre Auger Collaboration]. Nucl. Phys. Proc. Suppl., **165**, 45-49, 2007.

[126] H. BLÜMER *et al.* [Pierre Auger Collaboration]. Proc. of 26$^{\text{th}}$ Int. Cosmic Ray Conf., Salt Lake City, Utah, 1999.

[127] X. BERTOU *et al.* [Pierre Auger Collaboration]. Nucl. Instrum. Meth., **A568**, 839-846, 2006.

[128] D. ALLARD *et al.* [Pierre Auger Collaboration]. astro-ph/0510320 and Proc. of 29$^{\text{th}}$ Int. Cosmic Ray Conf., Pune, India, 2005.

[129] R. KNAPIK *et al.* [Pierre Auger Collaboration]. arXiv:0708.1924 [astro-ph] and Proc. of 30$^{\text{th}}$ Int. Cosmic Ray Conf., Merida, Mexico, 2007.

[130] H. GEMMEKE *et al.* IEEE Trans. Nucl. Sci., **47**, 371-375, 2000.

[131] J.T. BRACK, R. MEYHANDAN, G.J. HOFMAN, AND J. MATTHEWS. Astropart. Phys., **20**, 653-659, 2004.

[132] B. KEILHAUER *et al.* [Pierre Auger Collaboration]. astro-ph/0507275 and Proc. of 29$^{\text{th}}$ Int. Cosmic Ray Conf., Pune, India, 2005.

[133] B. KEILHAUER, J. BLÜMER, R. ENGEL, AND H.O. KLAGES. astro-ph/0511153 and Astropart. Phys., **25**, 259-268, 2006.

[134] B. WILCZYNSKA *et al.* Proc. of 30$^{\text{th}}$ Int. Cosmic Ray Conf., Merida, Mexico, 2007.

[135] B. KEILHAUER, J. BLÜMER, R. ENGEL, H.O. KLAGES, AND M. RISSE [Pierre Auger Collaboration]. astro-ph/0405048 and Astropart. Phys., **22**, 249-261, 2004.

[136] S.Y. BENZVI *et al.* [Pierre Auger Collaboration]. astro-ph/0609063 and Nucl. Instrum. Meth., **A574**, 171-184, 2007.

[137] B. FICK *et al.* [Pierre Auger Collaboration]. JINST, **1**, P11003, 2006.

[138] S.Y. BENZVI *et al.* [Pierre Auger Collaboration]. arXiv:0704.0303 [astro-ph] and Astropart. Phys., **28**, 312-320, 2007.

[139] S.Y. BENZVI *et al.* [Pierre Auger Collaboration]. arXiv:0706.1710 [astro-ph] and Proc. of 30th Int. Cosmic Ray Conf., Merida, Mexico, 2007.

[140] S. ARGIRO *et al.* arXiv:0707.1652 [astro-ph] and Nucl. Instrum. Meth., **A580**, 1485-1496, 2007.

[141] R. ULRICH, M. UNGER, F. SCHÜSSLER, R. ENGEL AND J. BLÜMER. *Mapping of FD camera pixels to electronic readout channels and its implementation in Offline - impact on data reconstruction and aperture simulations*, GAP2007-003, 2007.

[142] R. ULRICH, J. BLÜMER, R. ENGEL, F. SCHÜSSLER AND M. UNGER. *Implementation of the differential atmosphere in a curved geometry in Offline - Impact on data reconstruction and simulation*, GAP2007-007, 2007.

[143] C. BLEVE [Pierre Auger Collaboration]. arXiv:0706.1491 [astro-ph] and Proc. of 30th Int. Cosmic Ray Conf., Merida, Mexico, 2007.

[144] B. FICK [Pierre Auger Collaboration]. astro-ph/0308512 and Proc. of 28th Int. Cosmic Ray Conf., Tsukuba, Japan, 2003.

[145] B. EFRON *et al. An Introduction to the Bootstrap*. Chapman and Hall, 1986.

[146] M. KLEIFGES, A. MENSHIKOV, D. CHERNYAKHOVSKY, AND H. GEMMEKE. IEEE Trans. Nucl. Sci., **50**, 1204-1207, 2003.

[147] M. UNGER, B.R. DAWSON, R. ENGEL, F. SCHÜSSLER AND R. ULRICH. *Reconstruction of Longitudinal Profiles of Ultra-High Energy Cosmic Ray Showers from Fluorescence and Cherenkov Light Measurements*, to be published, 2007.

[148] M. UNGER, R. ENGEL, F. SCHÜSSLER, AND R. ULRICH. arXiv:0706.1501 [astro-ph] and Proc. of 30th Int. Cosmic Ray Conf., Merida, Mexico, 2007.

[149] M. UNGER. *ShowerProfile Reconstruction from Fluorescence and Cherenkov light*, GAP2006-010, 2006.

[150] J. ALVAREZ-MUÑIZ, E. MARQUES, R.A. VAZQUEZ, AND E. ZAS. Phys. Rev., **D67**, 101303, 2003.

[151] M. RISSE AND D. HECK. Astropart. Phys., **20**, 661, 2004.

[152] F. KAKIMOTO *et al.* Nucl. Instrum. Meth., **A372**, 527, 1996.

[153] M. NAGANO *et al.* Astropart. Phys., **22**, 235, 2004.

[154] M. GILLER, G. WIECZOREK, A. KACPERCZYK, H. STOJEK, AND W. TKACZYK. J. Phys., **G30**, 97-105, 2004.

[155] A.M. HILLAS. J. Phys., **G8**, 1461, 1982.

[156] A.M. HILLAS. J. Phys., **G8**, 1475, 1982.

[157] T.K. GAISSER AND A.M. HILLAS. Proc. of 15th Int. Cosmic Ray Conf., Plovdiv, Bulgaria, **8**, 353, 1977.

[158] T. PIEROG *et al.* Proc. of 29th Int. Cosmic Ray Conf., Pune, India, 2005.

[159] H.M.J. BARBOSA *et al.* Astropart. Phys., **22**, 159, 2004.

[160] S.Y. BENZVI *et al.* [Pierre Auger Collaboration]. astro-ph/0706.3236 and Proc. of 30th Int. Cosmic Ray Conf., Merida, Mexico, 2007.

[161] ROOT, the object-oriented data analysis framework. `http://root.cern.ch`.

[162] I.C. MARIŞ, F. SCHÜSSLER, R. ULRICH AND M. UNGER. *Data Summary Trees and Shower Visualization for Reconstructed Auger Events*, GAP2006-081, 2006.

[163] The Auger Observer. `http://observer.fzk.de`.

[164] L. PRADO *et al.* Nucl. Instrum. Meth., **A545**, 632-642, 2005.

[165] D. GORA *et al.* astro-ph/0505371 and Astropart. Phys., **24**, 484-494, 2006.

[166] R. M. BALTRUSAITIS *et al.* J. Phys., **G13**, 115-119, 1987.

[167] B.R. DAWSON, R.W. CLAY, J.R. PATTERSON, AND J.R. PRESCOTT. J. Phys., **G15**, 893-908, 1989.

[168] M. GILLER *et al.* Proc. of 28th Int. Cosmic Ray Conf., Tsukuba, Japan, 2003.

[169] B.R. DAWSON, M. GILLER AND G. WIECZOREK. Proc. of 30th Int. Cosmic Ray Conf., Merida, Mexico, 2007.

[170] J. RAUTENBERG. *Measurement of the inclined photon flux*, Auger Collaboration Meeting, Malargüe, 2006.

[171] R. CARUSO *et al.* [Pierre Auger Collaboration]. astro-ph/0507105 and Proc. of 29th Int. Cosmic Ray Conf., Pune, India, 2005.

[172] A. MENSHIKOV, M. KLEIFGES, AND H. GEMMEKE. IEEE Trans. Nucl. Sci., **50**, 1208-1213, 2003.

[173] Event reconstruction `Er`, CDAS software release v4r0. `http://www.auger.org.ar/CDAS`.

[174] I. LHENRY-YVON, E. PARIZOT, D. ALLARD, P. GHIA, J. AUBLIN. *The Lateral Trigger Probability function (LTP) for UHE showers detected by the PAO surface detector*, GAP2003-088, 2003.

[175] J. AUBLIN, D. ALLARD, I. LHENRY-YVON, E. PARIZOT. *The Lateral Trigger Probability (LTP) and the Lateral Distribution of Bins Over Threshold (LaDiBOT) as powerful tools to test MC simulations*, GAP2007-015, 2003.

[176] C. BONIFAZI AND P.L. GHIA. *Selection of data periods and calculation of the SD geometrical acceptance*, GAP2006-101, 2006.

[177] A. BIALAS, M. BLESZYNSKI, AND W. CZYZ. Nucl. Phys., **B111**, 461, 1976.

[178] W.R. NELSON, H. HIRAYAMA, AND D.W.O. ROGERS. SLAC-0265, 1985.

[179] F. SCHÜSSLER, J. BLÜMER, R. ENGEL, R. ULRICH, M. UNGER, D. VEBERIČ . *Investigation of AGN correlation and composition hypotheses with hybrid and stereo FD data*, GAP2007-094, 2007.

[180] F. JAMES AND M. ROOS. Comput. Phys. Commun., **10**, 343-367, 1975.

[181] J.A.J. MATTHEWS. private communication, 2007.

[182] M.D. HEALY [Pierre Auger Collaboration]. arXiv:0710.0025 [astro-ph] and Proc. of 30th Int. Cosmic Ray Conf., Merida, Mexico, 2007.

[183] O. KALASHEV, G. GELMINI, AND D. SEMIKOZ. arXiv:0706.3847 [astro-ph] and Proc. of 30th Int. Cosmic Ray Conf., Merida, Mexico, 2007.

[184] P. YOUNK. *Cosmic Rays at the Ankle.*, PhD-thesis, GAP2007-102, 2007.

[185] M. SETTIMO, L. PERRONE, G. CATALDI, I. DE MITRI, AND D. MARTELLO. *Parameterisation of the "Lateral Trigger Probability" functions at low energies.*, GAP2007-069, 2007.

[186] M. UNGER, R. ENGEL, F. SCHÜSSLER AND R. ULRICH. *Measurement of the average Xmax as a function of energy*, GAP2007-005, 2007.

[187] M. PROUZA [Pierre Auger Collaboration]. arXiv:0706.1719 [astro-ph] and Proc. of 30th Int. Cosmic Ray Conf., Merida, Mexico, 2007.

[188] F. SCHÜSSLER, R. ENGEL, M. UNGER, AND R. ULRICH. *Measurement of the UHECR spectrum based on Auger Hybrid data*, Auger Collaboration Meeting, Malargüe, 2006.

[189] F. SCHÜSSLER *et al. Auger Uptime*, `http://augerobserver.fzk.de/index.php?id=29`.

[190] M. UNGER. private communication, 2007.

[191] J. A. BELLIDO *et al.* [Pierre Auger Collaboration]. astro-ph/0507103 and Proc. of 29th Int. Cosmic Ray Conf., Pune, India, 2005.

[192] M. RISSE AND P. HOMOLA. astro-ph/0702632 and Mod. Phys. Lett., **A22**, 749-766, 2007.

[193] J. ABRAHAM *et al.* [Pierre Auger Collaboration]. astro-ph/0606619 and Astropart. Phys., **27**, 155-168, 2007.

[194] PHILIPP P. KRONBERG. Rept. Prog. Phys., **57**, 325-382, 1994.

[195] S. LEE, A. OLINTO, AND G. SIGL. astro-ph/9508088 and Astrophys. J., **455**, L21, 1995.

[196] K. DOLAG, D. GRASSO, V. SPRINGEL, AND I. TKACHEV. astro-ph/0410419 and JCAP, **0501**, 009, 2005.

[197] Y. XU, P.P. KRONBERG, S. HABIB, AND Q.W. DUFTON. astro-ph/0509826 and Astrophys. J., **637**, 19-26, 2006.

[198] T. STANEV. astro-ph/9607086 and Astrophys. J., **479**, 290, 1997.

[199] J. ALVAREZ-MUNIZ, R. ENGEL, AND T. STANEV. astro-ph/0112227 and Astrophys. J., **572**, 185-201, 2001.

[200] M. ROTH [Pierre Auger Collaboration]. arXiv:0706.2096 [astro-ph] and Proc. of 30th Int. Cosmic Ray Conf., Merida, Mexico, 2007.

[201] P. SOMMERS. Astropart. Phys., **14**, 271-286, 2001.

[202] K. KAMATA AND J. NISHIMURA. Prog. Theoret. Phys. Suppl., **6**, 93, 1958.

[203] K. GREISEN. in: Progress in Cosmic Ray Physics Vol III, ed. G. Wilson, 1956. Amsterdam: North-Holland.

[204] T. ANTONI *et al.* [KASCADE Collaboration]. astro-ph/0004233 and Astropart. Phys., **14**, 245-260, 2001.

[205] D. BARNHILL *et al.* [Pierre Auger Collaboration]. astro-ph/0507590 and Proc. of 29th Int. Cosmic Ray Conf., Pune, India, 2005.

[206] E. ZAS, F. HALZEN, AND T. STANEV. Phys. Rev., **D45**, 362-376, 1992.

[207] G. RODRIGUEZ-FERNANDEZ, E. ZAS, T. OHNUKI, A. TRIPATHI, D. BARNHILL, J. LEE, K. ARISAKA, W.E. SLATER. *Surface detector response using loopup tables based on GEATN4 simulations.*, GAP2004-045, 2004.

[208] L. CAZON, R.A. VAZQUEZ, A.A. WATSON, AND E. ZAS. astro-ph/0311223 and Astropart. Phys., **21**, 71-86, 2004.

[209] L. CAZON, R.A. VAZQUEZ, AND E. ZAS. astro-ph/0412338 and Astropart. Phys., **23**, 393-409, 2005.

Appendix

A Fast hybrid simulations

To generate realistic air shower events in the surface detector of the Pierre Auger Observatory without employing full air shower as well as detector Monte Carlo simulations, the `SdSimpleSim` tool is developed. `SdSimpleSim` is distributed as part of the Offline-framework. It is based on 1D shower particle number profiles at the shower axis, as they are provided for example by CONEX. The use of the CONEX simulation tool assures minimal computation time while providing realistic air showers profiles including shower-to-shower fluctuations. It is the task of `SdSimpleSim` to generate the lateral and timing information needed for the surface detector simulation.

Figure A.1 (left) shows the longitudinal profile of electrons $N_e(X)$, muons $N_\mu(X)$, gammas $N_\gamma(X)$ as well as muon production $\mathrm{d}N_\mu^{\mathrm{prod}}(X)/\mathrm{d}X$, as simulated by CONEX for a 10 EeV proton primary. The alignment of surface detector thanks with respect to an incoming air shower is illustrated in Figure A.1 (right). Each tank i sees the air shower at development stage X_i and is located in a radial distance of r_i from the shower axis. Parameterizations based on the NKG-distribution [202, 203]

$$f_{\mathrm{NKG}}(r,\ r_0,\ s) = \frac{1}{2\pi r_0^2}\,\frac{\Gamma(4.5-s)}{\Gamma(s)\Gamma(4.5-2s)}\left(\frac{r}{r_0}\right)^{s-2}\left(1+\frac{r}{r_0}\right)^{s-4.5} \tag{A.2}$$

have proven to be very useful to describe lateral particle densities [204, 205] for air showers at age $s_i = 3/(1+2X_{\mathrm{max}}/X_i)$ and radial scale r_0. For the purpose of `SdSimpleSim` it turns out to be a good approximation to use (A.2) to calculate particle densities as

$$\mathrm{e}^\pm,\gamma: \qquad \rho_{\mathrm{e}^-,\mathrm{e}^+,\gamma}(r_i,\ s_i) = N_{\mathrm{e}^-,\mathrm{e}^+,\gamma}(X_i)\cdot f_{\mathrm{NKG}}(r_i,\ r_{\mathrm{m}},\ s_i/2) \tag{A.3}$$

$$\mu: \qquad \rho_\mu(r_i,\ s_i) = N_\mu(X_i)\cdot f_{\mathrm{NKG}}(r_i,\ 350\,\mathrm{m},\ s_i)\,, \tag{A.4}$$

with the Moliere radius r_m. Combined with the total projected tank surface (as illustrated in Figure A.2, left)

$$A_{\mathrm{proj}} = \pi\, r_{\mathrm{tank}}^2\cos\theta + 2\,h_{\mathrm{tank}}\,r_{\mathrm{tank}}\sin\theta \tag{A.5}$$

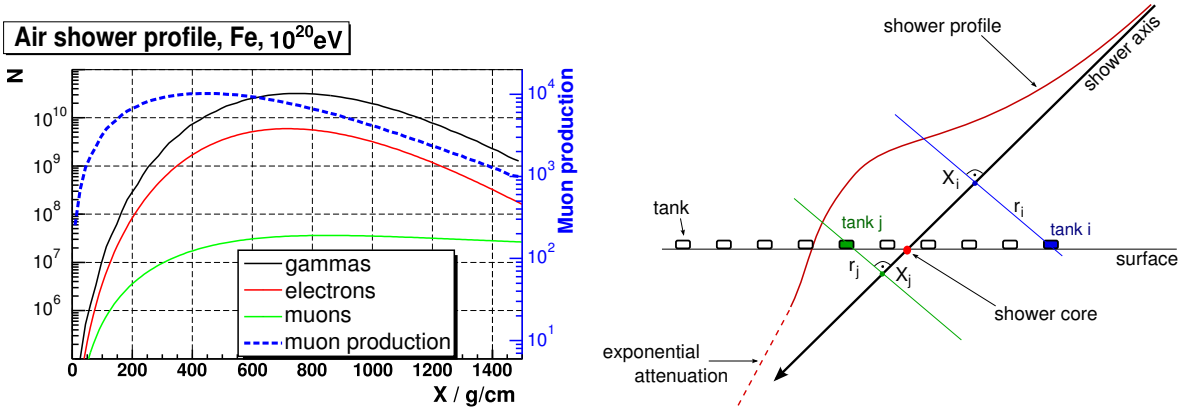


Figure A.1: Left panel: longitudinal particle distribution of an air shower. Right panel: schematic view of the geometry of the surface detector tanks with respect to an air shower.

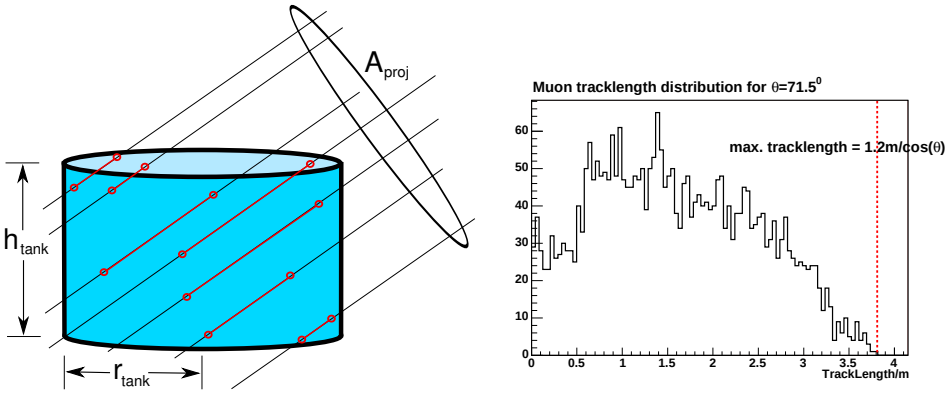


Figure A.2: Left panel: oblique muon particle tracks intersecting a tank. Right panel: tracklength distribution of penetrating muons hitting an Auger water tank at an angle of 71.5°.

we can sample the number of particles hitting tank i

$$e^{\pm}, \gamma: \qquad n^i_{e^-,e^+,\gamma} = \text{Poissonian}\left[A_{\text{proj}} \cdot \rho_{e^-,e^+,\gamma}(r_i, s_i)\right] \tag{A.6}$$

$$\mu: \qquad n^i_{\mu} = \text{Poissonian}\left[A_{\text{proj}} \cdot \rho_{\mu}(r_i, s_i)\right] . \tag{A.7}$$

A vertical muon intersection the tank is depositing exactly the energy corresponding to one *vertical equivalent muon* (VEM). Therefore an arbitrary muon trajectory j with a tracklength t_j deposits an energy equivalent to t_j/h_{tank} VEM. Figure A.2 (right) demonstrates the realistic sampling of the muon tracklength for muons hitting a tank with a zenith angle of 71.5°. In contrary electromagnetic particles are in good approximation releasing all of their energy within the water tank. The deposited electromagnetic energy is converted to a muon tracklength equivalent by the factor 5.25 m/GeV [206]. Due to absorption in the plastic walls of the tank, the energy of electrons is reduced by 0.29 MeV and the energy of gammas by 0.3 MeV [207]. In contrast to the muon, where only the geometric tracklength is important for the induced tank signal, for electromagnetic particles the tanks behaves as an efficient absorption calorimeter and thus the energy distribution of the particles is crucial. Therefore we

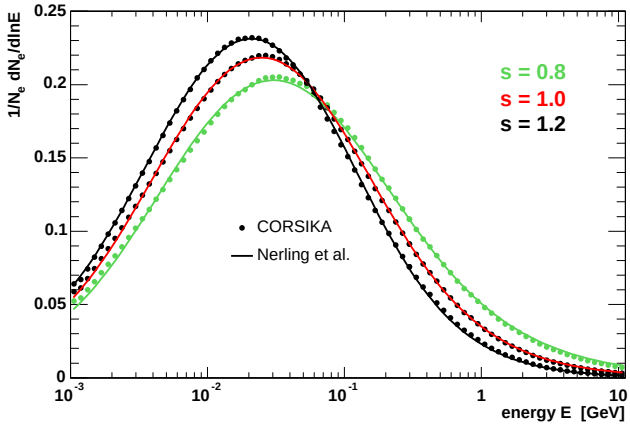


Figure A.3: Universal electron energy distribution as a function of the shower age s [80].

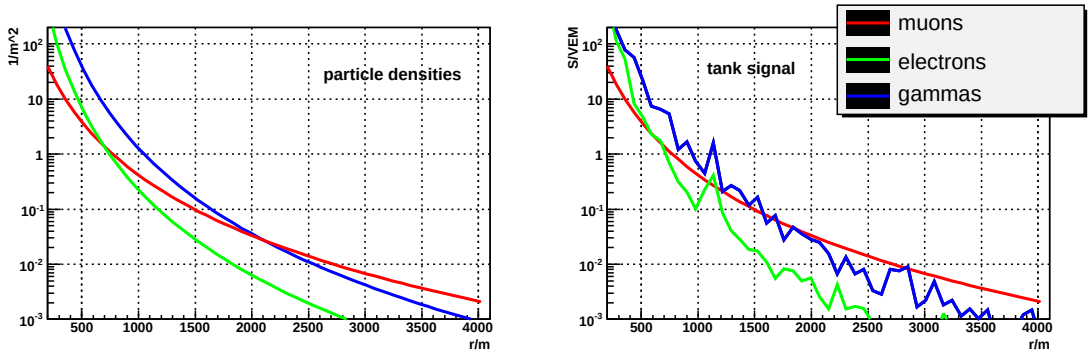


Figure A.4: Left panel: parameterization of the lateral particle densities. Right panel: corresponding tank signal after conversion to VEM. The fluctuations are due to Poissonian particle number fluctuations as well as sampling of the muon tracklength and the electromagnetic energy.

use the universality of electromagnetic energy spectra [80] to sample from a realistic energy distribution for electrons, positrons and gammas, specificly at the shower age s_i for tank i (see Figure A.3). Figure A.4 demonstrates the resulting lateral parameterization for a 10 EeV proton primary.

The arrival time of particles in tank i depends on the structure of the air shower as well as the timing resolution of the detector. Figure A.5 (left) illustrates the intersection of a typical shower front. Muons are traveling most directly from their production point to the observation level. They are almost not scattering or interacting, therefore they reach the detector slightly prior to other shower components. Muon timing was modeled based on the muon production profile along the shower axis [208, 209]. Figure A.5 (right) illustrates the underlying assumption. Muons are produced close to the shower axis according to the

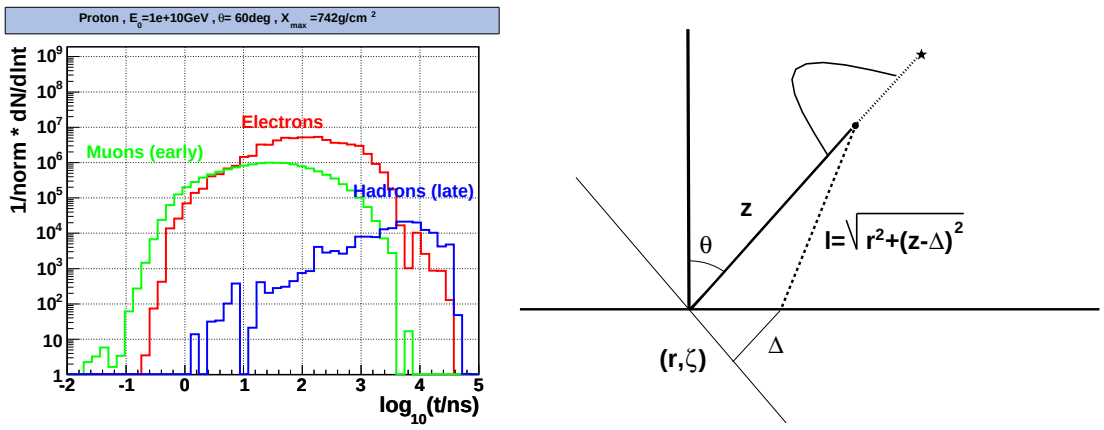


Figure A.5: Left panel: Structure of the shower front in terms of arrival times for the different shower components. Right panel: Illustration of the muon timing model from [208]. Muons are emitted close to the shower axis during the evolution of a shower.

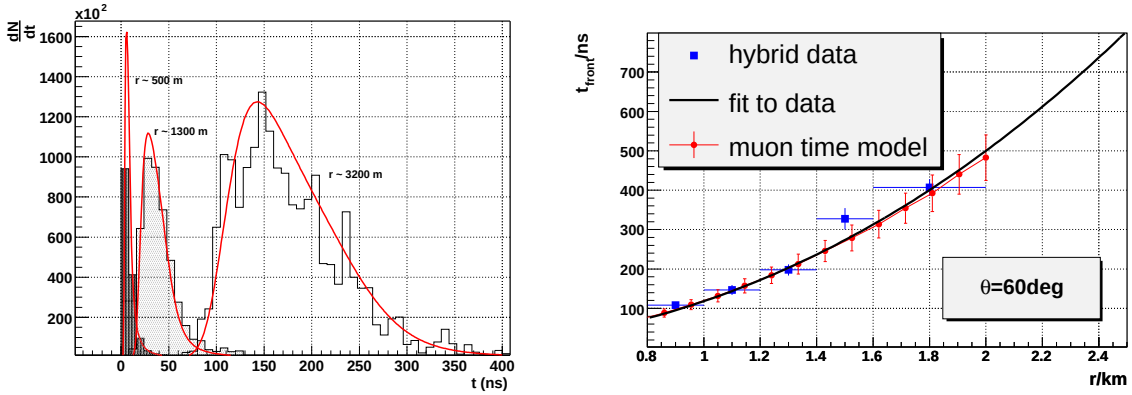


Figure A.6: Left panel: Predictions of the muon arrival time model compared to CORSIKA simulations [208]. Right panel: `SdSimpleSim` timing compared to Auger data.

muon production profile $\mathrm{d}N_\mu^{\mathrm{prod}}/\mathrm{d}X$. Most muons are produced by decaying pions. Muons are emitted at the shower axis with an angular distribution corresponding to their parent pions, which is determined by the pion $p_\perp$-distribution. Besides this purely geometric muon timing model, there are second order effects like the energy dependent speed of muons. Since most muons have energies above 1 GeV, these kinetic differences in muon timing are not large. Figure A.6 (left) compares the predictions of the resulting muon timing model to CORSIKA simulations. In addition to the air shower timing also the finite detector resolution has to be taken into account. The timing resolution of an Auger surface water Cherenkov tank can be measured with twin tanks. These are pairs of tanks positions very close to each other. Averaging over many air shower events the time resolution was deduced to be $\sigma_t = 26.84\,\mathrm{ns}$ from their response. This timing resolution is added as a Gaussian fluctuation to the muon time model prediction. Figure A.6 (right) shows a comparison of the resulting `SdSimpleSim` timing model to Auger data.

An important issue for hybrid simulations are noise tanks in the surface array. Since the hybrid detection efficiency is already very high at energies much below full efficiency of the surface detector many hybrid events only contain a single tank hit. Therefore a random noise tank can well be misidentified as the hybrid tank, which is in fact an important source of

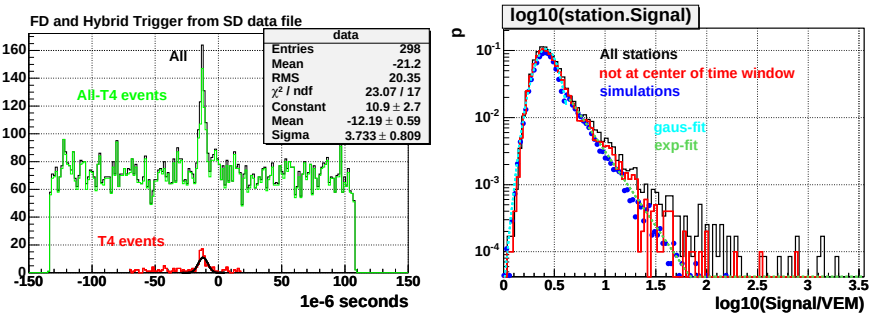


Figure A.7: Left panel: trigger time distribution of surface detector tanks from one night of hybrid data taking. Only "T4" events fulfill basic surface detector quality criteria. Right panel: Distribution of integrated signals of noise tanks.

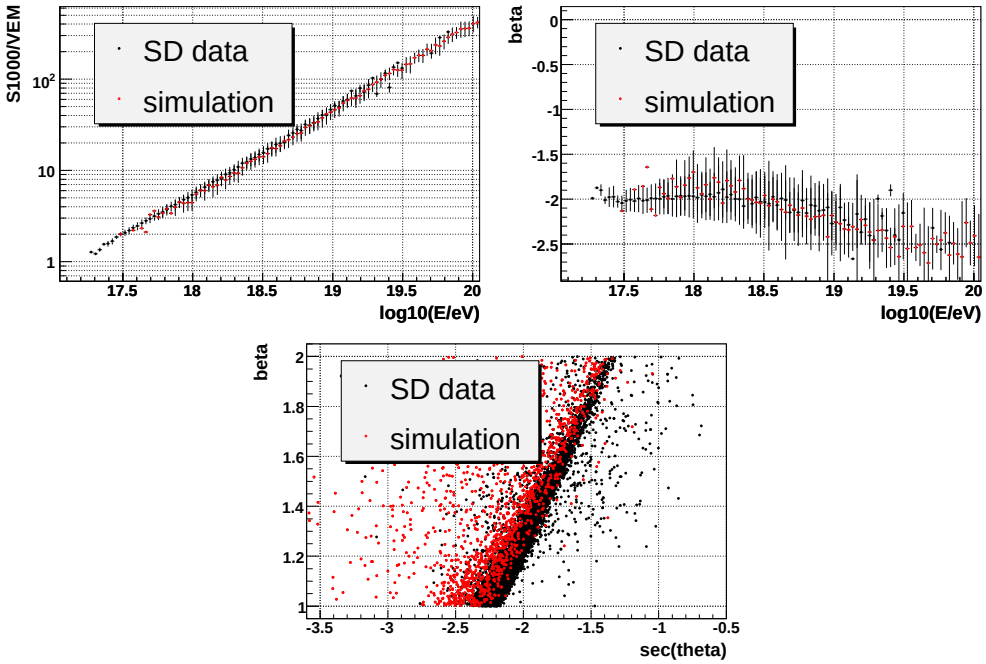


Figure A.8: Comparison of `SdSimpleSim` to data from the surface detector.

uncertainty for hybrid reconstruction. For every telescope trigger the surface detector array is readout, assuming a concentric timing evolution with the distance from the triggered eye. The time window for each tank is 100 μs. Since the T1 trigger rate per tank is self dynamically adjusted to yield 100 Hz, the average number of noise tanks can be easily calculated as

$$n_{\text{noise}} = n_{\text{tot}} \cdot 200\,\mu\text{s} \cdot 100\,\text{Hz} \approx 10 \;\text{ up to }\; 30\,, \tag{A.8}$$

where n_{tot} is the active number of tanks in the array. Figure A.7 (left) shows the time distribution of one night of hybrid data, where only the "T4" events also fulfill basic surface detector quality criteria. The distribution for "all-T4" are events that would not have triggered surface detector readout by themselves. These class of events is called *brass hybrid* and represents the largest fraction of all hybrid data. It is nicely visible how the real shower events are located in a small peak centered at about -20 ns, corresponding to the mean distance of shower cores from the telescope building, which was sending the trigger. All events which are building the flat background can be safely assumed to represent a typical minbias surface detector status. Figure A.7 (right) shows the distribution of integrated tank signals of these minbias tanks. The distribution can be parameterized by a Gaussian with an exponential tail to large signals. `SdSimpleSim` samples from this parameterization to generate realistic noise signals.

The overall performance of `SdSimpleSim` can be demonstrated by generating 10000 surface detector events for a realistic cosmic ray input spectrum with a flux $\propto E^{-2.7}$ and a zenith angle distribution from 0° to 60°. The simulation does only last for $\sim$ 1 h. A comparison to real SD data is given in Figure A.8. The main characteristics of the lateral signal distribution in the data like the energy estimator S_{1000} or the slope parameter β are resembled very well.

B ΔX-parameterizations

B.1 Energy dependence

B.1.1 SIBYLL

Table B.1: Complete list of parameterizations of the ΔX_1-model for SIBYLL.

$f_{10\,\mathrm{EeV}}$	N_{sim}	α [gcm^{-2}]	β [gcm^{-2}]	γ [g^{2}cm^4]	χ^2/ndf
17.0	9764	608.5±0.5	18.4±0.2	-3.8e-02±1.0e-02	153.6/72
17.5	7524	640.0±0.5	17.2±0.2	-4.0e-02±1.2e-02	116.1/63
18.0	8367	664.3±0.5	17.3±0.2	-3.2e-02±1.1e-02	127.7/65
18.5	5507	702.9±0.6	16.9±0.3	-6.8e-03±1.7e-02	100.5/58
19.0	9826	723.2±0.4	16.5±0.2	-4.4e-03±1.1e-02	97.6/59
19.5	2754	765.3±0.7	16.2±0.3	-1.8e-02±1.9e-02	51.4/50
20.0	8117	781.6±0.4	15.8±0.2	-1.9e-02±1.1e-02	68.1/59

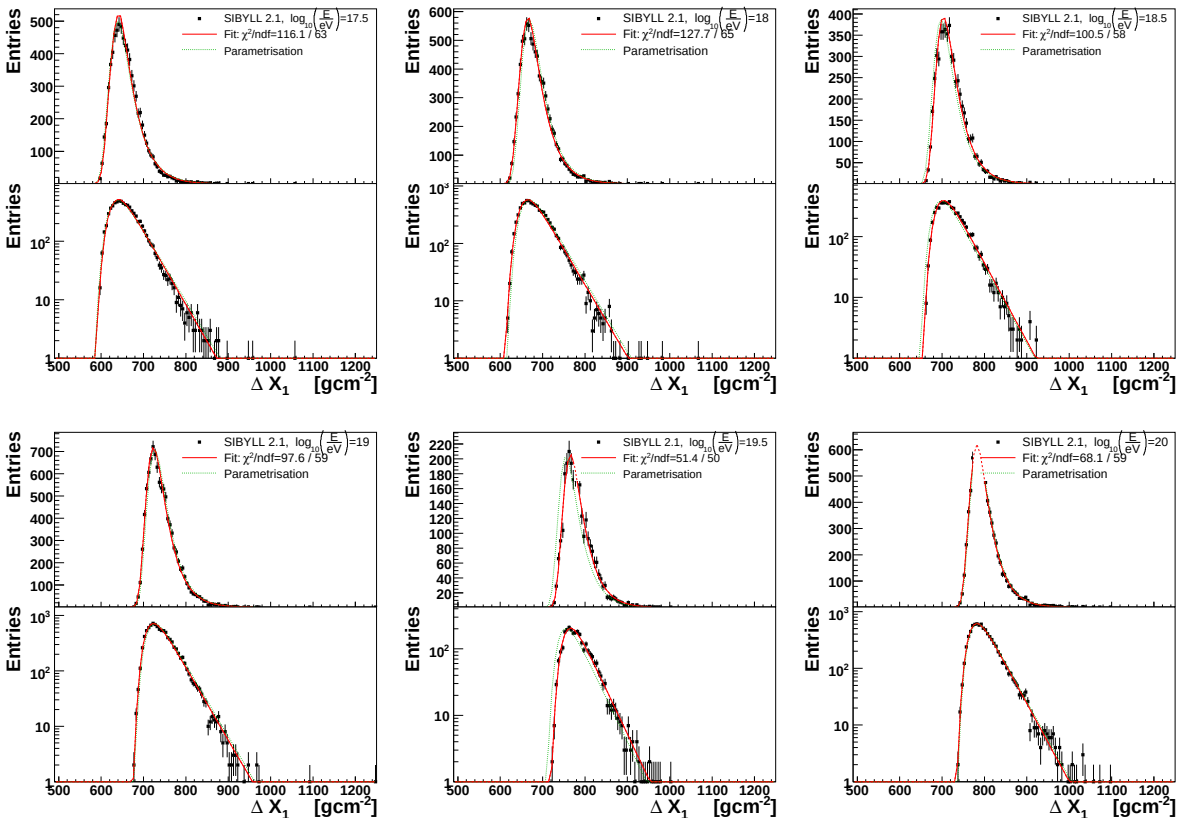


Figure B.9: Selection of fits and parameterizations of the ΔX_1-distributions for SIBYLL.

B.1.2 QGSJET01

Table B.2: Complete list of parameterizations of the ΔX_1-model for QGSJET01.

$f_{10\,\mathrm{EeV}}$	N_{sim}	α [gcm^{-2}]	β [gcm^{-2}]	γ [g^{2}cm^4]	χ^2/ndf
17.0	9647	590.4±0.4	19.4±0.2	5.7e-02±1.4e-02	148.0/77
17.5	9656	619.2±0.4	19.1±0.2	8.9e-02±1.6e-02	88.7/72
18.0	5774	648.0±0.5	19.5±0.3	1.5e-01±2.5e-02	97.1/71
18.5	2026	674.0±0.8	19.2±0.4	2.1e-01±4.9e-02	66.9/56
19.0	9645	701.8±0.3	19.8±0.2	1.9e-01±2.0e-02	111.3/72
19.5	389	726.7±1.9	21.6±1.1	2.8e-01±1.5e-01	46.0/46
20.0	7709	753.8±0.5	20.8±0.3	8.0e-02±2.0e-02	158.5/71

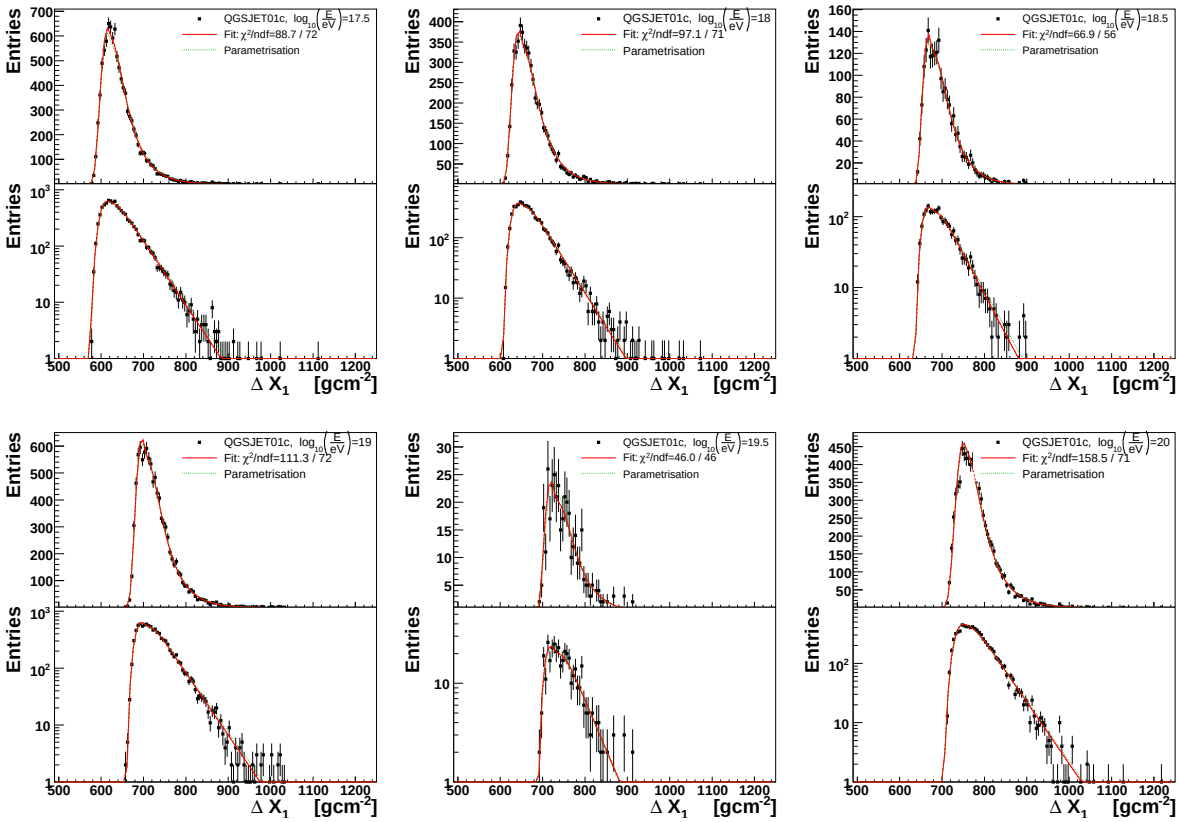


Figure B.10: Selection of fits and parameterizations of the ΔX_1-distributions for QGSJET01.

B.1.3 QGSJETII

Table B.3: Complete list of parameterizations of the ΔX_1-model for QGSJETII.

$f_{10\,\mathrm{EeV}}$	N_{sim}	α [gcm^{-2}]	β [gcm^{-2}]	γ [g^{2}cm^4]	χ^2/ndf
17.0	9496	611.3±0.4	17.0±0.2	-5.9e-02±8.1e-03	110.8/71
17.5	9639	641.6±0.4	15.7±0.2	-1.0e-01±6.5e-03	103.0/68
18.0	8923	670.9±0.4	15.2±0.2	-1.1e-01±6.5e-03	126.6/63
18.5	2228	696.1±0.9	14.8±0.4	-1.3e-01±1.1e-02	91.0/51
19.0	9695	724.8±0.4	15.3±0.2	-1.4e-01±5.0e-03	118.7/64
19.5	1835	748.4±1.2	15.7±0.5	-1.6e-01±1.1e-02	65.0/52
20.0	7812	771.4±0.6	16.1±0.2	-1.5e-01±6.3e-03	82.7/62

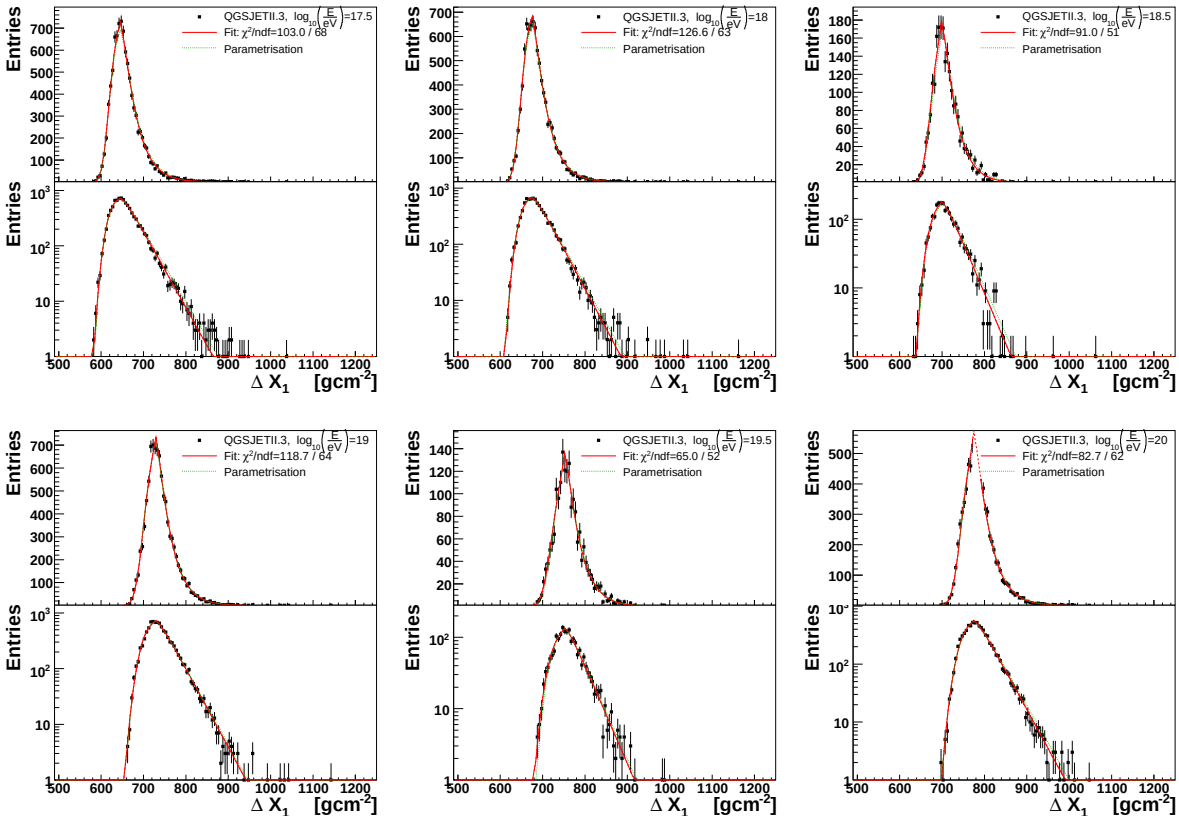


Figure B.11: Selection of fits and parameterizations of the ΔX_1-distributions for QGSJETII.

B.2 Cross section dependence

B.2.1 SIBYLL 10 EeV

Table B.4: Complete list of parameterizations of the ΔX_1-model for SIBYLL.

$f_{10\,\mathrm{EeV}}$	N_{sim}	α [gcm^{-2}]	β [gcm^{-2}]	γ [g^{2}cm^4]	χ^2/ndf
0.5	9832	744.6±0.5	23.3±0.2	4.8e-02±1.3e-02	157.5/95
0.6	9818	739.3±0.5	21.7±0.2	2.1e-02±1.2e-02	147.5/87
0.7	9815	734.0±0.4	19.7±0.2	-1.6e-03±1.1e-02	105.4/74
0.8	9830	729.5±0.4	18.5±0.2	-1.7e-02±1.0e-02	109.2/70
0.85	9823	727.3±0.4	18.3±0.2	-1.0e-02±1.1e-02	125.6/71
0.9	9827	725.3±0.4	17.4±0.2	-3.4e-02±9.7e-03	132.9/68
0.95	9821	723.5±0.4	17.1±0.2	-3.6e-02±9.5e-03	80.7/67
0.99	9828	722.6±0.3	16.6±0.2	-4.9e-02±8.9e-03	115.0/64
1.0	9825	721.5±0.4	16.3±0.2	-3.1e-02±9.4e-03	74.5/60
1.01	9809	721.7±0.4	16.3±0.2	-3.6e-02±9.4e-03	80.7/62
1.05	9810	720.0±0.4	16.1±0.2	-4.8e-02±9.1e-03	120.2/63
1.1	9821	717.2±0.4	16.1±0.2	-1.8e-02±1.1e-02	81.5/57
1.15	9825	717.0±0.4	15.8±0.2	-4.5e-02±9.2e-03	85.8/60
1.4	9808	709.4±0.3	14.6±0.2	-5.0e-02±8.6e-03	80.6/57
1.6	9817	704.3±0.3	13.9±0.2	-5.3e-02±8.6e-03	57.4/53
1.8	9835	700.0±0.3	13.8±0.2	-3.8e-02±9.2e-03	60.0/55
2.0	9822	696.5±0.3	13.6±0.1	-3.1e-02±9.4e-03	77.1/54

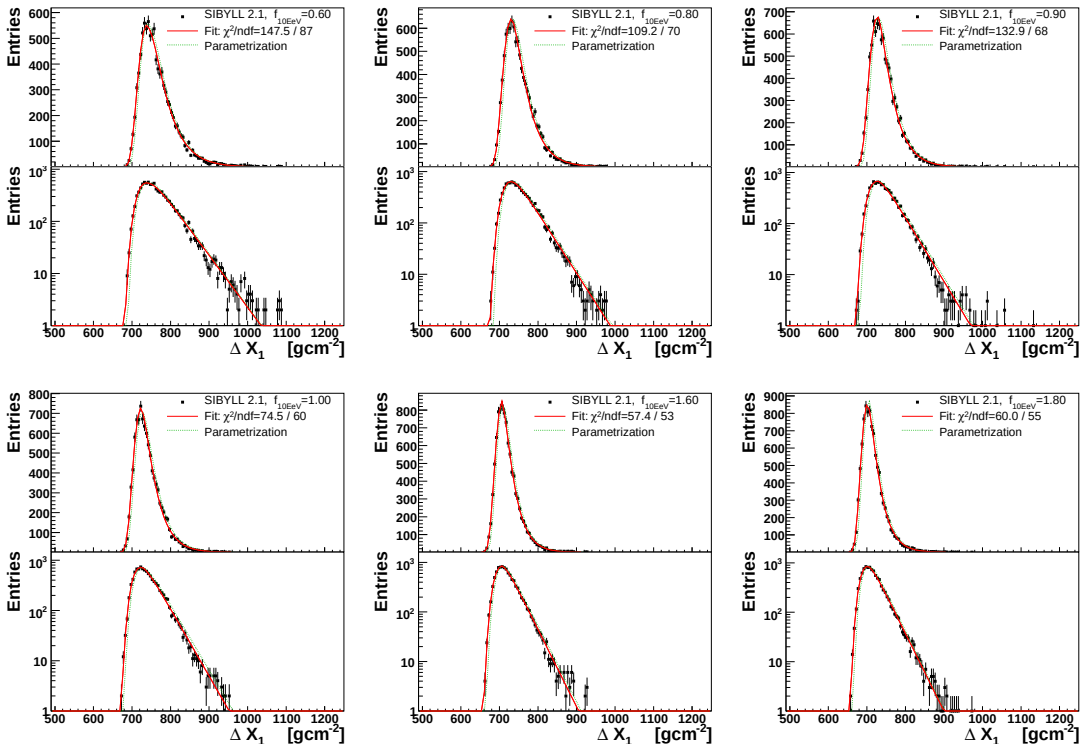


Figure B.12: Selection of fits and parameterizations of the ΔX_1-distributions for SIBYLL.

B.2.2 QGSJET01 10 EeV

Table B.5: Complete list of parameterizations of the ΔX_1-model for QGSJET01.

$f_{10\,\mathrm{EeV}}$	N_{sim}	α [gcm^{-2}]	β [gcm^{-2}]	γ [g^{2}cm^4]	χ^2/ndf
0.5	9658	717.4±0.4	27.6±0.3	2.7e-01±2.3e-02	218.4/105
0.6	9650	712.6±0.4	25.5±0.3	2.5e-01±2.3e-02	169.8/98
0.7	9686	709.6±0.4	23.8±0.2	1.8e-01±1.9e-02	125.7/87
0.8	9645	707.4±0.4	22.2±0.2	1.4e-01±1.8e-02	111.7/79
0.85	9676	706.1±0.4	21.6±0.2	1.1e-01±1.7e-02	95.2/77
0.9	9646	704.3±0.4	20.9±0.2	1.1e-01±1.7e-02	110.8/73
0.95	9629	702.5±0.4	21.1±0.2	1.5e-01±1.9e-02	92.2/75
0.99	9683	701.5±0.4	20.6±0.2	1.4e-01±1.8e-02	133.3/74
1.0	9703	701.7±0.4	19.8±0.2	8.3e-02±1.6e-02	92.8/72
1.01	9690	701.3±0.4	20.2±0.2	1.2e-01±1.8e-02	102.4/70
1.05	9680	700.9±0.4	19.7±0.2	7.9e-02±1.5e-02	121.3/74
1.1	9614	698.8±0.4	19.3±0.2	9.9e-02±1.6e-02	77.9/68
1.15	9643	698.5±0.4	19.3±0.2	1.0e-01±1.6e-02	115.1/69
1.4	9678	692.8±0.4	17.8±0.2	8.3e-02±1.6e-02	87.4/60
1.6	9653	690.2±0.3	16.9±0.2	4.7e-02±1.3e-02	90.7/60
1.8	9679	686.9±0.3	16.3±0.2	5.7e-02±1.4e-02	103.4/59
2.0	9646	684.4±0.3	15.6±0.2	5.1e-02±1.4e-02	107.7/57

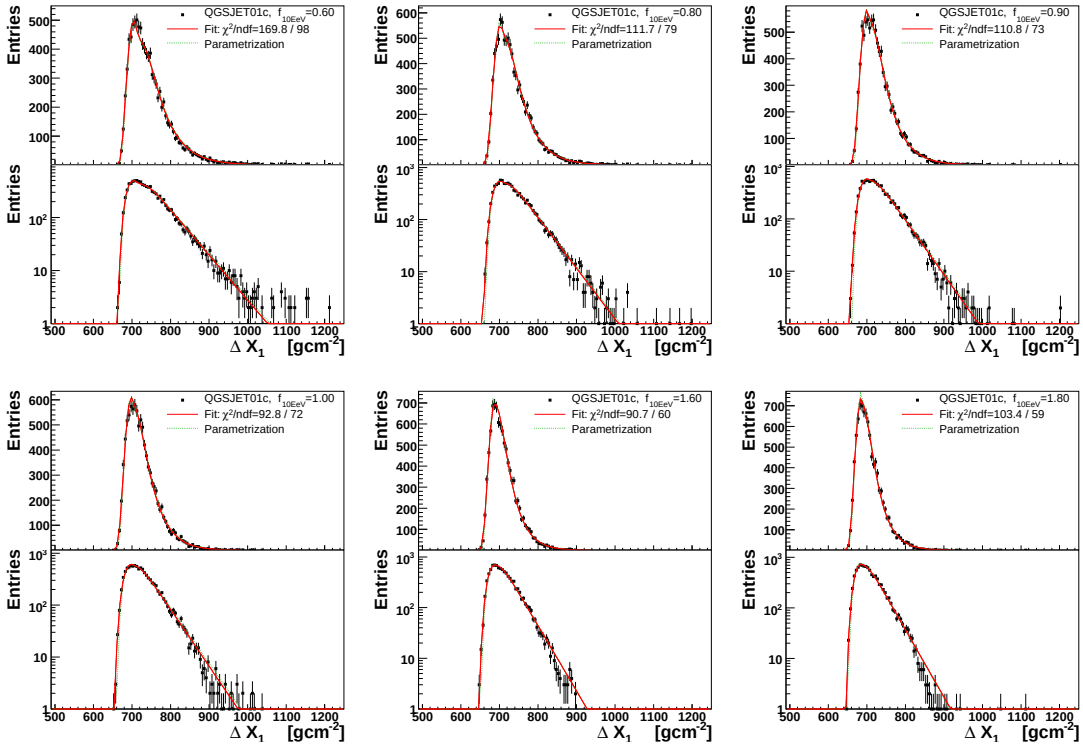


Figure B.13: Selection of fits and parameterizations of the ΔX_1-distributions for QGSJET01.

B.2.3 QGSJETII 10 EeV

Table B.6: Complete list of parameterizations of the ΔX_1-model for QGSJETII.

$f_{10\,\mathrm{EeV}}$	N_{sim}	α [gcm^{-2}]	β [gcm^{-2}]	γ [g^{2}cm^4]	χ^2/ndf
0.5	9687	739.8±0.5	21.9±0.3	-1.0e-01±6.4e-03	227.7/96
0.6	9701	736.5±0.6	20.2±0.2	-1.2e-01±6.2e-03	189.1/87
0.7	9636	733.7±0.5	18.2±0.2	-1.4e-01±5.1e-03	152.7/75
0.8	9613	729.9±0.5	17.0±0.2	-1.5e-01±4.8e-03	131.3/72
0.85	9651	728.3±0.4	16.7±0.2	-1.5e-01±4.7e-03	137.5/69
0.9	9650	727.0±0.7	16.4±0.2	-1.5e-01±6.2e-03	106.2/68
0.95	9587	725.8±0.5	15.9±0.2	-1.5e-01±4.9e-03	99.1/66
0.99	9685	724.9±0.5	15.9±0.2	-1.5e-01±4.9e-03	116.1/66
1.0	9688	725.0±0.5	15.7±0.2	-1.5e-01±4.7e-03	105.0/66
1.01	9639	724.1±0.5	15.6±0.2	-1.5e-01±4.7e-03	107.1/63
1.05	9681	723.0±0.4	14.9±0.2	-1.5e-01±4.5e-03	90.3/63
1.1	9664	722.7±0.3	14.5±0.2	-1.6e-01±3.8e-03	116.3/63
1.15	9699	719.7±0.5	15.2±0.2	-1.4e-01±5.3e-03	138.1/59
1.4	9576	714.4±0.4	13.8±0.2	-1.5e-01±4.9e-03	88.0/56
1.6	9581	711.0±0.4	12.8±0.2	-1.6e-01±4.7e-03	88.1/52
1.8	9667	708.2±0.3	12.2±0.2	-1.7e-01±4.2e-03	94.3/52
2.0	9654	704.0±0.4	12.5±0.2	-1.5e-01±4.8e-03	107.6/50

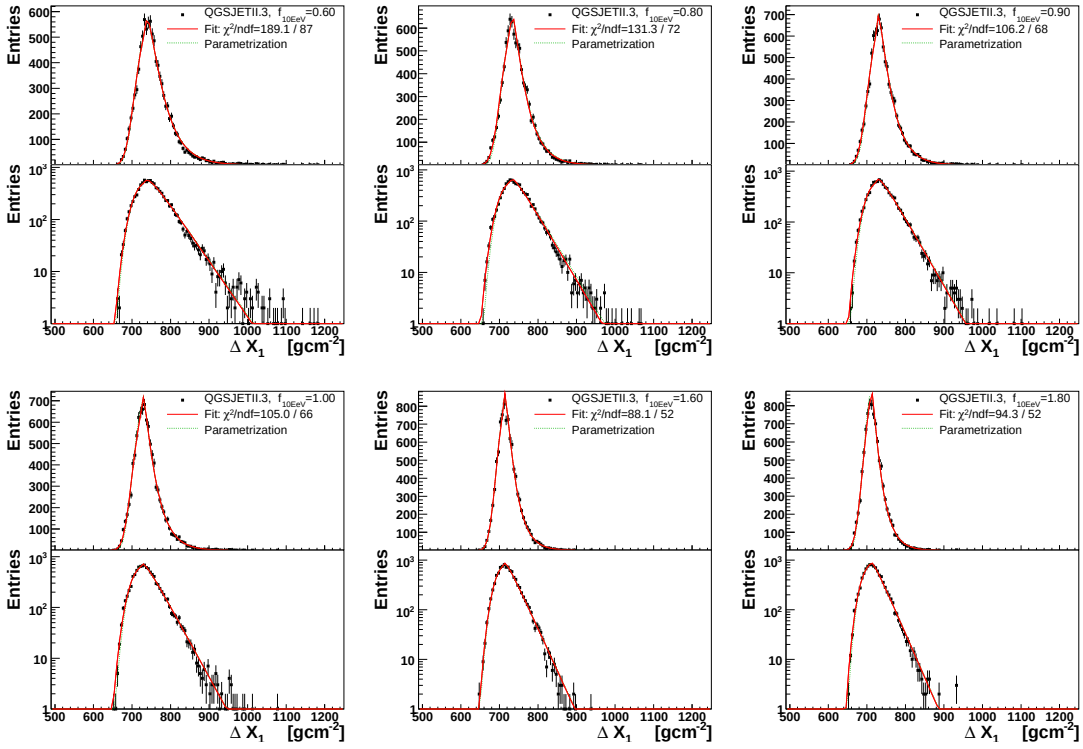


Figure B.14: Selection of fits and parameterizations of the ΔX_1-distributions for QGSJETII.

C Fits of the cross-section analysis

C.1 Analysis based on SIBYLL

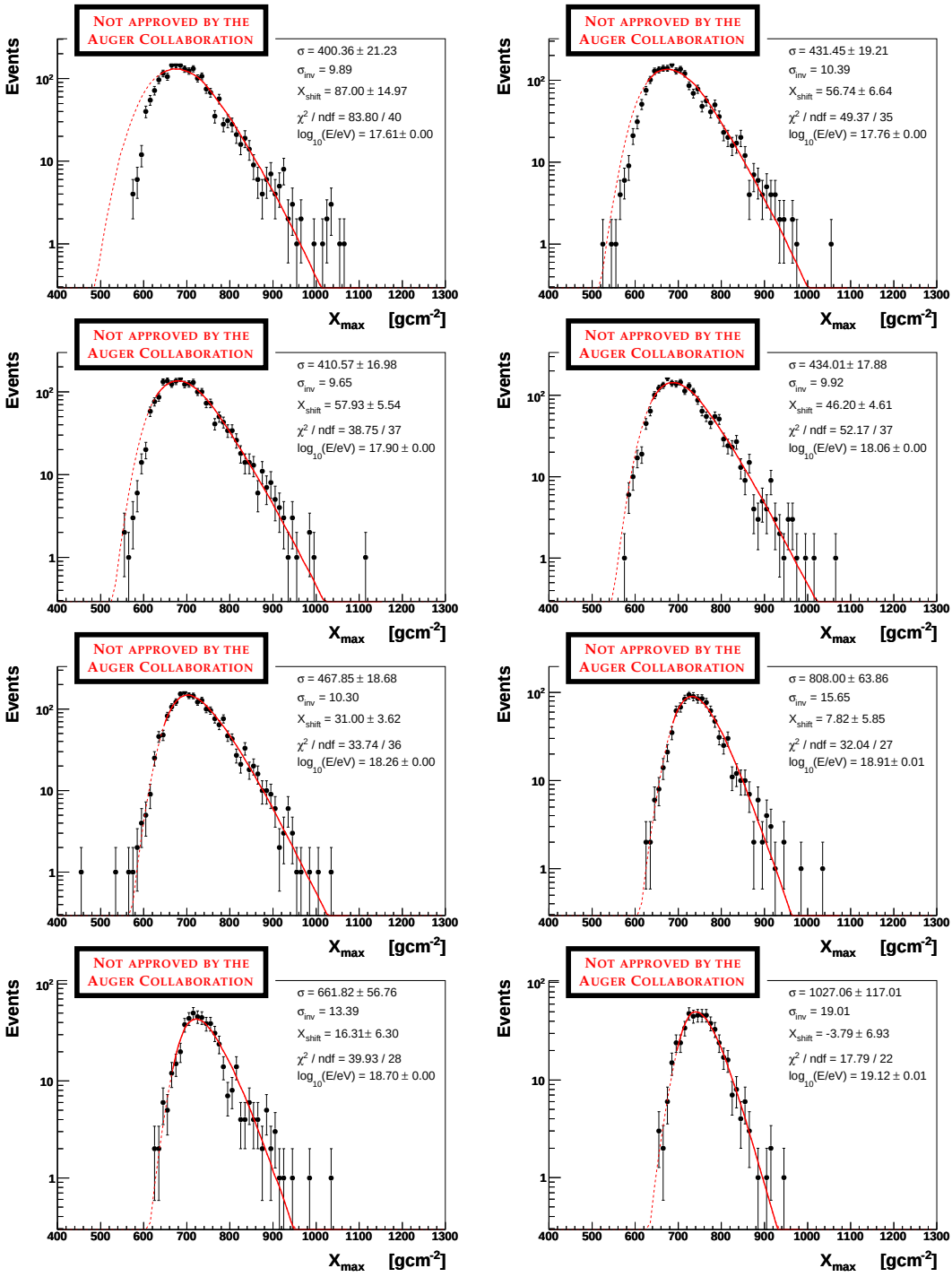


Figure C.15: Fits to the hybrid $X_{\max}$-distributions. *Note that the shown hybrid data are not corrected for detector acceptance effects. Read Section 5.3 for more details.*

C.2 Analysis based on QGSJET01

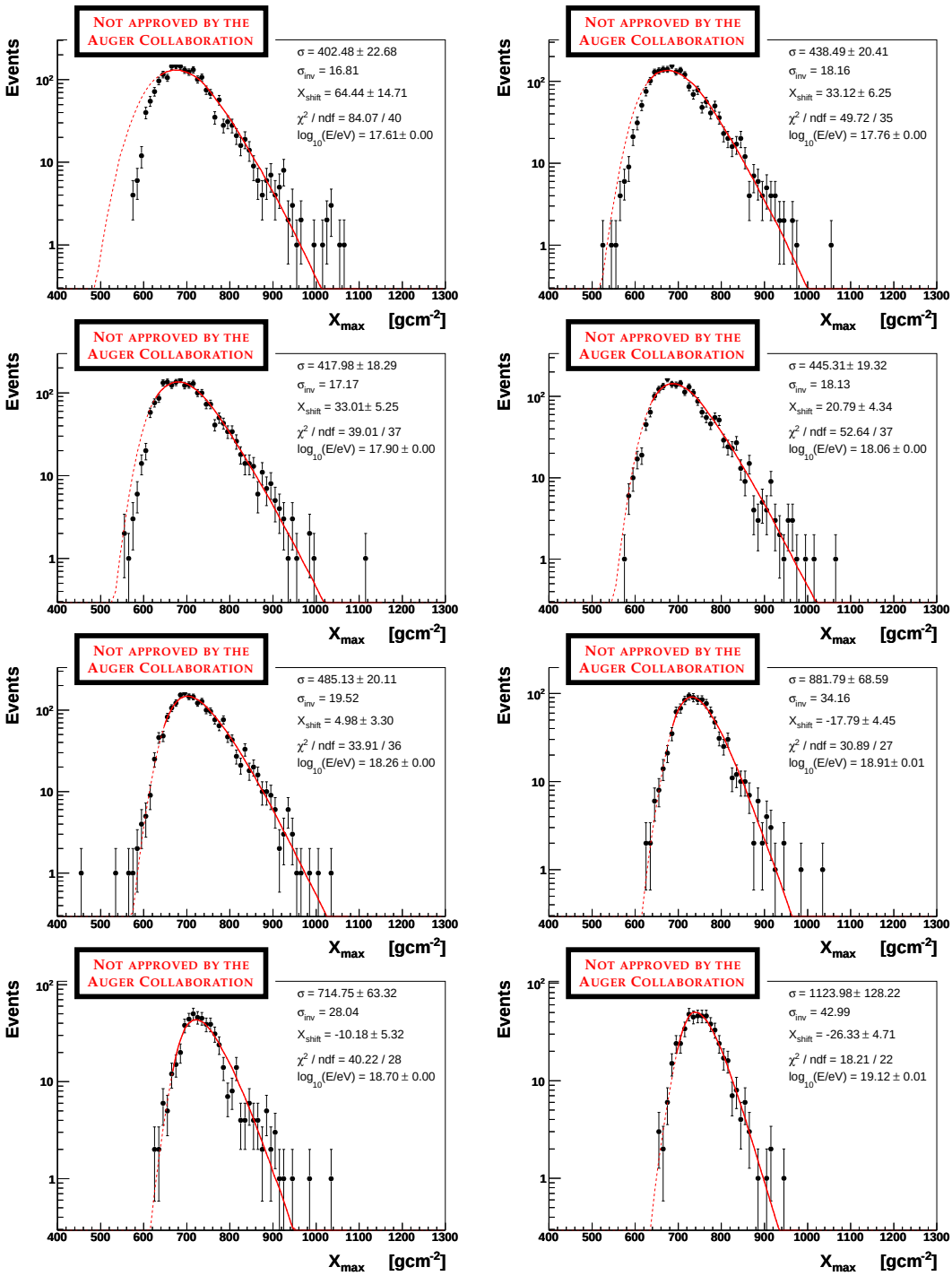


Figure C.16: Fits to the hybrid $X_{\max}$-distributions. *Note that the shown hybrid data are not corrected for detector acceptance effects. Read Section 5.3 for more details.*

C.2.1 Analysis based on QGSJETII

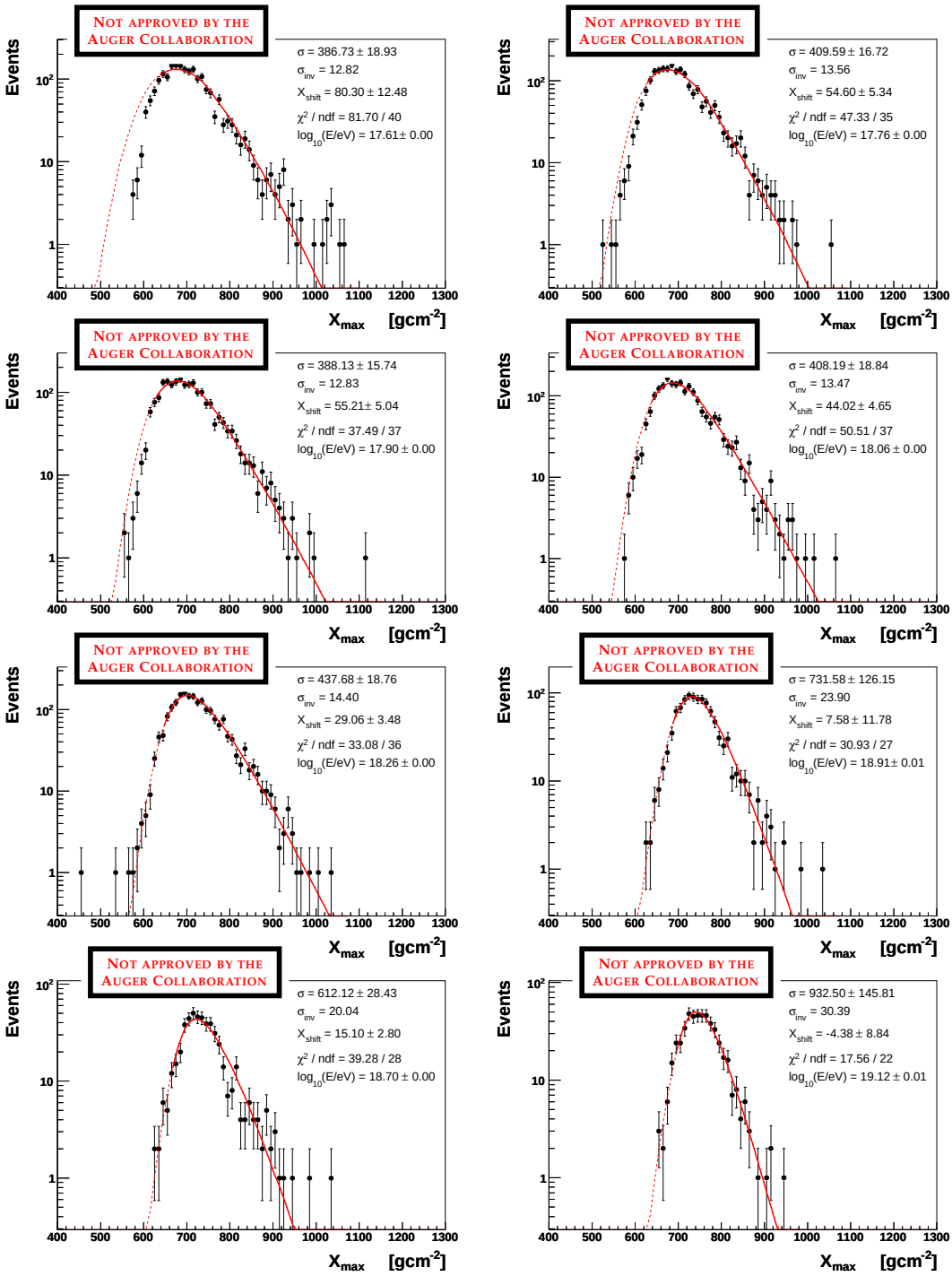


Figure C.17: Fits to the hybrid $X_{\max}$-distributions. *Note that the shown hybrid data are not corrected for detector acceptance effects. Read Section 5.3 for more details.*