

**Radio-detection of ultra-high energy cosmic rays:
multi-channel data acquisition and combined analysis of timing and
energy information**

Zur Erlangung des akademischen Grades eines
DOKTORS DER NATURWISSENSCHAFTEN (Dr. rer. nat.)

von der KIT-Fakultät für Physik des
Karlsruher Instituts für Technologie (KIT)

genehmigte
DISSERTATION

von
Dipl. Phys. Benedikt Hermann Zimmermann
aus Dernbach

Tag der mündlichen Prüfung: 21.07.2017

Referent: Prof. Dr. Marc Weber

Korreferent: Prof. Dr. Dr. h.c. Johannes Blümer

Betreuer: Dr. Matthias Kleifges

Betreuer: Dr. Tim Huege

Contents

1. Deutsche Zusammenfassung	1
2. Introduction	3
3. Cosmic ray physics at the highest energies	5
3.1. Properties of cosmic rays	5
3.1.1. Energy	6
3.1.2. Particle type	8
3.1.3. Arrival direction	9
3.2. Cosmic ray induced extensive air-showers	10
3.2.1. Heitler-Matthews and semi-superposition model	10
3.2.2. Monte Carlo Codes	12
3.2.3. Muon deficit	13
3.3. Radio emission of extensive air-showers	13
3.3.1. Macroscopic picture	13
3.3.2. Microscopic picture	15
3.3.3. Measurements and their interpretation	17
4. The Pierre Auger Observatory	21
4.1. The detectors	22
4.1.1. The Surface Detector (SD)	22
4.1.2. The Fluorescence Detector (FD)	23
4.1.3. The Auger Muon and Infill Ground Array (AMIGA)	24
4.1.4. Upgrade of the Pierre Auger Observatory	24
4.2. Radio detection with the Auger Engineering Radio Array	25
4.2.1. Scientific goals of AERA	26
4.2.2. Deployment stages of AERA	26
4.2.3. Properties of a single AERA station	27
4.2.4. The data acquisition	31
4.2.5. Time corrections via the beacon setup	31
4.2.6. Results and future	32
5. Improvement of the local data acquisition software	35
5.1. Controlled shutdown for extreme conditions	35
5.2. Reliability and sustainable development via memory protection	38

5.3.	Incorporation of the differential GPS measurements in the detector software	39
5.4.	Independently configurable readout for all input channels	40
5.4.1.	Physical motivation	40
5.4.2.	Adaptions with respect to the 2-channel readout	41
5.4.3.	Selftrigger performance	43
5.4.4.	Timing at AERA	47
5.4.5.	4-channel readout performance	48
5.5.	Conclusion	49
6.	Analysis of radio data in a global fit approach	51
6.1.	High level data from the baseline detectors	52
6.1.1.	SD reconstruction	52
6.1.2.	FD reconstruction	53
6.2.	The AERA data	53
6.2.1.	The Simulated and measured data set	54
6.3.	The standard analysis strategy	56
6.3.1.	Arrival time fit	56
6.3.2.	Radiation energy fluence fit	59
6.4.	The global regression approach	60
6.4.1.	The analytic description of the weighted uncorrelated χ^2 fit	60
6.4.2.	Testing the independence of the variables	62
6.5.	Implementation of the global regression	63
6.5.1.	The fit parameter	64
6.5.2.	Shared Geometry	64
6.5.3.	Staged fit	64
6.5.4.	Changes with respect to the standard reconstruction	65
6.5.5.	Adaption of the measurement uncertainties	65
6.6.	Validation of the module	65
6.6.1.	Arrival time fit	65
6.6.2.	Energy fluence fit	69
6.7.	Results of the global regression	72
6.7.1.	Reconstruction performance	72
6.7.2.	Reconstruction quality	74
6.7.3.	Application of the global fit to data	79
6.8.	Discussion	81
6.8.1.	Insensitivity of the radio wavefront to the distance to the shower maximum	81
6.8.2.	Insensitivity of the radio wavefront to the atmospheric depth of the shower maximum	82
7.	Conclusion	85
	Appendices	89
A.	The weighted mean and its uncertainty	89

B.	Arrival time fit $R_{\max}(\gamma)$ without Noise	90
C.	Application of the global fit to data	91
D.	Reconstruction settings for simulation data	92
D.1.	Modulesequence	92
D.2.	Bootstrap	94

Deutsche Zusammenfassung

Kontinuierlich treffen hochenergetische Teilchen der kosmischen Strahlung auf die Erde. Sie trägt, abhängig von der Höhe des Lebensraums, zu einem Viertel der jährlichen Strahlenbelastung von Menschen bei. In der Atmosphäre lösen hochenergetische Teilchen Kaskaden (sogenannte ausgedehnte Luftschauer) von Sekundärteilchen aus, welche den Boden erreichen können. Die geladenen Sekundärteilchen eines Luftschauers strahlen ein Radiosignal ab, wenn sie im Erdmagnetfeld abgelenkt werden. Dieses Signal kann mithilfe von großflächig verteilten Radioantennenstationen nachgewiesen werden. Das Auger Engineering Radio Array (AERA) ist das weltweit größte Radioexperiment zur Untersuchung hochenergetischer kosmischer Strahlung. Es ist Teil des Pierre Auger Observatoriums, welches in Argentinien gelegen ist.

Im Zuge dieser Arbeit wurde die Datennahmesoftware der AERA-Detektoren verbessert, um eine einstellbare Kombination der vier Eingangskanäle auslesen zu können. Dies wird durch den Wunsch motiviert, in der neusten Ausbaustufe AERA-III zusätzlich zu der Ost-West und Nord-Süd Polarisation nun auch den vertikalen Anteil des Radiosignals zu messen. Daher ist die Elektronik der AERA-III Stationen auf das Auslesen von dreipoligen Antennen ausgelegt. Gerade bei horizontalen Luftschauern, welche eine zu den Oberflächendetektoren komplementäre Information tragen, gewinnt dieser Polarisationsanteil des Radiosignals an Bedeutung. Schon ab einer Neigung von mehr als 54° strahlen Luftschauer bis zu zwei Drittel der Leistung vertikal polarisiert ab. Da die Rдиодetektion von Luftschauern durch das unvermeidliche galaktische Rauschen begrenzt ist, ist es sinnvoll, alle Signalkomponenten zu messen. Zusätzlich ermöglicht die neue Datennahmesoftware zum ersten Mal, die Kanäle mit niedriger Verstärkung bei den älteren Ausbaustufen auszulesen. Eine weitere Verbesserung wurde durch das Einführen eines neuen Betriebszustands, des sogenannten Schlafmodus, erreicht. Im argentinischen Winter schützt die Detektorelektronik die Batterien vor Tiefentladung, indem sie die Stromversorgung ab einer Spannungsschwelle von 10,9 V unterbricht. Der neue

Schlafmodus ermöglicht es, die Ausfallzeiten des Detektors während bewölkter Phasen zu steuern, um auch im Winter hochqualitative Daten in Koinzidenz mit den Fluoreszenzteleskopen aufzeichnen zu können. Mit dem Schlafmodus wird ein spontanes Abschalten einzelner Stationen bei Unterschreitung der Abschaltsschwelle vermieden, indem der gesamte Detektor bereits vorbeugend zum Aufladen in diesen Modus versetzt wird. Auf Stationsebene wurde die Genauigkeit der Ereigniszeitstempel durch die Verwendung präziserer GPS-Positionen der Stationen verbessert. Diese Verbesserungen betreffen nicht nur die Stationssoftware, sondern auch die FPGA-Firmware, das Betriebssystem der Stationen und auch die Mikrokontrollersoftware. Das endgültige Datenerfassungssystem wurde daher sorgfältig getestet und schließlich in den AERA Detektor integriert.

Im zweiten Teil dieser Arbeit wird der Vorteil einer multivariaten Beschreibung des Radiosignals durch ein Zeit- und Amplitudenmodell untersucht. Das Amplitudenmodell beschreibt die Energiedichte, welche von den verschiedenen Radiostationen in einem Ereignis gemessen wird, durch eine zweidimensionale laterale Verteilungsfunktion. Das Integral der angepassten Funktion ist sensitiv auf die Energie der kosmischen Strahlung und die Breite der Verteilung auf die Eindringtiefe des Schauers in der Atmosphäre. Das Ankunftszeitenmodell beschreibt die Wellenfront des Radiosignals durch eine Hyperbel. Der Öffnungswinkel der Wellenfront ist ebenfalls sensitiv auf die Eindringtiefe. Beide Modelle hängen über die Schauergeometrie voneinander ab. In dieser Arbeit wurden diese beiden funktionalen Modelle in einem neuen Analysemodul kombiniert. Die Ankunftszeiten- und Energiedichteverteilung wurden als unabhängig voneinander betrachtet und daher ein unkorrelierter Chi-Quadrat-Fit als Ansatz gewählt. Es wurden zusätzliche Modellunsicherheiten für eine korrekte Gewichtung der funktionalen Modelle eingeführt. Die Implementierung der Modelle wurde anhand der bereits existierenden univariaten Module validiert. Es stellt sich jedoch heraus, dass die multivariate Anpassung zu keiner verbesserten Rekonstruktion der Energie oder der Eindringtiefe des Schauers führt. Dies liegt zum Teil an der geringeren numerischen Stabilität der Anpassung, welche bis zu acht freie Parameter hat. Die limitierenderen Faktoren sind jedoch die größere Sensitivität des Ankunftszeitenmodells auf Rauschen und die intrinsisch geringere Sensitivität auf Luftschauerobservablen. Eine grobe Abschätzung zeigt, dass die Zeitauflösung wesentlich besser als die bisher erreichten 2 ns sein muss, um mithilfe des Wellenfrontmodells zumindest Eisenkerne und Protonen auf statistischer Basis als Primärteilchen unterscheiden zu können. Dieses Ergebnis ist unabhängig von der hier verwendeten Analyseverfahren des unkorrelierten Chi-Quadrat-Fit.

Zusammenfassend hat diese Arbeit den durchgängigen Betrieb des AERA Radiodetektors auch unter schwierigen Umweltbedingungen, wie zum Beispiel starker Bewölkung im Winter, gewährleistet. Durch die Softwareverbesserungen lässt sich das volle Potential des Detektors ausschöpfen. Gleichzeitig bereiten diese Verbesserungen die Elektronik für das Auslesen der nächsten Generation von Prototypstationen vor. Die Analyse der Zeitbestimmung auf Stationsebene und die multivariate Analyse zeigen, dass die erreichte Zeitauflösung nicht ausreicht, um von der gemessenen Ankunftszeit auf die Schauereindringtiefe schließen zu können. Zukünftige Radioexperimente brauchen eine bessere Realisierung der Zeitbestimmung auf Stationsebene.

CHAPTER 2

Introduction

Cosmic ray particles constantly impact the earth atmosphere. Some of them have energies well beyond the reach of human made particle accelerators. Thus, they offer the chance to probe elementary particle physics at energy scales inaccessible in laboratory experiments. However, the origin of these ultra high energy cosmic rays is yet unknown. Only the most violent astronomical sources are expected to be capable of accelerating particles to the extent of observed energies. By detecting cosmic rays on earth, one can learn about their origin, their acceleration mechanism and the environment they have passed.

High energy cosmic rays initiate cascades of charged secondary particles in the Earth's atmosphere. These air showers emit a radio signal which can be measured with radio detectors. The Auger Engineering Radio Array is the largest radio detector for ultra high energy cosmic ray physics, worldwide. It is located at the Pierre Auger Observatory in Argentina. The newest deployment stage AERA-III is optimised to measure the radiation from inclined air showers with a sparse antenna grid. As a consequence, the front-end electronics of the newest AERA detectors are designed for the readout of three-pole antennas which are also sensitive to the vertical component of the radio emission of air showers. The electronics of the former deployment stages AERA-I and AERA-II record the East-West and North-South component of the radio signal from air showers, but each with a high and low gain for a higher dynamic range. However, the contemporary data acquisition software can only read out two channels, i.e. the high gain channels of the AERA-I and AERA-II stations. In this work, the data acquisition software of the radio stations is adapted for an improved and more reliable radio detection, to uncover the full potential of AERA. The functionality of the new data acquisition system has been verified in the laboratory and used productively at two AERA stations.

Up to now, the AERA data has always been analysed step by step while focusing on a special kind of information (e.g. signal strength, signal time and signal polarisation) in each step. Already, there are

widely respected models to describe the signal time and the signal strength of the radio emission from air showers. The signal strength model describes the energy fluence measured at each radio detector by a two-dimensional lateral distribution function. The signal arrival time model assumes the wave-front of the radio signal to be hyperbolic. Both models share the sensitivity to air shower observables of interest (e.g. the primary particle type). This work clarifies the prospects of describing the measured radio signal by a combination of these models. The performance of the existing univariate analyses and the new multivariate analysis is examined in a simulation study taking the effects of realistic noise into account. The new multivariate analysis is also applied to measured data.

Cosmic ray physics at the highest energies

3.1. Properties of cosmic rays

Generally speaking, cosmic rays are accelerated stable particles like neutrinos, photons and ionised nuclei, which propagate through our universe. By detecting cosmic rays on earth, one can learn about their origin, their acceleration mechanism and the environment they have passed. As every particle type carries a different piece of information, many experiments each specialised in detecting a particular particle are needed. It has become clear that combining the numerous experimental findings in a so-called multi-messenger approach is the best way to progress in astronomy and astroparticle physics. However, the following section will focus on the special case of nuclei.

Depending on the energy of cosmic rays, two different detection strategies need to be employed: Cosmic rays with energies below 10^{14} eV are best detected directly by experimental setups carried on balloons or satellites avoiding their absorption in the atmosphere. These experiments can identify all particle properties with high accuracy but their size and thus their aperture is limited to the loading capacity of the aircraft. At high energies above 10^{14} eV to 10^{15} eV the particle flux becomes too low to allow for a direct detection with sufficient statistics. Instead, high energy cosmic rays can only be detected indirectly via the extensive air showers they produce in the atmosphere. Extensive air showers are the product of the diffractive interaction of the primary particle with an air nucleus high in the atmosphere which subsequently starts a hadronic and electromagnetic particle cascade of secondary particles, similar to particle showers in particle accelerator experiments. Only large experiments, which can cover areas of several square kilometres, are capable of studying cosmic ray physics at the highest energies with enough statistics. In the energy range from 10^{14} eV to 10^{15} eV both detection strategies can be applied.

In the following, I will shortly discuss the properties of cosmic rays, concentrating mostly on high energy particles. Their experimental detection is conducted with the aid of large ground based observatories.

3.1.1. Energy

Figure 3.1 shows the measured energy dependent all particle flux $J(E)$ with a log scaling on both axes. It is weighted by the cosmic ray energy $E^{2.5}$ to reveal deviations from the general trend. Both, the flux and energy of cosmic particles arriving at earth span a range of many decades. Cosmic rays with energies below 10^{10} eV are shielded by the solar magnetic field and thus the flux at earth is suppressed. Above this threshold, the flux $J(E)$ can be described by a broken power law

$$J(E) \propto E^\gamma. \quad (3.1)$$

Here, the energy spectrum can be divided in four intervals. The spectral index γ equals -2.7 in the first interval $10^{10} \text{ eV} \leq E \leq 3 \times 10^{15} \text{ eV}$. Then, after a discontinuity in the slope which is called the “knee”, the energy spectrum becomes softer with $\gamma = -3.1$ for $3 \times 10^{15} \text{ eV} < E < 10^{18} \text{ eV}$. Within the third energy interval $10^{18} \text{ eV} \leq E \leq 4 \times 10^{19} \text{ eV}$, the so-called “ankle”, the energy spectrum hardens before a steep falloff takes place. The reasons for the discontinuities in the slope are open questions. However, I will outline the most popular explanations which are consistent with the data: It is believed that below $O(10^{17} \text{ eV})$ only galactic sources contribute to the cosmic ray flux, because the diffusion time through typical extra-galactic magnetic fields becomes comparable to the age of the universe. Also, one expects galactic magnetic fields to be strong enough to contain such particles. Up until cosmic ray energies slightly above the knee, particle physics is experimentally well checked. Hence, one can be confident to note astrophysical effects in the energy spectrum. At latest, cosmic rays with energies above the ankle are expected to be of extra galactic origin, because they are no longer believed to be confined by the magnetic fields. Additionally, the galactic sources might reach their limit for acceleration. In such a scenario, the flattening of the flux in the ankle region is interpreted as the transition from galactic to extra-galactic cosmic rays [1]. So far, it is not clear if the strong suppression of the flux at energies of around 10^{20} eV is caused by the shrinking maximal distance at which sources can be seen due to the Greisen-Zatsepin-Kuzmin (GZK) effect [2, 3] or due to a maximum acceleration limit of extra-galactic cosmic ray sources or a combination of both effects. One should keep in mind that the momentarily observed falloff in the flux at earth does not necessarily rule out sources which are able to accelerate particles to even higher energies somewhere far away.

Independent of the considerations concerning the GZK-horizon and the diffusion times in the extra-galactic and galactic magnetic fields, even the astrophysical sources which are capable of accelerating UHECR to the measured energies of up to $O(10^{20} \text{ eV})$ have not been identified yet. The “Hillas Plot”(fig. 3.2) shows the relationship between the magnetic field strength and the object size (gyro radius) for different acceleration energies and possible astrophysical objects in a rough estimation. The Hillas criterion does not take any energy losses during the acceleration itself and the subsequent propagation through the universe into account. The essence of this simple picture is that only few and very violent sources are likely to accelerate cosmic rays to the energies measured at earth,

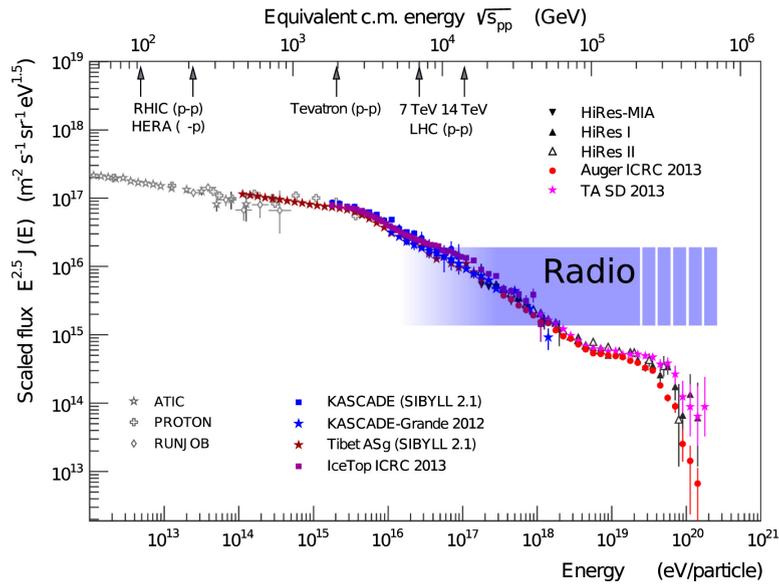


Figure 3.1.: Measured energy spectrum of cosmic rays. From [4] related to [5]

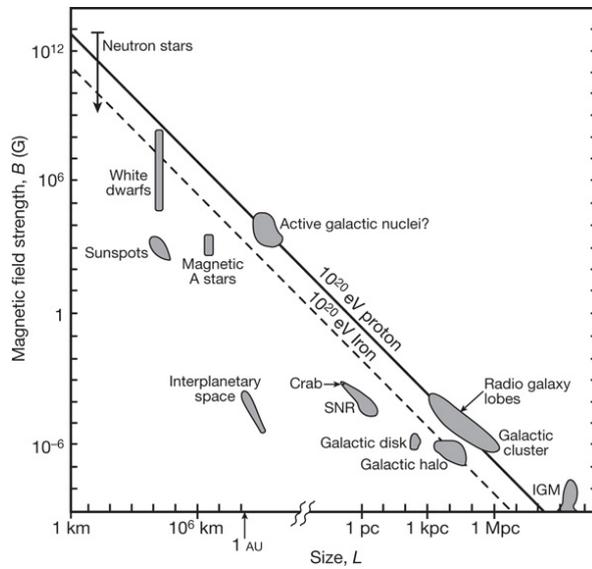


Figure 3.2.: The Hillas Plot [6].

even though the assumptions made in this plot are rather optimistic.

3.1.2. Particle type

At the highest energies, mostly nuclei contribute to the cosmic ray flux. The energy to charge ratio E/Z , which essentially is the rigidity for relativistic particles, specifies how these nuclei interact electro-magnetically during their acceleration in shocks and subsequently with the galactic and extra galactic magnetic fields while propagating through the universe. Therefore, the particle type is a very important property of cosmic rays to know. The composition of cosmic ray particles at given energies confines astrophysical source models on a statistical basis. An event-by-event identification of light nuclei could allow for the reconstruction of their flight paths and would point back to the sources themselves. Figure 3.3 from [7] shows the mean logarithmic mass distribution $\langle \ln A \rangle$ as measured by several experiments for the energy range 10^{15} eV to 10^{20} eV as grey band compared to predictions from several astrophysical models. The band includes the uncertainties and the spread of the measurements, and the use of different hadronic interaction models for the composition determination. The experimental data agree in a change of composition from a light to a heavier one in the energy range

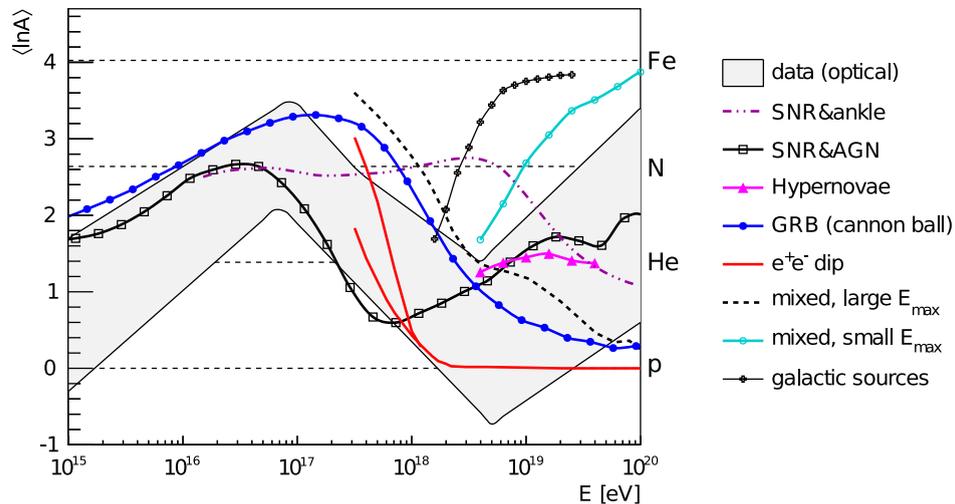


Figure 3.3.: Measured composition of cosmic rays and the source models [7]

from 10^{15} eV to 10^{17} eV above the knee. This matches the expectations for either a rigidity dependent escape of particles from our galaxy or the reach of the maximum rigidity, galactic accelerators can generate. Both mechanisms, which are solely based on rigidity, would lead to a repetition of spectral features (e.g. the knee) for different energies proportional to their atomic charge Z . This phenomenon is called Peter cycles and explains data measured with KASCADE well [8]. Above 10^{17} eV the picture gets indistinct. There is a discrepancy in the determined composition between different experiments. All optical measurements find a trend towards a lighter composition, while the surface detector measurements (e.g. Yakutsk) are compatible with a constant composition. The radio measurements of LOFAR even claims a composition dominated by light nuclei at around 10^{17} eV [9]. But still there is an overall agreement on a trend in composition towards lighter elements within 10^{17} eV to 10^{19} eV. Above 10^{19} eV both optical and particle detector measurements find an increase in the heavier particles fraction. But in this energy range, the errors due to the extrapolation of the hadronic interaction models and experimental uncertainties make the interpretation of the data yet difficult.

Independent of the specific properties all source models behave roughly similar for energies below

10^{19} eV and agree with experimental data (see figure 3.3). Above this energy, the models diverge in terms of the mass composition but also the data cannot constrain this phase-space yet. If the extra galactic cosmic ray flux is not dominated by few nearby sources, it is expected to solely consist of a mixture of protons and iron nuclei due to the GZK suppression and photodisintegration. In such a scenario, a large observatory (e.g. the Pierre Auger Observatory after its upgrade) will constrain the hadronic interaction and source models by discriminating the extreme particle groups at the highest energies.

3.1.3. Arrival direction

The arrival direction of cosmic rays is vital for the search of their origin. But, as the measured arrival direction of cosmic rays is a convolution of the source distribution, source type and propagation effects (like GZK suppression and rigidity dependent deflection), the interpretation of arrival directions is a non-trivial endeavour.

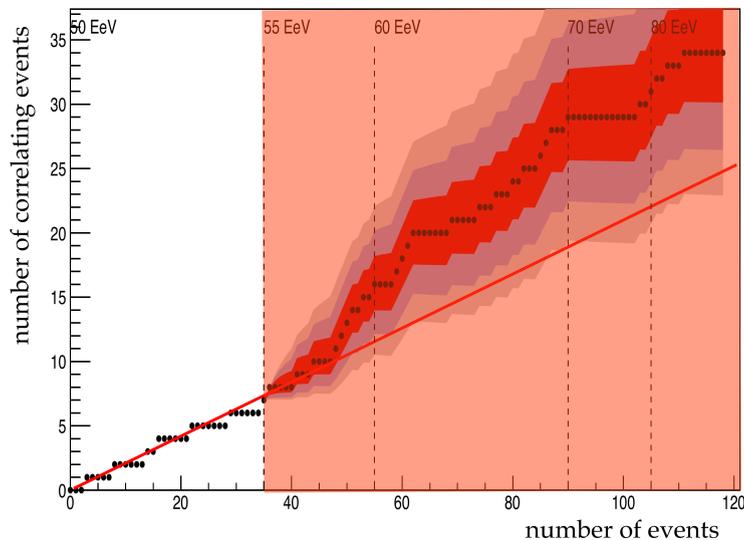


Figure 3.4.: Number of events correlating to objects in the Veron-Cetty Veron Catalog [10] over the total number of measured events by the Pierre Auger Observatory [11]. The events are ordered by their energy. The expectation for an isotropic distribution of arrival directions is shown by the solid line. The contours around the measurements denote the 1-, 2- and 3-sigma contours.

One way to analyse the arrival direction data is to compare it to the known positions of cosmic accelerators in the sky: Figure 3.4 shows the correlation of cosmic ray arrival directions measured by the Pierre Auger Observatory with active galactic nuclei of the Veron-Cetty Veron Catalog [10]. The best correlation is found in the energy interval 55 EeV to 70 EeV. Here, the number of events correlating with sources exceeds the expectation for an isotropic sky with a significance of 2–3 σ .

The most recent analyses search for over densities in the arrival direction distribution [cf. figures 3.5)&(3.6)]. Yet, no evident anisotropy in terms of a deviation of 5 sigma from an isotropic distribution have been found. While [12] probes the arrival directions for cosmic rays at the highest energies, [13] considers cosmic ray events in an energy range where GZK suppression does not constrain the visible horizon. Therefore, it is surprising to note the promising and strong hint for a

dipole moment for energies above 8 EeV in figure 3.6.

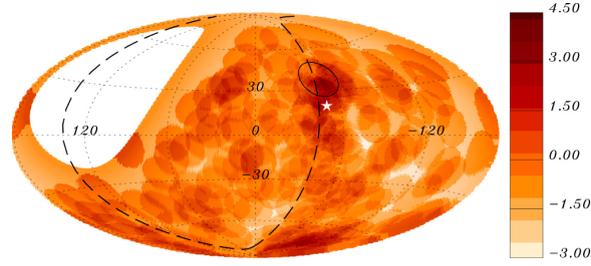


Figure 3.5.: Significances of over densities in 12° windows within galactic coordinates for events with energies above 54 EeV. The dashed line denotes the Super-Galactic Plane while the white star denotes the position of Centaurus A, from [12].

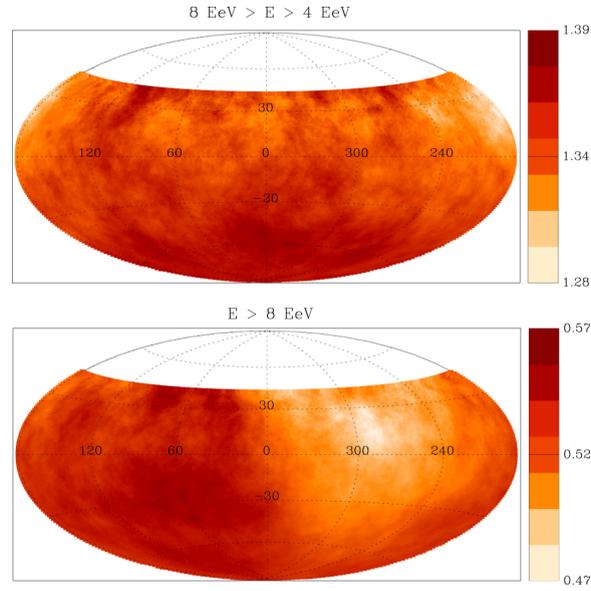


Figure 3.6.: The cosmic ray flux J in units of $\text{km}^{-2} \text{yr}^{-1} \text{sr}^{-1}$ in equatorial coordinates, from [13]. The relative flux variation w.r.t the lowest value only amounts to 8 % for energies $E \in 4 \text{ EeV} \text{ to } 8 \text{ EeV}$, while it reaches 21 % for $E > 8 \text{ EeV}$. Most of the considered cosmic ray events in this figure have energies well below the GZK suppression region.

3.2. Cosmic ray induced extensive air-showers

Highly energetic cosmic rays do not reach the surface of the earth. Instead they interact with air molecules in the atmosphere and produce secondary particles, which themselves interact again producing a particle cascade, called extensive air shower. The following chapter concentrates on the physics of the particle shower development in the air.

3.2.1. Heitler-Matthews and semi-superposition model

Figure 3.7 depicts a simple physics model for air showers, the Heitler-Matthews model [14]. One basic assumption behind the model is the repeated splitting of particles after the propagation through a constant amount of matter.

Let us consider the electromagnetic shower, originally discussed by Walter Heitler, first. In this simple picture, every electron (positron) e will emit a photon γ via bremsstrahlung at each splitting point n

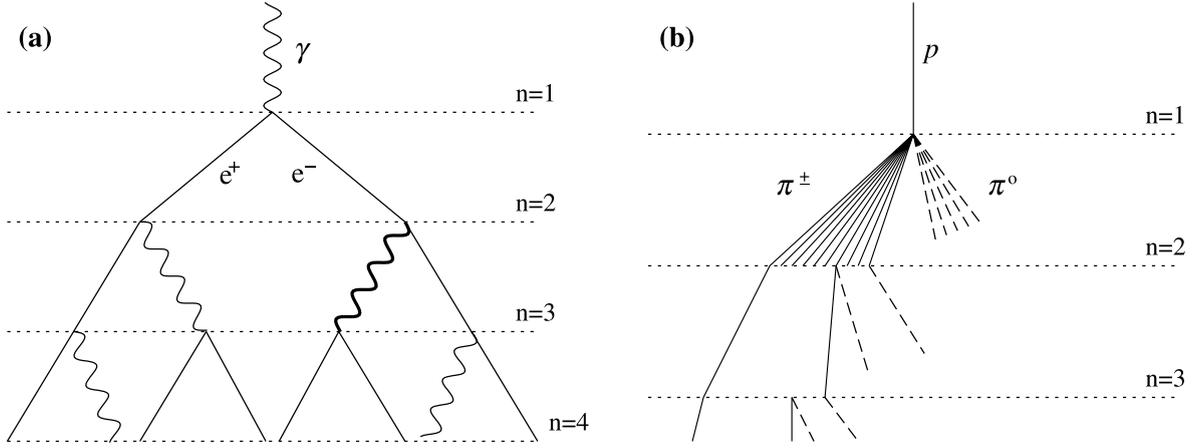


Figure 3.7.: The Heitler-Matthews model: The geomagnetic emission (left) produces a parallel polarisation in the shower plane, the charge excess emission (right) is radially polarised [14].

such that both final state particles carry one half of the parental lepton energy. The constant distance d between these splittings is defined by the radiation length $\lambda_r = 36.95 \text{ g cm}^{-2}$ via

$$d = \lambda_r \ln 2. \quad (3.2)$$

This is the characteristic length over which an electron would lose half of its energy due to radiation. Accordingly, each photon will produce an electron positron pair after traversing the same amount of matter. In this way the energy of the primary particle is uniformly distributed among all secondary particles at every splitting state n . The growing of the electromagnetic cascade ceases when the energy of the secondary particles drops below the critical energy $\xi_{\text{crit}}^e = 85 \text{ MeV}$. It is defined as the energy at which energy losses due to inelastic collision start to dominate over radiation losses. By fixing this critical energy, now, only the energy of the primary particle E_0 defines the whole electromagnetic shower:

The maximum number of splittings n_{crit} is given by

$$n_{\text{crit}} = \ln(E_0/\xi_{\text{crit}}^e)/\ln 2, \quad (3.3)$$

and the maximum number of particles in the electromagnetic cascade $N_{e,\text{max}}$ which is reached at this point is

$$N_{e,\text{max}} = 2^{n_{\text{crit}}} = \frac{E_0}{\xi_{\text{crit}}^e}. \quad (3.4)$$

But most important, the depth of the shower maximum X_{max} in this simple picture only depends on the energy:

$$X_{\text{max}} = \lambda_r \ln(E_0/\xi_{\text{crit}}^e) \quad (3.5)$$

This leads to an elongation rate

$$\Lambda = dX_{\text{max}}/d\log(E_0) \approx 85 \text{ g cm}^{-2} \quad (3.6)$$

for electromagnetic showers in air. In fact, the elongation rate in (3.6) is an upper limit, called the (*Linsley elongation rate theorem* [15]). Measurements of higher elongation rates are a hint for composition change in the cosmic ray flux.

The hadronic cascade in the Heitler-Matthews approach is modelled analogous. Again a splitting length is defined, which differs only by the longer strong radiation length of hadrons ($\lambda_l \approx 120 \text{ g cm}^{-2}$). At every hadronic interaction charged and neutral pions are produced with the relative abundance of 2/1 either by charged pions π^\pm or at the beginning by the leading baryon of the original cosmic ray, carrying equal energy fractions of the parental particle. The actual number of pions after each splitting as depicted in figure 3.7 (10 charged and 5 neutral pions) has been determined from simulation studies [14]. The neutral pions are assumed to instantaneously decay into two photons which start electromagnetic sub-showers as discussed above. In the process the electromagnetic cascade is fed by 1/3 of the hadronic energy budget with every hadronic interaction step. Analogous to the electromagnetic case, the hadronic cascade stops when the energy of the π^\pm drops below a critical energy ξ_{crit}^π of about 20 GeV. At this energy the interaction length of the pions exceeds their decay length and all charged pions are assumed to decay into muons. For a proton as primary particle, their number can be estimated to

$$N_{\mu,\text{max}}^p = \left(\frac{E_0}{\xi_{\text{crit}}^\pi} \right)^\beta, \quad \text{with } \beta \approx 0.85. \quad (3.7)$$

Hadronic particle cascades initiated by primary particles heavier than protons p can be described by semi-superposition. It states that a nucleus of mass number A can be described by A independent subatomic particles, sharing the energy of the nucleus E_0/A equally. These subatomic particles (neutrons and protons) independently interact with the air mirroring a superposition of A cascades of protons with energies E_0/A . The only modification to this superposition model is that the cross section of the primary particle scales with its atomic number with $A^{2/3}$ shifting the first interaction statistically to a higher place in the atmosphere. By this, [16] calculates the muon number to be

$$N_{\mu,\text{max}}^A(E) \approx N_{\mu,\text{max}}^p(E) A^{1-\beta}. \quad (3.8)$$

By measuring both $N_{\mu,\text{max}}^A(E)$ and $N_{e,\text{max}}^A(E)$ [cf. equations (3.4) & (3.7)] one can determine both, the energy and the primary particle type.

3.2.2. Monte Carlo Codes

Extensive air showers typically consist of billions of particles interacting stochastically. Thus, the analytic calculation of probability distributions of shower observables is impossible. Instead, the development of shower cascades is modelled by Monte Carlo simulations.

One such Monte Carlo simulation code is the well-known program CORSIKA [17] with the hadronic interaction model QGSJETII-04 [18]. It was used for the generating simulation data for this work and complemented by the program CoREAS [19] which simulates the radio emission using the endpoint formalism (cf. section 3.3.2) within CORSIKA.

3.2.3. Muon deficit

As discussed in section 3.2.1, the muon number at ground is an important air shower observable as it can be used in combination with the electron number to determine the particle type and its energy. However, there is a not understood discrepancy between the muon number predictions by the state of the art Monte Carlo simulation programs and measurements [20]. So far it is unclear, if this hints to new hadronic interaction physics at PeV centre of mass energies which are above the energy range in reach of the Large Hadron Collider. On the other hand, the total inelastic hadronic proton-proton cross section is in good agreement with the extrapolations of measurements at the Large Hadron Collider. Thus, one has to be careful when interpreting the composition studies at the highest energies (cf. section 3.1.2)

3.3. Radio emission of extensive air-showers

The radio emission of air showers can be described on different levels of abstraction. The most straight forward and accurate way is to understand the radio emission as the superposed emission from all charged particles accelerated within an air shower. Modern Monte Carlo programs are capable of performing this calculation. In order to understand the result of simulations and measurements qualitatively, a simplified picture of the emission mechanism is more helpful. This chapter will discuss a macroscopic and microscopic picture of radio emission and compare them with experimental results.

3.3.1. Macroscopic picture

The macroscopic picture abstracts the emission of billions of charged particles to the emission of time dependent macroscopic charges. The first description of this ansatz goes back to Kahn and Lerche [21], recent ones can be found in [22] and [23].

Emission mechanisms

Today there is a broad agreement that two leading mechanisms cause a particle cascade — particularly an extensive air shower — to emit a radio signal.

The most general emission mechanism is called *Askaryan effect* or *charge excess effect* [24]. Every particle cascade ionises the medium it traverses along its path. Simultaneously it collects a negative net charge at the shower front, satisfying the conservation of charge. The lateral extent of both the ionisation channel and the plasma disk at the shower front amounts to a few meters only, corresponding to the quick decrease of the lateral distribution of charged particles with the distance to the shower axis in electromagnetic cascades [25]. The longitudinal extend of the plasma disk is of the order of few meters, too. As the number of particles in a shower changes during its development, also the net charge excess in the shower front varies. Thus, this time dependent charge emits a radially polarised radio signal depicted in figure 3.8 on the right-hand side. Although the Askaryan effect is the most general emission type, its emission only accounts for about 10 % of the total radio signal for air showers.

The dominant emission mechanism — the *geomagnetic effect* — is related to the geomagnetic field of the earth. It separates the charged particles in an air shower via the Lorentz force:

$$\vec{F}_L = q(\vec{E} + \vec{v} \times \vec{B}) \quad \text{or} \quad (3.9)$$

$$F_L = qvB \sin \alpha, \quad \text{for } E = 0 \quad (3.10)$$

The geomagnetic field of the earth is then denoted with B , and v is the velocity of the relativistic particles travelling along the shower axis with charge q . The angle α between the magnetic field and the particle velocity (roughly parallel to the shower axis) is called geomagnetic angle. But the charged particles are not forced on helix trajectories emitting synchrotron radiation, as one might guess first. Instead, they start drifting with constant velocity \vec{v}_d in opposite directions, perpendicular to the direction of the shower axis \vec{v}/v and the magnetic field \vec{B}/B due to scattering with other particles and the surrounding air. Thus, the magnetic field induces a constant current of charges in the shower front. Once again, the convolution of the fluctuation of the particle number with the drift current creates a time dependent dipole which emits the linearly polarised coherent radiation depicted in (fig. 3.8 [left]).

The strength of the geomagnetic effect in (3.10) depends on the shower geometry (i.e. geomagnetic angle α) and is proportional to the earth magnetic field at the experiment site. Therefore, the relative contribution of the Askarayan and geomagnetic effect is also dependent on the shower geometry and the location of the experiment. It can be estimated by measuring the polarisation pattern.

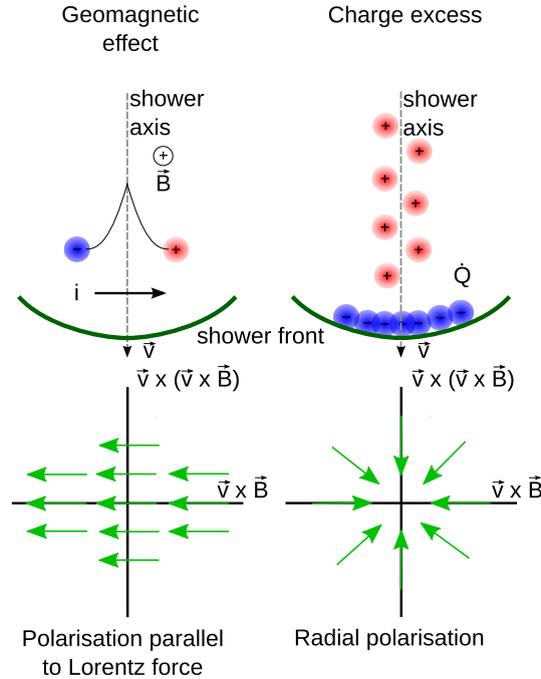


Figure 3.8.: Macroscopic emission models: The geomagnetic emission (left) produces a parallel polarisation in the shower plane, the charge excess emission (right) is radially polarised.

Coherence; length scales and time compression

The radio emission due to the two major emission mechanisms is coherent. This is an important feature, because therefore the observed radiation energy scales quadratically with the number of electrons

and thus (via eq. (3.4)) quadratically with the energy of the primary particle. As incoherent emission power scales only linearly with the cosmic ray energy, coherent emission always dominates incoherent emission above a certain energy threshold.

Radiation is coherent if the waves emitted by the individual particles do not acquire relative phase shifts exceeding $\pi/2$ on their way to the observer. For frequencies up to few MHz the coherence is caused by the length scales of the emission regions. The shower disc, which contains the negative charge excess and the geomagnetically induced electric dipole, has a thickness of the order of meters. Thus - depending under which angle one observes this disc - the radio emission is coherent up to few hundred MHz. Another effect to be taken into account, is the time compression of the radio signal, caused by the relativistic velocity of the charged particles. Since the charged relativistic particles travel with a higher velocity than the speed of light in the ambient medium, there is an specific angle under which one observes the emission from the whole shower development instantaneously. This angle is the Cherenkov angle. In its vicinity (typically at a distance of around 100 m from the shower axis for high energy air showers), the radio pulse becomes infinitesimal short (neglecting the finite source size) and thus, the upper frequency limit for coherent emission increases. The same holds for the received signal amplitude. Therefore, one can observe air showers at even higher frequencies, e.g. in the GHz regime near the Cherenkov ring on ground. Inside the Cherenkov cone the radio signal is inversely received in time w.r.t the emission time. I.e. one observes first the radiation emitted at ground and afterwards the one emitted higher in the air.

Relativistic radiation characteristics

The particles in an extensive air shower propagate with relativistic velocities. Hence, the emission of each particle can be described by the Liénard-Wiechert potential in the relativistic limit. The emission pattern of such particles is strongly forward beamed. Therefore, most of the radiation energy is emitted into a cone, defined by a small opening angle around the shower axis. The “illuminated” area on ground is sometimes called the radio footprint. The forward beamed nature of radio emission of extensive air showers helps on the one hand to detect the enhanced radio signal over the galactic radio background. On the other hand, it requires a small spacing of about 300 m between the antenna stations to detect near vertical showers coincidentally in three or more antennas [26]. The only known way to measure air showers with antennas on a large grid (e.g. 750 m) is to consider highly inclined air showers. Their radio footprint is enlarged both due to the projection and additionally due to the higher distance from the shower maximum to the observer on ground. This source-distance effect is confirmed by simulations (cf. fig. 3.10) and measurements (cf. fig. 3.15).

3.3.2. Microscopic picture

Nowadays, Monte Carlo simulation programs calculate the radio emission from air showers caused by the secondary particles in reasonable time. The simulation codes consider only the emissions of the individual particles instead of assuming macroscopic currents and particle distributions. In the following, the endpoint formalism is described as an example for a microscopic calculation of radio emission, because it is relevant for this work in particular (cf. section 3.2.2).

The endpoint formalism calculates the radio emission of each particle, separating its trajectory into segments of average velocities. At each segment border (endpoint), the particle under investigation is stopped and accelerated to its new average velocity instantaneously. The resulting electric field of this mostly destructive interference is propagated to the desired observation points and finally superimposed for all particles and all segment borders. The radio signal from the electromagnetic cascade is thus calculated from electrodynamics first principles. The calculation has no hidden parameters and the particle physics of the electromagnetic part of air showers is understood without major uncertainties. The limiting uncertainty is related to computation power. When performing thinned simulations, i.e. simulations which do not account for all particles of the cascade but instead use fewer particles with increased weights, this introduces artificial coherence. This is essential when simulating air showers initiated by high energy cosmic rays. These showers require the consideration of a vast number of leptons, which practically exceeds the commonly available computing power. Fortunately, this only affects predictions for ultra-high energy cosmic rays [4], where thinning artefacts indeed can exceed the Galactic noise floor. Nevertheless, the results of CoREAS simulations can be taken as absolute and robust benchmarks.

Figure 3.9 shows the simulation of the electric field amplitude from a vertical cosmic ray shower with CoREAS. The shower was initiated by an iron cosmic ray with 10^{17} eV energy. The asymmetry in the electric field is created by the superposition of the geomagnetic effect and the sub-leading charge excess effect. This is in good qualitative agreement with the macroscopic picture discussed above. The agreement of simulation, macroscopic models and measured data will be discussed in section 3.3.3. The good understanding of the radio emission and the radio measurements offers the solid basis to deduce cosmic ray properties from measured radio pulses.

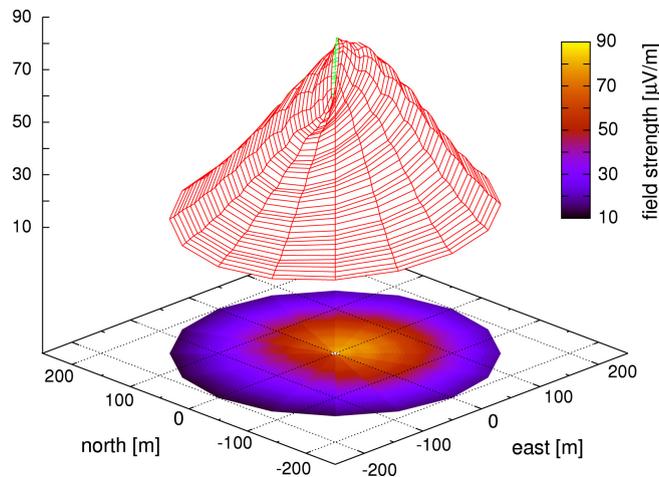


Figure 3.9.: Simulation of the electric field amplitude in the 40-80 MHz band from a vertical cosmic ray shower. The asymmetric shape is due to the different polarisation characteristics of the geomagnetic and charge excess emission [27].

Figure 3.10 confirms the source-distance effects, discussed at the end of section 3.3.1. It shows the simulated radiation energy for air showers of energy of 5×10^{18} eV at different zenith angles. This simulation predicts the measurability of radio signals which illuminate areas of tens of square kilometres for zenith angles of around 75° . Radio signals whose electric field strength exceeds $1\text{-}2 \mu\text{V m}^{-1} \text{MHz}^{-1}$ can be detected above the galactic noise floor with the radio technique [26]. Not only

the sheer size of the displayed footprint is impressive. Highly inclined showers reach the ground at a late stage of their development. Hence, mostly only muons arrive at ground and can be measured by particle detectors. The radio technique offers the possibility to measure the energy within the electromagnetic cascade and therefore the electron number. For this reason, horizontal air showers are well suited to perform composition studies in a combined analysis of data simultaneously recorded with a radio and particle detector.

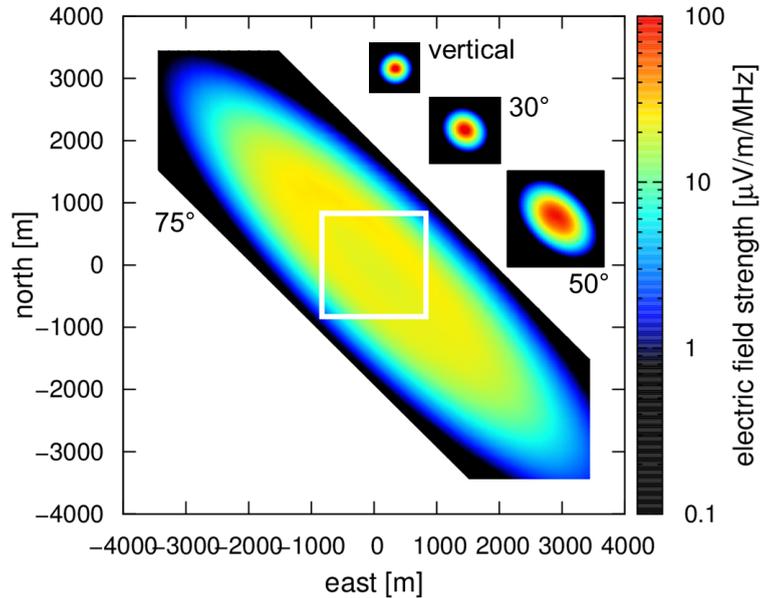


Figure 3.10.: The radio footprint of the same air shower for different zenith angles. The frequency range considered in the simulation is 30 MHz to 80 MHz and the air shower energy amounts to 5×10^{18} eV. Electric field strength of one to two $\mu\text{V m}^{-1} \text{MHz}^{-1}$ typically exceed the galactic noise and can be detected with radio experiments [26].

3.3.3. Measurements and their interpretation

The radio signal from extensive air showers is coherent for frequencies from the MHz regime up to a few GHz. Fortunately, the earth's atmosphere is transparent for electromagnetic waves in this frequency band. Thus, there exist experiments for both MHz and GHz radiation.

Besides AERA, the Low Frequency Array (LOFAR) is an example for a MHz radio detector. The high density of the LOFAR antennas allows precise studies of emission mechanisms and shower front analyses (cf. fig. 3.11 & 3.12). The rather sparse array of AERA makes observations with this high resolution impossible. The measurement and simulation in figure 3.11 coincide nicely.

A study on AERA data [30] confirmed the Monte Carlo calculations as well as the macroscopic models more generally. It determines for the first time the mean ratio a of the charge excess to the geomagnetic effect to be 14% for the site of the Pierre Auger Observatory. Figure 3.13 shows the azimuthal polarisation angle of the measured electric fields on ground $\phi_p(\text{me.})$ and their prediction by simulations $\phi_p(\text{pr.})$ for different values of a . The lowest reduced χ^2 and the highest Pearson correlation coefficient, respectively, are found for a relative contribution of $a = 14\%$ of the charge excess (right panel). The assumption of a purely geomagnetic emission ($a = 0$) is disfavoured by a roughly ten times larger reduced χ^2 and a lower Pearson correlation coefficient (left panel). Every mark in this figure

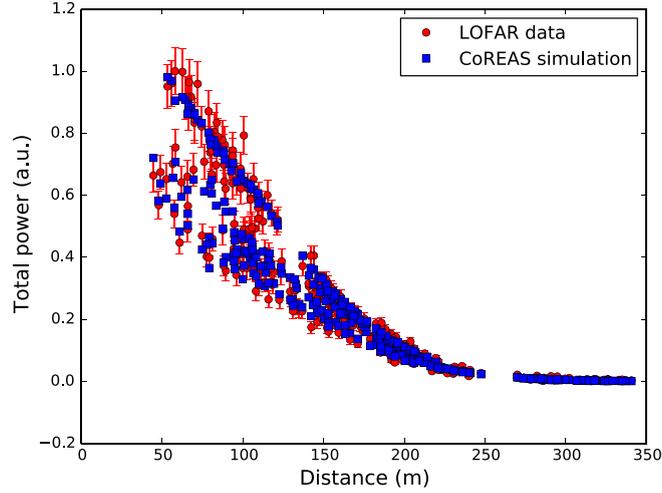
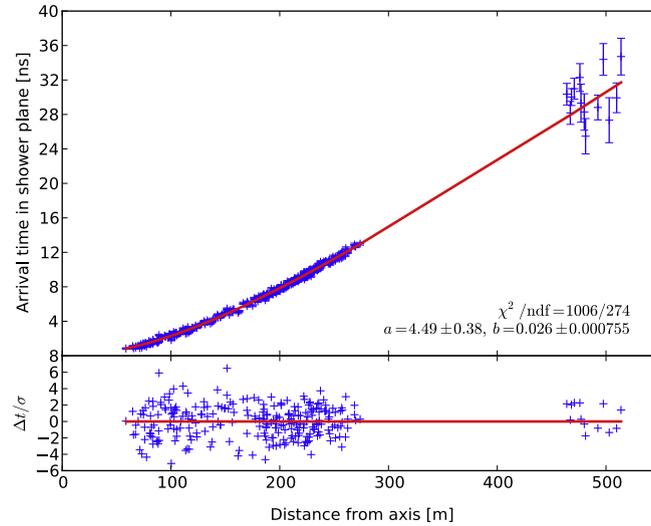


Figure 3.11.: Comparison of an example event measured by LOFAR and a corresponding CoREAS simulation. The measured spread in the total power reflects asymmetries in the shower plane due to the superposition of charge excess and geomagnetic emission mechanism. Experimental data and simulation agree well, even for this unprecedented experimental resolution [28].



(a) Hyperbolic fit

Figure 3.12.: The arrival time distribution of one example event measured by LOFAR as a function of the distance to the shower axis. A hyperbola describes the data best [29]

corresponds to a measured event by a particular radio antenna and 17 measured events contribute to this study. Still, the statistic for this analysis is quite low.

Another feature of the radio emission is its forward beamed nature. Due to the time compression of the signal, one can observe coherent radiation in the GHz regime, provided the detector is located in the vicinity of the Cherenkov ring. The measurements of the CROME (Cosmic-Ray Observation via Microwave Emission) experiment [31] confirm this in good agreement with predictions by CoREAS (cf. figure 3.14) For all 31 considered events, the positions of the antennas which passed the quality cuts are plotted relative to the shower core. In addition, this figure shows the simulated signal strength in the measured frequency bandwidth at ground for a typical air shower with a maximum depth $X_{\max} = 658 \text{ g cm}^{-2}$. Taking into account a polarisation and viewing angle study, the authors of [31]

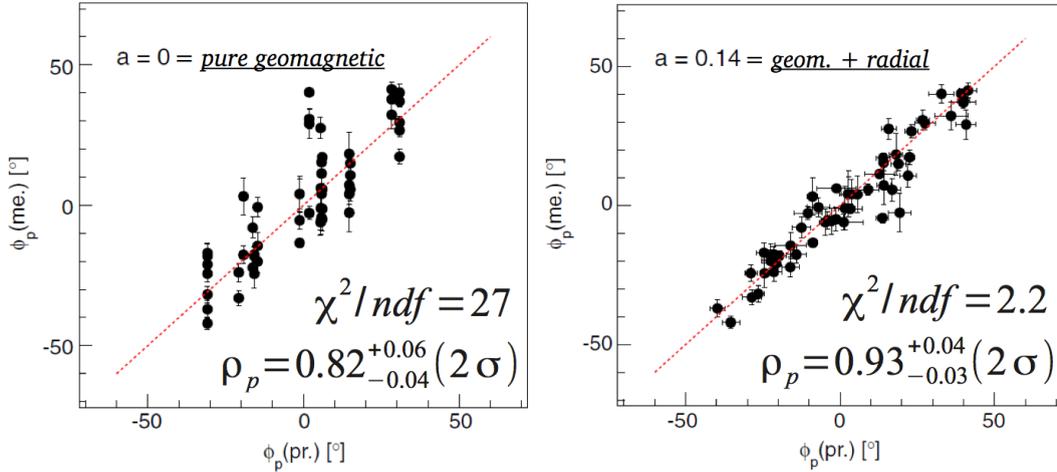


Figure 3.13.: The measured versus the predicted polarisation angle for pure geometric emission (left) and a contribution of 14 % from the Askaryan effect (right). The fit of a correlation is optimal for the latter case [30].

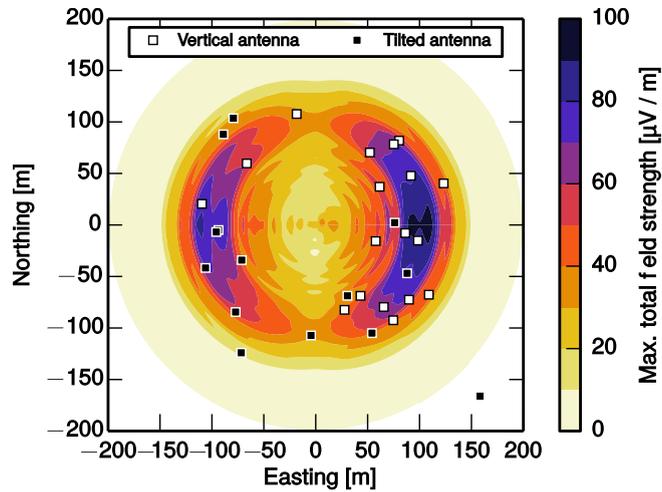


Figure 3.14.: Comparison of the radio signal in the GHz regime measured with CROME against CoREAS simulations. The rectangles denote the positions relative to the shower core (0, 0) m at which microwave signals have been detected. The colour contours indicate the maximum total field strength prediction at ground level for a typical vertical air shower ($X_{\max} = 658 \text{ g cm}^{-2}$) [31]

draw three conclusions: a) the measurement is compatible with the signature of GHz radiation which is collimated in the cone about the Cherenkov angle; b) looking at the vertical events (white marks) in the printed figure, the pure number of measurements is in agreement with the asymmetry due to the superposition of charge excess and geomagnetic effect; c) the measurement — especially the polarisation study not shown here — disfavors the presumed isotropic emission due to molecular bremsstrahlung.

A last case is the measurement of horizontal air showers which have extremely large radio footprints on ground. A direct measurement of such an event has been performed with AERA and is shown in figure 3.15. The reconstructed energy of this air shower is about 13 EeV and the zenith angle of the shower axis is 74° . The lateral distribution function has been sampled with 69 radio antennas to a distance of 1200 m from the shower axis. This translates to a measurable footprint on ground with a major axis of roughly 8.7 km.

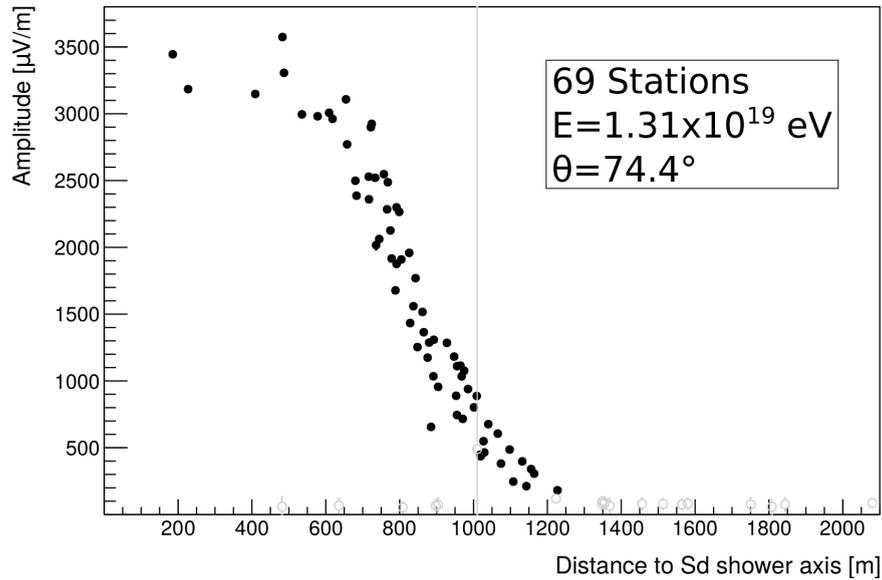


Figure 3.15.: The LDF of an air shower with a zenith angle of about 74° and an energy of about 13 EeV measured with AERA. 69 radio stations have seen a signal, even at distances of up to 1.2 km from the shower axis [32].

Finally, it is worth to mention the existence of radio experiments for the observation of transition radiation and radiation in dense media. One example is the balloon experiment ANITA [33], which searches for neutrinos skimming the arctic ice and thus producing showers. The last example is the T-510 experiment [34], which tested and confirmed the emission models for air showers in the laboratory.

The Pierre Auger Observatory

The Pierre Auger Observatory [35] is the world's largest experimental setup for the detection of ultra-high energy cosmic rays. Its construction has been motivated by a considerable controversy about the existence of a suppression of the cosmic ray flux [36] at the highest energies which for instance is expected due to GZK suppression. Therefore, the Pierre Auger Observatory is designed to measure the flux, mass composition and arrival direction distribution from 10^{18} eV to the highest energies with high statistical significance over the whole southern sky. The observatory covers an area of 3000 km^2 and combines several instruments for a precise measurement and the possibility of cross-calibration to a hybrid setup. The baseline detectors are the surface detector (SD) and the fluorescence detector (FD),

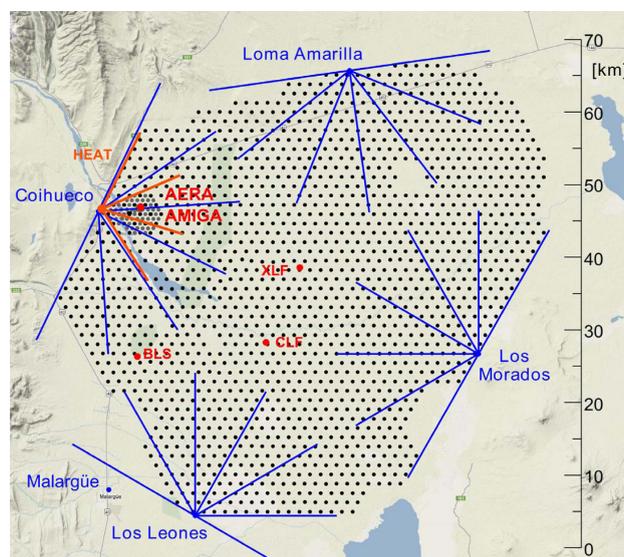


Figure 4.1.: Layout of the Pierre Auger Observatory: the black dots show the locations of the water-Cherenkov tanks, the blue lines indicate the field of view of the fluorescence telescopes at the sites Los Leones, Los Morados, Loma Amarilla and Coihueco.

which provide complementary information on the air showers. The surface detector measures the lateral distribution of the air showers while the fluorescence detector measures the longitudinal shower development. At clear and dark nights, all events above 10^{19} eV are detected with SD and FD in coincidence. The surface array features a uniform coverage in right ascension on a daily basis [37]. Figure 4.1 shows a map of the Pierre Auger Observatory with its baseline detectors, which is located on a high plain in western Argentina. Today, the Pierre Auger Observatory also operates the radio detector AERA and the extensions HEAT and Amiga. However, in the following I will discuss only the detector systems directly affecting this work. A short outline of the analysis methods applied to the recorded data can be found in section 6.1.

4.1. The detectors

4.1.1. The Surface Detector (SD)

The surface detector consists of 1660 water-Cherenkov detectors covering an area of 3000 km^2 . 1600 detectors are deployed on a 1.5 km triangular grid. This configuration reaches its full efficiency for air showers with energies above 3 EeV . In the vicinity of AERA, the spacing is reduced to 750 m with 60 additional detectors in order to lower the detection threshold to 3×10^{17} eV. This low energy extension of the surface detector (the “infill array”) is discussed in section 4.1.3.

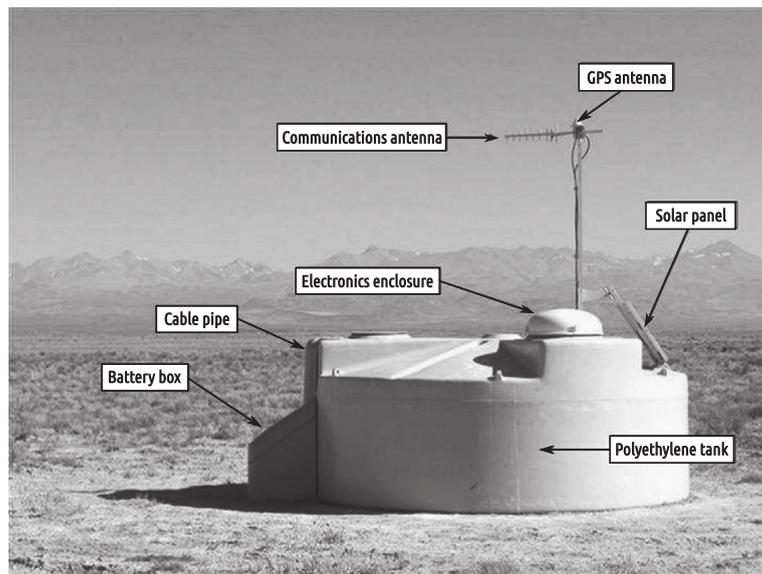


Figure 4.2.: The main components of a water-Cherenkov detector which has been deployed at the Pierre Auger Observatory [35].

Each water tank is filled with 12 m^3 of purified water. Superluminal charged particles emit ultraviolet Cherenkov light in water. This light is detected by three photo-multipliers after the reflection on the tank liner. Even photons can be detected due to the sufficient height of 1.2 m, which allows for electron-positron pair production and subsequent detection. Thus, the surface detector measures a slice of the air shower at ground.

All tanks are self-contained, i.e. they operate autonomously. The electric power needed by the

electronics is provided by a solar-panel in combination with a battery. The data is transmitted via a custom-made wireless communication system to the central data acquisition system (CDAS). A time-stamp of 10 ns accuracy is derived from a GPS receiver (Motorola Oncore MU+) and associated to every event [35] on the perimeter of the array. A schematic view of a water-Cherenkov detector is shown in figure 4.2.

4.1.2. The Fluorescence Detector (FD)

The Pierre Auger Observatory operates 24 fluorescence telescopes which overview the surface detector from the four sites “Loma Amarilla”, “Coihueco”, “Los Leones” and “Los Morados” (cf. fig. 4.1). Air shower excite nitrogen molecules in the atmosphere which in turn emit ultraviolet light isotropically when relaxing to the ground state. The detectors follow the design of a Schmidt telescope (cf. fig. 4.3): The light enters through an aperture system of 2.2 m diameter. It passes an UV transparent filter and is focused by a segmented spherical mirror focuses the light on the camera. The pixels of the camera consist of photo-multiplier tubes which are arranged in a matrix of 20 columns and 22 rows. Each telescope has a field of view of about $30^\circ \times 30^\circ$ in the zenith and azimuth. Thus, a building with six telescopes covers a view of 180° in azimuth. Three additional telescopes are installed in special enclosures at Coihueco. These enclosures as a whole can be vertically tilted by 29° . As a consequence, air showers from less energetic primary particles get in the field of view. That is mainly because the shower maximum is logarithmically shifted upwards for decreasing energies (cf. eq. (3.5)). Thus, these telescopes lower the hybrid detection threshold to 10^{17} eV. The extension is called the High Elevation Auger Telescopes (HEAT).

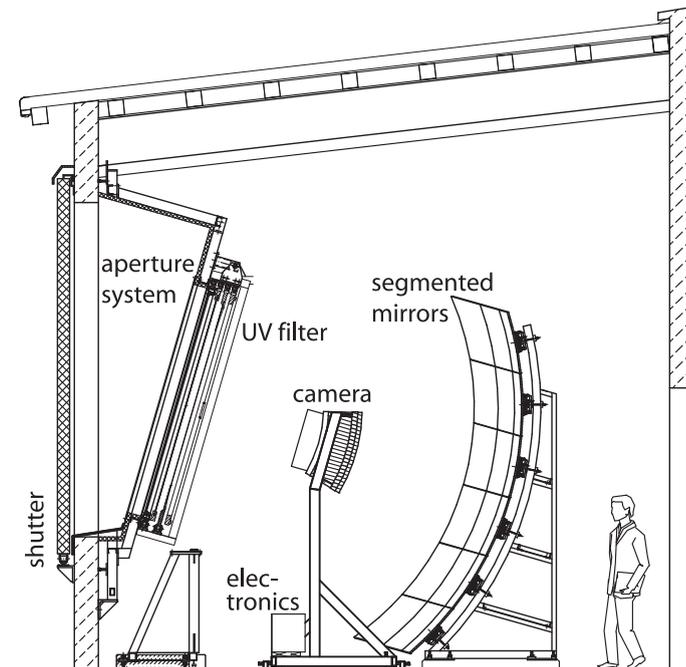


Figure 4.3.: Sketch of a fluorescence telescope at the Pierre Auger Observatory. Each fluorescence building typically hosts six telescopes. [35].

The intensity of the light emitted along the longitudinal air shower development is proportional to the number of charged particles mostly in the electromagnetic part of the cascade. Hence, the fluorescence detector provides a near-calorimetric measurement of the primary particle energy. One major systematic uncertainty is the fluorescence yield, i.e. the proportionality factor between the emitted light and the energy deposited in the atmosphere. This yield depends on the actual atmospheric conditions which are therefore carefully monitored at the Pierre Auger Observatory. Neutrinos and high energy muons can carry a significant part ($O(10\%)$) of the energy of the air shower without stimulating light emission in the atmosphere. The only way to correct for that “invisible energy” is to estimate it with Monte Carlo air shower simulations. The cosmic ray energy, determined by the fluorescence detector, has an absolute uncertainty of $\sim 14\%$ and a statistical uncertainty of 7.6% . The fluorescence detector is used to calibrate the energy reconstruction of the surface detector.

Additionally, the fluorescence telescopes directly observe the depth of the shower maximum X_{\max} which is the most sensitive mass estimator. The statistical resolution of X_{\max} is 20 g cm^{-2} .

The FD measurement suffers the most from the limited duty cycle of 15% . Measurements are only feasible in clear nights when the moon is not in the field of view.

4.1.3. The Auger Muon and Infill Ground Array (AMIGA)

The sub-detector system AMIGA is an enhancement to the surface detector. Within the regular surface detector array, a denser array of water-Cherenkov detectors with 750 m spacing has already been build near the fluorescence telescope site Coihueco. It spans an area of 23.5 km^2 . Additionally, it is planned to bury a 10 m^2 scintillator next to each WCD in a depth of 1.3 m . Figure 4.4 shows an AMIGA

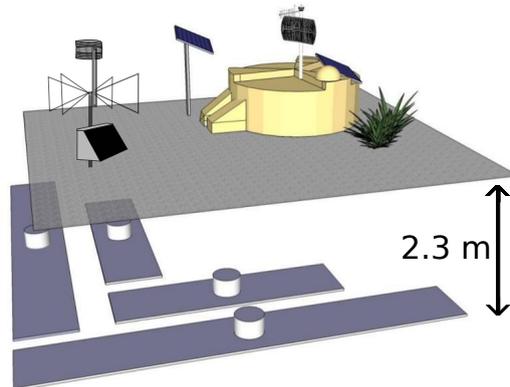


Figure 4.4.: Sketch of a prototype AMIGA station and an AERA station. The muon detectors of regular AMIGA stations will be buried 2.3 m underground [38].

prototype station. Its scintillators have been buried at a greater depth (2.25 m) and the area they cover is bigger (30 m^2) compared to the baseline concept. The infill array is fully efficient for air showers with energies above $3 \times 10^{17} \text{ eV}$ and zenith angles below 55° .

4.1.4. Upgrade of the Pierre Auger Observatory

As already discussed in chapter 3.1.2, a precise determination of the primary particle type at the highest energies is needed to identify sources. The Pierre Auger Collaboration addresses this challenge and

has started a major upgrade of its instruments known as “AugerPrime” [39].

The concrete goals of the upgrade are:

- Measurement of the mass composition at the highest energies and identification of the reasons for the flux suppression. Up to know, it is not clear, if the suppression at the highest energies is due to the GZK effect or due to the acceleration limit inherent to extra-galactic cosmic ray sources (cf. section 3.1.1).
- To be sensitive to a proton fraction of 10 %. This will clarify the prospects of proton astronomy for the future.
- The study of the multi-particle production in extensive air showers. This is related to the search of Lorentz-invariance violations and new physics at energies not in reach with accelerator experiments.

In order to meet these aims, the Pierre Auger Collaboration plans the following upgrade steps:

- The most challenging effort is the upgrade of the water-Cherenkov detectors with new electronics and a surface scintillator detector (SSD). The combined measurement with the water-Cherenkov detectors surface scintillator detectors of the upgraded surface detector will allow to determine the electron to muon ratio and thus to discriminate the primary particle type. The new electronics will be capable of processing the data of both sub-detectors (WCD and SSD). Additionally, it will provide enough computational power for running improved trigger algorithms and monitoring.
- The SD infill area will be equipped with an underground muon detector in order to independently cross check and tune WCD and SSD measurements. The completed AMIGA detector will serve for this.
- The data taking period of the fluorescence detector will be extended to periods with a higher night sky background. The duty cycle of FD is expected to increase by 50 %.

The upgrade is scheduled for completion in two years, and the operation of the upgraded Pierre Auger Observatory will be extended until the end of the year 2024.

4.2. Radio detection with the Auger Engineering Radio Array

The Pierre Auger Collaboration explores the radio technique as a complementary detection method for air showers with a rich R&D program. Four different radio sub-detectors exist at the observatory. Three rather small setups are dedicated to the radio detection at GHz frequencies (AMBER, MIDAS and EASIER). The fourth one — the Auger Engineering Radio Array (AERA) – is the world’s largest antenna array for air shower detection. The radio detector covers with 153 autonomous antenna stations an area of about 17 km² and operates in the frequency band of 30 to 80 MHz. The unique experimental setup of the observatory provides ideal conditions for radio measurements. The co-located surface detector (SD) and the nearby fluorescence telescopes at Coihueco allow for common analyses as well

as for external triggering, which is one operation mode of AERA.

4.2.1. Scientific goals of AERA

The scientific goals of the AERA project are the following:

- The radio emission from air showers is to be calibrated including sub-dominant emission techniques.
- The capabilities of the radio technique is to be demonstrated by measuring the characteristic air shower properties (energy, arrival direction and composition) with the precision needed for a physical application.
- The AERA project aims to elucidate the transition from galactic to extra galactic cosmic rays by performing composition measurements in the energy range from 0.3 to 5 EeV.
- The radio array is to be used to research and develop the concepts for next generation cost effective large radio detectors.

Today, the radio emission mechanism is well understood, also due to results from AERA (cf. fig. 3.3). The energy and arrival direction of air showers are routinely determined by radio measurements. The precision achieved with the radio technique is meanwhile comparable to the conventional ones. Several trigger modes and antenna types have been tested with AERA as part of R&D activities. The direct comparison of observables determined with different instruments of the observatory allows a calibration of the detection techniques against each other. Open questions are to what precision one can determine the mass composition of cosmic rays with radio and if the energy spectrum measured by AERA is in agreement with the one from FD and SD.

4.2.2. Deployment stages of AERA

AERA has been deployed in three phases:

- The dense array of 24 stations (AERA-I) has been build in the autumn of 2010. The radio stations of this phase are deployed on a regular grid with 150 m spacing. Simulation studies [40] expect an energy threshold of 2×10^{17} eV for this dense array.
- The second stage (AERA-II) has been constructed in 2014 and extends the array by 100 additional antennas. Due to the larger of 250 m to 375 m, the energy threshold increases to 10^{18} eV [40].
- In March 2015 the radio detector has been completed with the deployment of AERA-III. Today, AERA consists of 153 antennas. AERA-III exhibits the largest antenna spacing of 750 m. The even larger spacing increases the statistics of high energy air showers measurements and the covered area allows to sample the large footprints of horizontal air showers for the first time (cf. fig 3.15).

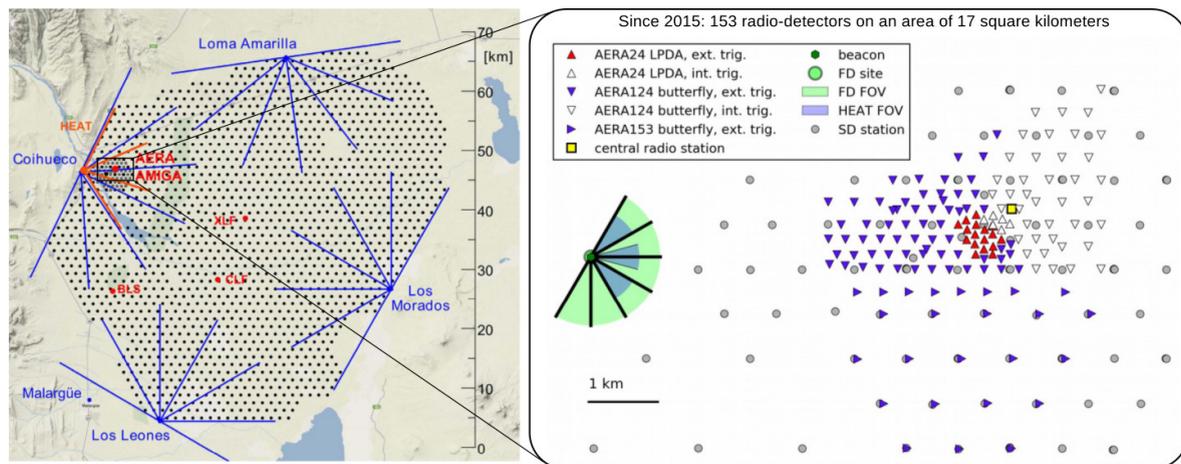


Figure 4.5.: The positions of the AERA radio stations at the Pierre Auger Observatory.

Figure 4.5 shows the positions of the AERA radio stations at the Pierre Auger Observatory and their mutual distances. The stations are additionally distinguished on their ability for being externally triggered by the surface and fluorescence detector. It takes a few seconds for the central data acquisition system of the observatory to identify cosmic ray events and to distribute their time-stamps to the radio stations. Only a part of the radio stations feature front-end electronics with a deep enough memory to buffer the measured data for the required time.

4.2.3. Properties of a single AERA station

Components of the radio stations

The AERA stations operate autonomously in the remote and rural environment of the Argentinean Pampa. The climatic conditions are extreme in this area, e.g. the temperature differences between day and night can exceed 20° Celsius. The strong ultraviolet radiation at the altitude of 1500 m leads to faster fatigue of materials, especially of plastics. Each journey to an individual antenna is time consuming and laborious. Therefore, the need for maintenance on site has to be reduced by all means. The design of the AERA stations (depicted in figure 4.6) meets the correspondingly high requirements on mechanics and electronics:

The heart of every AERA station is the radio antenna, i.e. the antenna which receives the radio signal emitted by extensive air showers. For AERA-I stations a log-periodic dipole antenna (LPDA) is used and for the newer stations a bowtie (Butterfly) antenna has been chosen. The electronics, battery and charge control unit of each detector is protected by a sand- and rainproof RFI-tight Faraday cage. The solar module provides the power for the electrical components of the detector, i.e. the front-end electronics, the low noise pre-amplifier (LNA), the communications system and the charge controller itself. The power supply system has been optimised from phase one to the phases two and three. Phase-I stations are equipped with solar panels of 240 W_p peak power and two batteries a 165 Wh act as buffers for nights and overcast periods. The phase II and III stations use only one solar panel providing 140 W_p and one battery with a capacity of 170 Wh. One major conceptual change between AERA-I and II concerns the communication. AERA-I stations are connected via optical fibres to a

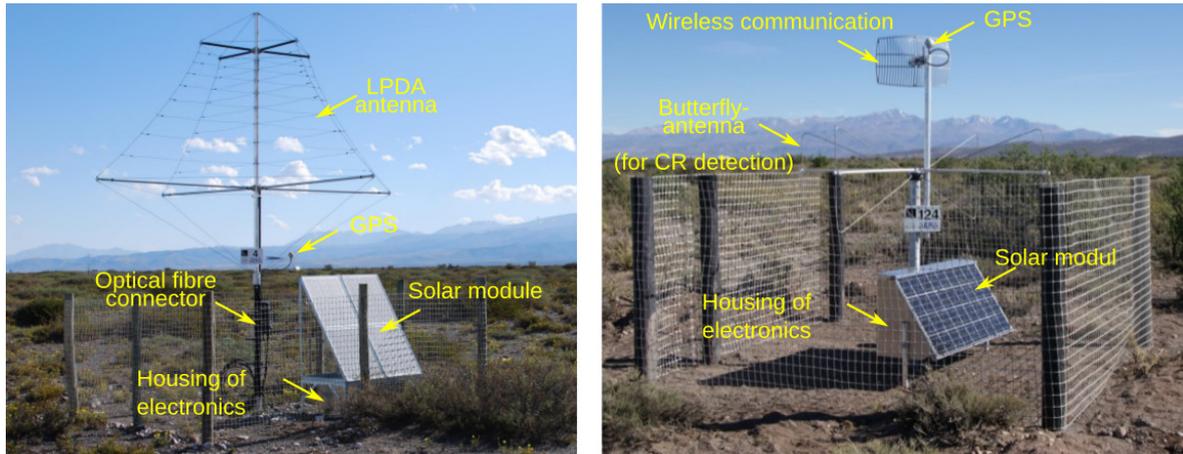


Figure 4.6.: The components of an AERA station. The design has been improved from phase I (left) to the phases II/III (right).

central relay whereas phase-II stations feature a commercial wireless communication system. This is an important development step towards large scale radio detection. The differences between AERA phase II and III stations affect only the front-end electronics, which are in phase III are capable of recording 4 individual channels simultaneously instead of 2 channels with a high and a low gain for phase I and II stations.

The analogue signal chain

The emitted radio signal is received by either LPDA or butterfly antennas. Figure 4.7 shows the antenna pattern and the frequency spectrum for both antenna types. The H-plane is vertical to the antenna arm. The E-plane is perpendicular to the H-plane and parallel to the antenna arm. Thus, the gain vanishes per definition for the horizontal direction in the E-Plane. LPDA antennas are barely sensitive

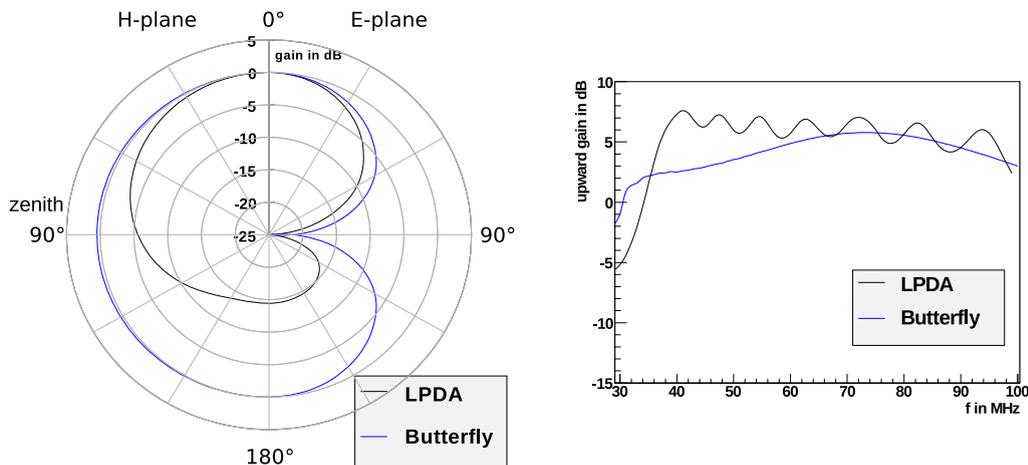


Figure 4.7.: The radiation pattern of the LPDA and butterfly antenna at 60 MHz in free space depending on the zenith angle. The gain is normalised to the vertical sensitivity (left). The frequency response for the upward direction. The gain of the butterfly has been shifted by -30 dB (right) [41].

to ground conditions. The spectrum of the LPDA antenna is rather flat and is not affected to signals reflected from ground. These good electric properties are however accompanied by a high complexity,

a limited suitability for mass production and quite high costs. The butterfly antenna is robust, simple and cheap. Its frequency spectrum is not as flat as the one of an LPDA. The butterfly antenna was selected for AERA phase II and III. Its sensitivity to the ground introduces an additional uncertainty to analyses. The performance of the butterfly antenna is nevertheless a good compromise for a precise radio detection of cosmic rays.

Radio pulses which are induced by cosmic rays are short in time and possess a broad frequency spectrum. Therefore, one aims for a wide band detection since the signal energy scales linearly with the bandwidth, while white noise scales only as the square root of the bandwidth. Unfortunately, there are large anthropogenic and atmospheric background contributions in the short wave band below 30 MHz and the FM band above 80 MHz. Radio detection in the MHz regime is thus limited to the frequencies in between. For the best signal quality, the received signal is first amplified by a low noise amplifier directly at the antenna. A subsequent filter amplifier suppresses frequencies outside the desired bandwidth. Figure 4.8 shows exemplary the transmission curves of all produced phase-III filter amplifiers. The older phase-I and II filter amplifiers have only two channels as input. There, each channel is split and either amplified by 20 dB in the transmission range or not. Hence, AERA-I and II stations record only two channels but each with a high and a low gain for a higher dynamic range.

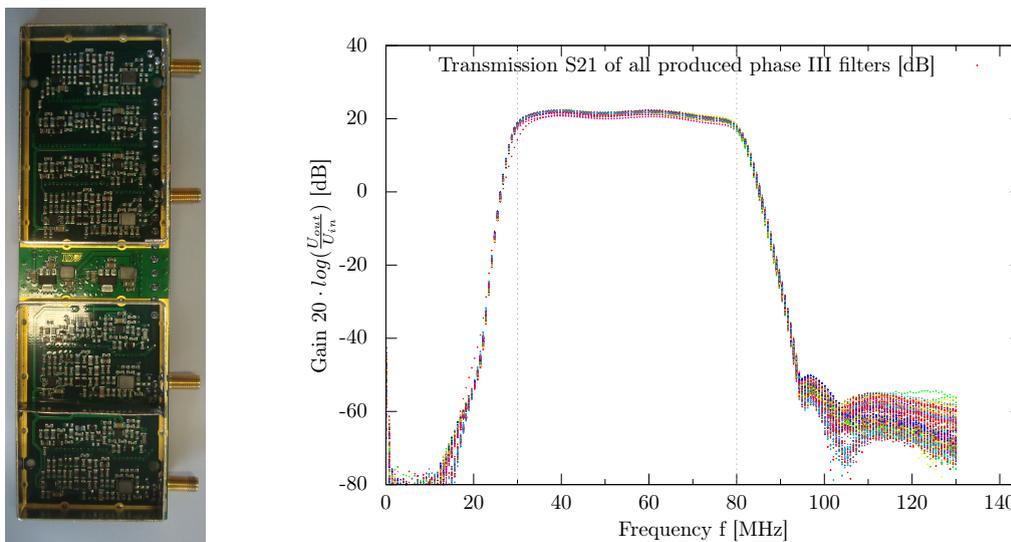


Figure 4.8.: The phase III analogue filter amplifier provide for 4 input channels an amplification of 20 dB in the transmission range (30 MHz to 80 MHz) and an attenuation of -60 dB outside.

The front-end electronics

The front-end electronics (cf. fig. 4.9) samples the analogue data, generates an hardware trigger based on the data stream and communicates with the central data acquisition system. For this purpose four 12 bit ADCs (Analog Devices AD9626) sample the analogue signal received from the filter amplifier at 180 MHz. The resulting digital stream is online processed by an FPGA. The signals are buffered in SODIMM memories of 4 GB size, large enough to store the data of two channels for roughly 7 s. A micro-controller monitors the board temperatures, supply voltages and exchanges status information with the charge controller. All communication media (optical fibre or commercial wireless) are connected to the digital board via an Ethernet interface. For AERA-II and III stations, the filter

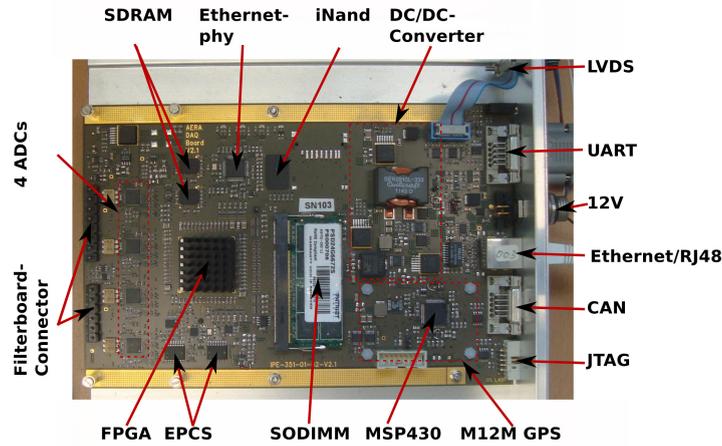


Figure 4.9.: AERA-II front-end electronics

amplifier is directly mounted and interfaced to the motherboard via a pin header. The LNAs are also supplied with 6 V phantom voltage by the front-end electronics via the filter amplifier and the signal cables. For AERA-I stations, the filter is a separate unit which is supplied with 6 V externally.

Firmware

All data acquisition tasks which are critical in terms of time are performed by logic directly implemented in the FPGA. Uncritical tasks as for example the communication with the central data acquisition system are performed by a local data acquisition program which is written in the high level programming language C. It runs on the embedded NIOS CPU in the context of a μ Clinux operating system. Figure 4.10 schematically shows the operating mode of the FPGA. The main modules are the data

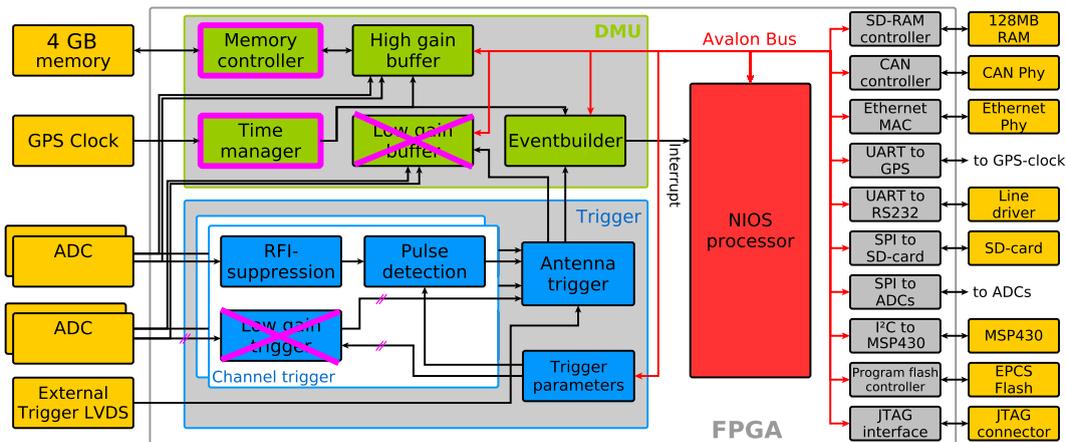


Figure 4.10.: Block diagram of the AERA FPGA firmware. The adaptations for read-out of 4 independent channels in the DMU, trigger and input part are magenta. Adapted from [41]

management unit (DMU), the trigger logic and the central processor unit. The DMU and trigger logic are purpose-built modules optimised for radio detection. Nevertheless, the system can be used to record any electrical signal which matches the input voltage range of the ADCs. In this sense, only the exchangeable analogue part restricts the front-end electronics to radio detection. The central processor

is connected to these modules via an Avalon bus and thus, they are visible from the μ Clinux operating system. The eventbuilder and GPS clock can initiate interrupts to start special routines of the local acquisition program.

All modules marked magenta in figure 4.10 have been firmware adapted for the recording of four channels.

4.2.4. The data acquisition

The data acquisition of AERA is split into two parts:

Local data acquisition program

If the trigger logic of the station identifies a cosmic ray candidate in the continuously examined radio data, the local data acquisition program sends the trigger time to the central AERA data acquisition (ADA). On request, the local station can send data from its 4 GB deep buffer to the AERA data acquisition. Besides reading out externally triggered data, the AERA data acquisition system uses this mechanism to request every 100 s background data. Additionally, all local radio stations send periodically monitoring data (e.g. different supply voltages, momentary trigger rates, etc.) to the AERA data acquisition.

AERA data acquisition

The central data acquisition of AERA runs on servers which are located in the fluorescence detector building Coihueco (cf. fig. 4.5). The task of this acquisition system is to identify coincident station triggers, to reject noise events by their arrival time distribution at the radio detectors and to request the readout of event data. The system also accepts triggers from the central data acquisition system of the observatory (CDAS) which is responsible for the whole Auger Observatory. If either the surface detector, the Infill Ground Array of AMIGA or the fluorescence telescopes located at Coihueco register a cosmic ray event, the AERA data acquisition system requests the read-out of the radio detectors which are capable of being externally triggered. Furthermore, the AERA data acquisition system can control the operation modes of the radio detectors. It can, for instance, reboot or set the detectors into a dedicated power save mode.

The resulting data is stored as binary data in a first step. Subsequently it is converted to the ROOT data format and merged with the data from the surface detector, fluorescence telescopes and AMIGA. The data analysis presented in this work uses only externally triggered and merged data.

4.2.5. Time corrections via the beacon setup

A precise relative timing among the individual radio detectors is crucial for the deduction of cosmic ray properties from the arrival time distribution of the radio signal. For radio interferometry or the determination of the shower maximum from the arrival times a precision of about 1 ns is needed. The manufacturer of the “M12M Timing Oncore” GPS receivers used in most of the AERA front-end electronics claims an accuracy of 2 ns. Absolute measurements find at least a precision of 3 ns [42]

which is compatible with the findings in this work. Unfortunately, a high precision on station level does not necessary lead to a high relative precision between individual detectors. For this purpose, a high time accuracy or at least a high precision with a stable offset to the true absolute time is needed. In order to analyse and correct for relative drifts of the individual clocks a beacon analysis has been performed [43]. For this, a beacon transmitter is mounted at the communication tower at Coihueco. It emits four sine waves with the frequencies 58.887 MHz, 61.523 MHz, 68.555 MHz and 71.191 MHz which are recorded by the AERA radio detectors. The superposition of these waves is a characteristic beacon beat which is recorded by each radio detector with a periodicity of about 1100 ns. The beacon

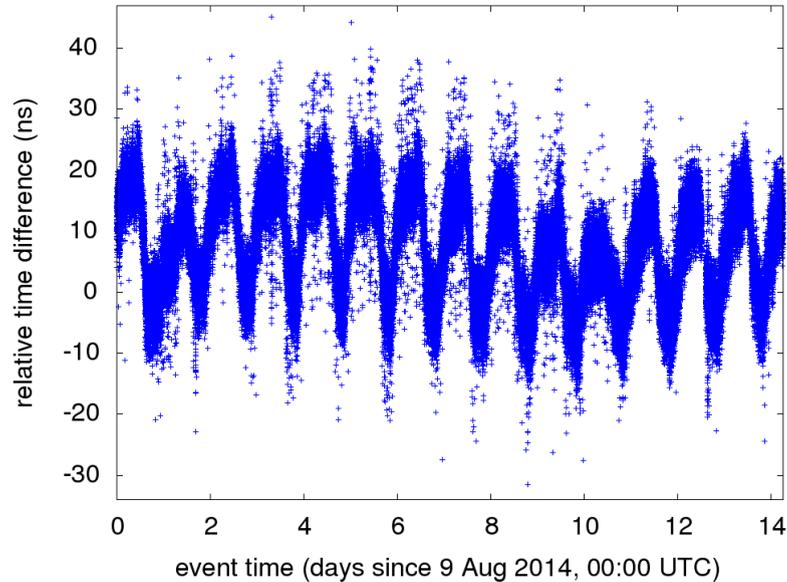


Figure 4.11.: Example of the relative timing between two AERA stations determined with the beacon method [43]

beat is expected to arrive at each radio station with a constant offset t_{geo} according to the distance between the emitter and the station. Ideally, the cross-correlated signal from all stations i would be maximal after shifting their signal in the time domain by $t_{\text{geo},i}$. In reality, one needs to shift each signal trace by an additional not constant factor $t_{\text{jitter},i}$ in order to reach the maximal cross-correlation. This additional shift can be interpreted as the convolution of the GPS precision with an additional GPS clock jitter. Figure 4.11 shows the relative time difference $t_{\text{jitter},j} - t_{\text{jitter},i}$ between two typical stations (j,i) over a time period of 14 days. The short term stability of this difference is of the order of a few nanoseconds corresponding to the high precision of the GPS receivers. But the absolute difference, which depends on the accuracy of the individual time measurements, deviates by a several tens of nanoseconds. The jitter depicted in figure 4.11 drifts on a daily basis.

4.2.6. Results and future

Some results from the radio measurements with AERA have already been shown in section 3.3. There, the measurement of the relative contribution of the Askarayan and geomagnetic effect is presented. Also, the first measured radio footprint with an extent of several kilometres on ground is discussed there.

Universal energy estimation

The definition of a universal energy estimator for radio measurements is an important contribution to astroparticle physics: It is defined by the integral over the energy fluence $f(r)$ (cf. equation (6.12)) divided by $\sin^2(\alpha)$. The division eliminates the dependence of the geomagnetic on the geomagnetic angle (cf. equation (3.10)). Figure 4.12 shows the resulting energy estimator. It shows the cosmic ray energy measured by the Pierre Auger surface detector versus the corrected radiation energy in the 30 MHz to 80 MHz frequency band. The power-law fit relates the radiation energy to the cosmic ray

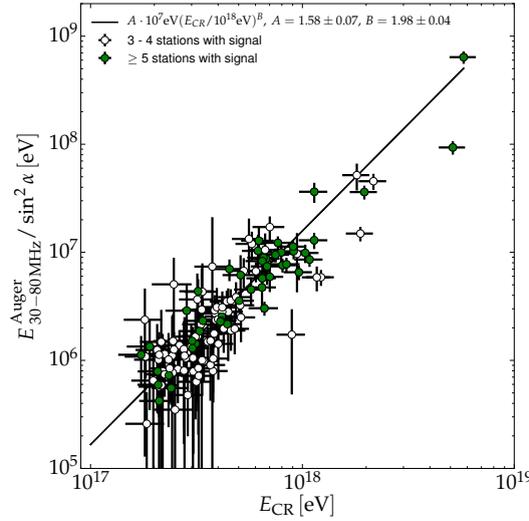


Figure 4.12.: Estimator of the radiation energy in the 30 to 80 MHz band versus cosmic ray energy determined by the surface array. A power-law fit confirms the quadratic scaling, the different colours distinguish events with different number of radio stations [44].

energy via the energy scale of the Pierre Auger Observatory:

$$E_{30-80 \text{ MHz}} = [15.8 \pm 0.7 (\text{stat}) \pm 6.7 (\text{syst})] \text{ MeV} \times \left(\sin \alpha \frac{E_{\text{CR}}}{10^{18} \text{ eV}} \frac{B_{\text{Earth}}}{0.24 \text{ G}} \right)^2. \quad (4.1)$$

The quadratic scaling of the corrected radiation energy with the cosmic ray energy as expected from coherent emission is verified by the fit. Typical energy estimator for particle detector use the fitted signal at an optimal distance (cf. section 6.1). These estimators depend on various parameters, like the detector spacing and the altitude of the experimental site. Thus, these energy estimators are specific for every experiment and the results are difficult to compare. Whereas the energy estimator in equation (4.1) only requires that the bulk of radio emission has already been emitted before the shower reaches the ground.

The method yields a combined uncertainty of AERA and SD of 24 % for events with more than four stations with a signal above the background noise.

Mass composition of cosmic rays

The AERA group investigates several analyses to determine the primary particle mass. The atmospheric depth X_{\max} and the electron-to-muon ratio (cf. section 3.2.1 & 3.3.2) are commonly used to deduce the primary mass. Until now, the best results are obtained by comparing the measured data to with a library of radio footprints generated by Monte Carlo simulations. Then atmospheric depth X_{\max} is taken from the simulation which describes the measurement best. The comparison of such a radio reconstruction (RD) to the reconstruction from air-fluorescence measurements (FD) is shown in figure 4.13. The method yields a combined precision of about 50 g cm^{-2} . The generation of the shower

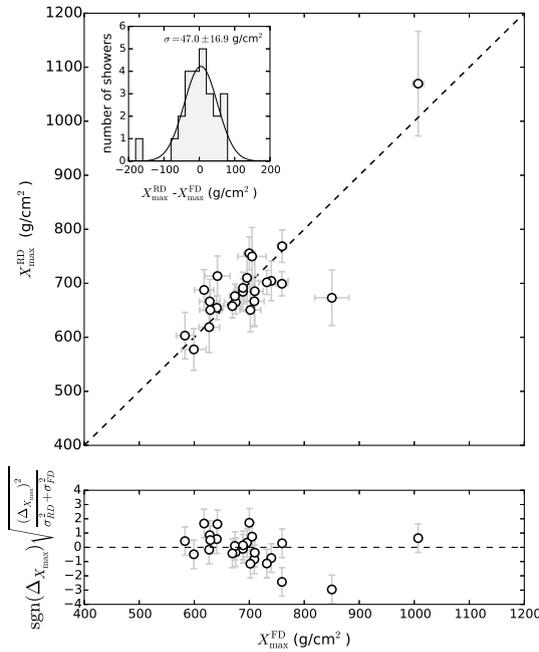


Figure 4.13.: Comparison of X_{\max} measured with FD versus the determination with radio. The lower plot shows the residuals weighted with the errors [45,46].

libraries is computing intensive. This drawback is balanced the universality of the approach which does not depend on any models for the lateral signal distribution. Alternative approaches use parameters sensitive to the geometrical distance R_{\max} from the impact point on ground to the shower maximum. In a next step, the atmospheric depth X_{\max} of the shower maximum is deduced from the geometrical distance R_{\max} by taking into account the density profile of the atmosphere (cf. section 6.7.2). The lateral distribution function width of the ratio signal and the slope of the wavefront are commonly used for such analyses. In recent analysis [46] on AERA data, a combined uncertainty of about 60 g cm^{-2} has been achieved for RD and FD measurements by relating the lateral distribution function width of the ratio signal to the atmospheric depth of the shower maximum.

These lateral distribution function analysis and the wavefront fit are discussed in chapter 6 in more detail.

Improvement of the local data acquisition software

In this work, the software and the operating system of the local radio stations has been adapted for an improved and more reliable radio detection.

5.1. Controlled shutdown for extreme conditions

The power supply system of the AERA-II/III detectors has been optimised for a low power consumption with the experience gained during the operation of AERA phase I. The solar modules of phase II/III stations generate 42 % less electrical power w.r.t. phase I stations and the battery capacity is reduced by 48 % (cf. section 4.2.3). The optimised power supply system provides sufficient power most of the time. However, power shortages have been observed during extreme weather periods and after short winter days. The micro-controller of the electronic can predict these power shortages through the decrease of the input voltage at the front-end electronics. If the input voltage falls below 10.9 V, the micro-controller switches off the supply voltages for the rest of the detector to protect the battery from exhaustive discharge. After the battery protection mode is entered, the micro-controller requires an input voltage of 12.9 V to reactivate the supply-voltages for the local station. In the battery protection mode, the communication system of the stations is switched off and no monitoring data is send to the database. Figure 5.1 shows the resulting hysteresis regulation of the battery protection.

The protection of the batteries from damage due to total discharge is vital for the autonomously operating detectors. But as a matter of fact, it affects the physics goals of the setup. The most important events are the ones measured in coincidence with the fluorescence telescopes. The winter period with its long nights provides ideal conditions for the fluorescence technique, but the lower irradiation drives the power supply system to its limits. The radio stations enter the battery protection mode at individual times during the night and drop out off the measurement. The resulting volatile "holes" in the aperture

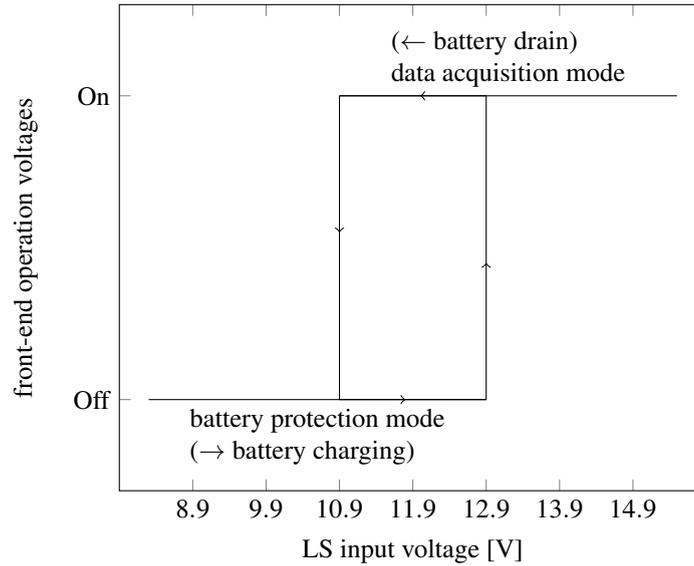


Figure 5.1.: The hysteresis control of the battery protection mode for the AERA stations via the input voltage applied to the front-end electronics

of the radio detector hamper the data acquisition during the winter period and significantly reduce the number of events measured in coincidence of FD and AERA. Figure 5.2 shows the continuous battery discharge in the Argentinean winter 2015. It manifests itself in the decrease of the minimal voltage per day (i.e. the baseline of the data shown in fig. 5.2 & 5.3). On the 26th of July, the combination of short days and bad weather forces the radio station 68 into the battery protection mode.

In order to avoid such situations, a controlled shut down (sleep mode) of the whole array is needed to recharge the batteries before the protection mode is entered. The best recharge performance is achieved, when except for the power-saving micro-controller with about 0.7 W all other functions are deactivated.

Within this work, the local station software has been extended to accept a new type of control message (“LS_PAUSE”) from the central acquisition system. The general structure of the AERA command messages is described in [42]. The LS_PAUSE message includes a start and an end time given in GPS seconds as parameters. From these given times, each radio detector calculates the shutdown duration and stops the data acquisition at the start time. The minimal shutdown time lasts 120 s. LS_PAUSE requests for shorter periods will result in this minimal shutdown time. The maximal sleep time is constrained to 32767 min (about 23 days) by the micro-controller. LS_PAUSE requests for longer periods will be truncated to this maximal sleep time. Then, a special command which passes the shutdown time is send to the micro-controller. Finally, the micro-controller switches off all supply voltages for the FPGA, the front-end electronics and the communication system. The duration of the sleep time is counted down its internal oscillator. The station stays — in contrast to the battery protection mode — in the sleep mode independent of the input voltage. At the end of the sleep period, the micro-controller reactivates all supply voltages and the radio station starts taking data again. The communication between the micro-controller and the CPU is performed via a serial Inter-Integrated Circuit (I²C) connection. The firmware of the FPGA has been slightly adapted to provide this interface.

On 22nd of July, the firmware, the micro-controller software and the local data acquisition program

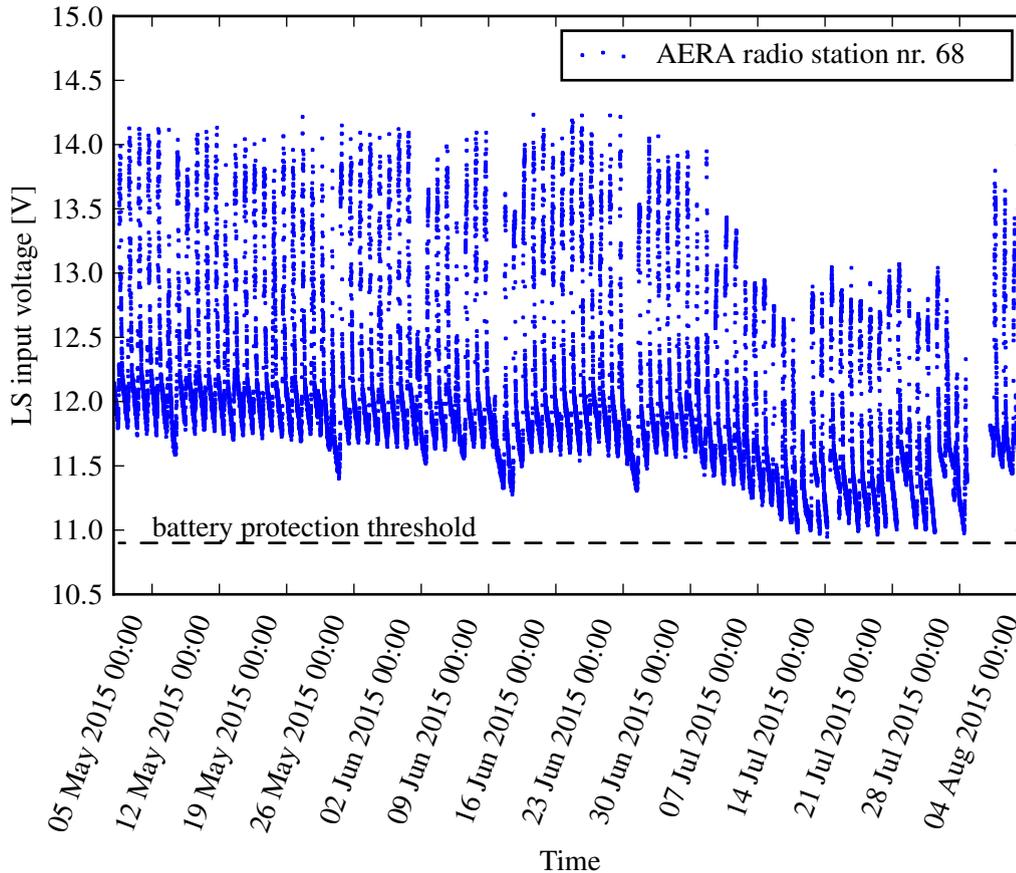


Figure 5.2.: The input voltage of the local radio station 68 measured during the Argentinean winter 2015.

have been updated from remote successfully. From the 28th to the 31st of July, the batteries of all AERA-II and III stations have been recharged by applying the sleep mode. Figure 5.3 shows the LS input voltage for the radio station 68 during two weeks of challenging weather conditions. The software update, a battery protection period and finally the application of the sleep mode are evident as time periods without data. The measured voltage of 12.5 V after the battery protection mode is no contradiction to the required voltage of at least 12.9 V by the hysteresis control for enabling the operation mode. The input voltage is measured without any significant load during the protection mode, while the first transmitted monitoring data is measured when all power consuming parts of the station are already activated.

Following the sleep mode period, the minimal daily battery voltage has increased by 0.5 V which is a good recovery considering that the maximal difference in the voltage baseline during the winter accounts to only 1 V. But the more important improvement is that all detectors have been recharged simultaneously leading to a defined detector aperture at all times (cf. fig. 5.2). The new sleep mode can be used, to recover the batteries during the day to ensure continuous and coincident data taking with all radio stations and FD under extreme weather conditions.

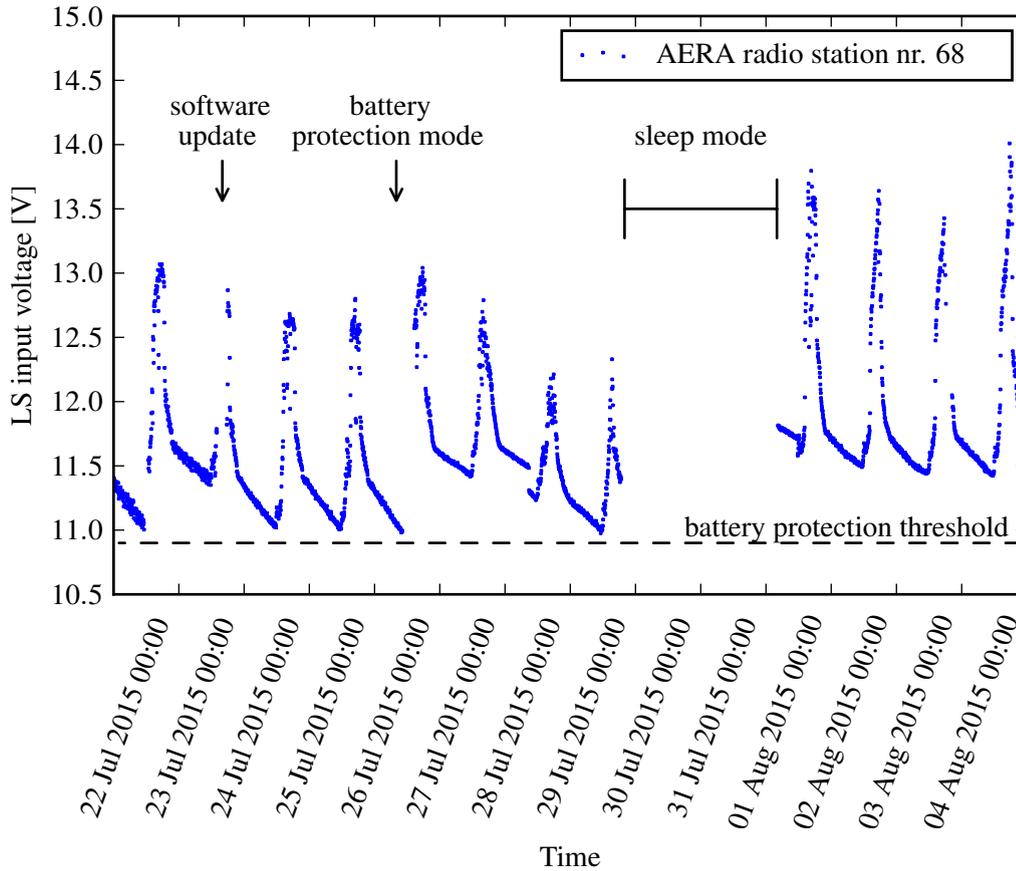


Figure 5.3.: The input voltage of the local radio station 68 measured during the Argentinean winter 2015.

5.2. Reliability and sustainable development via memory protection

In the first firmware designs for AERA, the NIOS softcore CPU has been implemented without a memory management unit (MMU). A memory management unit is a hardware element which is related to a central processor unit. It maps the physical addresses of the main memory on a virtual address space which in general is larger than the physical one. In this way, unsegmented memory blocks with contiguous virtual address spaces can be assigned to the operating system and each application program. The task of the MMU within the virtual memory concept is to prevent processes from accessing addresses beyond their own address space. The MMU throws an exception upon each attempted memory violation which can then be handled by the operating system. Generally, embedded systems do not feature a MMU because the application programs for such devices are intended to be simple. The local data acquisition program in contrast is a complex and concurrent program with six threads. When the development of the firmware and software for AERA started, no proprietary FPGA design for the NIOS processor with MMU was available. Today, implementing an MMU is highly desirable to guarantee a sustainable development of the application software. Because, for new personal it is impossible to know every detail of the memory layout which is an essential precondition for the maintenance and the further development of an application program on a system without memory protection. Additionally, a MMU will prevent defect code to corrupt the data measured by the detectors.

The new custom made firmware for the improved data acquisition software now features a MMU. Thus, also a new uClinux operation system which can cope with a MMU has been configured and customised to the embedded frontend electronics. Finally, the software has been adapted to the new memory management, too. With the aid of the MMU, small errors — e.g. a buffer overflow due to the include of an inappropriate header — have been identified and fixed.

5.3. Incorporation of the differential GPS measurements in the detector software

The positions of the AERA radio detectors have been determined with a precision of about 10 cm by differential GPS measurements in two independent campaigns in 2014 and 2015. In this work, the local data acquisition software has been adapted to program the GPS module at every measurement start with these positions for consistency. After that, the GPS receivers are set into the “position hold” mode. The manufacturer of the “M12M Timing Oncore” claims an accuracy of 2 ns. The programmed position and the “position hold mode” are checked periodically to ensure that the data is taken with the highest possible accuracy. If one of these conditions is not fulfilled, the data acquisition from that local station is aborted. By changing the reference position for the timestamp calculation, one expects a time shift Δt proportional to the vertical position displacement Δz [47]:

$$\Delta t \approx c_z \Delta z \quad \text{with } c_z = 1.94 \quad (5.1)$$

Figure 5.4 illustrates the influence of the used GPS positions on the timing of the radio detectors. It shows the relative timing offset between two stations. A relative time offset of zero implies that the timestamps of both stations have the same offset w.r.t. the expected timestamp by the beacon analysis [43] (cf. sec. 4.2.5). Thus, the figure displays how the timing of the two stations drifts w.r.t. each other. On the 08th of August 2014, the GPS positions in the GPS modules have been reprogrammed with the positions of the differential GPS survey 2014. A significant jump of the timestamps according to sudden change of the relative programmed GPS height is visible in the blue data of figure 5.4. As a cross check, one can adapt the timing before the position programming to the one afterwards by retroactively applying equation (5.1) to the older data. The result is plotted as red crosses in figure 5.4. The visible continuous drifts of the order of 10 ns have already been reported in [42] and [43]. In order to achieve a relative timing accuracy of a few nanoseconds, a beacon calibration is additionally performed which has been checked independently by analysing the transmitted status-messages from air-planes crossing the location of AERA [43].

On the one hand, figure 5.4 proves that the beacon calibration reliably corrects time offsets due to wrong GPS positions. But on the other hand, it becomes clear that the use of the accurate GPS position is vital for a precise relative timing between the individual detectors. Thus, the high level trigger algorithms of the AERA data acquisition benefit from the programming of the GPS. In addition, this also holds for measurements which cannot be corrected by the beacon method.

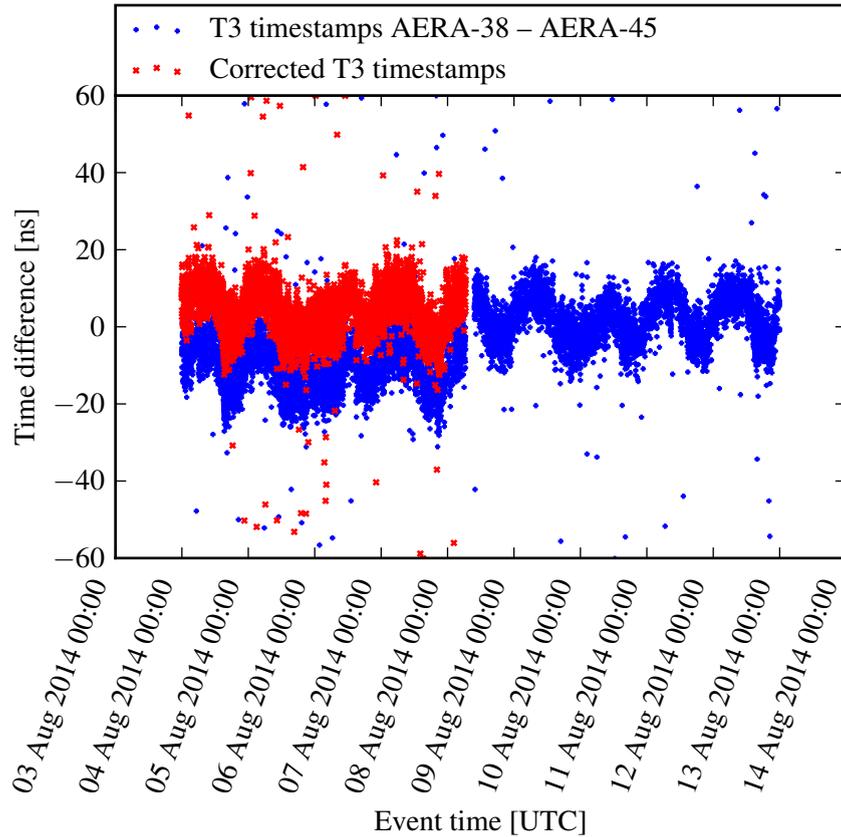


Figure 5.4.: Time offset between station #38 and #45 as determined by the beacon analysis (blue dots). On the 08th of August 2014, the position in the receivers was reprogrammed to reflect the best measurement of the position which is visible as a jump. Taking into account the change in position a correction of the timing for times before the change is possible (red crosses) [48].

5.4. Independently configurable readout for all input channels

The AERA front-end electronics can simultaneously digitise four input channels (cf. section 4.2.3). AERA-I and AERA-II stations record the East-West and North-South component of the radio signal from air showers but each with a high and low gain for a higher dynamic range. Only the high gain data is buffered in the 4 GB SODIMM memory. The low gain data is only stored when a saturation of the high gain data is registered. The readout of the low gain data was not implemented in the data acquisition software.

The new front-end electronics for the AERA-III stations is designed to record up to four signals with the same amplification. Therefore the data acquisition software needed to be improved to store and read out a custom number of channels between one and four. In this section, the differences between the new 4 channel software and the old 2 channel software are discussed. Also, the proper functioning of the new setup is verified.

5.4.1. Physical motivation

The dominant radio emission mechanism for air showers is the geomagnetic emission. The radio wave polarisation is parallel to the acting Lorentz force. Thus, the signal is predominantly polarised

perpendicular to the direction of the earth magnetic field \vec{B} . In figure 5.5, the red circle within the green plane denotes the polarisation phase space for a pure geomagnetic emission.

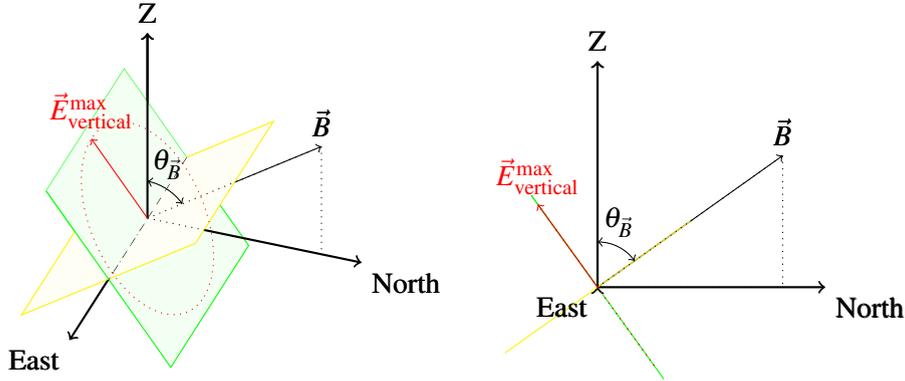


Figure 5.5.: Polarisation considerations for the AERA site: The polarisation vector of the geomagnetic emission is perpendicular to the Earth's magnetic field. Neglecting the Askarayan effect, all electric field vectors \vec{E} reside in the green plane. The electric field $\vec{E}_{\text{vertical}}^{\text{max}}$ with the maximal possible vertical content is emitted from showers with arrival directions parallel to the yellow plane.

The magnetic field is inclined by $\theta_{\vec{B}} = 54.4^\circ$ at the Pierre Auger Observatory. For simplicity, the azimuthal direction is assumed to be North $\phi = 87.3^\circ \approx 90^\circ$ in figure 5.5. A projection of the red circle on the ground corresponds to the electric field component measured with two-dimensional antennas in the North-East plane. For pure geomagnetic emission the North-South polarised signal is rather small in comparison with the East-West component. Here, the emission from the Askarayan effect plays a significant role.

Air showers which arrive parallel to the yellow plain in figure 5.5 emit signals whose electric polarisation from the geomagnetic effect $\vec{E}_{\text{vertical}}^{\text{max}}$ has the maximal possible vertical component. The vertical signal power amounts to roughly two thirds of the total emission ($\sin^2 \theta_{\vec{B}} \approx 66\% = 2/3$). Emission from the Askarayan effect is radially polarised w.r.t the shower axis and independent of the earth magnetic field. Thus, the horizontal polarised fraction of the total signal power depends on the inclination θ_{S} of the shower only. It is directly proportional to $0.5 * \sin^2(\theta_{\text{S}})$ and lost to horizontal radio antennas. The factor 0.5 reflects that only one specific polarisation direction (along the major axis of the footprint) gets rotated in the horizontal direction when inclining a shower. Therefore, only half of a radially polarised total signal is affected.

The challenge for measuring the vertical signals is the high sensitivity of such antennas to man-made noise from the horizon. However, it is highly desirable to measure all three polarisations of the radio emission. Hence, the AERA-III front-end electronics has been designed to record up to four independent channels.

5.4.2. Adaptions with respect to the 2-channel readout

The independent readout of 4 channels requires changes in the programmable hardware of the FPGA and in the local data acquisition software:

- With a flexible number of recorded input channels, the period during which data is buffered

in the 4 GB memory becomes variable, too. If only one channel is recorded, the RAM buffer can hold up to 14 s of data. As a consequence, the coding of the event time stamps requires at least 4 bit and the former 3 bit representation needed to be updated. In the new 4-channel data acquisition firmware the GPS second representation of timestamps is enhanced to even 16 bit. The local acquisition software processes the full GPS second as 32 bit data type. The same holds for the driver, which conducts the communication between application programs running on the NIOS CPU and the FPGA firmware.

- The 2-channel version determined the value of the current second by interpretation of the status messages from the GPS module. The messages are transferred via a serial RS232 connection. Experience showed that this message transfer has to be considered as unreliable. If the transfer fails, the information is lost and an error prone exception handling is performed.

To avoid this problem, the timing in the 4-channel acquisition software does no longer depend on GPS messages. A global variable as time counter in the local acquisition program is set from the GPS status message at start time, once. Both GPS second representations are incremented simultaneously with every pulse per second (PPS) signal from the GPS module. Additionally, the Data Management unit (DMU) second (cf. sec 4.2.3) is compared to the software counter at each PPS occurrence. If they differ, the data acquisition is aborted. For the future, one can think of reprogramming the DMU second at this point. The GPS second is only used for cross-checks. If its value differs from the DMU second three times in a row, the data acquisition is aborted, too. If the GPS message cannot be read, the sawtooth parameter (see sec. 5.4.3) also gets lost and the timestamp accuracy is deteriorated. The 4-channel software cannot correct this problem, but flags this imprecise data in the data header.

- Most firmware changes affect the Data Management Unit (cf. fig. 4.10). The flexible number of recorded channels need a flexible data storage in the 4 GB RAM module. Therefore the write routines in the FPGA firmware and the read routines in the driver needed extensive adaptations.
- The old firmware generated interrupt requests for every T2 trigger which was inefficient use of CPU time. The new firmware generates interrupts periodically, which are then read out in one go. This new mode lowers the interrupt load and allows for a higher trigger throughput. The period length is currently set to 33 ms. This also implies that the triggers are handled with a delay of 33 ms but this is uncritical due to the long duration of data buffering of at least 3.5 s
- The new FPGA firmware and driver software is changed to support the read-out of an arbitrary number of data samples per event.

Thus, the upgrade to the 4-channel setup affects almost all vital parts of the detector software except for the micro-controller firmware and the trigger module of the FPGA. Therefore, the functionality of the whole data acquisition system needed verification after the upgrade.

5.4.3. Selftrigger performance

The independent read-out of the channels has required adaptations in the timestamp calculations and in the internal representation of the time (see above). The timing of the new firm- and software is tested as described in the following: Short pulses (as depicted in figure 5.6) are simultaneously fed to several frontend electronics through a 1-to-N distributor network. The signal has a full duration at half of the maximum of roughly 6 ns and, thus, it is short enough to emulate the radio emission from cosmic rays. The FPGA logic triggers, if the recorded signal exceeds the fixed threshold of 1000 ADC-counts. The resulting GPS time of the T2 timestamp is calculated by the local data acquisition program and then send to the central data acquisition. The received T2 timestamps were recorded for subsequent analysis. The T2 timestamps are truncated to 28 bit before they are send to the AERA data acquisition. Therefore, their resolution is restricted to 4 ns, independent of the hardware limitations.

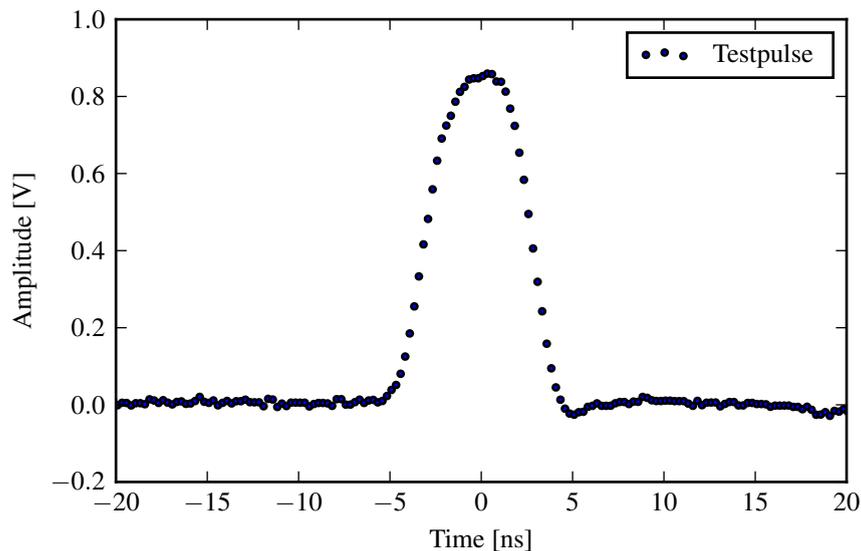


Figure 5.6.: The input signal used for testing the performance of the 4-channel detector software.

A timestamp consists of two parts: a) the absolute GPS second and b) a GPS subsecond (in units of ns). They are stored and treated separately in the firmware and software. While the GPS second internally corresponds to a counter value which is incremented upon each PPS signal, the GPS subsecond is computed from the number of the 180 MHz FPGA clock cycles between two PPS. The precision of the PPS signals is restricted to about ± 25 ns by quantisation errors, due to the limited clock frequency on the GPS module. In order to increase the time resolution, the GPS module sends the estimated delay of the next PPS every second. By applying this correction (called sawtooth parameter) it is possible to reconstruct timestamps of the former second with a precision of 2 ns. In general, the T2 timestamps for selftriggers are evaluated and sent to the central DAQ within the second they occur. As a result, T2 timestamps can deviate by up to 25 ns from the true signal time and have a 5.5 ns resolution due to the 180 MHz frequency of the FPGA clock. Their values are multiples of 4 ns, due to the subsequent cutting of the value (see above).

Figure 5.7 shows compares T2 timestamps generated with different data acquisition soft- and firmware.

The green distribution shows the time difference between the new 4-channel acquisition (LS #302) with the old 2-channel acquisition (LS #101). The yellow distribution compares the measurements of a pure 2-channel setup, while the red one corresponds to a pure 4-channel setup. The 2-channel software revision is 575 and the firmware revision 232. This is the actual detector setup used in Argentina. During the last years, this firmware has not been changed with respect to the timing. The 4-channel setup uses the software revision 606 and the firmware revision 302.

The lower plot in figure 5.7 shows constant deviations of minus 8 s with a relative frequency of the order of 10^{-5} for the established 2-channel version. This deviation is explained by the unreliable serial connection to the GPS module (cf. section 5.4.2) and an incorrect error handling in the 2-channel acquisition software. It affects the T2 timestamps of the self trigger, but it should also lead in rare cases to a wrong timing in the externally triggered data. This is fixed by adopting the time calculation of the 4-channel version. The new 4-channel version shows an accurate performance. A previous firmware revision 301 introduced a 1 s deviation with a probability of 26.5×10^{-6} which affected self-triggered data only. It has been used on AERA stations 41 and 42 from the 7th until the 18th of March 2017. On the 18th of March both stations have been updated with the properly working firmware revision 302. All externally triggered events which have been recorded with the 4-channel version have correct timestamps.

The upper panel of figure 5.7 shows the nanosecond deviation of the timestamps. Their values are multiples of 4 ns, due to the 4 ns resolution during transmission (see above). In addition, the rolling mean, rolling standard deviation and the T2 data are depicted in figures 5.8 & 5.9 for the 2- and 4-channel version. The rolling quantities have been computed for 10000 samples. After the computation, the plotted data has been thinned by a factor of 1000 which mainly results in a smaller scatter of the pure data (blue). Both versions have a short term precision of the relative T2 timing between the independent stations of 7 ns. It is denoted by the rolling standard deviation in figure 5.8 & 5.9. But the unsteady rolling mean limits the use of the T2 timestamps for comparisons between stations. It broadens the nanosecond distributions in figure 5.7 and shows that their mean is not meaningful. It is in agreement with the findings of the beacon analysis discussed in section 4.2.5 and compatible with the precision of 7 ns for the self-triggered timestamps found in [42].

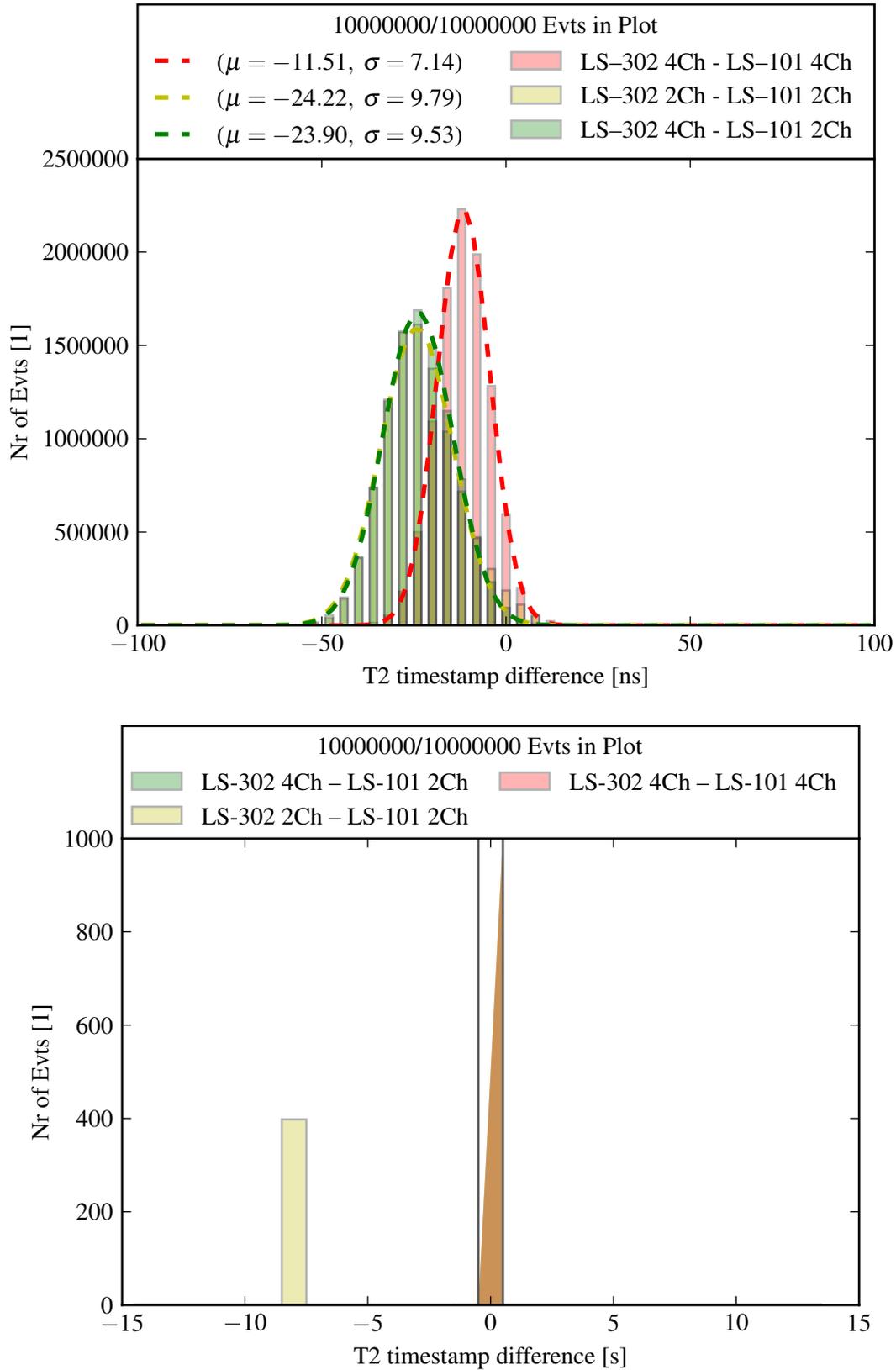


Figure 5.7.: Second and nanosecond deviation of timestamps measured with different data acquisition setups. The FPGA firmware revision of the 4-channel setup is 302.

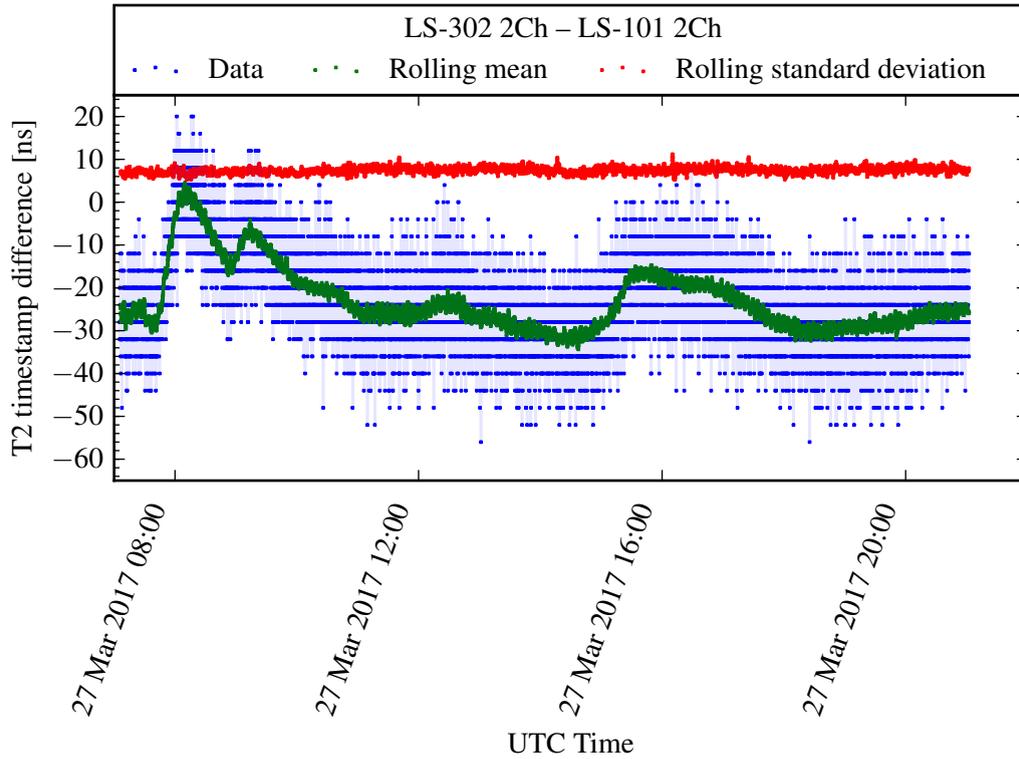


Figure 5.8.: T2 timestamps, their rolling mean and their rolling standard deviation for the 2-channel setup. The rolling quantities have been computed from 10000 samples each. The T2 data has been thinned by a factor of 1000, afterwards.

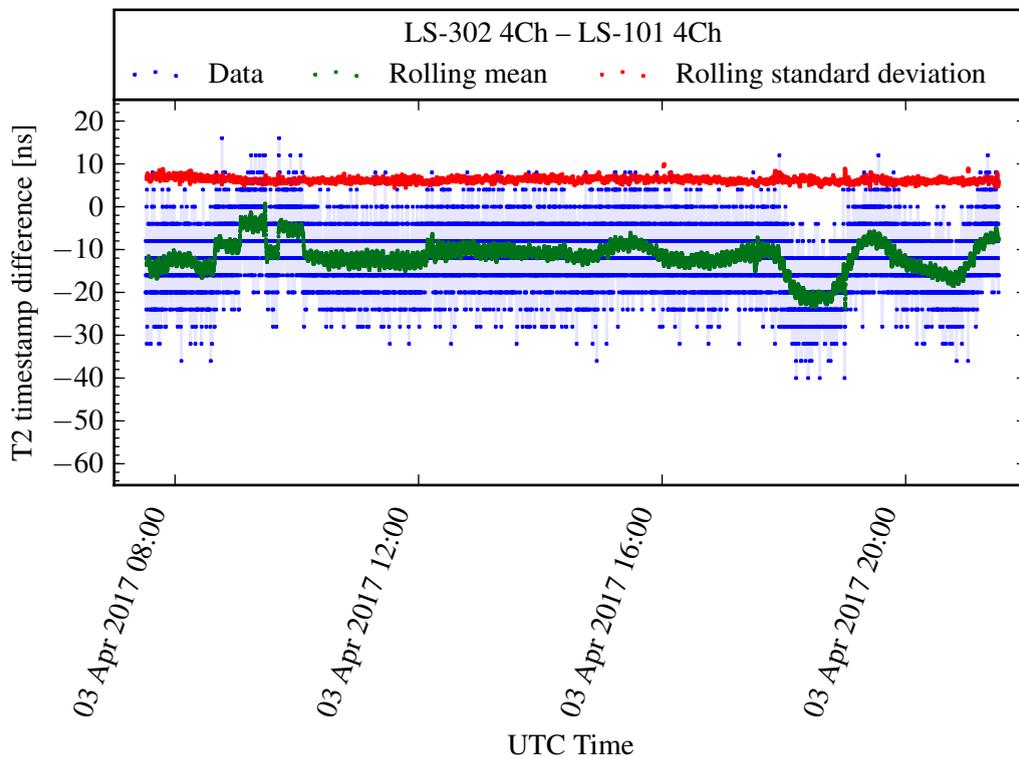


Figure 5.9.: T2 timestamps, their rolling mean and their rolling standard deviation for the new 4-channel setup. The rolling quantities have been computed from 10000 samples each. The T2 data has been thinned by a factor of 1000, afterwards.

In conclusion, the results imply that high level trigger algorithms for AERA need to be designed for processing timestamps with a limited accuracy of a few tens of nanoseconds. However, the read out data can be corrected by the beacon analysis and the relative timing accuracy for subsequent physics analyses is of the order of about 2 ns. In my opinion, it is nevertheless desirable to achieve a better state-of-the-art realisation of the distributed clocks for the next generation of large scale radio detectors.

5.4.4. Timing at AERA

On the 18th of March the 4 channel version has been installed on the AERA station 41 and 42 for testing purposes. The stations participated in the regular data acquisition. Figure 5.10 shows the offset of the stations with respect to station 11 which has been chosen as reference stations in the beacon calibration. For comparison, also station 1 with 2 channel software is shown. The output of the beacon analysis is displayed in transparent colours while the rolling mean is denoted by opaque points. The rolling mean is calculated from data windows with 36 entries. At least 28 entries need to contain valid data entries (instead of e.g. “not a number” entries). The analysis has performed on periodically triggered data only, which are taken in 100 s intervals by AERA. Therefore, the rolling mean generally averages over 1 h of data.

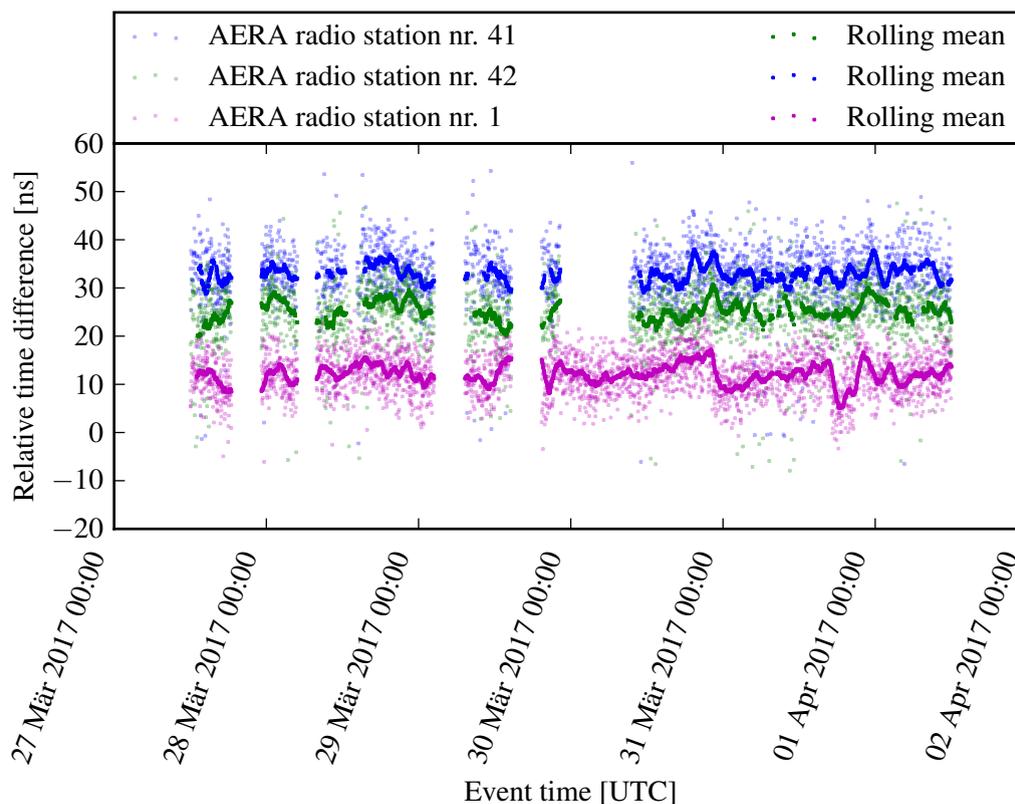


Figure 5.10.: Relative timestamp offsets from periodically triggered AERA data as determined by the beacon correction w.r.t. the reference station 11. The data from AERA stations 41 and 42 are taken with the new 4 channel software. AERA station 1 was equipped with the 2 channel software for the analysed data. The rolling mean is calculated for 36 data points which in general corresponds to averaging over 1 hour of data. The beacon correction failed for the periods without data.

Figure 5.11 displays the timing offsets of the same data as in figure 5.10 as a histogram. The time difference w.r.t. station 11 is nearly constant for each station and the spread accounts to roughly 5 ns

for the 4 channel data from AERA station 41 & 42 and the 2 channel data from AERA station 1. The unexpectedly small jitters of the GPS clocks (cf. section 4.2.5) are explained from strong interfering radio signals affecting the beacon analysis, which have been observed during the displayed measurement time.

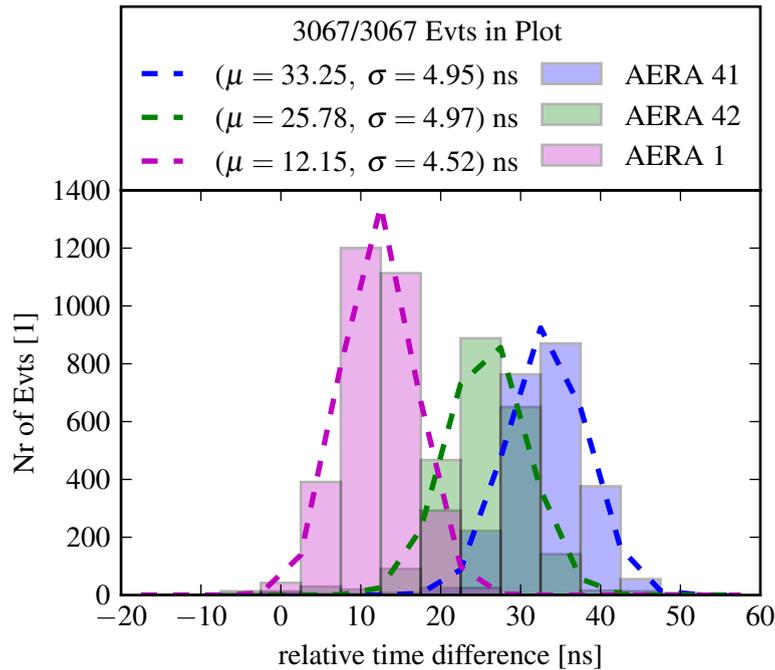


Figure 5.11.: Histogram of the relative timestamp offsets from periodically triggered AERA data as determined by the beacon correction w.r.t. the reference station 11. The data from AERA stations 41 and 42 are taken with the new 4 channel software. AERA station 1 was equipped with the 2 channel software during the displayed time.

One can conclude from figures 5.10 & 5.11 that the timing of the new 4 channel version works as well as the one of the 2 channel version. The 4 channel version shows stability concerning the work in the field and it can be used in future, productively.

5.4.5. 4-channel readout performance

Read out of test patterns

The FPGA programmable hardware of the 4-channel version includes a module to test the Data Management Unit with test patterns. With the aid of this module, one can inject signal ramps at the input of the DMU logic. The data displayed in figure 5.12 has been created by injecting the same signal to the four input channels of the DMU module, but shifted by different offsets. Depending on the number of recorded channels, the DMU writes the data in different structures into the 4 GB RAM memory. The RAM is read out by a custom made driver, which has been developed at the IPE.

Figure 5.12 shows that all these processes work reliably and without any mixing up of data from different channels. It has been triggered externally. The sample number 5106 corresponds to the GPS second 1 174 737 605.999 999 997 . The 1-PPS signal occurs at sample number 5107 and the new ramp is started for all channels one sample later. This plot proves that the data write and read routines work

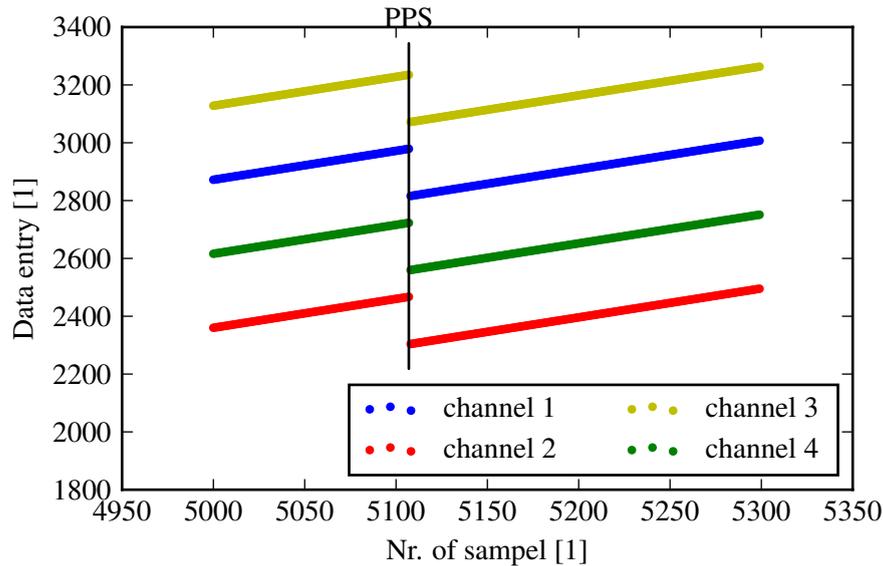


Figure 5.12.: Test of the DMU and driver: The same signal pattern is injected to all 4 channels but with different offsets. Mix-ups of the data by either the FPGA write or the driver read routine would become apparent in this plot. The vertical line marks the sample in which the PPS occurs. Each sample resembles a period of 5.5 ns.

properly for all possible channel configurations.

5.5. Conclusion

In this work, the detector software for the AERA front-end electronics has been improved to read out a adjustable set of the four input channels. This new functionality can be used to measure the horizontal component of the radio signal in addition to the vertical component by equipping the AERA-III stations – as is planned – with three-pole antennas in future. At the detector site, two thirds of the total emitted signal power can be vertically polarised for air showers with an inclination above 54° and thus is lost to the present detector. Since radio measurements are limited by the unavoidable galactic noise floor, it is desirable to record all signal components. The necessary adaptations affect almost all vital parts of the detector software, especially the timing. The functionality of the new data acquisition system has been verified in the laboratory and used productively at two AERA stations. The timing has been improved on station level by using the precise detector positions from a GPS survey for the timestamp calculation. However, the relative timing precision between individual stations without beacon correction is limited to about 10 ns due to clock jitters. The new sleep mode can be used to control the down-time of the detector during overcast periods. This allows for the acquisition of high quality data in coincidence with the fluorescence telescope throughout the Argentinean winter.

The sum of these improvements recovers the full potential of the detector and prepares the front-end electronics for the next generation of prototype stations (e.g. three pole antennas).

Analysis of radio data in a global fit approach

Performing a global fit on data is the only way to determine correlated shower parameters consistently from different types of measurements. So far, AERA data was always analysed step by step: In each step only a special type of signal information (e.g. signal strength, signal time, signal polarisation) was used. Since the parameters of the functional models for the individual steps are correlated, it is highly desirable to determine them in a global regression. The independent and stepwise regression of the correlated models on sub-phase spaces of the measurement is an approximation, but not the optimal description of the air shower.

Within the following chapter, the advantages of a global regression ansatz of the existing models for the arrival time and energy fluence is investigated. Both models are well-respected in the field of cosmic ray radio detection (cf. [4], [49]). After the AERA data has been calibrated against SD measurements, the parameters of the energy fluence fit have been used to determine the energy estimation of cosmic rays [44, 50] (c.f. section 4.2.6). The hyperbolic arrival time model has been adopted by several experiments to describe the data successfully [51], [29], [52]. Simulation studies have shown a relation between the hyperbola slope and the geometrical distance from ground to the shower maximum [51]. First promising attempts to reconstruct the atmospheric depth of the shower maximum with AERA data have been performed in [52]. A detailed description of the standalone models can be found in sections 6.3.1 and 6.3.2.

While the parameters of the functional models are strongly correlated, the experimental observables arrival time and energy fluence themselves are considered to be independent in the regression (cf. section 6.4.2). The weighted uncorrelated chi-square fit is the natural method to describe such a scenario and it is discussed in this work (cf. section 6.5).

6.1. High level data from the baseline detectors

Due to its hybrid measurement technique, the Pierre Auger Observatory provides ideal conditions for radio measurements (cf. section 4). If available, the information from the fluorescence detector, surface detector and the muon counter measurements are merged with the AERA data. The resulting data set allows for unique hybrid analyses. The standard air shower reconstruction from externally triggered AERA measurements even relies on a previous precise surface detector reconstruction.

6.1.1. SD reconstruction

The surface detector data consists of signal traces from 3 photo-multiplier tubes which are normalised to the expected signal for a muon traversing a water-Cherenkov tank vertically. The resulting signal unit is called ‘‘Vertical Equivalent Muon’’ (VEM). The arrival direction of the air showers is obtained by fitting a signal time model to the start times of the measured signals. For measurements with less than five water-Cherenkov detectors, a plane signal front is assumed. For measurements with a higher station multiplicity, the air shower front is approximated by an inflating sphere model. The four free parameters of the fit are the virtual origin \vec{x}_{sh} and the start time t_0 of the shower development [35]. The sphere model is defined as:

$$c(t_i - t_0) = |\vec{x}_{\text{sh}} - \vec{x}_i| \quad (6.1)$$

Here, \vec{x}_i denotes position of the water-Cherenkov detectors on ground and t_i the measured signal times. Figure 6.1 illustrates this model. The energy of air shower is estimated from the lateral distribution

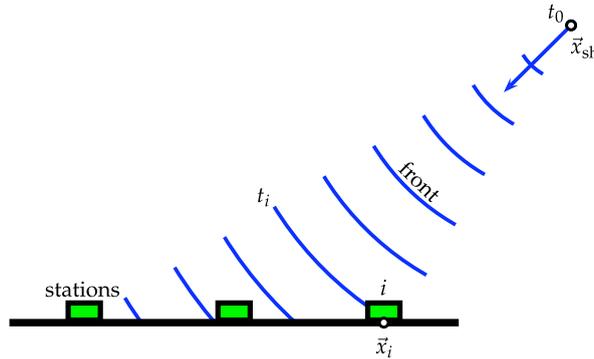


Figure 6.1.: Description of the shower front evolution by an inflating sphere model [35].

function of the shower size $S(r)$:

$$S(r) = S(r_{\text{opt}}) \left(\frac{r}{r_{\text{opt}}} \right)^\beta \left(\frac{r + r_1}{r_{\text{opt}} + r_1} \right)^{\beta + \gamma} \quad (6.2)$$

Equation (6.2) is an modified Nishimura-Kamata-Greisen function [53, 54]. The distance to the shower axis is given by r . The distance r_{opt} where the shower size $S(r)$ fluctuates least due to different shower developments depends on the detector spacing. For the regular station spacing of 1500 m r_{opt} is 1000 m, for the 750 m spacing r_{opt} is 450 m. Depending on the station multiplicity, the shape parameter β and γ are fitted or parametrised values are used. $S(r_{\text{opt}})$ is a fitted scaling parameter. After a correction for

its zenith dependence (called the constant intensity cut $f_{\text{CIC}}(\theta)$) this parameter is used as an estimator for the cosmic ray energy:

$$E_{\text{SD}} = AS(r_{\text{opt}})^B \quad (6.3)$$

A and B are calibrated with the absolute energy scale of the fluorescence detector measurements.

6.1.2. FD reconstruction

First, the shower plane which contains the shower axis and the telescope is reconstructed from the viewing direction of the participating pixels in the photo-multiplier array (cf. 4.1.2). The arrival direction of the shower within this shower-detector plane is determined from the signal times. The precision of the directional reconstruction is improved by taking into account the time of the water-Cherenkov detector with the highest signal [35]. The standard fluorescence detector reconstruction is therefore a hybrid reconstruction. With the reconstructed geometry, the integrated photo multiplier signal can be fitted with a Gaisser-Hillas function:

$$f_{\text{GH}}(X) = \left(\frac{dE}{dX} \right)_{\text{max}} \left(\frac{X - X_0}{X_{\text{max}} - X_0} \right)^{(X_{\text{max}} - X_0)} e^{(X_{\text{max}} - X)/\lambda} \quad (6.4)$$

There, X denotes the atmospheric slant depth of the fluorescence light emission. X_{max} is atmospheric depth of the shower maximum. Depending on the measured track length and the signal strength, the shape parameter X_0 and λ are fitted or constrained to their average value. For the reconstruction of the Gaisser-Hillas light profile, the atmospheric condition, the fluorescence yield and the invisible energy need to be taken into account (cf. section 4.1.2). The achieved X_{max} resolution is 20 g cm^{-2} .

The cosmic ray energy can be determined from the integral over the fitted Gaisser-Hillas function. The energy resolution is 7.6 % and the systematic uncertainty 14 % [35].

6.2. The AERA data

In this work, externally triggered radio data is analysed. Only a subset of the AERA detector can be triggered externally (cf. section 4.2.4). The following section considers data from such AERA stations only:

The raw event info of AERA consists of digitised radio signal traces. The possible signal range is 0 to 4095 ADC counts (12 bit). The time resolution is 5.5 ns. For a better time resolution, the raw traces are subsequently up-sampled by a factor of 4 in the analysis, a calibration is applied and the resulting signal is cleaned from contributions of the four mono-frequent beacon signals. Figure 6.2 shows an example of the resulting reconstructed electric-field trace. All radio quantities are calculated from such \vec{E} -field traces. In my analysis, two quantities are of special interest:

Energy fluence:

The energy fluence at a station is determined by summing up the squared signal magnitude over a time window of 200 ns around the signal peak (which is determined from the maximum of a

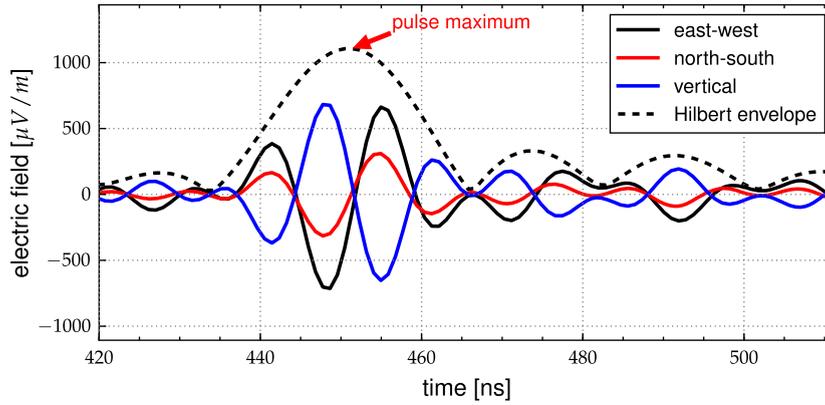


Figure 6.2.: A typical reconstructed radio signal as emitted by an air shower measured by an AERA radio station. The dashed line denotes the Hilbert envelope of the total signal. As an exception to the standard reconstruction, the signal has been up-sampled by a factor of five [44].

fitted Hilbert envelope). Its value is corrected for noise contributions by subtracting the signal intensity determined in a dedicated noise window [44].

Arrival time:

In this work, the peak position of the Hilbert envelope, relative to the recorded signal window, is referred to as “(signal) arrival time” of a certain station. It corresponds to the quantity “eSignalTime” in the Offline analysis framework [55]. The absolute GPS time of a measurement is the sum of the arrival time and the start time of the signal trace.

6.2.1. The Simulated and measured data set

For every event, which has been measured in coincidence of AERA and SD, CoREAS simulations with a proton and an iron as primary particle have been performed. The arrival direction, energy and impact point, reconstructed from the SD measurement have been used as input parameters of the simulations. Therefore, the simulation has the same energy and arrival time distribution as CR. The primary particles of the simulations used here assume iron as primary particle. This is not realistic, as the cosmic ray flux has a mixed composition in the energy range of AERA measurements (cf. figures 6.3 & 3.3). However, a iron composition minimises the shower to shower fluctuations and is — as a best case scenario — a good starting point. The underlying data have been measured between June 2013 and March 2015. In this work, only events which have been measured with more than five signal stations are considered. Figure 6.3 displays the arrival direction and energy distribution of the Monte Carlo simulation set and hence, the SD cosmic ray reconstruction of the data for this subset of 1304 events. Vertical events and events arriving from north are suppressed by the requirement of five signal stations. Vertical showers emit the radio signal into a small area on the ground and thus are detected by fewer stations. The radio signal of air showers arriving from north is suppressed due to the small geomagnetic angle (cf. section 3.3). The geomagnetic field \vec{B} is inclined by 54° and the azimuthal direction is 90° at the site of the Pierre Auger Observatory.

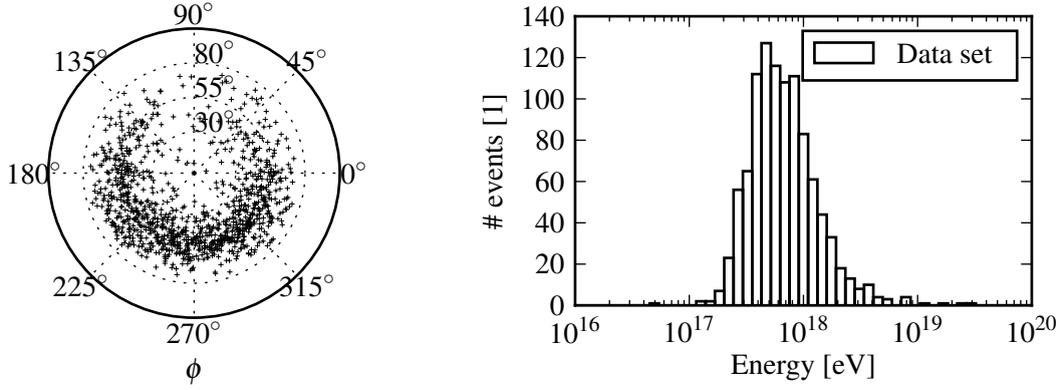


Figure 6.3.: The arrival directions (left panel) and energy distribution (right panel) of the 1304 simulated and measured events, which are considered in this work. The magnetic north is defined as $\phi = 90^\circ$.

Adding Noise to simulated signal traces

In order to work with most realistic simulations, noise is added to the simulated signal traces, which has been measured by the detector periodically. When an event is simulated at a specific time the simulation takes into account the noise level which has been recorded in a time interval of ± 10 minutes around the event time. In addition, the start time of the simulated signal traces is shifted by a delay time which is Gaussian distributed with a width of 2 ns. In this way, the jitter of the GPS clocks after applying the beacon timing correction is modelled.

Therefore, the uncertainty on the signal time is modelled with two components in the analysis framework Offline: a fixed uncertainty of 2 ns and a volatile uncertainty which depends on the signal to noise ratio. Figure 6.4 illustrates the assumed dependence of the signal time uncertainty on the signal to noise ratio. The signal time uncertainty of signal stations (blue marks) ranges from 2 ns to 3 ns. Here, the signal to noise ratio (SNR) is defined as the squared ratio of maximum of the Hilbert envelope over the Root Mean Square (RMS) of the noise.

$$\text{RMS} = \sqrt{\frac{1}{n} \sum_{i=0}^{n-1} E_{i,\text{noise}}^2} \quad (6.5)$$

$$\text{SNR} = \left(\frac{E_{\text{Hilbert, max}}}{\text{RMS}_{\text{noise}}} \right)^2 \quad (6.6)$$

Thus, the signal to noise ratio is defined as a power fraction. If the mean of the noise signal is zero, the RMS equals the standard deviation of the noise signal. Stations which have recorded a signal with a signal to noise ratio above 10 are referred to as signal stations.

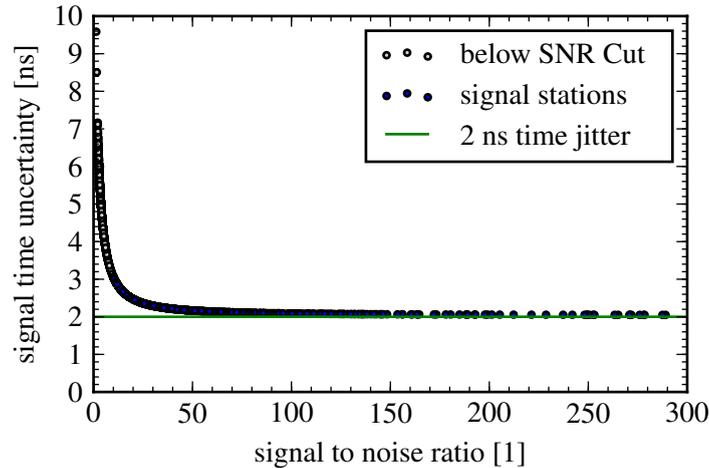


Figure 6.4.: The signal time uncertainty used for the weighting of the time fit. It consists of the constant error of 2 ns due to the time jitter of the GPS clocks and a volatile error depending on the signal to noise ratio. The data has been thinned by a factor of 50 for the stations below the SNR cut and a factor of 10 for the signal stations.

6.3. The standard analysis strategy

The standard analysis pipeline for AERA data is called the “RdObserver”. It is implemented in the software framework Offline. The “RdObserver” is designed for the analysis of externally triggered data (cf. section 4.2.4) with the merged information from all sub-detectors of the Pierre Auger Observatory. The availability of a successful air shower reconstruction from SD data is an essential condition for the “RdObserver”, as it acts as the starting point for the radio reconstruction. When dealing with simulation data, the standard pipeline is defined by the Offline configuration files of the “RdSimObserver”. The module sequences and settings used in this work can be found in the appendix D.

The analysis algorithms are implemented in the Offline framework as distinct modules. These modules can access the event data, which are either real measured or simulated event data. Additionally, the modules have read access to the detector description, i.e. the detector configuration, the corresponding calibration data (e.g. the response of the antenna, ...) and monitoring data. On the basis of these two types of information, the modules can calculate high level observables and store them in the event data, which thus is the only one-directional communication link between the subsequently executed analysis modules. Therefore, the correct way to combine two existing analyses is to rewrite them in one new module. The Offline module, which has been implemented in the course of this work is named “RdGlobalFit”.

But before discussing the new module, the older univariate implementations are described:

6.3.1. Arrival time fit

It is generally accepted that the wavefront of the radio signal from extensive air showers can be well described by a hyperbola [4, 29, 49, 51]. Strictly speaking, the wavefront is the surface perpendicular to the signal velocity. However, radio experiments only access the arrival time of the signal and the definition of this (e.g. pulse maximum, crossing of a given signal threshold, ..) introduces ambiguity.

Furthermore, these definition depend on detector properties like the bandwidth of the measurement, etc. As a consequence, one should compare different experimental measurements of the radio wavefront with great care.

Following [29], the signal arrival time t at an observer with distance d to the shower-axis can be described by a hyperbola $t(r)$. In a space-type coordinate system (ct, d) , where c denotes the speed of light, a generic hyperbola is defined by

$$\frac{c^2 t^2}{c^2 b^2} + \frac{d^2}{\gamma^2} = 1. \quad (6.7)$$

Here, cb denotes the semi major axis while γ is the conjugate axis or semi minor axis. The angle between the asymptotes and the shower-plane is given by $\rho = \arctan(cb/\gamma)$.

Equation (6.7) can be solved for the arrival time and shifted by minus b such that the vertex lies on the d -axis. Furthermore, the detector position can have an offset $z_{i,\text{sp}}$ to the shower plane which needs to be corrected for. The additional parameter t_0 takes into account that the time of the shower impact on ground is unknown in the measurement. Instead, the recorded signal trace is relative to a quite arbitrary time window defined by the SD reconstruction.

$$t_i = b \sqrt{1 + \frac{d^2}{\gamma^2}} - b - z_{i,\text{sp}} + t_0. \quad (6.8)$$

Equation (6.8) takes all of this into account. It describes the model for the expected arrival time t at a detector station i . Figure 6.5 depicts this scenario.

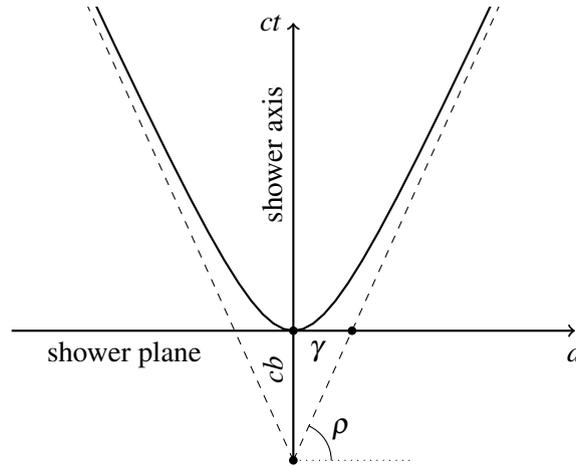


Figure 6.5.: Drawing of a hyperbola as described by equation (6.8). The offset t_0 is set to zero and all observers are assumed to be located along the shower plane ($z_{i,\text{sp}} = 0$).

This model is inconsistent to some extent, as a wavefront is expected to propagate perpendicular to its surface. This surface has prior been identified with the arrival time ct . Therefore, the arrival time of this wavefront cannot be described by the distance of the wavefront to the ground ct itself. However, one can ignore this problem by simply abstaining from calling the arrival time distribution a wavefront.

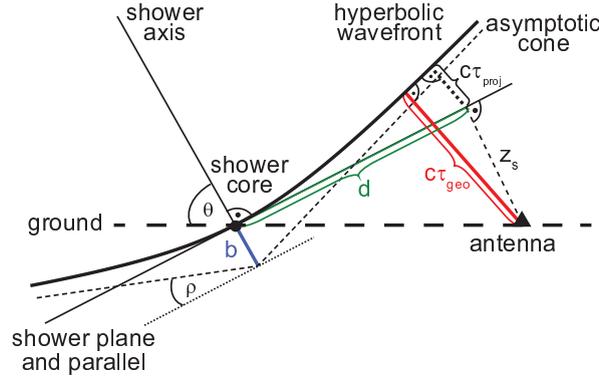


Figure 6.6.: Sketch of a shower front with the parameters involved in the fit of a hyperbolic wavefront from [51].

In contrast, [51] assumes the wavefront to be hyperbolic (cf. figure 6.6) and defines an arrival time model from this:

$$c\tau_{\text{geo}}(d, z_s) = \sqrt{(d \sin \rho)^2 + (cb)^2} + z_s \cos \rho + cb \quad (\text{time model from [51]}) \quad (6.9)$$

Expressing γ via ρ and setting t_0 to zero, equation (6.8) can be rearranged to:

$$ct_i = \sqrt{(cb)^2 + d^2 \tan^2(\rho)} - cb - z_{i,\text{sp}} \quad (\text{time model used in this work}) \quad (6.10)$$

Typical values for the angle ρ are about $\rho = 1^\circ$. For such small angles the approximations $\tan^2(\rho) \approx \sin^2(\rho)$ and $\cos^2(\rho) \approx 1$ are accurate to the order of permille. Hence, equations (6.10) and (6.9) are equivalent, taking into account the ambiguity of the signs. In other words, postulating either the wavefront or the arrival time to follow a hyperbolic distribution leads to the same result as long as the angle ρ of the hyperbola is small. Furthermore, the questionable identification of the arrival time with the wavefront becomes valid under this conditions. Thus, it is a matter of taste to start arguing from the measured data or from the theoretical model for the experimentally inaccessible wavefront.

However, the full model also depends due to the special choice of the shower coordinate system on the shower geometry, i.e. the impact point (X, Y) and arrival direction (zenith θ and azimuth ϕ). The height Z of the impact point is no free parameter. In Argentina, it is defined by the profile of the landscape $Z(X, Y)$. Also the relative position \vec{r}_i of the antennas i to the shower axis and to the shower plane are defined by the geometry. The major axis b is fixed to 9 ns. This reduces the hyperbola fit to an effective cone fit, where the slope of the asymptotes defines the curved distribution around the vertex and vice versa. But, the sensitivity to the shower maximum is increased by this [51, 52]. Hence, the free parameters for the implementation in Offline are

$$t_i = t_i \left(\underbrace{\vec{r}_i}_{\text{station position}} ; \underbrace{X, Y, \theta, \phi, \gamma, t_0}_{6 \text{ free parameters}} ; \underbrace{b = 9 \text{ ns}}_{1 \text{ fixed parameter}} \right). \quad (6.11)$$

In this representation, the minor axis γ describes directly the slope of the asymptotic cone and therefore the angle ρ of the hyperbola. Therefore, the minor axis γ is in the following also referred to as ‘‘slope parameter’’.

6.3.2. Radiation energy fluence fit

A distinct feature of the radio signal from air showers is the asymmetric shape of the lateral distribution. As discussed in chapter 3.3, this asymmetry is a result of the superposition of the geomagnetic effect and the charge excess effect. Therefore, the radio signal at ground cannot be well described by a simple radially symmetric function as is done for e.g. the lateral distribution of electrons and muons via the Nishimura-Katama-Greisen [53] function.

Based on [56], the energy fluence (f or ε) measured at AERA is well described by a superposition of two Gaussian functions of the functional form [44, 50]:

$$\varepsilon(\vec{r}) = f(\vec{r}) = A \left[\exp\left(-\frac{(\vec{r} - \vec{r}_{\text{core}} + C_1 \vec{e}_{\vec{v} \times \vec{B}})^2}{\sigma^2}\right) - C_0 \exp\left(-\frac{(\vec{r} - \vec{r}_{\text{core}} + C_2 \vec{e}_{\vec{v} \times \vec{B}})^2}{(C_3 e^{C_4 \sigma})^2}\right) \right] \quad (6.12)$$

All coordinates are in the shower plane. \vec{r} denotes the projection of the station position on the shower

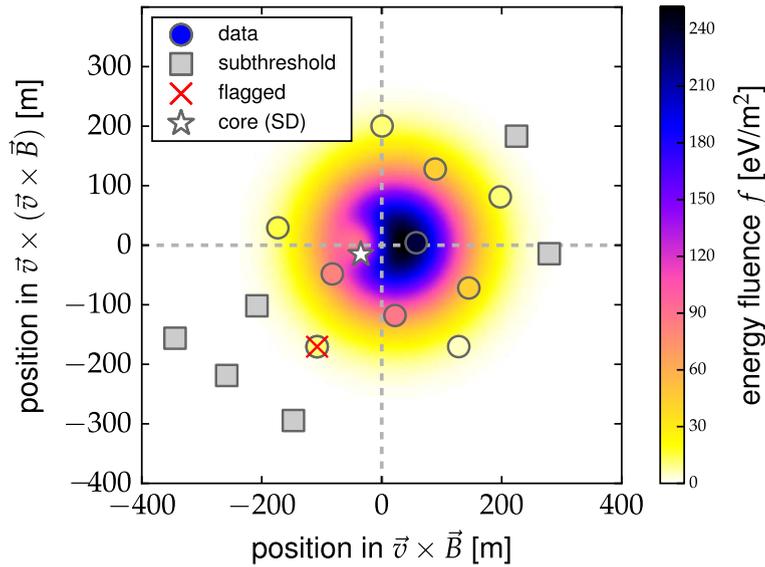


Figure 6.7.: The two-dimensional LDF of a measured event in the shower plane. Therefore the reconstructed radio core is per definition located at (0, 0) m. The measured air shower arrived with an energy of 0.75 EeV under 37 ° inclination. The colour within the circles encodes the measured energy fluence while the coloured area around is determined by the fit [44].

plane. The core position (or impact point), i.e. the intersection of the shower axis with the shower plane, is denoted by \vec{r}_{core} . The core position \vec{r}_{core} , the distribution width σ , and the amplitude A are the free parameters of the model. The constants C_0 , C_1 , C_2 , C_3 and C_4 have been determined from simulations for the best description of the AERA data.

This LDF model has already been implemented in a chi-square fit in the `Offline` module “Rd2dLDFitter” [46]. Figure 6.7 shows the fit result to an example event measured in coincidence of AERA and SD.

6.4. The global regression approach

The global regression of the arrival time and energy fluence is performed as weighted uncorrelated chi-square fit.

6.4.1. The analytic description of the weighted uncorrelated χ^2 fit

In this section, the analytic regression method called “weighted uncorrelated chi-square” is shortly discussed, by means of the global regression problem as an example. The reasoning follows [57] closely.

Consider an observation vector \vec{L} which holds the measured data. The global regression problem has two types of observations:

- The signal arrival time \vec{L}_t at each station i relative to the signal window.
- The energy fluence \vec{L}_ϵ measured at each station i .

The free parameters of the model are collected in the parameter vector \vec{X} . The dimension of the observation vector is $dim(\vec{L}) = dim(\vec{L}_t) + dim(\vec{L}_\epsilon) = 2i$, where i is the number of signal stations. Signal stations are stations, which have recorded a signal above a certain threshold and have not been rejected due to an unexpected waveform, known hardware problems etc. The regression is exclusively based on the measurements of these stations. The functional model f_k tries to describe each observation L_k by the parameters \vec{X} .

$$L_k + v_k = f_k(\vec{X}) \quad (6.13)$$

The mismatch of the observations L_k and the model description f_k is called the improvement v_k . The best choice of parameters \vec{X} minimises the improvement v_k . In our case, there are two types of functional models:

- The hyperbolic arrival time model f_t , which is defined by equation (6.8).
- The lateral distribution function, which is defined by equation (6.12).

Again, the dimension of the model vector $dim(\vec{f})$ is $2i$. Consequently, also the improvement vector has the length $2i$ and can be split in an energy fluence part and a time part ($\vec{v} = (\vec{v}_\epsilon, \vec{v}_t)$).

Trying to solve the problem analytically, the first step would be to linearise it by expanding the functional model into a Taylor series around a supporting point X_0 :

$$\vec{f}(\vec{X}_0 + \vec{x}) \approx \vec{f}(\vec{X}_0) + A\vec{x} \quad (6.14)$$

where A is the coefficient matrix or Jacobi matrix

$$A_{k,l} = \frac{\partial f_k}{\partial x_l} \quad (6.15)$$

Thus, one introduces the reduced observation vector \vec{L} and the reduced parameter vector x via

$$\vec{L} \approx \vec{L}_0 + \vec{l} \quad (6.16)$$

$$\vec{X} \approx \vec{X}_0 + \vec{x}, \quad (6.17)$$

with the abbreviation $\vec{L}_0 = \vec{f}(\vec{X}_0)$. Considering an event with n signal stations, the coefficient matrix of our problem reads with $\vec{X} = (X, Y, Z, \phi, \theta, A, \sigma, \gamma, t_0)$ as parameter vector:

$$A = \begin{pmatrix} \frac{\partial f_{i,1}}{\partial X} & \frac{\partial f_{i,1}}{\partial Y} & \frac{\partial f_{i,1}}{\partial Z} & \frac{\partial f_{i,1}}{\partial \phi} & \frac{\partial f_{i,1}}{\partial \theta} & \frac{\partial f_{i,1}}{\partial A} & \frac{\partial f_{i,1}}{\partial \sigma} & \frac{\partial f_{i,1}}{\partial \gamma} & \frac{\partial f_{i,1}}{\partial t_0} \\ \vdots & \vdots \\ \frac{\partial f_{i,n}}{\partial X} & \frac{\partial f_{i,n}}{\partial Y} & \frac{\partial f_{i,n}}{\partial Z} & \frac{\partial f_{i,n}}{\partial \phi} & \frac{\partial f_{i,n}}{\partial \theta} & \frac{\partial f_{i,n}}{\partial A} & \frac{\partial f_{i,n}}{\partial \sigma} & \frac{\partial f_{i,n}}{\partial \gamma} & \frac{\partial f_{i,n}}{\partial t_0} \\ \frac{\partial f_{\varepsilon,1}}{\partial X} & \frac{\partial f_{\varepsilon,1}}{\partial Y} & \frac{\partial f_{\varepsilon,1}}{\partial Z} & \frac{\partial f_{\varepsilon,1}}{\partial \phi} & \frac{\partial f_{\varepsilon,1}}{\partial \theta} & \frac{\partial f_{\varepsilon,1}}{\partial A} & \frac{\partial f_{\varepsilon,1}}{\partial \sigma} & \frac{\partial f_{\varepsilon,1}}{\partial \gamma} & \frac{\partial f_{\varepsilon,1}}{\partial t_0} \\ \vdots & \vdots \\ \frac{\partial f_{\varepsilon,n}}{\partial X} & \frac{\partial f_{\varepsilon,n}}{\partial Y} & \frac{\partial f_{\varepsilon,n}}{\partial Z} & \frac{\partial f_{\varepsilon,n}}{\partial \phi} & \frac{\partial f_{\varepsilon,n}}{\partial \theta} & \frac{\partial f_{\varepsilon,n}}{\partial A} & \frac{\partial f_{\varepsilon,n}}{\partial \sigma} & \frac{\partial f_{\varepsilon,n}}{\partial \gamma} & \frac{\partial f_{\varepsilon,n}}{\partial t_0} \end{pmatrix} \quad (6.18)$$

$$= \begin{pmatrix} \frac{\partial f_{i,1}}{\partial X} & \frac{\partial f_{i,1}}{\partial Y} & \frac{\partial f_{i,1}}{\partial Z} & \frac{\partial f_{i,1}}{\partial \phi} & \frac{\partial f_{i,1}}{\partial \theta} & 0 & 0 & \frac{\partial f_{i,1}}{\partial \gamma} & \frac{\partial f_{i,1}}{\partial t_0} \\ \vdots & \vdots \\ \frac{\partial f_{i,n}}{\partial X} & \frac{\partial f_{i,n}}{\partial Y} & \frac{\partial f_{i,n}}{\partial Z} & \frac{\partial f_{i,n}}{\partial \phi} & \frac{\partial f_{i,n}}{\partial \theta} & 0 & 0 & \frac{\partial f_{i,n}}{\partial \gamma} & \frac{\partial f_{i,n}}{\partial t_0} \\ \frac{\partial f_{\varepsilon,1}}{\partial X} & \frac{\partial f_{\varepsilon,1}}{\partial Y} & \frac{\partial f_{\varepsilon,1}}{\partial Z} & \frac{\partial f_{\varepsilon,1}}{\partial \phi} & \frac{\partial f_{\varepsilon,1}}{\partial \theta} & \frac{\partial f_{\varepsilon,1}}{\partial A} & \frac{\partial f_{\varepsilon,1}}{\partial \sigma} & 0 & 0 \\ \vdots & \vdots \\ \frac{\partial f_{\varepsilon,n}}{\partial X} & \frac{\partial f_{\varepsilon,n}}{\partial Y} & \frac{\partial f_{\varepsilon,n}}{\partial Z} & \frac{\partial f_{\varepsilon,n}}{\partial \phi} & \frac{\partial f_{\varepsilon,n}}{\partial \theta} & \frac{\partial f_{\varepsilon,n}}{\partial A} & \frac{\partial f_{\varepsilon,n}}{\partial \sigma} & 0 & 0 \end{pmatrix} \quad (6.19)$$

Starting from the basic idea that the observations \vec{L} can be described by the functional model \vec{f} , one can use the Taylor series to write down an approximated relation between the reduced quantities \vec{x} and \vec{l} :

$$\vec{L} + \vec{v} = \vec{f}(\vec{X}) \quad (6.20)$$

$$\vec{L} + \vec{v} \approx \vec{f}(\vec{X}_0) + A\vec{x} \quad (6.21)$$

$$\vec{l} + \vec{v} \approx A\vec{x} \quad (6.22)$$

In this case the general minimisation criterion χ^2 of chi-square regression reads:

$$\chi^2 = \vec{v}^T \Sigma_{ll}^{-1} \vec{v} \quad (6.23)$$

where Σ_{ll} is the covariance matrix.

$$\Sigma_{ll} = \begin{pmatrix} \sigma_\varepsilon^2 & \rho_{\varepsilon,t} \sigma_\varepsilon \sigma_t \\ \rho_{t,\varepsilon} \sigma_\varepsilon \sigma_t & \sigma_t^2 \end{pmatrix} \quad (6.24)$$

The covariance matrix Σ_{ll} is assumed to be diagonal at the moment. Therefore, the observations are considered to be independent measurements. The validity of this assumption is studied in section 6.4.2.

The minimisation criterion for this stochastic model reads:

$$\chi^2 = \begin{pmatrix} \vec{v}_\varepsilon \\ \vec{v}_t \end{pmatrix} \begin{pmatrix} 1/\sigma_\varepsilon^2 & 0 \\ 0 & 1/\sigma_t^2 \end{pmatrix} \begin{pmatrix} \vec{v}_\varepsilon \\ \vec{v}_t \end{pmatrix} \quad (6.25)$$

$$= \sum_n \frac{v_{\varepsilon,n}^2}{\sigma_{\varepsilon,n}^2} + \sum_n \frac{v_{t,n}^2}{\sigma_{t,n}^2} \quad (6.26)$$

The RdGlobalFit module uses the numerical minimisation tool “MINUIT”. Therefore, one can formulate the problem directly with the observation vector and parameter vector. According to the equation (6.26), the implemented minimisation criterion reads:

$$\chi^2 = \sum_n \left(\frac{L_{\varepsilon,n} - f_\varepsilon(X)}{\sigma_{\varepsilon,n}} \right)^2 + \sum_n \left(\frac{L_{t,n} - f_t(X)}{\sigma_{t,n}} \right)^2 \quad (6.27)$$

The MINUIT routines then try to find the minimum by varying the free parameters.

6.4.2. Testing the independence of the variables

Figure 6.8 displays the reconstructed energy fluence versus the reconstructed arrival time of simulated events. The underlying events are from the simulation library for the “RdSimulationObserver”. The primary particle type is iron. Events are considered only, if at least 5 stations have recorded a signal with an signal to noise ratio larger than 10. The simulated signal traces have been blurred in time by a Gaussian distributed jitter with a standard deviation of 2 ns as described in section 6.2.1. Additionally, measured noise from periodically triggered AERA data has been added to the signal traces. Data from stations which have been rejected in the “RdSimulationObserver” or have no signal (i.e. their signal-to-noise ratio is below 10) are disregarded. The whole data are plotted with blue dots. However,

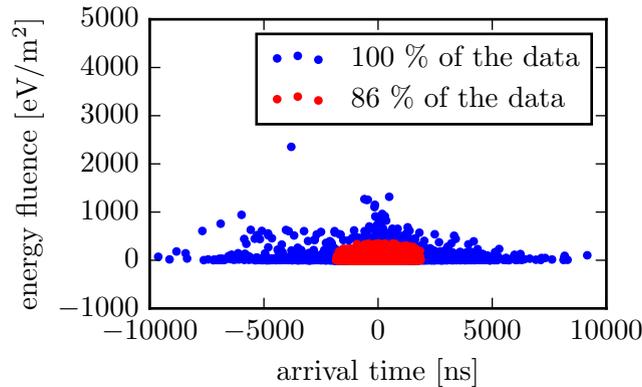


Figure 6.8.: Scatter plot of the reconstructed signal arrival time versus the reconstructed energy fluence of events of the simulation library RdSimObserver. The primary particle type is iron for all events. Measured noise and a time jitter of 2 ns has been taken into account to the CoREAS simulation afterwards.

86 % of the data, which are denoted by red dots, are restricted to a small rectangular sub phase-space. The data are governed by cutting away the upper 1 % quantile of the energy fluence and the lower and upper 7 % quantile of the arrival time. The resulting data are homogeneous enough for a chi-square contingency hypothesis test of the independence of the arrival time and energy fluence. Therefore, a

contingency table of the 86 % sub-data-set has been created. The data set is sliced 45 times in both variables equidistantly. The result is displayed in figure 6.9 on the left panel. By eye, there is no

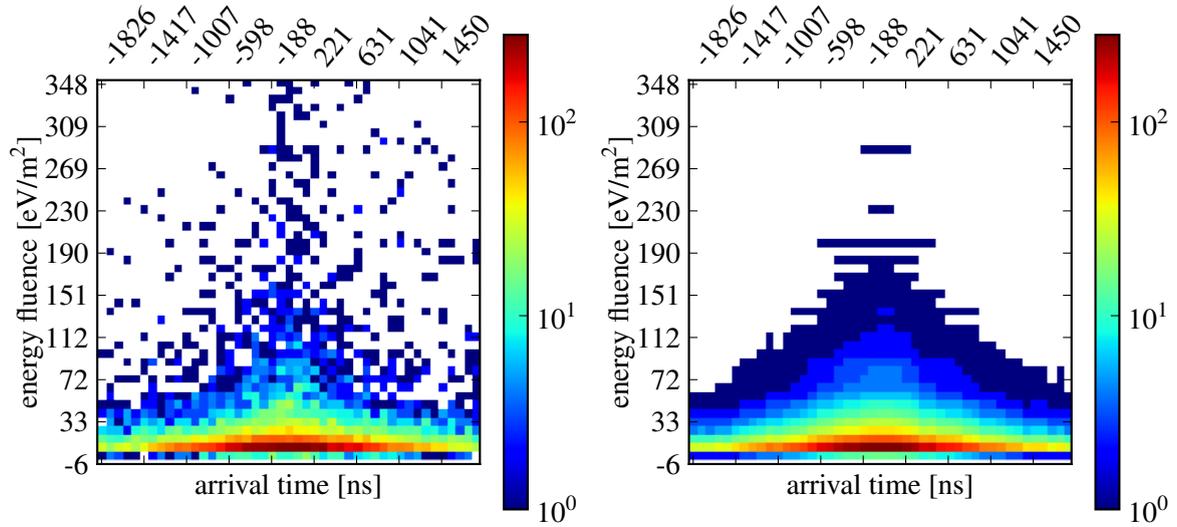


Figure 6.9.: A contingency test of the arrival time and energy fluence data. Corresponding to the red data in figure 6.8 86 % of the data has been binned (left panel). The fitted chi-square statistic is shown (right panel). The independence of the observed frequencies is favoured with a p -value of 0.97

visible dependence between the arrival time and the energy fluence. The independence of the variables manifests in the fact that each row follows the same probability density function. The same holds for the columns. The only change between different rows or columns is a constant scaling factor. One should note that the blue data in figure 6.8 follows the two-dimensional triangular structure which is apparent in figure 6.9. Therefore, also these data seems to follow the frequencies observed in the contingency table. The right panel of figure 6.9 shows the results from a fit of a chi-square statistic to the data assuming the arrival time and energy fluence to be independent random variables. This hypothesis is favoured by a p -value of 0.97.

As a consequence, the choice of an uncorrelated chi-square fit for the global regression is well-justified for at least 86 % of the data. In addition, this result retroactively justifies the presently used step-wise analysis strategy, because it requires the independence of timing and energy fluence, too.

6.5. Implementation of the global regression

The module “RdGlobalfit” is implemented as a module in Offline. It performs the simultaneous minimisation of equations (6.8) and (6.12) in a weighted chi-square approach (equation (6.26)) using the MINUIT library. Equations (6.8) and (6.12) share the shower geometry and the sensitivity to the shower maximum as common parameter. Therefore, the functional models are highly correlated.

6.5.1. The fit parameter

The global functional model of the arrival time fit t and the energy fluence fit ε has 15 parameters:

$$t = t(\underbrace{\vec{r}_i}_{\text{station position}} ; \underbrace{X, Y, \theta, \phi, \gamma, t_0}_{6 \text{ free parameters}}; \underbrace{b, Z}_{2 \text{ fixed parameter}}) \quad (6.28)$$

$$\varepsilon = \varepsilon(\underbrace{\vec{r}_i}_{\text{station position}} ; \underbrace{X, Y, \theta, \phi, A, \sigma_+}_{6 \text{ free parameters}}; \underbrace{C_0, C_1, C_2, C_3, C_4, Z}_{6 \text{ fixed parameter}}) \quad (6.29)$$

All free parameters can be fixed in the configuration files. Even the fixed parameters can be released, which can compromise the performance of the stepwise regression. The start values for the arrival direction (θ, ϕ) and impact point (X, Y, Z) can be taken from the distinct previous reconstructions (e.g. from radio or surface detector). For an univariate fit, the direction in (6.29) is typically also fixed.

The impact time t_0 and the impact point (X, Y, Z) are degenerate. Therefore, the altitude Z is fixed to the mean height of all active stations in an event. Technically speaking, the start value of the impact point is propagated along the preliminary shower axis to the point, where it crosses the x-y plane of the coordinate system with the origin at the radio “RdBaryCenter”. The naming “barycentre” is unfortunately misleading as the “RdBaryCenter” is in fact the simple mean of all signal station positions. The global regression will search the impact point in the x-y plane of the defined coordinate system. The search plane of the core can be customised by setting a different “CoordinateOrigin” in the “RdEventInitializer” module.

6.5.2. Shared Geometry

Both functional models (t and ε) are sensitive to the geometry. In the “RdGlobalFit” module, the geometry is not included explicitly in the fit functions. Instead, for every parameter set (X, Y, θ, ϕ) a temporal coordinate system is defined at the point (X, Y) with the z-axis perpendicular to the surface of the earth. In the next step all station positions \vec{r}_i are transformed in the shower plane with the origin in (X, Y) and the axis defined by (θ, ϕ) .

Then equation (6.26) is calculated, setting X and Y to zero. The result is returned to MINUIT. In the next minimisation step, again all station quantities are transformed into a new shower plane and the solution of (6.26) is returned to MINUIT until convergence is reached and the best set of parameters is found within the numerical uncertainty.

6.5.3. Staged fit

The eight degrees of freedom make it difficult to minimise (6.26) in one step. Instead, the minimum is found in five steps to improve the fit stability. The basic idea is to find good start values with the robust energy fluence fit before applying the arrival time fit. If a functional model of the fit is disabled, steps which fit only parameters of this model are skipped. If the parameters are set fixed in the configuration file, they stay fixed all the time:

Step 1 Pre-fit of the lateral distribution function: only parameter A and σ_+ are free.

Step 2 The full energy fluence fit is performed, but direction is fixed.

Step 3 Only the impact time t_0 is a free parameter.

Step 4 The arrival time model is fit to the data while the impact point is still taken from the energy fluence fit (step 2).

Step 5 In the final stage, all parameters are released according to the configuration file setting.

The shower impact time t_0 is always restricted to the interval of ± 30 ns around the mean signal time.

6.5.4. Changes with respect to the standard reconstruction

The “RdTopDownStationSelector” module rejects in a loop with the “RdPlaneFit” module stations in order to find a set of measured signal times which can be described by a plane wave with a chi-square probability higher than 0.05. Thus, this mechanism rejects stations which are sensitive to a hyperbolic wavefront. Therefore, this loop is commented out in the module sequence used in this work.

6.5.5. Adaption of the measurement uncertainties

The correct determination of the measurements uncertainty is vital for the global regression approach. A univariate fit will find the minimum even if the measurement uncertainties are all wrong by a constant factor. In the global fit (equation (6.26)) the uncertainties define the relative weight of the functional models. Without any modifications, the mean of the chi-square over the number of degrees (NDF) distributions are 1.65 for the arrival time fit and 1.79 for the energy fluence fit. These values have been determined after applying the cuts used in section 6.6. This corresponds to underestimating the sum of the model uncertainty and measurement uncertainty by the square root of these values. Thus, the uncertainties have been enhanced by these factors for a better balance of the functional models.

6.6. Validation of the module

In the following, the implementations of the energy fluence fit and arrival time fit in the “RdGlobalFit” module is checked against the implementations in the univariate analysis modules. Therefore, the “RdGlobalFit” module is configured such that it also performs only the corresponding univariate fit.

6.6.1. Arrival time fit

The events used in this comparison are selected on the basis of the reconstruction result of the “RdGlobalFit”. Table 6.1 shows the applied chain of selection criteria. For every stage, the total number of events, which fulfil the criteria up to this point are given. The total fraction w.r.t. the full number of simulated events is shown in the third column. The fourth column displays the fraction of rejected events at this stage w.r.t. the number of events before its application.

Selection criteria	Nr. of events	total fraction	rejected events
All events of the simulation library	5292	100 %	
Number of stations with pulse ≥ 7	715	14 %	-86 %
Inclination of shower $\leq 55^\circ$	356	7 %	-50 %
Valid minimum (reported by MINUIT)	250	5 %	-30 %
Slope parameter not at limit ($51 \text{ m} \leq \gamma \leq 299 \text{ m}$)	168	3 %	-33 %
p-value of fit > 0.001	168	3 %	0 %
Relative slope parameter error $\Delta\gamma/\gamma < 0.4$	132	2 %	-21 %

Table 6.1.: Presentation of the quality cuts and their effect on the data set. The last column displays the relative loss of events w.r.t. the preceding line.

In principle, one can divide the cuts into three classes:

- The first class reduces the events to the number of events, which belong to the standard reconstruction regime (i.e. inclination $\theta \leq 55^\circ$) and theoretically can be described by the functional model (nr. of stations ≥ 7). This fraction amounts only to (7 %) for a hyperbolic wavefront model.
- The second class of cuts selects successfully reconstructed events. The implemented functional model in the “RdGlobalFit” can describe only half of the standard reconstruction events (168/356) with a satisfactory p-Value > 0.001 . The “RdHyperbolicWavefrontFit” (revision 28871) is more robust and can in principle reconstruct 70 % (i.e. 248 events) of the events for the standard reconstruction. The higher robustness is due to the fact that in contrast to the “RdGlobalFit”, the “RdHyperbolicWavefrontFit” does not fit all model parameters. Instead it does a core search: It fits only a model for the direction, time offset and slope parameter $t(\theta, \phi, \gamma, t_0)$ for many different core positions (X, Y) . Then, the fit with the best minimum according to an M-estimator is taken as a result [58]. This method is a more robust solution suited for the univariate ansatz. By design, the core search method cannot provide an error estimation for the core position and its resolution is limited by the number of generated impact points. The fact that many events are reconstructed with an slope parameter at the fit limit shows that the arrival time fit suffers from numerical problems.
- The last cut on the relative error of the reconstructed slope parameter γ of the arrival time distribution is only motivated by the error distribution. It is performed to increase the sensitivity to the geometrical distance to the shower maximum R_{\max} .

In the following the reconstructed subset of 132 events is considered. The mean of the χ^2/NDF distribution is 1.09 despite the reweighting described in previous section 6.5.5. This is caused by a minor change in the event set which survives the cuts. Especially the p-value cuts are sensitive to the assumed uncertainty.

Reconstruction of the shower geometry

The reconstruction deviation of the impact point w.r.t. the Monte Carlo truth has been calculated by propagating the core along the arrival direction to the X-Y-plane of the Monte Carlo coordinate system.

This special coordinate system has the true impact point as origin. The z-axis is perpendicularly oriented to the earth surface. The quantity $\Delta R = \sqrt{\Delta X^2 + \Delta Y^2}$ is a commonly used quality criterion for the reconstruction. However, one should keep in mind that it does not denote the minimal distance of the true impact point to the reconstructed axis which can be smaller. But, in combination with the solid angle between the reconstructed and true direction, this method provides an intuitive criterion for the quality of the reconstructed geometry.

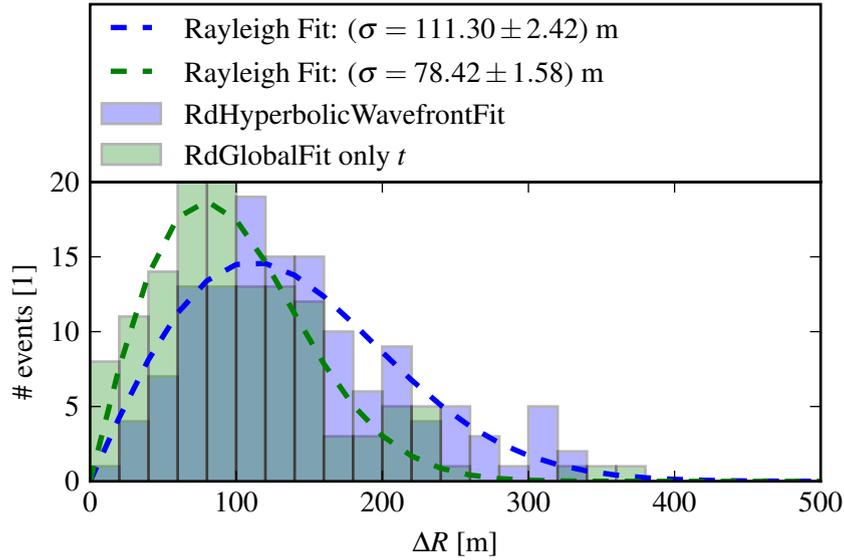


Figure 6.10.: Deviation of the reconstructed impact point to the truth in the x-y-plane of the Monte Carlo coordinate system (see text). A Rayleigh distribution is fitted to the data. The width parameter σ describes the standard deviation along the x-axis and the y-axis (ΔX , ΔY) assuming the variables to be independent and equally distributed.

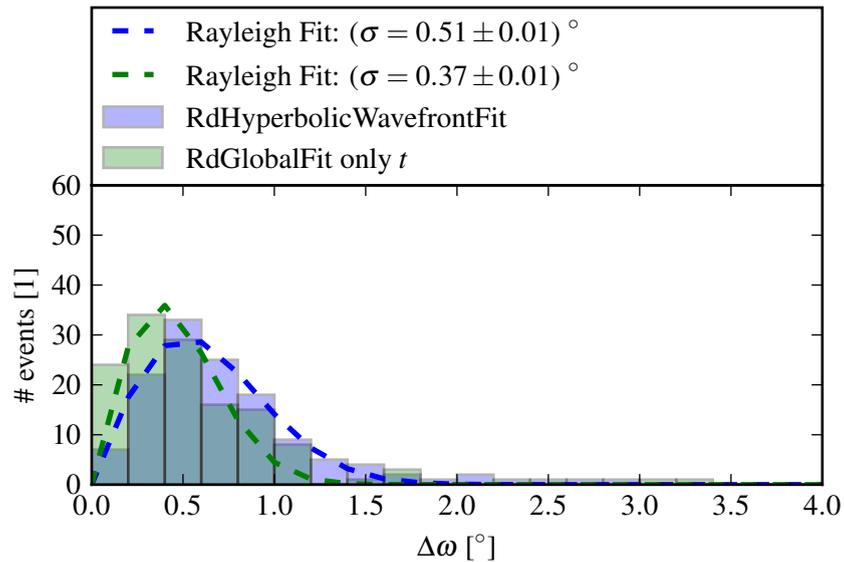


Figure 6.11.: Comparison of the direction reconstruction performance of the “RdGlobalFit” module against the “RdHyperbolicWavefrontFit” module. The opening angle between the reconstructed and true arrival direction is described by a Rayleigh function.

Figure 6.10 shows the distance of the reconstructed impact point to the truth in the x-y-plane of the Monte Carlo coordinate system. The angle included between the reconstructed directions is displayed

in figure 6.11. Assuming an equal reconstruction quality for the core in the stochastically independent ΔX and ΔY direction, the deviation in the shower plane ΔR can be described by a Rayleigh distribution. Also, small angular differences can be fitted with a Rayleigh function. The fit parameter σ describes directly the standard deviation of the independent random variables (ΔX , ΔY).

The reconstruction quality of the arrival time module of the “RdGlobalFit” reconstructs the air shower geometry significantly better than the more robust standalone implementation.

Sensitivity to the distance to the shower maximum

The minor axis γ of the hyperbolic arrival time model (cf. section 6.3.1) describes the slope of the asymptotic cone. Thus, it describes directly the opening angle of the hyperbola which is sensitive to the geometrical distance R_{\max} from the shower impact point on ground to the shower maximum. The

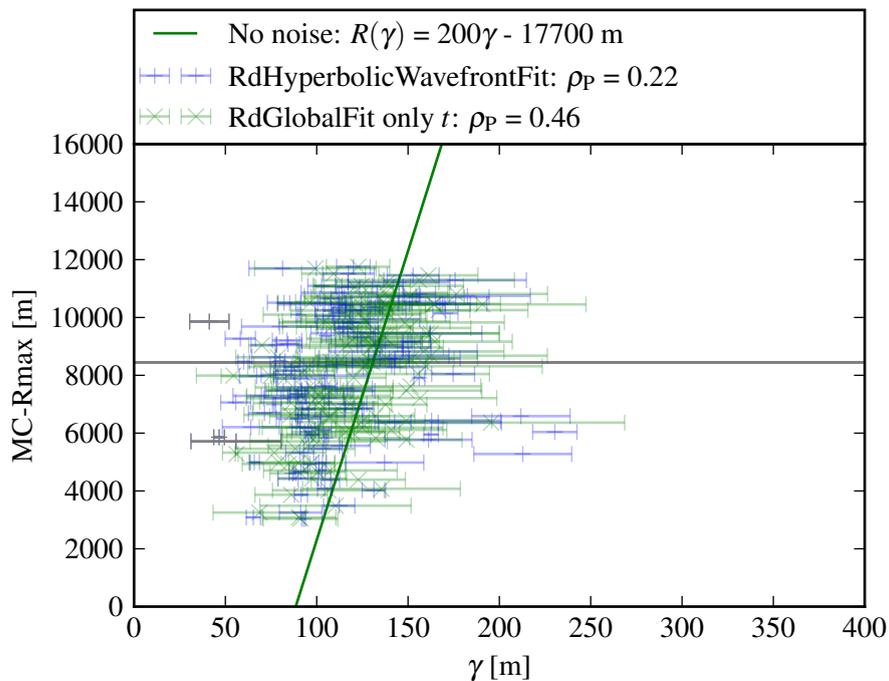


Figure 6.12.: The sensitivity of the slope parameter γ to the geometrical distance R_{\max} from ground to the shower maximum. For this subset, the Pearson correlation coefficient ρ_P of the slope parameter γ and R_{\max} has a higher value for the reconstruction with the “RdGlobalFit” than for the “RdHyperbolicWavefrontFit “. Six events (grey) from the “RdHyperbolicWavefrontFitter” have not been used for the calculation of the correlation. Additionally, the linear relation $R_{\max}(\gamma)$ from Monte Carlo simulations without noise is shown (cf. fig. B.1).

relation between these two quantities is linear in the first approximation. However, under realistic conditions (i.e. with noise added to the simulated signal traces) the correlation diminishes. The Pearson correlation coefficient ρ_P is a measure for the correlation of two variables (a, b). It is defined as the ratio of the covariance over the individual standard deviations:

$$\rho_P = \frac{\text{cov}(a, b)}{\sigma_a \sigma_b} \quad (6.30)$$

Figure 6.12 shows the relation between the reconstructed slope parameter γ and the distance to the shower maximum R_{\max} . Six of the events, which have been reconstructed with the

“RdHyperbolicWavefrontFitter” do not fulfil the quality criteria for the slope parameter γ (cf. table 6.1). They have not been used for the calculation of the correlation. They are marked grey in the illustration. The correlation between γ and R_{\max} is low for both reconstructions. Still, for the given data set, the quantities reconstructed with the “RdGlobalFit” exhibits a higher correlation coefficient ($\rho_P = 0.46$) than the ones reconstructed with the “RdHyperbolicWavefrontFitter” ($\rho_P = 0.22$). The expected linear relation

$$R(\gamma) = 200\gamma - 17700\text{m} \quad (6.31)$$

as determined from simulations without added noise and with a small Gaussian jitter of 0.1 ns (cf. figure B.1) is also shown in figure 6.12 as line for comparison.

6.6.2. Energy fluence fit

The reasoning in here is the same as in section 6.6.1. The fit of the energy fluence with the “RdGlobalFit” and with the “Rd2dLDFitter” are implemented very similar. The complete neglect of stations with a signal below threshold in the “RdGlobalFit” is one difference to the univariate implementation. Additionally, the start values of the impact point parameters are not varied in the global regression implementation.

Table 6.2 shows the cuts which have been applied to the reconstruction results from energy fluence fit of the “RdGlobalFit” module. The direction is set to the reconstruction result from the plane wave

Selection criteria	Nr. of events	total fraction	rejected events
All events of the simulation library	5292	100 %	
Number of stations with pulse ≥ 5	1304	25 %	-75 %
Inclination of shower $\leq 55^\circ$	852	16 %	-35 %
Valid minimum (reported by MINUIT)	844	16 %	-1 %
Width not at limit ($51\text{ m} \leq \sigma \leq 299\text{ m}$)	750	14 %	-11 %
p-value of fit > 0.001	719	14 %	-4 %
Relative width error $\Delta\sigma/\sigma < 0.1$	310	6 %	-57 %
(Relative width error $\Delta\sigma/\sigma < 0.4$)	621	12 %	-14 %

Table 6.2.: Presentation of the quality cuts and their effect on the data set. The last column displays the relative loss of events w.r.t. the preceding line. The last line shows the result for a relaxed cut on the relative lateral distribution reconstruction error.

fit. Due too fewer parameters (X, Y, A, σ), the energy fluence fit needs fewer signal stations for the minimisation. A perfect model could in principle describe 16 % of the simulated data of the standard reconstruction (inclination $\theta < 55^\circ$). Thus, the energy fluence fit can a priori process more than twice the fraction of the events compared with the arrival time fit. Moreover, the regression leads for 85 % (719 of 852) of the events to results which are compatible with the functional model (p-value > 0.001). Thus, the reconstruction capability is significantly higher than the one of the arrival time model (47 %, 168 of 356 events, cf. table 6.1). In total, the energy fluence fit reconstructs more than twice the number of recorded events w.r.t. the arrival time fit due to the better reconstruction capability and the lower number of required stations. The lower number of signal stations is additionally favoured by the standard reconstruction cut on the shower inclination.

The resulting data-set is reduced by more than a factor of two, if one applies a sharp cut on the relative error of the reconstructed lateral distribution width $\Delta\sigma/\sigma < 0.1$. In contrast to the cut on the slope parameter γ , the cut on the lateral distribution width σ is not justified by the error distribution. Here, it is chosen for consistency with the results given in [46], only.

However, the selection of such a high quality data set, which contains only 6 % of the recorded events contradicts the reasoning of having a duty-cycle of nearly 100 % with radio measurements.

The mean of the χ^2/NDF distribution is 0.996 due to the reweighting of the uncertainties described in section 6.5.5. The lower sensitivity of the fit to changes in the weighting confirms the higher robustness of the fit compared to the time fit.

Reconstruction of the shower impact point

The reconstruction quality of the shower impact point is similar for the energy fluence fit modules ‘‘RdGlobalFit’’ and ‘‘Rd2dLDFitter’’. The standard deviation in the x and y direction is roughly 10 m (cf. figure 6.13). For a detailed description of the plotted observable R refer to section 6.6.1.

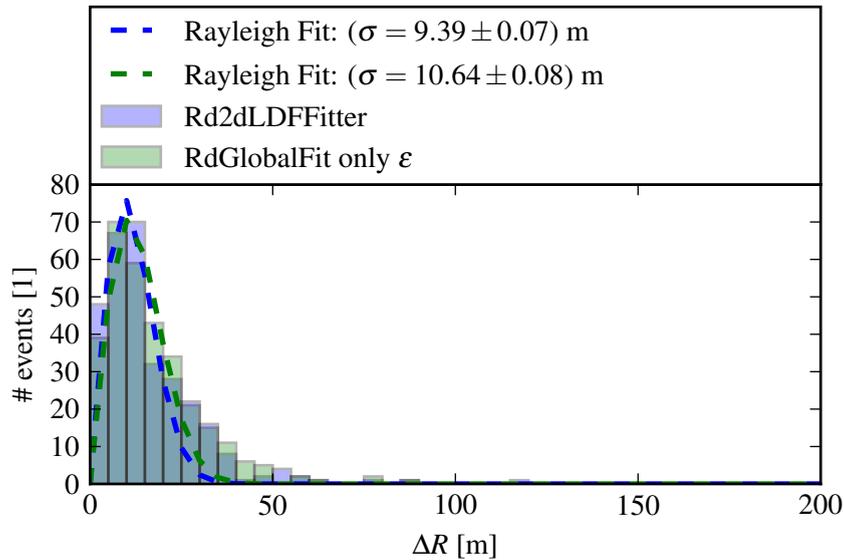


Figure 6.13.: Deviation of the reconstructed impact point to the truth in the x - y -plane of the Monte Carlo coordinate system (cf section 6.6.1). A Rayleigh distribution is fitted to the data. Here, the width parameter σ describes the standard deviation along the x -axis and the y -axis (ΔX , ΔY) assuming the variables to be independent and equally distributed.

Energy reconstruction

The radiation energy [44,50] in the measured frequency band can be calculated by integrating the lateral distribution function over the shower plane (x, y). The energy of the primary particle is proportional to the square-root of the radiation energy. For AERA measurements, the relation implemented in all Offline modules reads:

$$E_{CR} = 2.3137 \left(\frac{\int dx dy \varepsilon}{\sin^2(\alpha_B)} \right)^{0.506} \quad (6.32)$$

which solves equation (4.1) from [44] within the uncertainties. Figure 6.14 plots the results for both modules. The relative shift by -25% reflects the fact that equation (6.32) has been determined for real

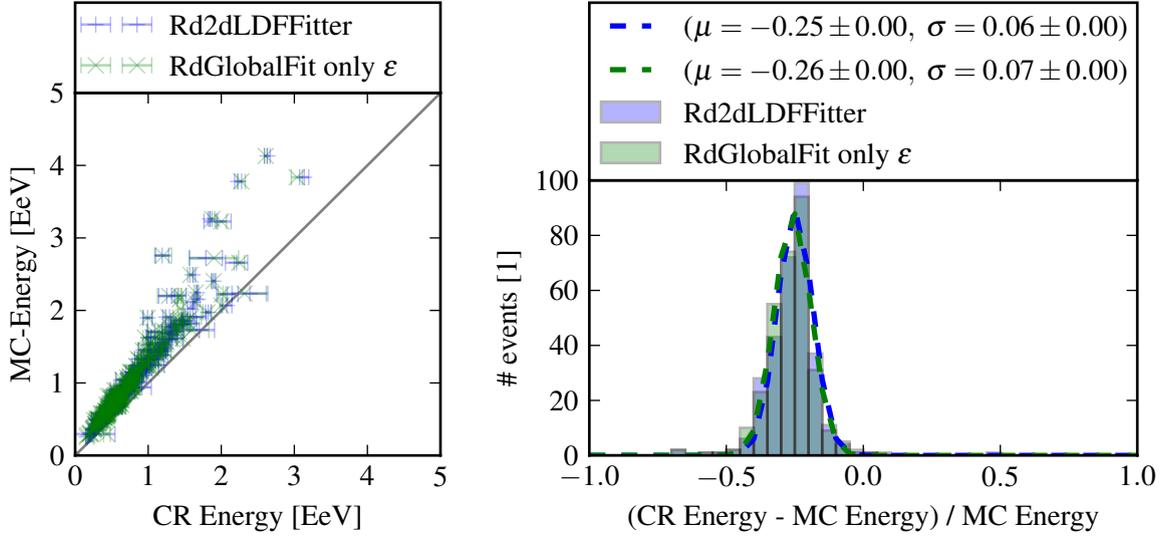


Figure 6.14.: Comparison of the reconstructed cosmic ray energy versus the true Monte Carlo energy. The energy calibration from [44, 50] (eq. (6.32)) has been used for the energy reconstruction.

data. The number of muons in measurements and Monte Carlo simulations (cf. 3.2.3) differs. This leads to a discrepancy between the energy reconstruction with the surface detector for measured and simulated data.

The proper energy resolution is taken from a comparison to the reconstructed energy shifted by a factor of 1.37%. The mean energy deviation of the global regression to the Monte Carlo truth is then zero and the right fitted energy resolution σ in the right panel amounts to 9%.

Sensitivity to the distance to the shower maximum

The Pearson correlation coefficient ρ_P of the lateral distribution width σ_+ and the geometrical distance to the shower maximum is comparable for the implementation of the energy fluence fit in the “RdGlobalFit” ($\rho_P = 0.94$) and the “Rd2dLDFitter” module ($\rho_P = 0.94$). Two events, which have been reconstructed with the “Rd2dLDFitter” module but do not fulfil the quality cuts on the lateral distribution with σ_+ , have been ignored in the calculation of the correlation. Figure 6.15 displays the true distance to the shower maximum R_{\max} and the lateral distribution width σ_+ in a scatter plot. Additionally, the Pearson correlations and the result of a weighted linear fit are displayed.

The data can be well described by the linear relation:

$$R_{\max} = (65 \pm 1)\sigma_+ - (3400 \pm 200) \text{ m} \quad (6.33)$$

This relation is in agreement with the results from the application of the “Rd2dLDFitter” on the “RdSimObserver” simulation set [46]:

$$R_{\max} = (65.0 \pm 0.8)\sigma_+ - (3129 \pm 121) \text{ m} \quad (6.34)$$

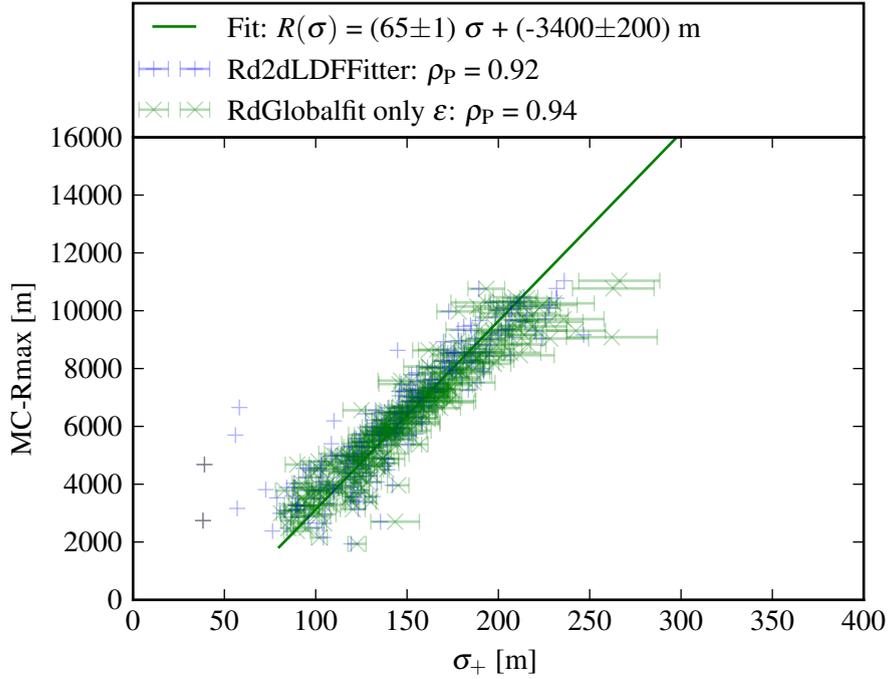


Figure 6.15.: Correlation between the width σ_+ of the fitted lateral distribution function and the geometrical distance R_{\max} from the impact point on ground to the air shower maximum. The Pearson correlation coefficient ρ_p is calculated for the results from the Rd2dLDFitter module and the RdGlobalfit module ignoring the grey points from the Rd2dLDFitter. The green line denotes the result of a weighted linear fit to the RdGlobalfit reconstruction.

6.7. Results of the global regression

In the following, the prospect of the global regression (cf. section 6.5) is examined. Therefore, the reconstruction results from the new module “RdGlobalFit” are compared when only turning on the arrival time fit (t), or the energy fluence fit (ε), or when performing the global regression ($t + \varepsilon$).

6.7.1. Reconstruction performance

The cuts on the reconstruction results are the same as described in tables 6.1 and 6.2. The only difference to the validation is the relaxation of the cut on the relative uncertainty of the lateral distribution width $\Delta\sigma_+/\sigma_+$ from 0.1 to 0.4. The cuts on the results of the global regression are shown in table 6.3. In

Selection criteria	Nr. of events	total fraction	rejected events
All events of the simulation library	5292	100 %	
Number of stations with pulse ≥ 5	1304	25 %	−75 %
Inclination of shower $\leq 55^\circ$	852	16 %	−35 %
Valid minimum (reported by MINUIT)	556	11 %	−35 %
Slope parameter not at limit ($51 \text{ m} \leq \gamma \leq 299 \text{ m}$)	392	7 %	−29 %
Width not at limit ($51 \text{ m} \leq \sigma \leq 299 \text{ m}$)	349	7 %	−11 %
p-value of fit > 0.001	297	6 %	−15 %
Relative slope parameter error $\Delta\gamma/\gamma < 0.4$	211	4 %	−29 %
Relative width error $\Delta\sigma/\sigma < 0.4$	197	4 %	−7 %

Table 6.3.: Presentation of the quality cuts and their effect on the data set. The last column displays the relative loss of events w.r.t. the preceding line. The last line shows the result for a relaxed cut on the relative lateral distribution reconstruction error.

principle, the global regression needs only 5 signal stations for the fit of 8 model parameter due to the independence of the arrival time and energy fluence measurements. Thus, a perfect functional model could describe the same amount of data as the energy fluence fit (16 %) with an additional sensitivity to the arrival direction. But the increased model complexity and the higher number of simultaneously fitted parameters leads to a reduced reconstruction capability in comparison to the simple energy fit. The “RdGlobalFit” converges for only 65 % of standard reconstruction events. Moreover, a significant amount of events are reconstructed such that the slope parameter γ and lateral distribution width σ are at the limits, which have been set to $\{\sigma, \gamma\} \in [50, 300]$ m. Finally, 197 events remain after applying the quality cuts. This is a high number compared to the arrival time fit (132 events), but also significantly lower number with respect to the energy fluence fit (621 events).

Figure 6.16 displays the connection between the sets of reconstructed events when fitting the arrival time and energy fluence models to the data individually and simultaneously. In total 649 events

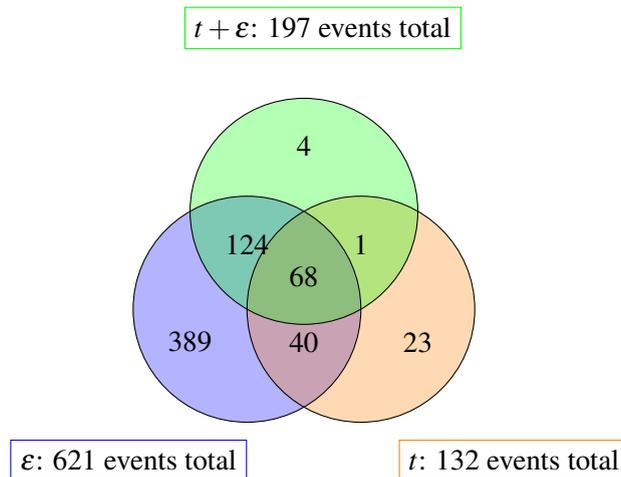


Figure 6.16.: Illustration of the number of events reconstructed with different modules of the RdGlobalFit activated. A valid reconstruction with the energy fluence fit (ϵ) only is a precondition for the success of the global regression ($t + \epsilon$).

are displayed. The energy fluence fit (ϵ) reconstructs 96 % of these events and 60 % (389/651) are exclusively reconstructed by this fit. Furthermore, the successful reconstruction with the energy fit is an essential precondition for the global regression to work. Only 3 % of the events which have been reconstructed with the global regression analysis have not been reconstructed by the energy fluence fit alone. The picture changes when looking at the intersection of the events reconstructed with the global regression approach with the events reconstructed with the arrival time model. Slightly more than half 69/132 of the events which can be reconstructed by the arrival time model are reconstructed by the global fit. Especially the 40 events which can be reconstructed by both standalone regressions individually but not with the combined global fit are conspicuous, as they should be perfect candidates for the global regression. Therefore, they are discussed here in more detail: The fit converges for 22/40 events with valid parameters for σ and γ . Only 3 events out of these 22 can be described by the functional model, i.e. the p-value exceeds 0.001. These remaining events are finally rejected by the quality cuts for σ and γ . Thus, the 40 events are the manifestation of the lower reconstruction capability of the global regression, probably due to the higher dimension of the parameter space and resulting numerical problems.

Therefore, the higher number of reconstructed events with the global regression w.r.t. the time fit is mostly due to the 124 events which are reconstructed by the energy fluence fit and the global fit. Two thirds of these events (84/124) have been recorded by less than seven stations. Thus they are can not be reconstructed by the arrival time fit, per se. Interestingly, the other third (40/125) could have been reconstructed by a perfect and robust arrival time model, but it is only reconstructed by the global regression.

In conclusion, the performance of the global fit is derogated by numerical instabilities due to its higher complexity.

6.7.2. Reconstruction quality

The following plots are based on the 68 events reconstructable by all three methods (c.f. fig. 6.16). Despite the low statistics, this allows for a good of the reconstruction quality of the functional models.

Geometry

Figure 6.17 shows the deviation $\Delta R = \sqrt{\Delta X^2 + \Delta Y^2}$ of the reconstructed impact point from the Monte Carlo truth. The precision of the reconstruction with the global regression is compatible to the reconstruction with the energy fluence fit within the 95 %-confidence interval according to the spread σ of the fitted Rayleigh distribution. The expected improvement of the reconstruction from a weighted

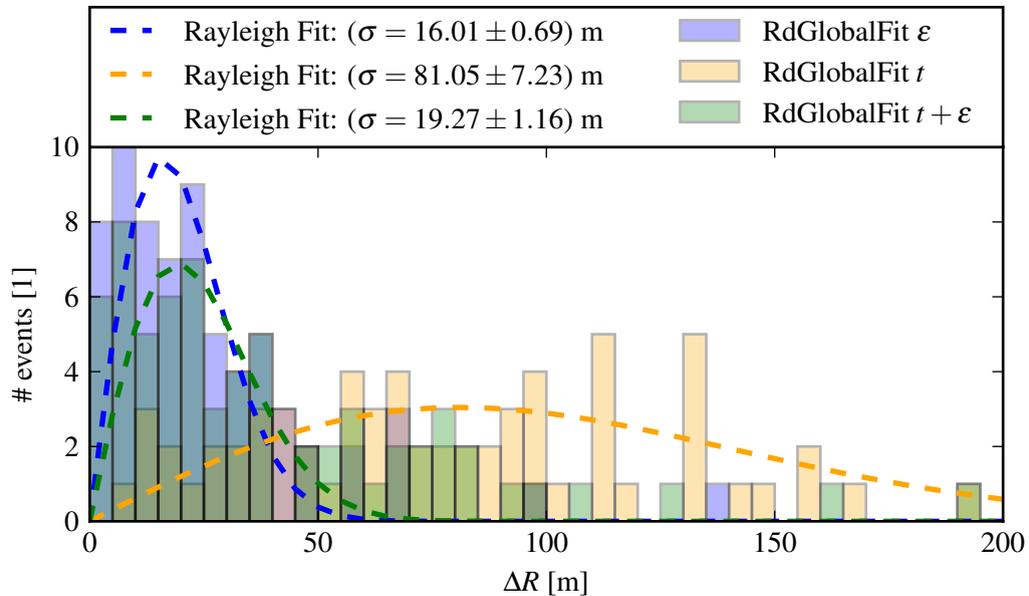


Figure 6.17.: Deviation of the reconstructed impact point to the truth in the x-y-plane of the Monte Carlo coordinate system (cf section 6.6.1). A Rayleigh distribution is fitted to the data. Here, the width parameter σ describes the standard deviation along the x-axis and the y-axis (ΔX , ΔY) assuming the variables to be independent and equally distributed.

fit can be estimated from equation (A.5). The uncertainty of the individual core reconstruction

$$\sigma_R = \sqrt{2}\sigma \quad (6.35)$$

can be calculated from the widths of the Rayleigh distribution σ . The uncertainty fraction n in equation (A.5) is therefore:

$$n = 81/16 \quad (6.36)$$

The quotient of the weighted mean uncertainty $\sigma_{\bar{R}}$ to the energy fluence reconstruction uncertainty $\sqrt{2}\sigma_R$ is thus

$$\frac{\sigma_{\bar{R}}}{\sqrt{2}\sigma_R} = \frac{n}{\sqrt{1+n^2}} \approx 0.98 \quad (6.37)$$

The expected improvement of 2 % corresponds to roughly 0.4 m and cannot be resolved. However, the number of outliers with $\Delta R > 50$ m in the global fit reconstruction is increased w.r.t. the pure energy fluence fit. These outliers are not properly described by the Rayleigh distribution and may be a result of the lower fit stability of the global regression.

The direction reconstruction, on the contrary, benefits from additional information. Figure 6.18 compares the angular deviations $\Delta\omega$ of the reconstructed shower axes with the global regression and with the arrival time fit. In the e_θ and e_ϕ direction, the angular standard deviation of the reconstruction

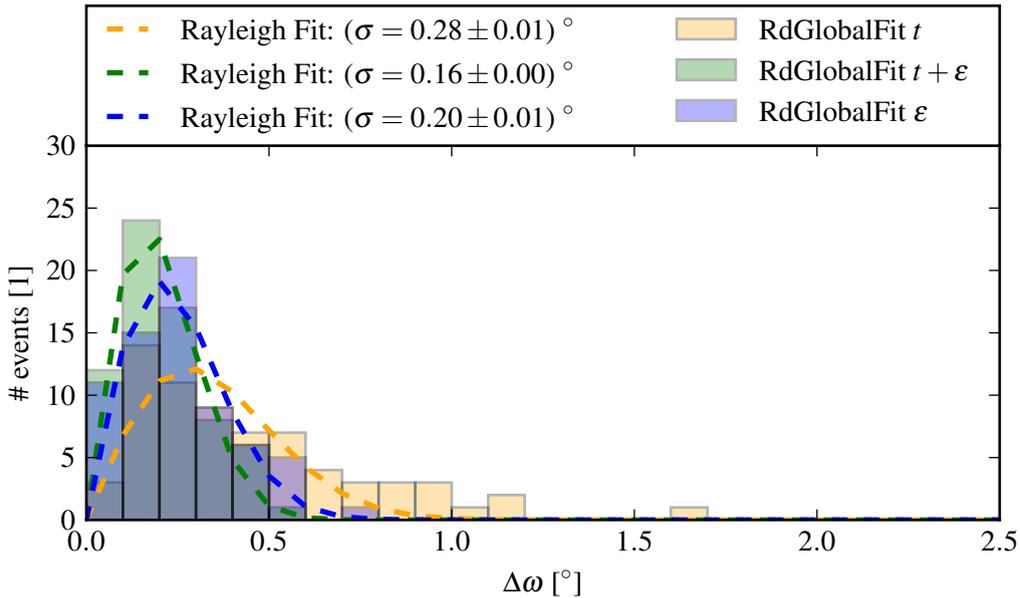


Figure 6.18.: Comparison of the direction reconstruction performance of the fit methods. The opening angle between the reconstructed and true arrival direction is described by a Rayleigh function.

with the global regression approach is almost halved w.r.t. the pure arrival time fit. The achieved precision, however, is better than required for typical applications in astroparticle physics except for gamma astronomy.

Energy reconstruction

The arrival time fit is not sensitive to the cosmic ray energy. Therefore, the global regression approach is not expected to improve the sensitivity to the cosmic ray energy. The compatible standard deviation σ

of the energy reconstruction in figure 6.19 confirms this. The -26% offset of the energy calibration is of not meaningful in this context (cf. section 6.6.2). The energy resolution of the corrected CR-Energy (i.e. shifted by the factor $1/0.74$ such that the mean deviation vanishes for the global regression) is 8% .

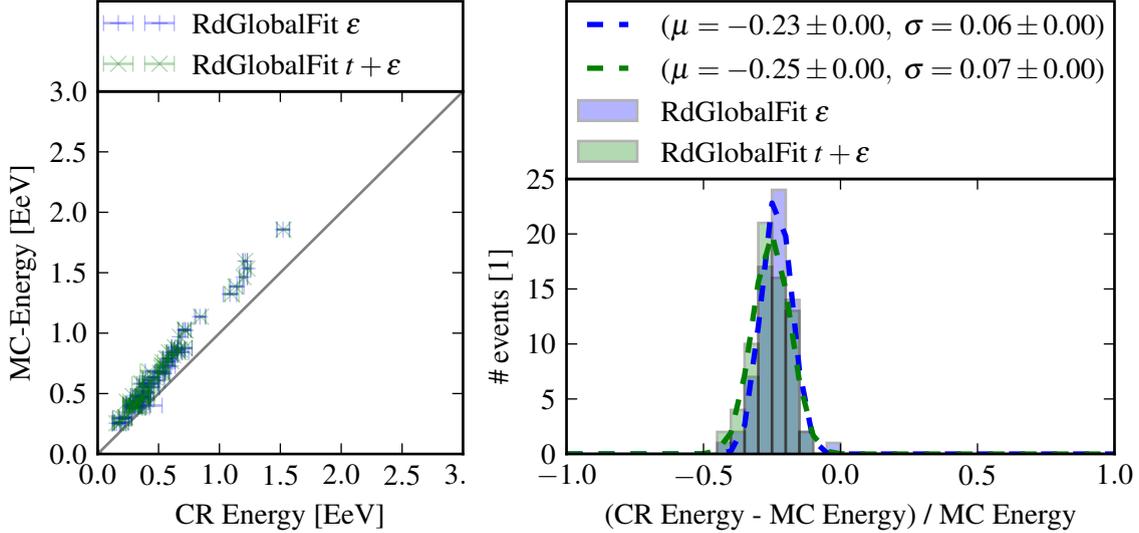


Figure 6.19.: Comparison of the reconstructed cosmic ray energy versus the true Monte Carlo energy. The energy calibration from [44, 50] (eq. (6.32)) has been used for the energy reconstruction.

Distance to the shower maximum

The atmospheric depth X_{\max} of the shower maximum can be calculated from the geometrical height h of the shower maximum above sea level assuming an exponential air density [59]:

$$X_{\max} = a_i + b_i e^{-h/c_i} \quad (6.38)$$

In this model, the atmosphere is vertically divided in four slices i . Depending on the altitude h of the shower maximum, the parameters a_i , b_i and c_i take different values. For the following reconstruction, the US standard atmosphere [60] parametrisation from [59] is adopted:

Layer	altitude [km a.s.l.]	a	b	c
1	0.0 - 7.0	-149.801663	1183.6071	954248.34
2	7.0 - 11.4	-57.932486	1143.0425	800005.34
3	11.4 - 37.0	0.63631894	1322.9748	629568.93
4	37.0 - 100.0	4.35453690e-4	655.67307	737521.77

Table 6.4.: The parametrisation of the US standard atmosphere model from [59].

Figure 6.20 displays a check of this method for the discussed 68 events. The distance to the shower maximum $X_{\max}(R_{\max, \text{MC}}, \theta_{\text{MC}})$ is reconstructed from the true geometric distance to the shower maximum $R_{\max, \text{MC}}$, and the true inclination θ_{MC} assuming the altitude of the impact point to be 1564m above sea level. The reconstructed shower maximum is compared to the Monte Carlo

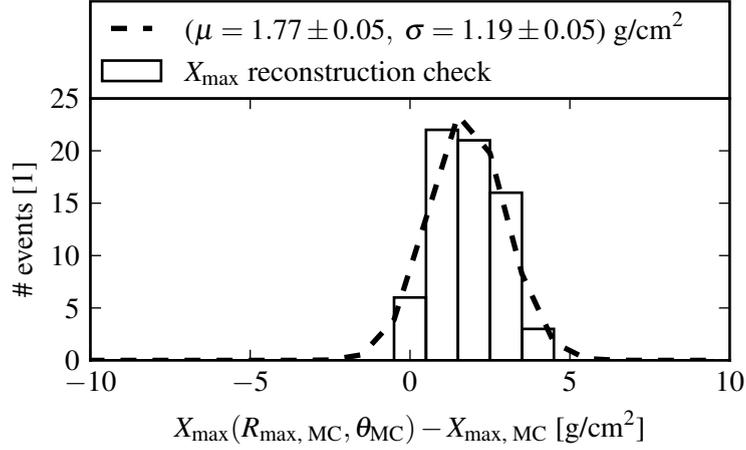


Figure 6.20.: Crosscheck of the $X_{\max}(h(R_{\max}, \theta))$ reconstruction with equation (6.38). The atmospheric depth reconstructed from the true values $X_{\max}(R_{\max, MC}, \theta_{MC})$ is compared to the true atmospheric depth $X_{\max, MC}$.

truth. The reconstructed value is systematically 1.8 g cm^{-2} higher with respect to the Monte Carlo truth and has a precision of 1.2 g cm^{-2} . This additional uncertainty is negligible in comparison to the uncertainties discussed in the following.

Reconstruction with the lateral distribution function: Figure 6.21 shows the reconstruction of the geometrical distance to the shower maximum R_{\max} from the width of the lateral distribution function σ_+ adopting equation (6.33). The geometrical distance to the shower maximum is

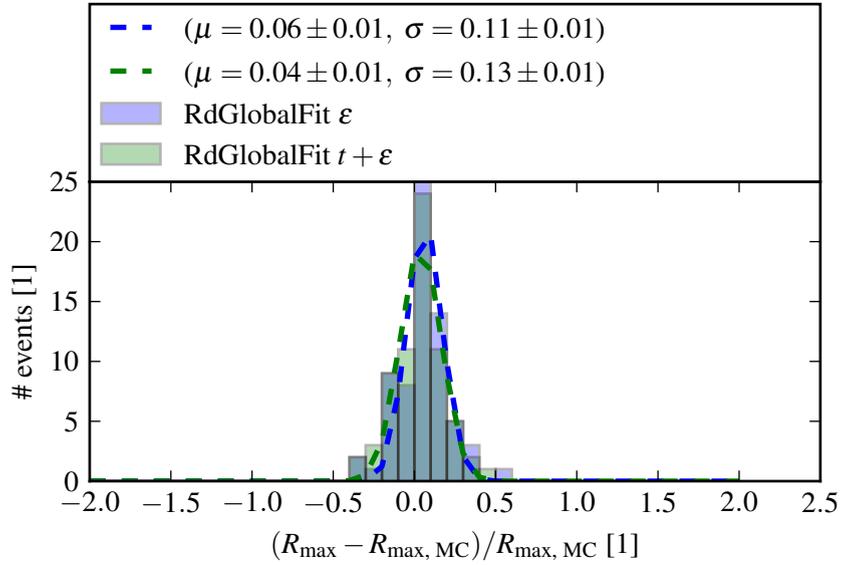


Figure 6.21.: The relative deviation from the distance to the shower maximum R_{\max} to the true value $R_{\max, MC}$.

reconstructed with a high precision (11 %) but its value is generally 4 % larger w.r.t. the Monte Carlo truth. Both methods show the same precision within the uncertainties. This is expected, because the width parameter σ_+ has already been shown to be similar by comparing the energy resolution.

The atmospheric depth of the shower maximum X_{\max} as determined from R_{\max} of the energy

fluence distribution is depicted in figure 6.22. The spread of the reconstructed atmospheric

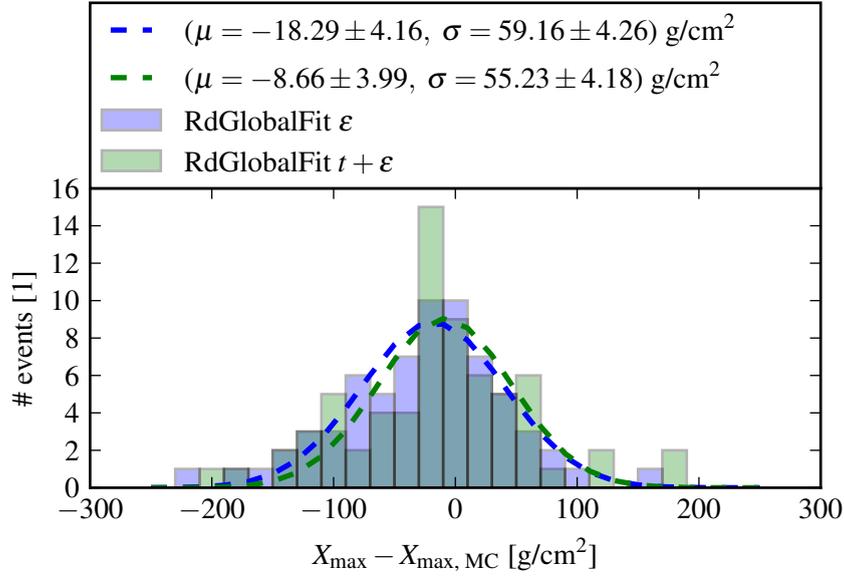


Figure 6.22.: The absolute deviation from the reconstructed atmospheric depth of the shower maximum X_{\max} to the true $X_{\max, MC}$.

depth agrees with the reconstruction results [$\mu = 0.3 \pm 1.7$, $\sigma_+ = 53.5 \pm 1.7$] g cm^{-2}] from the application of the “Rd2dLDFitter” module to simulations [46]. The bias μ has a relative uncertainty of about 50 % and can be adjusted in the range $\mu \in [-9, 7]$ g cm^{-2} by varying the offset of relation (6.33) within its uncertainty in the interval [-3200,-3400] m. Therefore, the bias of the X_{\max} is not significant.

Reconstruction from the arrival time distribution: Figure 6.23 shows the relative reconstruction quality of the R_{\max} reconstruction. The global regression and the standalone method are mostly

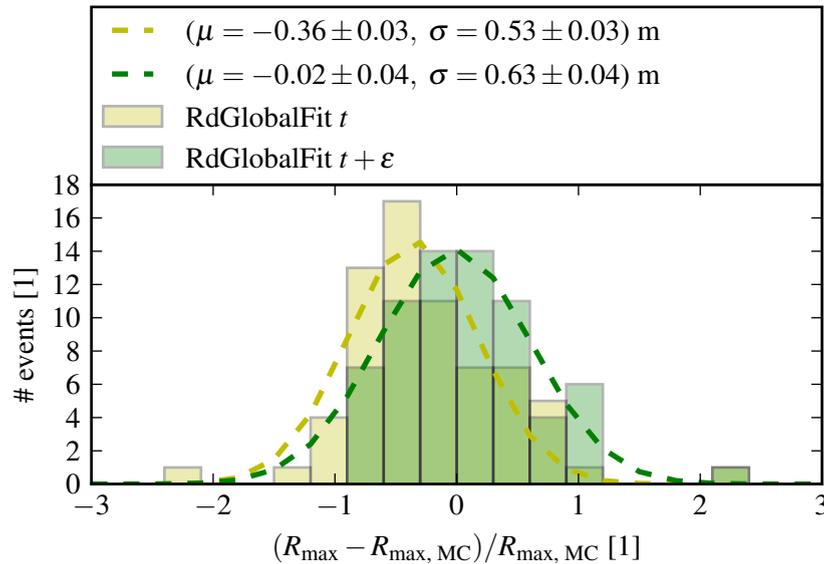


Figure 6.23.: The relative deviation from the distance to the shower maximum R_{\max} to the true value $R_{\max, MC}$.

insensitive to the distance to the shower maximum with relative uncertainties of about 60 %.

The reason for this is discussed in the following section 6.8. Thus, the reconstruction of the atmospheric depth using the slope parameter γ (cf. figure 6.24) fails due to the inaccurate R_{\max} reconstruction.

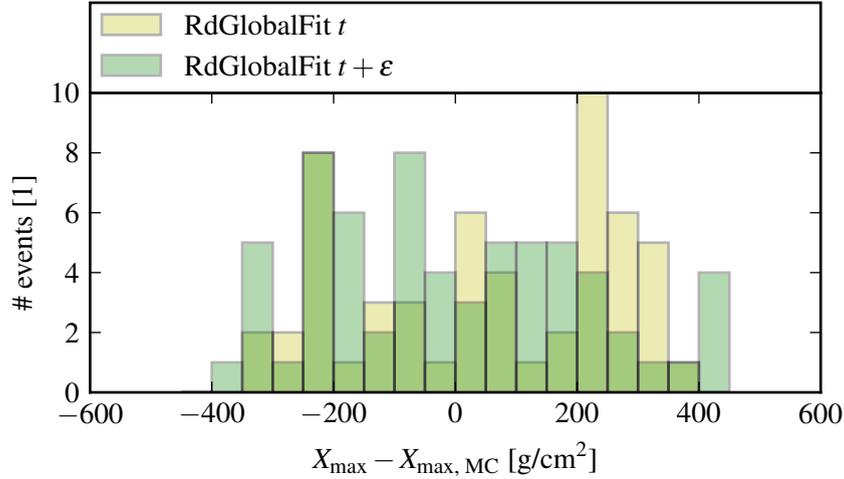


Figure 6.24: The absolute deviation from the reconstructed atmospheric depth of the shower maximum X_{\max} to the true $X_{\max, MC}$.

6.7.3. Application of the global fit to data

The ‘‘RdGlobalFit’’ was applied to data measured in the period from December 2011 to March 2015 with the same quality cuts as described in 6.7.1. The air shower inclination is restricted by the fitted value instead of the Monte Carlo truth. The signal has been upsampled by a factor of 8 for a precise time resolution. Recorded data without a valid beacon correction is discarded. The data set contains 1489 events which have been measured with more than four signal stations. 190 events fulfil all quality criteria.

Figure 6.25 compares the radio energy reconstruction E_{RD} of the ‘‘Rd2dLDFFitter’’ module and the ‘‘RdGlobalFit’’ to the surface detector energy reconstruction E_{SD} . The reconstruction with the ‘‘Rd2dLDFFitter’’ and the ‘‘RdGlobalFit’’ are equal within the uncertainties. This is expected from the previous simulation studies (cf. section 6.6.2). The combined uncertainty of AERA and SD of 24 % confirms the results from [50] (cf. section 4.2.6). For some events, the energy reconstructed with the radio technique exceeds the SD results by more than 100 %. These outliers show the typical signature of measurements during thunderstorms [61]. A scatter plot of the events (figure C.1) is shown in the appendix.

Figure 6.26 shows the comparison of the atmospheric depth reconstructed with the ‘‘RdGlobalFit’’ with the results from the air-fluorescence method. The atmospheric depth from the fluorescence detectors is required to have a relative uncertainty $X_{\max}/\Delta X_{\max}$ better than 0.5. This leaves only 10 events for the comparison. Thus, the statistic is too low for conclusions. However, the difference between the FD and RD reconstruction is shown in figure C.2 for completeness.

The application to data confirms the results of the simulation studies: The global fit reconstructs the

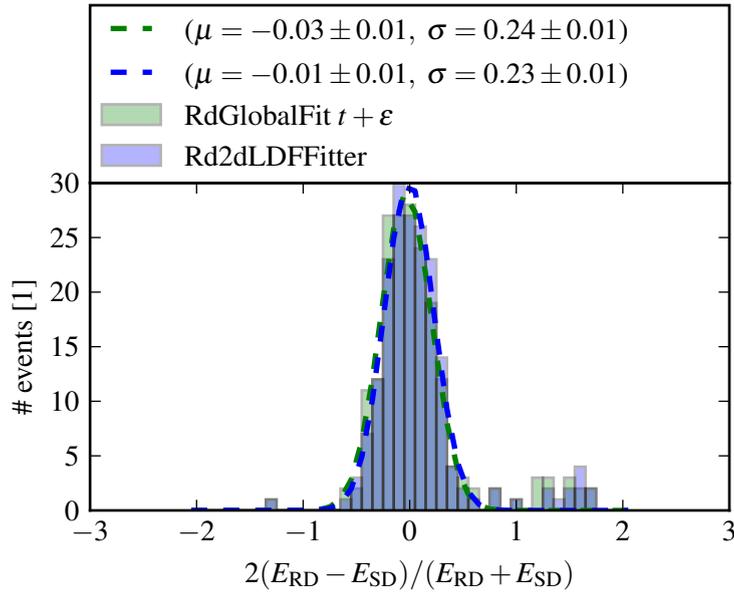


Figure 6.25.: Deviation of the cosmic ray energy reconstructed from radio (RD) and surface detector (SD) measurements with respect to the mean of both energy reconstructions. The radio reconstruction has been performed with the multivariate “RdGlobalFit” module and the univariate “Rd2dLDFitter” module.

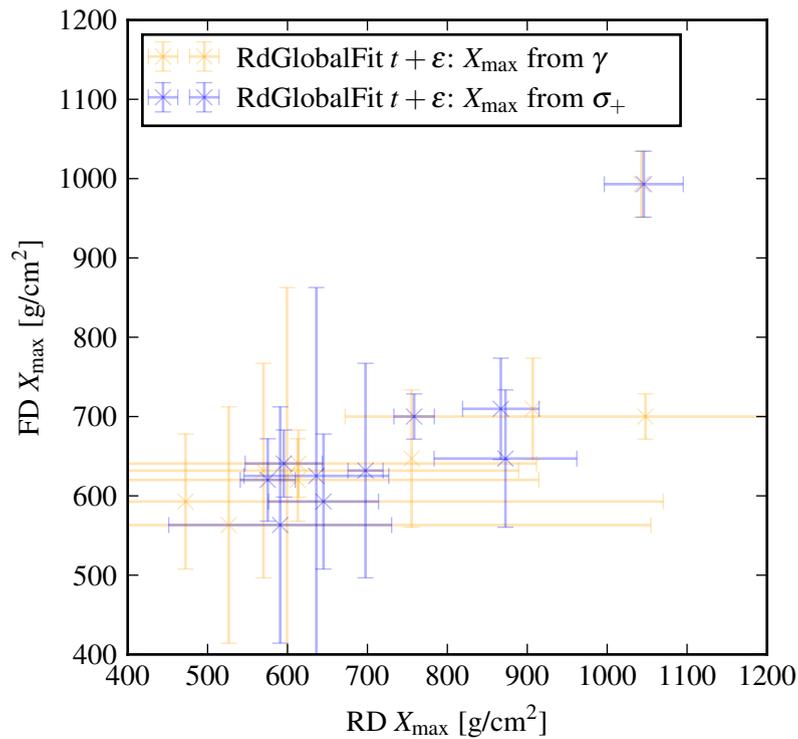


Figure 6.26.: Comparison of the atmospheric depth as reconstructed from AERA measurements (RD) and air-fluorescence measurement (FD). For the radio measurements, the atmospheric depth has been determined from the width σ_+ of the lateral energy fluence distribution and from the slope γ of the arrival time distribution.

parameters from the energy fluence fit — e.g. the cosmic ray energy — with the same precision as the standalone method. But, the higher instability of the fit reduces the reconstruction capability. As a consequence, the global fit reconstructs too few events for a meaningful analysis of the X_{\max} sensitivity.

6.8. Discussion

For the first time, the analysis presented in this work takes the effects from noise and the remaining jitter after the beacon correction into account. For AERA, the latter accounts for at least two thirds of the total time uncertainty. It turns out that the global fit approach does not lead to an improved reconstruction under these realistic conditions: First, the reconstruction performance is reduced due to fit instabilities which emerge from the larger number of degrees of freedom (cf. section 6.7.1). Only 197 events can be reconstructed in a simultaneous fit of the arrival time and energy fluence model. The default approach which determines the air shower observables from the independent and subsequent fit of a plane wave model and the energy fluence model reconstructs 621 events. Secondly, the quality of the reconstruction is not improved by combining the time fit and the energy fluence fit. In the presence of noise, the reconstruction quality of the shower impact point from the time fit is five times less compared with the reconstruction from the energy fluence. The minimal improvement expected from the combination of the information in a weighted mean is lower than the resolution (cf. sec 6.7.2). Only the arrival direction reconstruction profits from the global fit. In future, this finding can be used in the independent and stepwise fit approach by fitting a generic hyperbola with a fixed slope $\gamma \approx 120$ m instead of a plane wave model. The most surprising finding is the insensitivity of the wavefront slope to the primary particle type for realistic simulations and measurements. This result does not meet the expectations at the beginning of this work. However, the promising simulation studies in [51] have been performed without taking any effects of noise into account. Probably, [52] neglects the dominant timing uncertainty from the clock jitter after applying the beacon correction.

In the following sections 6.8.1 and 6.8.2 the (in)sensitivity to the primary particle type is discussed in more detail.

6.8.1. Insensitivity of the radio wavefront to the distance to the shower maximum

Figure 6.27 shows the expected arrival time t as a function of the distance to the shower axis d for two different distances to the shower maximum R_{\max} in the left panel. The slope parameter γ is calculated according to the fit in figure 6.12 via

$$\gamma(R_{\max}) = (R_{\max} + 17700)/200. \quad (6.39)$$

The expected arrival time is calculated from equation (6.8) setting $z_{i,\text{sp}}$ and t_0 to zero. The chosen distances to the shower maximum R_{\max} enclose the possible values of the data set and thus define the phase space of measurable arrival time distributions. The right panel shows the difference of arrival time distributions for these extreme cases. If one assumes a timing uncertainty of 2.5 ns at station level, only arrival time measurements from stations at a minimal distance (> 160 m) to the shower axis differ more than their combined uncertainty of 3.5 ns. Figure 6.28 shows the distances of the signal stations to the shower axis for different event selections. The left panel shows stations from all events with at least 5 signal stations. In the right panel, the shower axis of all events is additionally less inclined than 55° (c.f. table 6.2). Between a quarter to a half of all arrival time measurements are performed at distances below 160 m to the shower maximum. These measurements are insensitive to the geometrical distance to the shower maximum R_{\max} and therefore insensitive to X_{\max} .

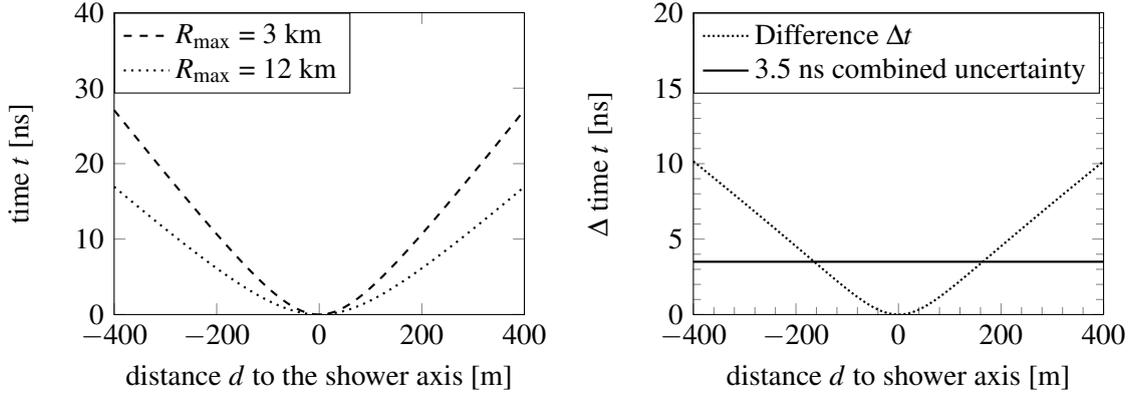


Figure 6.27.: The expected phase space of the arrival time measurement from figure 6.12. The expected arrival times for the two extreme distances to the shower maximum $R_{\max} \in \{3, 12\}$ km (left) and their subtraction (right). Assuming 2.5 ns timing uncertainty per measurement, only stations more than 200 m away from the shower axis can distinguish these extreme scenarios.

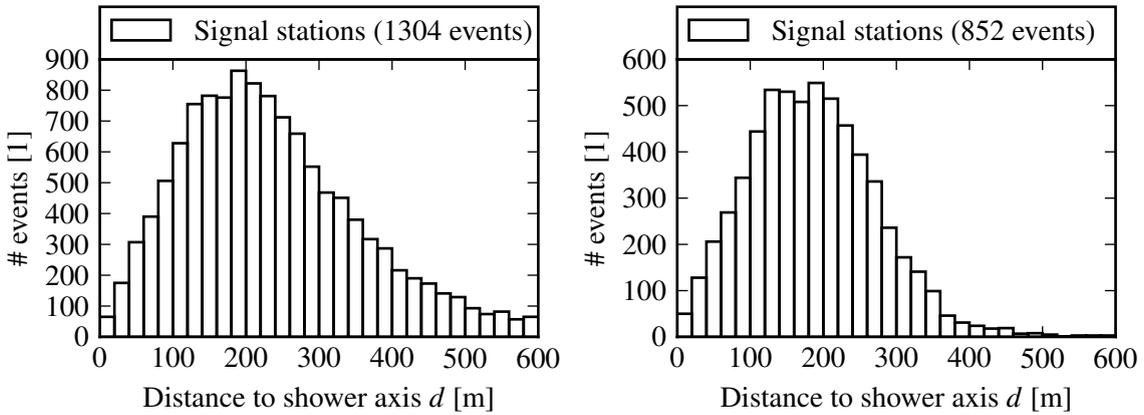


Figure 6.28.: The distances to the shower axis for events with at least 5 signal stations (left) and additionally inclinations below 55° (right). The quantiles are summarised in table 6.5

θ			25 %	50 %	75 %
$\leq 90^\circ$	1304 events	12299 stations	146 m	222 m	319 m
$\leq 55^\circ$	852 events	6154 stations	124 m	182 m	244 m

Table 6.5.: The quantiles of the distances of the signal stations to the shower axis for the distributions shown in figure 6.28. Requiring events with at least 5 signal stations leaves 1304 events with 12299 arrival time measurements from signal stations in total. The additional exclusion of horizontal showers leaves 852 events with 6154 independent measurements.

6.8.2. Insensitivity of the radio wavefront to the atmospheric depth of the shower maximum

The picture gets even worse when estimating the time resolution needed for a statistical discrimination of the primary mass: Figure 6.29 shows an overview on measurements of the mean atmospheric depth X_{\max} [49] for the energy range $10^{16.5}$ eV to $10^{19.8}$ eV. Additionally, the Monte Carlo predictions for X_{\max} are shown as determined from the atmospheric depth of the maximal energy deposit dE/dX for a pure proton and iron composition. The typical energy of air showers measured with AERA is about 5×10^{17} eV $\approx 10^{17.7}$ eV. The prediction for the mean atmospheric depth at $10^{17.7}$ eV, read from figure 6.29, is 606 g cm^{-2} for iron and 714 g cm^{-2} for proton when trusting the QGSJet-II.04 model.

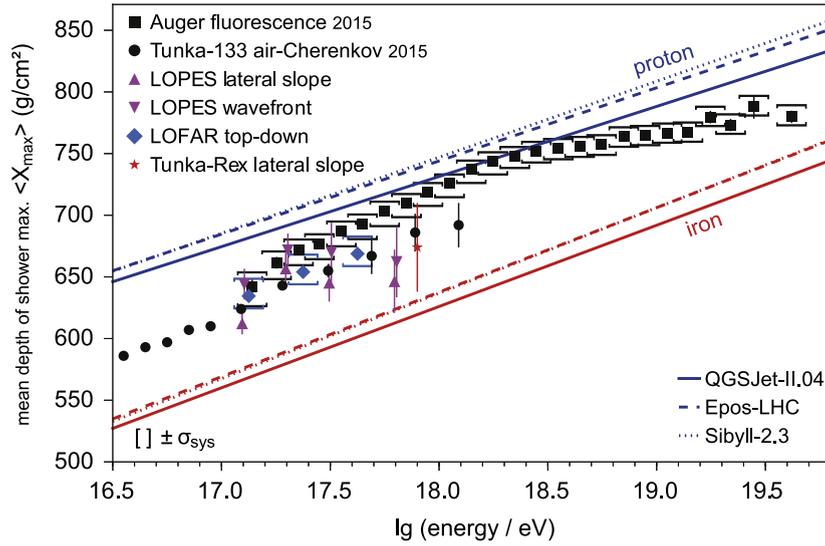


Figure 6.29.: Overview on different measurements of the mean X_{\max} . The predictions proton (blue) and iron primaries (red) by CORSIKA simulations are shown for different hadronic interaction models [49].

The geometrical distance to the shower maximum R_{\max} corresponding to these atmospheric depths X_{\max} can be determined from the atmospheric model (eq. (6.38)). The relation between R_{\max} and

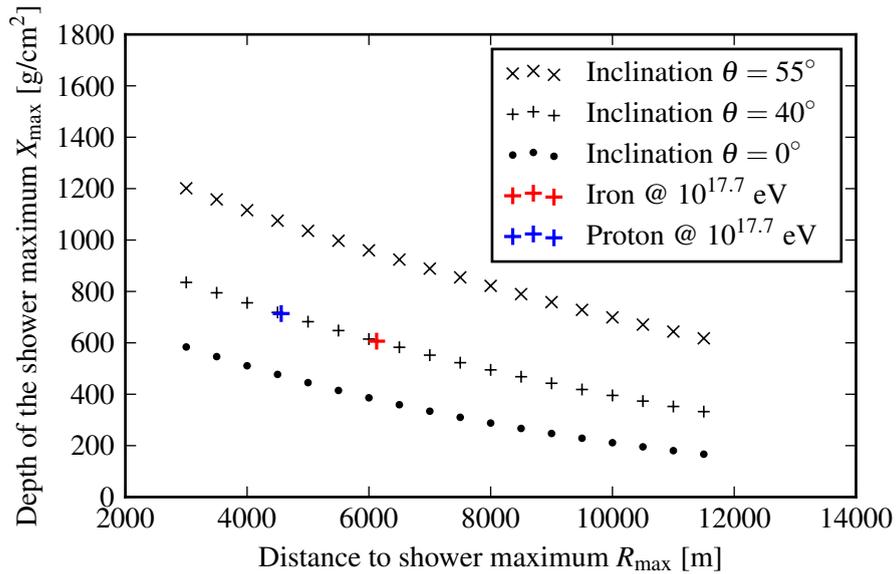


Figure 6.30.: The relation between R_{\max} and X_{\max} for different air shower inclinations θ as determined from equation (6.38). This allows to read of the geometrical distances R_{\max} for an air shower initiated by a CR with an energy of $10^{17.7}$ eV and an inclination of 40° assuming the primary particle to be a proton (blue) or an iron nucleus (red).

X_{\max} is shown for a vertical shower (inclination 0°), a typical radio measurement (inclination 40°) and the maximal inclination allowed in the standard reconstruction (inclination 55°) in figure 6.30. The blue cross denotes the expected mean distance to the shower maximum of 4.560 m for an air shower initiated by a proton with $10^{17.7}$ eV and an inclination of 40° . For an iron induced shower, the mean distance to the shower maximum is estimated to be 6127 m. Finally, these values can be used to plot the expected arrival time distributions in figure 6.31, as discussed in section 6.8.1. In this scenario, only stations at a distance to the shower axis exceeding 600 m are sensitive to the extreme compositions

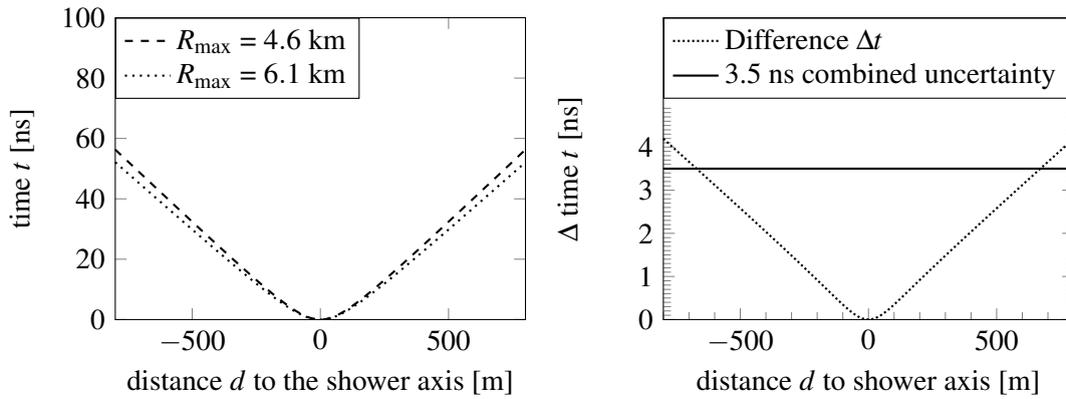


Figure 6.31: The expected arrival time distributions of the radio signal from an air shower initiated by an iron nucleus ($R_{\max} = 6.1$ km) and a proton ($R_{\max} = 4.6$ km) (left) and their subtraction (right). The CR energy is assumed to be $10^{17.7}$ eV and the inclination of the shower $\theta = 40^\circ$. These showers can only be distinguished by stations at distances to the shower axis exceeding 600 m.

of pure proton and pure iron. According to figure 6.28, there are practically no measurements for air showers less inclined than 55° .

As a consequence, the arrival time fit cannot be used for composition studies with AERA. The right panel in figure 6.31 implies that the time resolution needs to be substantially better than 1 ns to distinguish air showers initiated by protons from showers initiated by iron nuclei. However, the arrival time cannot be described by a simple hyperbolic wavefront in the sub-nanosecond regime anymore [51]. Instead, the wavefront becomes radially asymmetric. Thus, a higher timing does not necessarily guarantee a sensitivity of the hyperbolic wavefront fit to the primary mass (i.e. at least an X_{\max} resolution better than 100 g m^{-2}). Thus, it remains an open question if even the precise wavefront measurements of LOFAR (cf. figure 3.12) can be used for composition studies.

Cosmic ray particles constantly impact the earth atmosphere. They constitute up to a quarter of the annual radiation exposure of human beings depending on the altitude of the living environment. High energy cosmic rays produce cascades of secondary particles in the atmosphere which can reach the ground. One detection method of such air showers exploits the coherent radio emission from the charged secondary particles. In recent years, there has been huge progress in the field of radio detection, mostly driven by the availability of modern digital readout electronics and an improved theoretical understanding of the radio emission mechanisms. The Auger Engineering Radio Array is the largest modern radio detector worldwide for ultra-high energy cosmic ray physics and is located at the Pierre Auger Observatory in Argentina. It has been constructed to investigate the prospects of the radio technique for both, the large scale standalone detection of ultra-high energy cosmic rays and the combined detection with other detector techniques.

In this work, the detector software of the individual AERA radio detectors has been improved to read out a custom number of input channels. Thus, the detector software also meets the requirements of the newest deployment stage AERA III. It has been optimised to measure radiation from inclined air showers with a sparse antenna grid. Therefore, the front-end electronics have been designed for the readout of three-pole antennas which are sensitive to the horizontal as well as the vertical component of the radio emission. At the detector site, the vertical signal can amount to up to two thirds of the total emitted power for air showers with an inclination above 54° . Considering the unavoidable signal threshold due to the galactic radio noise floor, it is desirable to record the complete signal.

Moreover, the new customisable detector software can read out the low-gain channels of the former deployment stage, avoiding saturated data from high energy or near-vertical events. Generally, the new electronics can be used to read out data from any generic detector which matches the impedance and voltage input range of the electronics and can be meaningfully digitised with the 180 MHz sampling

rate of the electronics.

In the past winters, the photovoltaic power supply systems of the AERA stations regularly have reached their limits during overcast periods. This has caused individual detector stations to erratically stop the data acquisition, and the data quality consequently has deteriorated. The ageing of the batteries gives reason to expect such problematic periods more often in future. To control the inevitable detector down-times, a new detector operation mode (“sleep mode”) has been implemented. It can be used to shut down the whole detector array for a defined time to recharge the batteries. Now, the acquisition of high quality data in coincidence with the fluorescence telescopes can be scheduled during wintertime. With the new detector software, the precise GPS position of the detectors is now used for the time determination at each individual detector. This has corrected the relative timing by several nanoseconds, which is a significant change. Also, the calculation of the timestamps has been improved in the software. The data acquisition program is now capable of noticing and flagging recorded events, whose timestamps could only be calculated with reduced precision.

The improvements described above affect not only the detector software, but also the FPGA firmware, the operating system and the micro-controller firmware. The final system has been tested carefully and has been finally integrated in the AERA detector from remote.

The second part of this thesis clarifies the prospects of describing the measured radio signal by a combination of an amplitude model and a signal time model. The amplitude model describes the energy fluence measured at each radio detector by a two-dimensional lateral distribution function. The integral over the fitted distribution is sensitive to the energy of the cosmic ray, while its width is sensitive to the depth of the shower maximum. The signal arrival time model assumes the wave-front of the radio signal to be hyperbolic. The opening angle of the hyperbola is sensitive to the shower maximum. Both models are linked by their sensitivity to the shower geometry.

Assuming the independence of the measured energy fluence and arrival time distribution, a weighted uncorrelated chi-square fit is the natural approach to determine the model parameters.

For the first time, a multivariate analysis module has been developed for these functional models. It is designed such that it can easily be extended with other models in future. Thus, the implementation of the already existing univariate standalone modules has been crosschecked. To obtain a correct weighting of the models, the error calculation of the individual analyses has been revised. This effort has led to the introduction of additional model uncertainties.

However, it turned out that the combination does not lead to an improved reconstruction the energy or the depth of the air shower maximum. One problem is certainly the numerical stability of the multi-dimensional fit, which can have up to eight degrees of freedom. However, my analysis identifies the intrinsically lower predictive power of the arrival time model and its higher sensitivity to noise in comparison to the amplitude model as the main limiting factors. The low predictive power of the arrival time model is caused by the low time resolution in the presence of realistic noise. A rough estimate shows that the required time resolution for a meaningful X_{\max} -sensitivity needs to be better than the actual detector resolution of 2 ns. This result is independent of the rather simple approach to the multivariate problem. Hence, an improvement of the result by applying a more sophisticated regression ansatz is found to be unlikely.

In conclusion, this work has ensured a continuous operation of the AERA radio detector even under

challenging environmental conditions like overcast weather periods. The software adaptations have uncovered the full potential of the detector and have prepared the front-end electronics for the readout of the next generation of prototype stations (e.g. three-pole antennas) to come. The analysis of the detector timing and multivariate analysis of the data showed that the timing of the current detector is not good enough for an arrival time sensitivity to the shower maximum. Future radio detectors will need a better realisation of the distributed clocks.

Appendices

A. The weighted mean and its uncertainty

The weighted mean \bar{x} of an observable x_i which has been determined with different methods i is given by:

$$\bar{x} = \frac{\sum \frac{1}{(\Delta x_i)^2} x_i}{\sum \frac{1}{(\Delta x_i)^2}} = \frac{\sum w_i x_i}{\sum w_i} \quad (\text{A.1})$$

The uncertainties of the individual observables are denoted by Δx_i and are used as weights $w_i = 1/(\Delta x_i)^2$. In this thesis, the different methods $i = \{t, \varepsilon\}$ are the arrival time fit t and the energy fluence fit ε . The squared uncertainty of the weighed mean can be written as

$$(\Delta \bar{x})^2 = \frac{1}{\sum w_i} \quad (\text{A.2})$$

$$= \frac{(\Delta x_t)^2 (\Delta x_\varepsilon)^2}{(\Delta x_t)^2 + (\Delta x_\varepsilon)^2} \quad (\text{A.3})$$

$$= \frac{n^2}{(1+n^2)} (\Delta x_\varepsilon)^2, \quad (\text{A.4})$$

with the replacement $\Delta x_t = n \Delta x_\varepsilon$. Thus, the uncertainty of observables which have been calculated from weighted mean of the results from the energy fluence fit and the arrival time fit can be expressed by the ratio n of the individual uncertainties, and the uncertainty from energy fluence fit:

$$\Delta \bar{x} = \frac{n}{\sqrt{1+n^2}} \Delta x_\varepsilon \quad (\text{A.5})$$

Figure A.1 shows the ratio $\Delta \bar{x}/\Delta x_\varepsilon$ for $n > 1$. For $n < 1$ one would normalise everything to the arrival time fit t and would get the same result. For equal individual uncertainties $\Delta x_t = \Delta x_\varepsilon$, the combined uncertainty is $\Delta \bar{x} = 1/\sqrt{2} \Delta x_\varepsilon$. This is expected because it corresponds to increasing the statistics by a factor of two. For different uncertainties, the weighted mean and its uncertainty is dominated by the observable with the lower individual uncertainty (i.e. in this example the energy fluence fit).

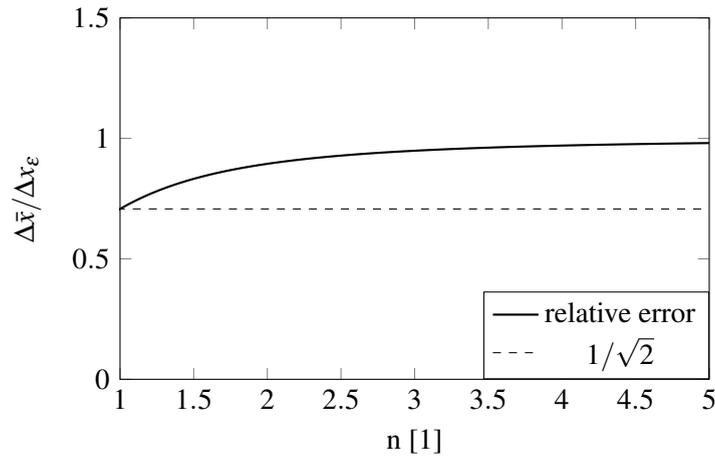


Figure A.1.: Relative uncertainty of the weighted mean w.r.t the uncertainty from the energy fluence fit as a function of the ratio of the individual uncertainties $n = \Delta x_t / \Delta x_\epsilon$. The benefit from a data combination is maximal for equal errors $n = 1$. The combined error behaves like doubling the statistics and is reduced by the factor $1/\sqrt{2}$. For $n > 1$, the error and the result is dominated by the better method energy fluence fit.

B. Arrival time fit $R_{\max}(\gamma)$ without Noise

Figure B.1 shows the sensitivity of the slope parameter γ to the geometrical distance R_{\max} from ground to the shower maximum. A small uncertainty of 0.1 ns has been added to the timing but no measured noise has been added to the signal traces.

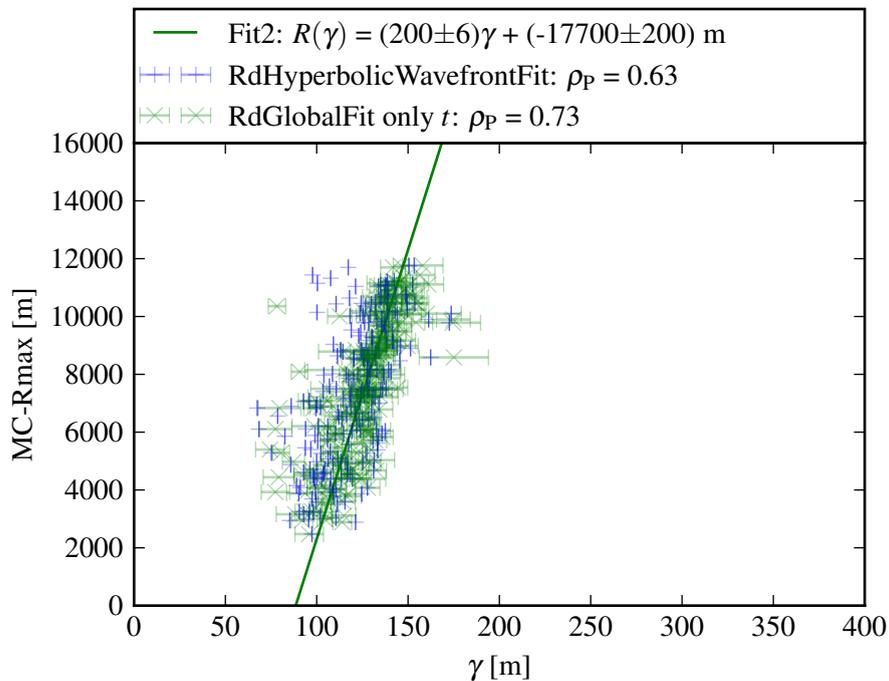


Figure B.1.: The sensitivity of the slope parameter γ to the geometrical distance R_{\max} from ground to the shower maximum. The Pearson correlation coefficient ρ_P of the slope parameter γ and R_{\max} has a higher value for the reconstruction with the “RdGlobalFit” than for the “RdHyperbolicWavefrontFit”. Additionally, the result $R_{\max}(\gamma)$ of a weighted linear fit is shown.

C. Application of the global fit to data

The following plots show results of the application of the ‘‘RdGlobalFit’’ module to real data measured with AERA. The results are discussed in section 6.7.3.

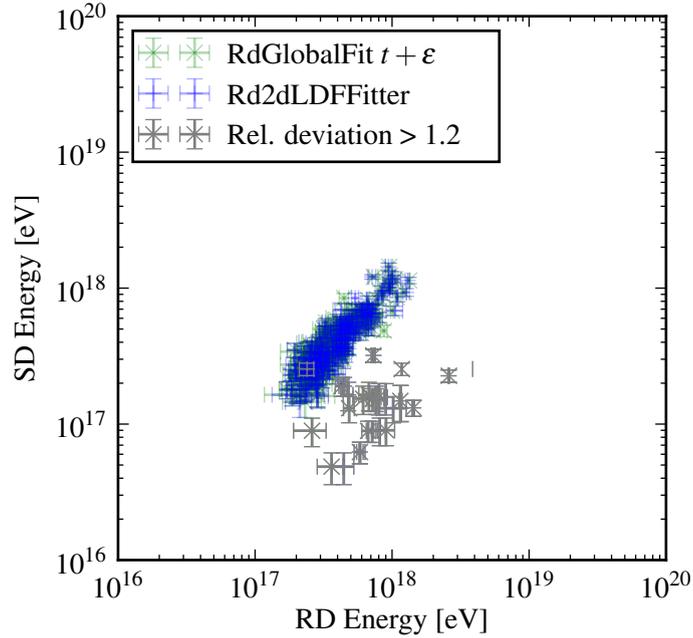


Figure C.1.: Comparison of the cosmic ray energy reconstructed from AERA measurements (RD) and air-fluorescence measurement (FD).

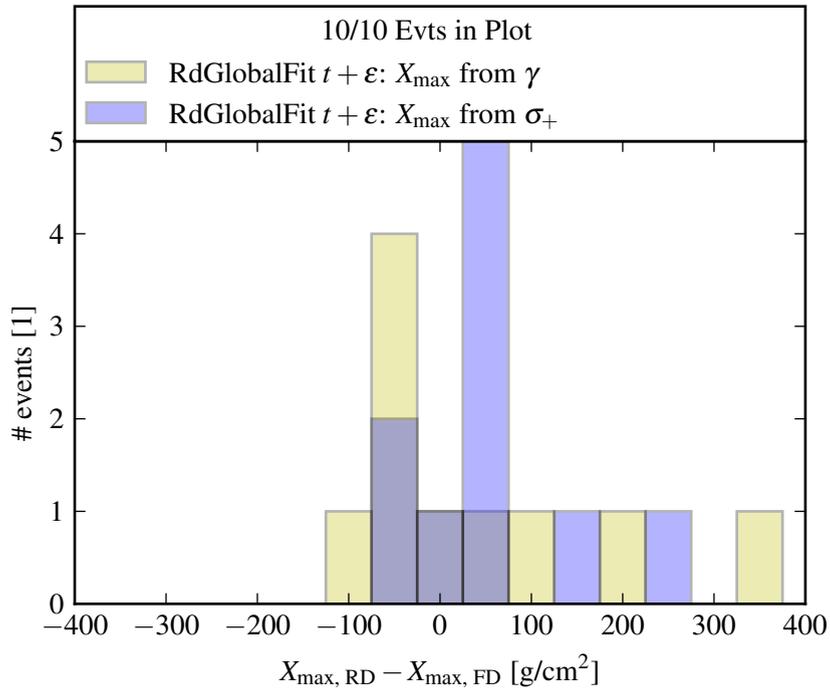


Figure C.2.: The absolute deviation of the atmospheric depth reconstructed from radio (RD) from the air-fluorescence reconstruction (FD). For the radio measurements, the atmospheric depth has been determined from the width σ_+ of the lateral energy fluence distribution and from the slope γ of the arrival time distribution.

D. Reconstruction settings for simulation data

The following files show the most important configurations used in the simulation studies with Offline.

D.1. Modulesequence

```
<!-- A standard sequence for processing simulated air showers including a radio reconstruction a la RdObserver. -->
<sequenceFile xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
              xsi:noNamespaceSchemaLocation='@SCHEMALOCATION@/ModuleSequence.xsd'>

  <enableTiming/>
  <moduleControl>

    <loop numTimes="unbounded" pushEventToStack="yes">
      <module> EventFileReaderOG          </module>
      <loop numTimes="1" pushEventToStack="yes">
        <module> RdStationAssociator      </module>
        <module> EventGeneratorOG        </module>

        <loop numTimes="unbounded" pushEventToStack="no">
          <module> CachedShowerRegeneratorOG </module>
        <!-- <module> G4TankSimulatorOG      </module> -->
          <module> TabulatedTankSimulatorKG </module>
        </loop>

        <module> SdSimulationCalibrationFillerOG </module>
        <module> SdPMTSimulatorOG              </module>
        <module> SdFilterFADCSimulatorMTU      </module>
        <module> SdBaselineSimulatorOG         </module>
        <module> TankTriggerSimulatorOG        </module>
        <module> TankGPSSimulatorOG           </module>
        <module> CentralTriggerSimulatorXb     </module>
        <module> CentralTriggerEventBuilderOG </module>
        <module> EventBuilderOG                </module>

        <module> EventCheckerOG                </module>
        <module> SdCalibratorOG                </module>
        <module> SdEventSelectorOG             </module>
        <module> SdMonteCarloEventSelectorOG   </module>

        <module> SdPlaneFitOG                  </module>
        <module> LDFFinderKG                   </module>
        <module> SdEventPosteriorSelectorOG    </module>

        <!-- radio reconstruction -->
        <module> RdAntennaStationToChannelConverter </module>
        <module> RdChannelResponseIncorporator    </module>
        <module> RdChannelResampler               </module>
        <module> RdChannelTimeSeriesClipper      </module>
        <module> RdChannelVoltageToADCConverter  </module>
        <module> RdChannelNoiseImporter          </module>

        <module> RdEventInitializer              </module>
        <module> RdTimeJitterAdder               </module> <!-- add this module if you want to add time jitters -->
        <module> RdStationRejector              </module>
        <module> RdChannelADCToVoltageConverter  </module>
        <module> RdChannelSelector               </module>
        <module> RdChannelPedestalRemover       </module>
        <module> RdChannelResponseIncorporator   </module>
      </loop>
    </loop>
  </moduleControl>
</sequenceFile>
```

```
<module> RdChannelTimeSeriesTaperer      </module>
<module> RdChannelBandstopFilter         </module>
<module> RdChannelUpsampler              </module>
<module> RdChannelRiseTimeCalculator     </module>

<module> RdAntennaChannelToStationConverter </module>
<module> RdStationSignalReconstructor    </module>
<module> RdStationEFieldVectorCalculator </module>

<module> RdClusterFinder                  </module>
<module> RdPlaneFit                       </module>

<module> RdStationRiseTimeCalculator     </module>

<module> GlobalFit                        </module>
<module> Rd2dLDFitter                    </module>
<module> RdHyperbolicWavefrontFit        </module>
  <module> RdEventPostSelector            </module>

  <module> RdStationTimeSeriesWindowCutter </module>
  <module> RdStationTimeSeriesTaperer     </module>
  <module> RecDataWriterNG                </module>
</loop>
</loop>
</moduleControl>
</sequenceFile>
```

D.2. Bootstrap

```
[obeytabs]
<?xml version="1.0" encoding="iso-8859-1"?>

<!DOCTYPE bootstrap [
  <!-- Path to your own configuration files -->
  <!ENTITY userConfigFiles '.'>
  <!-- alias for example Sd and Radio config files -->
  <!ENTITY standardFdIdealDetConfig SYSTEM '@CONFIGDIR@/standardFdIdealDetConfig.xml'>
  <!ENTITY standardSdIdealDetConfig SYSTEM '@CONFIGDIR@/standardSdIdealDetConfig.xml'>
  <!ENTITY standardSdSimModuleConfig SYSTEM '@CONFIGDIR@/standardSdSimModuleConfig.xml'>
  <!ENTITY standardSdRecModuleConfig SYSTEM '@CONFIGDIR@/standardSdRecModuleConfig.xml'>
  <!ENTITY exampleRConfig SYSTEM '@CONFIGDIR@/exampleRConfig.xml'>
]>

<bootstrap xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:noNamespaceSchemaLocation='@SCHEMALOCATION@/bootstrap.xsd'
  xmlns:xlink="http://www.auger.org/schema/types">

  &standardFdIdealDetConfig;
  &standardSdIdealDetConfig;
  &standardSdSimModuleConfig;
  &standardSdRecModuleConfig;
  &exampleRConfig;

  <centralConfig>

    <!-- ===== database servers KA ===== -->

    <configLink
      id      = "databaseServers"
      type    = "XML"
      xlink:href = "&userConfigFiles;/databaseServers.xml"/>

    <!-- ===== Rest ===== -->

    <configLink
      id      = "ModuleSequence"
      type    = "XML"
      xlink:href = "&userConfigFiles;/ModuleSequence.xml"/>

    <configLink
      id      = "EventFileReader"
      type    = "XML"
      xlink:href = "&userConfigFiles;/EventFileReader.xml"/>

    <configLink
      id      = "EventGenerator"
      type    = "XML"
      xlink:href = "&userConfigFiles;/EventGenerator.xml"/>

    <configLink
      id      = "LDFFinder"
      type    = "XML"
      xlink:href = "@CONFIGDIR@/LDFFinderKG_Infill.xml"/>

    <configLink
      id      = "SdEventSelector"
      type    = "XML"
```

```

    xlink:href = "@CONFIGDIR@/SdEventSelector_Infill.xml"/>

<configLink
  id      = "SStationListXMLManager"
  type    = "XML"
  xlink:href = "@CONFIGDIR@/SIdealStationListInfill.xml" />

</centralConfig>

<parameterOverrides>

  <!-- Set seeds to 0 for a random seed (results will not be reproducible then). -->
  <configLink id="RandomEngineRegistry">
    <RandomEngineRegistry>
      <DetectorSeed> 100 </DetectorSeed>
      <PhysicsSeed> 200 </PhysicsSeed>
    </RandomEngineRegistry>
  </configLink>

  <configLink id="RdStationAssociator">
    <RdStationAssociator>
      <MaximumAllowedDistance unit="m"> 10. </MaximumAllowedDistance>
      <ExcludedStationIds> 187 191 198 207 185 245 246 252 257 </ExcludedStationIds>
    </RdStationAssociator>
  </configLink>

  <configLink id="RecDataWriter">
    <RecDataWriter>
<SaveRadio> 0 </SaveRadio>
      <rootOutput>
        <outputFileName> ADST.root </outputFileName>
      </rootOutput>
    </RecDataWriter>
  </configLink>

  <configLink id="RdChannelResponseIncorporator">
    <RdChannelResponseIncorporator>
      <ForwardResponseOnFirstCall> 1 </ForwardResponseOnFirstCall>
    </RdChannelResponseIncorporator>
  </configLink>

  <configLink id="RdEventInitializer">
    <RdEventInitializer>
      <infoLevel> 2 </infoLevel>
      <SetRdShowerDirectionTo> SD </SetRdShowerDirectionTo>
      <SetCoordinateOriginTo> RdBaryCenter </SetCoordinateOriginTo>
      <UseUserSpecifiedWindows> 1 </UseUserSpecifiedWindows>
      <NoiseWindowStart unit="ns"> 6000.0 </NoiseWindowStart>
      <NoiseWindowStop unit="ns"> 8000.0</NoiseWindowStop>
      <SignalSearchWindowStart unit="ns"> 1500.0 </SignalSearchWindowStart>
      <SignalSearchWindowStop unit="ns"> 2500.0 </SignalSearchWindowStop>
      <ReferenceCorePosition> SD </ReferenceCorePosition>
      <ReferenceDirection> SD </ReferenceDirection>
    </RdEventInitializer>
  </configLink>

  <configLink id="RdTimeJitterAdder">
    <RdTimeJitterAdder>
      <sigma unit="ns"> 2.00 </sigma>

```

```

    </RdTimeJitterAdder>
</configLink>

<configLink id="RdChannelTimeSeriesClipper">
  <RdChannelTimeSeriesClipper>
    <ClippingMethod> RemoveSamplesFromBack </ClippingMethod>
    <NumberOfSamplesToKeep> 2048 </NumberOfSamplesToKeep>
    <NumberOfSamplesToRemove> 0 </NumberOfSamplesToRemove>
  </RdChannelTimeSeriesClipper>
</configLink>

<configLink id="RdChannelTimeSeriesTaperer">
  <RdChannelTimeSeriesTaperer>
    <RelativeWindowWidthOnEachSide> 0.03 </RelativeWindowWidthOnEachSide>
  </RdChannelTimeSeriesTaperer>
</configLink>

<configLink id="RdChannelSelector">
  <RdChannelSelector>
    <UseSelectedChannels> 1 </UseSelectedChannels>
    <SelectedChannels> 1 2 </SelectedChannels>
  </RdChannelSelector>
</configLink>

<configLink id="RdAntennaChannelToStationConverter">
  <RdAntennaChannelToStationConverter>
    <infoLevel> 1 </infoLevel>
    <UsedDirection> SdReconstruction </UsedDirection> <!-- set to McTruth for best case scenario -->
  </RdAntennaChannelToStationConverter>
</configLink>

<configLink id="RdChannelBandstopFilter">
  <RdChannelBandstopFilter>
    <UseOnlineBandstopFilter> 1 </UseOnlineBandstopFilter>
    <OnlineBandstopFilterSpectrumBaseline> Median </OnlineBandstopFilterSpectrumBaseline>
    <OnlineBandstopFilterDeviationFactor> 3.0 </OnlineBandstopFilterDeviationFactor>
  </RdChannelBandstopFilter>
</configLink>

<configLink id="RdChannelUpsampler">
  <RdChannelUpsampler>
    <UpsamplingFactor> 4 </UpsamplingFactor>
  </RdChannelUpsampler>
</configLink>

<configLink id="RdStationSignalReconstructor">
  <RdStationSignalReconstructor>
    <InfoLevel> 1 </InfoLevel>
    <MinSignalToNoise> 10 </MinSignalToNoise>
    <VectorialComponent> 4 </VectorialComponent>
    <useEnvelope> 1 </useEnvelope>
    <SignalDef> Peak </SignalDef>
    <NoiseDef> RMS </NoiseDef>
    <SignalToNoiseDef> RatioOfSquares </SignalToNoiseDef>
  </RdStationSignalReconstructor>
</configLink>

<configLink id="RdTopDownStationSelector">
  <RdTopDownStationSelector>
    <infoLevel> 2 </infoLevel>

```

```

    <minNumberOfStations> 4 </minNumberOfStations>
    <usedGeometry> SD </usedGeometry>
    <stopAtFirstStation> 0 </stopAtFirstStation>
    <chiSquareProbabilityCutValue> 0.05 </chiSquareProbabilityCutValue>
  </RdTopDownStationSelector>
</configLink>

<configLink id="RdClusterFinder">
  <RdClusterFinder>
    <RejectLonelyStations> 1 </RejectLonelyStations>
    <LonelyIfNoneInDistance unit="meter"> 400 </LonelyIfNoneInDistance>
    <LonelyIfOneInDistance unit="meter"> 800 </LonelyIfOneInDistance>
  </RdClusterFinder>
</configLink>

<configLink id="RdStationEFieldVectorCalculator">
  <RdStationEFieldVectorCalculator>
    <InfoLevel> 1 </InfoLevel>
    <MaximumAngleToLorentzForce unit="degree"> 55.0 </MaximumAngleToLorentzForce>
  </RdStationEFieldVectorCalculator>
</configLink>

<configLink id="RdStationTimeSeriesWindowCutter">
  <RdStationTimeSeriesWindowCutter>
    <WindowSize unit="ns"> 2000 </WindowSize>
  </RdStationTimeSeriesWindowCutter>
</configLink>

<configLink id="SdEventSelector">
  <SdEventSelector>
    <EnableBottomUpSelection> 0 </EnableBottomUpSelection>
  </SdEventSelector>
</configLink>

<configLink id="GlobalFit">
  <GlobalFit>
    <InfoLevel> 4 </InfoLevel>
    <FitConfig>
      <!-- If 1: Fit the arrival-time distribution (wavefrontfit) -->
      <fitArrivalTime> 1 </fitArrivalTime>
      <!-- If 1: Fit the two dimensional double gaussian LDF -->
      <fitTwoDLDF> 1 </fitTwoDLDF>

      <fitGammaAndSigmaPlusIndependently> 1 </fitGammaAndSigmaPlusIndependently>
      <!-- save contour data of CoreFit in file CoreContour.dat Its very slow!-->
      <SaveCoreContourData> 0 </SaveCoreContourData>
      <iterateFitConfig> 0 </iterateFitConfig>
      <printMinuitScans> 0 </printMinuitScans>
      <useCS> REF </useCS>
    </FitConfig>

    <GlobalFitData>
      <Aplus>
        <isFix> 0 </isFix>
        <startValue> 10 </startValue>
        <startError> 5 </startError>
        <lowLimit> -125.5 </lowLimit>
        <highLimit> 125.5 </highLimit>
      </Aplus>
    </GlobalFitData>
  </GlobalFit>
</configLink>

```

```

<core_x>
  <isFix> 0 </isFix>
  <startValueFrom> SD </startValueFrom>
  <startValue> 25</startValue>
  <startError> 20 </startError>
  <lowLimit> -125.5 </lowLimit>
  <highLimit> 125.5 </highLimit>
</core_x>

<core_y>
  <isFix> 0 </isFix>
  <startValueFrom> SD </startValueFrom>
  <startValue> -4.5 </startValue>
  <startError> 20 </startError>
  <lowLimit> -125.50 </lowLimit>
  <highLimit> 125.50 </highLimit>
</core_y>

<gamma>
  <isFix> 0 </isFix>
  <hasLimits> 1 </hasLimits>
  <lowLimit> 50 </lowLimit>
  <highLimit> 300 </highLimit>
  <startValue> 125 </startValue>
  <startError> 10 </startError>
</gamma>

<b>
  <isFix> 1 </isFix>
  <startValue> 9 </startValue>
  <startError> 0.1 </startError>
</b>

<arrivalDirection_theta>
  <isFix> 0 </isFix>
  <hasLimits> 1 </hasLimits>
  <startValueFrom> RD </startValueFrom>
  <startError> 0.02 </startError>
  <lowLimit> 0 </lowLimit>
  <highLimit> 1.570796326794896619231321691639751442098584699687552910487472296153 </highLimit>
</arrivalDirection_theta>

<arrivalDirection_phi>
  <isFix> 0 </isFix>
  <hasLimits> 1 </hasLimits>
  <startValueFrom> RD </startValueFrom>
  <startError> 0.01 </startError>
  <lowLimit> -3.14159265358979323846 </lowLimit>
  <highLimit> 3.14159265358979323846 </highLimit>
  <!-- <lowLimit> 0 </lowLimit> -->
  <!-- <highLimit> 6.283185307179586476925286766559005768394338798750211641949889184615 </highLimit> -->
</arrivalDirection_phi>

<t0>
  <isFix> 0 </isFix>
  <startValue> 0 </startValue>
  <startError> 3 </startError>
</t0>

```

```

    <sigmaPlus>
      <isFix> 0 </isFix>
      <hasLimits> 1 </hasLimits>
      <lowLimit> 50 </lowLimit>
      <highLimit> 300 </highLimit>
      <startValue> 100 </startValue>
      <startError> 10 </startError>
    </sigmaPlus>

    <showerMax>
      <isFix> 1 </isFix>
      <startValue> 13339 </startValue>
      <startError> 500 </startError>
      <lowLimit> 0 </lowLimit>
      <highLimit> 50000 </highLimit>
    </showerMax>

  </GlobalFitData>
</GlobalFit>
</configLink>

<configLink id="RdChannelNoiseImporter">
  <RdChannelNoiseImporter>
    <infoLevel> 5 </infoLevel>
    <NoiseFileSelection> Automatically </NoiseFileSelection>
    <NoiseFilePath> /cr/aera01/public/RdObserver/RdObs_v1r3/periodic_daily_selected </NoiseFilePath>
    <FilenamePrefix> sel_ </FilenamePrefix>
    <TimeIntervalInFile> Day </TimeIntervalInFile>
    <ExpectSubDirectories> 0 </ExpectSubDirectories>
    <EvtSelInNoiseFile> ByTimestamp </EvtSelInNoiseFile>
    <MaxTimeDifference unit="minute"> 10 </MaxTimeDifference>
  </RdChannelNoiseImporter>
</configLink>

<configLink id="RdHyperbolicWavefrontFit">
  <RdHyperbolicWavefrontFit>
    <UsedCore> SD </UsedCore>
    <DoCoreSearch> 1 </DoCoreSearch>
    <InfoLevel> 4 </InfoLevel>
  </RdHyperbolicWavefrontFit>
</configLink>
</parameterOverrides>

</bootstrap>

```


Bibliography

- [1] G. Giacinti, M. Kachelrieß, *et al.*, “Cosmic ray anisotropy as signature for the transition from galactic to extragalactic cosmic rays,” *Journal of Cosmology and Astroparticle Physics* 2012 no. 07, (2012) 031. <http://stacks.iop.org/1475-7516/2012/i=07/a=031>.
- [2] K. Greisen, “End to the Cosmic-Ray Spectrum?,” *Phys. Rev. Lett.* 16 (1966) 748–750. <http://link.aps.org/doi/10.1103/PhysRevLett.16.748>.
- [3] G. T. Zatsepin and V. A. Kuzmin, “Upper limit of the spectrum of cosmic rays,” *JETP Lett.* 4 (1966) 78–80. [Pisma Zh. Eksp. Teor. Fiz.4,114(1966)].
- [4] T. Huege, “Radio detection of cosmic ray air showers in the digital era,” *Physics Reports* 620 (2016) 1 – 52.
- [5] R. Engel, D. Heck, *et al.*, “Extensive Air Showers and Hadronic Interactions at High Energy,” *Annual Review of Nuclear and Particle Science* 61 no. 1, (2011) 467–489. <http://dx.doi.org/10.1146/annurev.nucl.012809.104544>.
- [6] P. M. Bauleo and J. Rodriguez Martino, “The dawn of the particle astronomy era in ultra-high-energy cosmic rays,” *Nature* 458N7240 (2009) 847–851.
- [7] K.-H. Kampert and M. Unger, “Measurements of the cosmic ray composition with air shower experiments,” *Astroparticle Physics* 35 no. 10, (2012) 660 – 678. <http://www.sciencedirect.com/science/article/pii/S0927650512000382>.
- [8] KASCADE-Grande Collaboration, W. D. Apel, J. C. Arteaga-Velázquez, *et al.*, “Kneelike Structure in the Spectrum of the Heavy Component of Cosmic Rays Observed with KASCADE-Grande,” *Phys. Rev. Lett.* 107 (2011) 171104. <http://link.aps.org/doi/10.1103/PhysRevLett.107.171104>.
- [9] S. Buitink, A. Corstanje, *et al.*, “A large light-mass component of cosmic rays at $10^{17} - 10^{17.5}$ electronvolts from radio observations,” *Nature* 531 no. 7592, (2016) 70–73. <http://dx.doi.org/10.1038/nature16976>. Letter.
- [10] M.-P. Véron-Cetty and P. Véron, “A catalogue of quasars and active nuclei: 12th edition,” *aap* 455 (2006) 773–777.

- [11] A. Letessier-Selvon, “Highlights from the Pierre Auger Observatory,” *Brazilian Journal of Physics* 44 no. 5, (2014) 560–570. <http://dx.doi.org/10.1007/s13538-014-0218-6>.
- [12] Pierre Auger Collaboration, A. Aab *et al.*, “Searches for Anisotropies in the Arrival Directions of the Highest Energy Cosmic Rays Detected by the Pierre Auger Observatory,” *Astrophys. J.* 804 no. 1, (2015) 15, arXiv:1411.6111 [astro-ph.HE].
- [13] Pierre Auger Collaboration, A. Aab *et al.*, “Large Scale Distribution of Ultra High Energy Cosmic Rays Detected at the Pierre Auger Observatory With Zenith Angles up to 80°,” *Astrophys. J.* 802 no. 2, (2015) 111, arXiv:1411.6953 [astro-ph.HE].
- [14] J. Matthews, “A Heitler model of extensive air showers,” *Astroparticle Physics* 22 no. 5–6, (2005) 387 – 397.
<http://www.sciencedirect.com/science/article/pii/S0927650504001598>.
- [15] J. Linsley, “Structure of large air showers at depth 834 G/sq cm. III-Applications,” in *International Cosmic Ray Conference*, vol. 12, pp. 89–96. 1977.
- [16] G. Sigl, *Astroparticle Physics: Theory and Phenomenology*. Atlantis Press, 2017.
- [17] D. Heck, G. Schatz, *et al.*, “CORSIKA: A Monte Carlo code to simulate extensive air showers,”
- [18] S. Ostapchenko, “QGSJET-II: Results for extensive air showers,” *Nucl. Phys. Proc. Suppl.* 151 (2006) 147–150, arXiv:astro-ph/0412591 [astro-ph].
- [19] M. Ludwig and T. Huege, “REAS3: Monte Carlo simulations of radio emission from cosmic ray air showers using an “end-point” formalism,” *Astroparticle Physics* 34 no. 6, (2011) 438 – 446.
<http://www.sciencedirect.com/science/article/pii/S0927650510002094>.
- [20] J. Allen, “Interpretation of the signals produced by showers from cosmic rays of 1019 eV observed in the surface detectors of the Pierre Auger Observatory,” in *Proc. of the International Cosmic Ray Conference*. 2011.
- [21] F. D. Kahn and I. Lerche, “Radiation from Cosmic Ray Air Showers,” *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 289 no. 1417, (1966) 206–213. <http://rspa.royalsocietypublishing.org/content/289/1417/206>.
- [22] O. Scholten and *et al.*, “Macroscopic model of geomagnetic-radiation from air showers,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 604 no. 1–2, Supplement, (2009) S24 – S26.
- [23] N. N. Kalmykov and *et al.*, “Macroscopic model of radio emission from extensive air showers,” *Physics of Atomic Nuclei* 74 no. 7, (2011) 1019–1031.
- [24] G. A. Askar’yan, “Coherent Radio Emission from Cosmic Showers in Air and in Dense Media,” *Soviet Journal of Experimental and Theoretical Physics* 21 (1965) 658.

- [25] S. Lafebre, R. Engel, *et al.*, “Universality of electron–positron distributions in extensive air showers,” *Astroparticle Physics* 31 no. 3, (2009) 243 – 254.
<http://www.sciencedirect.com/science/article/pii/S0927650509000292>.
- [26] T. Huege and A. Haungs, “Radio detection of cosmic rays: present and future,” *JPS Conf. Proc.* 9 (2016) 010018, arXiv:1507.07769 [astro-ph.IM].
- [27] T. Huege and *et al.*, “Simulating radio emission from air showers with CoREAS,” *AIP Conference Proceedings* 1535 no. 1, (2013) 128–132.
- [28] S. Buitink and *et al.*, “Method for high precision reconstruction of air shower X_{\max} using two-dimensional radio intensity profiles,” *Phys. Rev. D* 90 (2014) 082003.
- [29] A. Corstanje, P. Schellart, *et al.*, “The shape of the radio wavefront of extensive air showers as measured with LOFAR,” *Astroparticle Physics* 61 (2015) 22 – 31.
<http://www.sciencedirect.com/science/article/pii/S0927650514000851>.
- [30] The Pierre Auger Collaboration, “Probing the radio emission from air showers with polarization measurements,” *Phys. Rev. D* 89 (2014) 052002.
- [31] R. Šmída, F. Werner, *et al.*, “First Experimental Characterization of Microwave Emission from Cosmic Ray Air Showers,” *Phys. Rev. Lett.* 113 (2014) 221101.
<http://link.aps.org/doi/10.1103/PhysRevLett.113.221101>.
- [32] O. Kambeitz and for the Pierre Auger Collaboration, “Measurement of horizontal air showers with the Auger Engineering Radio Array,” *EPJ Web Conf.* (submitted 2016) .
- [33] S. Hoover, J. Nam, *et al.*, “Observation of Ultrahigh-Energy Cosmic Rays with the ANITA Balloon-Borne Radio Interferometer,” *Phys. Rev. Lett.* 105 (2010) 151101.
<https://link.aps.org/doi/10.1103/PhysRevLett.105.151101>.
- [34] T-510 Collaboration, K. Belov *et al.*, “Accelerator measurements of magnetically-induced radio emission from particle cascades with applications to cosmic-ray air showers,” *Phys. Rev. Lett.* 116 no. 14, (2016) 141103, arXiv:1507.07296 [astro-ph.IM].
- [35] The Pierre Auger Collaboration, “The Pierre Auger Cosmic Ray Observatory,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 798 (2015) 172 – 213.
- [36] M. Takeda, N. Sakaki, *et al.*, “Energy determination in the Akeno Giant Air Shower Array experiment,” *Astroparticle Physics* 19 no. 4, (2003) 447 – 462.
<http://www.sciencedirect.com/science/article/pii/S0927650502002438>.
- [37] J. Abraham, M. Aglietta, *et al.*, “Properties and performance of the prototype instrument for the Pierre Auger Observatory,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 523 no. 1–2, (2004) 50 – 95.

- <http://www.sciencedirect.com/science/article/pii/S0168900203033497>.
- [38] E. M. Holt and Pierre Auger Collaboration, “The Auger Engineering Radio Array and multi-hybrid cosmic ray detection,” *Journal of Physics: Conference Series* 718 no. 5, (2016) 052019. <http://stacks.iop.org/1742-6596/718/i=5/a=052019>.
- [39] Pierre Auger Collaboration, A. Aab *et al.*, “The Pierre Auger Observatory Upgrade - Preliminary Design Report,” arXiv:1604.03637 [astro-ph.IM].
- [40] The AERA group, “AERA proposal for the construction of the 20 km² Auger Engineering Radio Array at the Southern Auger Observatory,” *The Pierre Auger Observatory, internal note, GAP-Note 2009-172* (2009) .
- [41] A. Schmidt, *Realization of a Self-triggered Detector for the Radio Emission of Cosmic Rays*. PhD thesis, Karlsruher Institut für Technologie (KIT), 2012.
<http://digbib.ubka.uni-karlsruhe.de/volltexte/1000030957>. Karlsruhe, KIT, Diss., 2012.
- [42] C. F. Rühle, *Entwicklung eines schnellen eingebetteten Systems zur Radiodetektion kosmischer Strahlung*. PhD thesis, Karlsruher Institut für Technologie (KIT), 2014.
<http://digbib.ubka.uni-karlsruhe.de/volltexte/1000040424>. Karlsruhe, KIT, Diss., 2014.
- [43] A. Aab, P. Abreu, *et al.*, “Nanosecond-level time synchronization of autonomous radio detector stations for extensive air showers,” *Journal of Instrumentation* 11 no. 01, (2016) P01018.
<http://stacks.iop.org/1748-0221/11/i=01/a=P01018>.
- [44] The Pierre Auger Collaboration, “Energy estimation of cosmic rays with the Engineering Radio Array of the Pierre Auger Observatory,” *Phys. Rev. D* 93 (2016) 122005.
- [45] Pierre Auger Collaboration, F. Gaté, “ X_{\max} reconstruction from amplitude information with AERA,” *EPJ Web Conf.* 135 (2017) 01007, arXiv:1609.06510 [astro-ph.IM].
- [46] J. Schulz, *Cosmic Radiation Reconstruction of Cosmic Ray Properties from Radio Emission of Extensive Air Showers*. PhD thesis, Radboud University Nijmegen, 2016.
<http://hdl.handle.net/2066/151781>.
- [47] The Pierre Auger Collaboration, “GPS-Time-Shift-and-Variance,” *The Pierre Auger Observatory, internal note, GAP-Note 2014-001* (2014) .
- [48] F. Schröder private communication, 2017.
- [49] F. G. Schröder, “Radio detection of cosmic-ray air showers and high-energy neutrinos,” *Progress in Particle and Nuclear Physics* 93 (2017) 1 – 68.
<http://www.sciencedirect.com/science/article/pii/S0146641016300758>.
- [50] The Pierre Auger Collaboration, “Measurement of the Radiation Energy in the Radio Signal of

- Extensive Air Showers as a Universal Estimator of Cosmic-Ray Energy,” *Phys. Rev. Lett.* 116 (2016) 241101.
- [51] W. Apel and et al., “The wavefront of the radio signal emitted by cosmic ray air showers,” *Journal of Cosmology and Astroparticle Physics* 2014 no. 09, (2014) 025.
- [52] Pierre Auger Collaboration, Q. D. Hasankiadeh, “Advanced Reconstruction Strategies for the Auger Engineering Radio Array,” arXiv:1705.06230 [astro-ph.IM].
- [53] K. Greisen, “Cosmic ray showers,” *Ann. Rev. Nucl. Part. Sci.* 10 (1960) 63–108.
- [54] K. Kamata and J. Nishimura, “The Lateral and the Angular Structure Functions of Electron Showers,” *Progress of Theoretical Physics Supplement* 6 (1958) 93.
<http://dx.doi.org/10.1143/PTPS.6.93>.
- [55] S. Argirò, S. Barroso, *et al.*, “The offline software framework of the Pierre Auger Observatory,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 580 no. 3, (2007) 1485 – 1496.
<http://www.sciencedirect.com/science/article/pii/S0168900207014106>.
- [56] A. Nelles and et al., “A parameterization for the radio emission of air showers as predicted by CoREAS simulations and applied to LOFAR measurements,” *Astroparticle Physics* 60 (2015) 13 – 24.
- [57] W. Niemeier, *Ausgleichsrechnung : statistische Auswertemethoden*. De Gruyter Lehrbuch. de Gruyter, Berlin [u.a.], 2008.
http://deposit.d-nb.de/cgi-bin/dokserv?id=3018921&prov=M&dok_var=1&dok_ext=htm; <http://www.reference-global.com/isbn/978-3-11-020678-4>.
- [58] Q. D. Hasankiadeh private communication, 2016.
- [59] B. G. Keilhauer, “Investigation of atmospheric effects on the development of extensive air showers and their detection with the Pierre Auger Observatory,” tech. rep., Karlsruhe, 2004.
- [60] U. S. C. on Extension to the Standard Atmosphere, *U.S. standard atmosphere, 1976*. National Oceanic and Amospheric [sic] Administration : for sale by the Supt. of Docs., U.S. Govt. Print. Off., 1976.
- [61] J. Neuser, *Cosmic Rays and the Atmospheric Electric Field - Reconstruction and Data Analysis of Radio Emission from Air Showers at the Auger Engineering Radio Array*. PhD thesis, Bergische Universität Wuppertal, 2015.
<http://elpub.bib.uni-wuppertal.de/servlets/DocumentServlet?id=5560>.