THE UNIVERSITY
of ADELAIDE

A Magnetic Spectrometer Analysis Method for Ultra High Energy Cosmic Ray Data

Benjamin James Whelan

B.Sc. (Honours), Physics

A thesis submitted to the University of Adelaide in fulfilment of the requirements for the degree of Doctor of Philosophy.

School of Chemistry & Physics

Department of Physics

November 2011

Contents

Abstract

Since their discovery at the beginning of the 20^{th} Century, cosmic rays have been the subject of much research. However, fundamental questions, such as the origin of the most energetic of these particles, await an answer. The unambiguous identification of cosmic ray sources is made difficult by the deflections that these charged particles suffer in their propagation through cosmic magnetic fields. This problem is compounded by the extremely low flux of the highest energy cosmic rays at Earth, the arrival directions of which may be expected to retain information about their point of origin.

However, the advent of ultra high energy cosmic ray detectors possessing collecting areas of thousands of square kilometres, such as the Pierre Auger Observatory, means that the number of cosmic rays detected at the highest energies may be sufficient to enable directional cosmic ray astronomy in the near future. The 'magnetic spectrometer analysis', described in this thesis, is designed for the analysis of ultra high energy cosmic ray data sets. The analysis is designed to identify energy ordering in the arrival directions of a cosmic ray data set, and reconstruct the source directions of such events.

A brief history of the discovery of cosmic rays is presented in Chapter 1, along with an introduction to the physics of extensive air showers and methods of detecting them. The current knowledge of the properties of the cosmic ray flux at ultra high energies is reviewed in Chapter 2, and a summary of ultra high energy cosmic ray detectors, both past and present, is presented in Chapter 3.

The propagation of cosmic rays through magnetic fields, and the methods of measuring those fields, is briefly reviewed in Chapter 4. In addition, the Galactic magnetic field models that are used for the production of simulated cosmic ray data sets are described there.

Numerical integration is employed in the generation of the simulated data sets, and the method of doing so, as well as considerations that must be made for such simulations, is described in Chapter 5.

The magnetic spectrometer analysis method is introduced in Chapter 6 in addition to a discussion of related analyses found in the literature. The results of applying the magnetic spectrometer analysis to simulated cosmic ray data sets are presented in Chapter 7.

Finally, the results of an application of the magnetic spectrometer analysis to real data collected by the Pierre Auger Observatory are described in Chapter 8, followed by a discussion of those results and concluding remarks.

Declaration of Originality

I, Benjamin James Whelan, certify that this work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library catalogue and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Signed: ..... **Date:** .....

Acknowledgements

I would like to thank my supervisors, Roger Clay and Bruce Dawson, for the help and guidance they have offered me over the course of my Ph.D. They have taught me much, and I am grateful for the enthusiasm and interest toward my research they have shown during my studies. In addition, their eternal patience and calming influence was invaluable on the many occasions when I felt things were all getting to be too much.

I would also like to thank the other members of the High Energy Astrophysics research group at the University of Adelaide, both past and present. They have made my time as a Ph.D. student an interesting and enlightening one, and have been ever willing to offer computing expertise or help tackle a difficult problem. Particular thanks go to my office mate, Kerri Budd, who brings a cheerful mood to the office every morning. She has always been ready to talk over any issues I have had and provide useful input. In addition, her willingness to offer a fresh set of eyes whenever required was much appreciated.

The Pierre Auger Collaboration (www.auger.org/admin/Collaborators/author_list_alphabetical.html) has been an exciting group of people to work with, and I thank the collaboration members for the input and ideas that they have offered towards this work. The opportunity to work with a cutting edge detector, and be exposed to so many different areas of the field, has given me an excellent start to cosmic ray research.

I was fortunate enough to be awarded the Ferry Postgraduate Scholarship, for which I am extremely grateful. Computing facilities provided by eResearch SA were used extensively throughout my research, which allowed me to generate much of the data presented in this thesis.

My family and friends have been a constant source of support over the course of my studies. In particular, I extend thanks to my parents, Marg and Ken, and my sister, Kate. They have always encouraged me to do whatever makes me happy, and I couldn't ask for more.

Last, but most certainly not least, is my wife, Alex. An unwavering source of love, she has kept me going through even the most stressful periods of my Ph.D. It's an incredible feeling to have her believe in me, and I can't imagine having gotten this far without her.

Chapter 1

Introduction to Cosmic Rays and Extensive Air Showers

1.1 The Discovery of Cosmic Rays and Extensive Air Showers

The term 'cosmic rays', coined by R.A. Millikan in the 1920s [1, 2], refers to the extraterrestrial source of radiation that causes ionisation in the Earth's atmosphere. These charged particles - mainly protons and atomic nuclei - which produce the ionisation are the most energetic known to science, and have been the subject of intense study for close to a century. Cosmic ray (CR) science as a distinct field can be traced back to the historic balloon flights of Victor Hess in 1912 [3]. However, to find the first hints of the existence of these mysterious particles it is necessary to refer to experiments roughly two centuries old.

In Paris in 1785, Coulomb reported on his experiments into electrically charged bodies. During these experiments he had noticed that electrostatically-charged bodies would slowly leak their accumulated charge away. Careful insulation did not prevent this process, and he was unable to satisfactorily explain the reason for it [1].

A century later, physicists such as J. Elster, H.F. Geitel and C.T.R. Wilson were studying this phenomenon, which was termed 'dark current' [4, 5]. Careful isolation of electroscopes, sealed in airtight containers, from all known forms of radiation such as gamma rays and X-rays did not prevent their discharge. Even shielding the detector with lead as Rutherford did, while reducing the rate of discharge, did not stop it entirely [2]. It was apparent that ions were being formed in the air around the electroscopes, causing it to act as a conductor; the reason for this could not be found.

Complicating matters was the fact that - unknown to the researchers - the radiation they measured was of both terrestrial and extraterrestrial origin; Wilson's experiments in railway tunnels did not produce an observable difference in the rate of ionisation to that of an ordinary room. He had conjectured that the source of the ionisation could originate outside the Earth's atmosphere [1, 5] but the agreement between Wilson's experiments from inside the railway tunnels and outside seemed to eliminate the possibility of such a cosmic origin. Both Wilson

and Geitel concluded that the ionisation was an inherent property of the atmosphere [1, 2]. This apparently unexplainable origin led the ionisation to be termed 'spontaneous' [1, 5].

Uncertainty as to the origin of the radiation was compounded by the fact that many of the experiments were inaccurate and their results not reproducible. Atmospheric effects and other environmental factors complicated measurements and made it difficult to obtain conclusive results. As an example, several studies measured a considerable diurnal variation of the ionisation, hinting at an extraterrestrial source, but this could not be reproduced in subsequent experiments. This lack of consistency was fortunate in a way, as it spurred on further research to ascertain the source of the ionisation [2].

The breakthrough came in 1912, when Viktor Hess reported on his high-altitude balloon flights undertaken in that same year [3]. Carrying aloft three ionisation detectors designed by T. Wulf, Hess recorded measurements of the amount of ionisation within the detectors throughout each ascent - some of which reached over 5 km in altitude. At the beginning of each flight, the ionisation decreased as expected. This was due to the influence of terrestrial radiation diminishing. However, a marked increase in the ionisation came with higher altitudes, which Hess attributed to a penetrating radiation of extraterrestrial origin [1, 2, 3, 6]. He termed the source of the ionisation 'rays from on high' (*Höhenstrahlung)* [5]. This was the first convincing evidence of the 'spontaneous' ionisation having a cosmic origin, and the discovery earned Hess the 1936 Nobel Prize [1, 5, 7].

Even greater heights were reached by Werner Kolhörster in 1913 and 1914 in which he undertook flights to altitudes of over 9000 m. The measurements made during those ascents confirmed the results of Hess and demonstrated that the amount of ionisation continued to increase to such altitudes [1, 2].

In the 1920s, unconvinced by the explanation of radiation from space being the cause of the observed ionisation, Robert Millikan endeavoured to make accurate measurements with very low background ionisation. In a series of measurements in 1925, Millikan and Cameron lowered detectors into Californian lakes in an attempt to measure the absorption coefficient of the radiation. Contrary to Millikan's expectations, these results were consistent with the attenuation of extraterrestrial radiation with distance from the top of the atmosphere, with no additional radiation source required. It was after this work that Millikan first referred to the source of the ionisation as 'cosmic rays' [1, 2, 6].

Jumping forward to the 1930s, we come to the discovery of what would be termed extensive air showers (EAS). In 1938, Pierre Auger and Roland Maze conducted an experiment employing multiple Geiger counters to measure temporal coincidences of particle hits. Varying the distances between the detectors showed that coincidences were apparent out to distances of 20 metres [6]. In later experiments conducted at higher altitudes, coincidences were observed out to 300 m. These coincidences were concluded to be due to the arrival of a great number of particles from a common source, and the theory of multiplicative cascades was compatible with the observations. While the details of the original theory were erroneous, Auger cor-

rectly concluded from these experiments that the particles initiating the observed 'extensive air showers' (EAS) were in possession of energies of at least 10^{15} eV [8]. This value was far beyond the maximum known particle energies at the time and 10^8 times larger than accelerators could then produce - begging the question of their origin, one which is still awaiting a definitive answer today [1, 6]. The world's largest cosmic ray detector, the Pierre Auger Observatory (see Chapter 3), recognises its namesake for his important contributions to the field of particle astrophysics.

In the century following the discovery of cosmic rays, much has been learned about these elusive particles. Technological advances have allowed the construction of larger and more sensitive detectors, reaching sizes of thousands of square kilometres. However, fundamental questions such as the origin of cosmic rays remain a mystery, despite some compelling observations. It is hoped that the current and future CR detectors will have the accuracy and statistical gathering power to resolve some of these remaining questions.

1.2 Extensive Air Showers and Detection Techniques

Direct detection of cosmic rays is viable at low energies, where the CR flux at Earth is large enough (see Chapter 2) that satellite-borne detectors can record significant numbers of events. These detectors typically consist of several components which measure the basic properties of energy, mass and charge, together allowing the identification of particle species. This ability to distinguish the cosmic ray species on an event-by-event basis is highly desirable, although such a detector has a limited energy range imposed by the detector's physical dimensions [7].

However, above an energy of around 10^{14} eV, the cosmic ray flux becomes so small that indirect measurements must be employed [9]. At such energies, the properties of the extensive air showers produced in the Earth's atmosphere by incoming cosmic rays are measured instead. Thus direct detection gives way to the study of various EAS observables, from which the nature of the particle that initiated the shower is inferred.

Indirect detection is possible due to the large dimensions of the EAS which are produced in the Earth's atmosphere by incoming cosmic ray particles. The footprint at ground level can reach many square kilometres in area at high energies, increasing the effective aperture of a detector by requiring only that it lies within this footprint. In addition, light is produced along the path of the shower (see Section 1.2.1.3) which can be observed by specialised detectors from large distances. In this way, the atmosphere is used as a detection medium and, by building a detector of sufficient size for the energy range of interest, useful numbers of events can be measured even at extremely large energies.

This section serves to outline the processes involved in the development of extensive air showers as well as to introduce some of their general properties. In addition, two widely-used methods for the detection of ultra-high energy cosmic rays - ground arrays and atmospheric fluorescence - are discussed.

1.2.1 Formation and Morphology of Extensive Air Showers

Extensive air showers begin with the interaction between an incoming cosmic ray (hereafter referred to as the primary particle) and an air molecule in the Earth's atmosphere. Through a series of particle interactions and decays, a cascade of particles develops, which travels along the axis of the primary particle's motion. The leading edge of the cascade is known as the shower front, which is a thin disc-like structure which moves along the shower axis at close to the speed of light while spreading out laterally through as a result of particle interactions and Coulomb scattering. The number of particles ('shower size') in an EAS rapidly multiplies after the first interaction, building to a maximum number after which the shower size decreases due to energy losses to the surrounding environment [10]. It is through such interactions and energy losses that extensive air showers are detected.

It is appropriate to introduce here a quantity which will be used repeatedly in the following discussion. Measured in dimensions of length, the longitudinal development of an EAS depends on the local atmospheric density, and consequently any description of an EAS in terms of length would also require information about the atmospheric conditions in which it was observed. A more useful quantity for describing shower development is atmospheric depth, X. Atmospheric depth is the integral of the atmospheric density from the altitude of interest, h, to the top of the atmosphere (taken to be at infinity) along the axis of shower propagation:

$$X = \int_h^\infty \rho(l) dl \tag{1.1}$$

Most commonly expressed in units of $\mathrm{g\,cm^{-2}}$, X allows greater ease of comparison between showers observed in different locations or during different atmospheric conditions. Use of atmospheric depth does not completely decouple shower development from atmospheric density, however, as the relative likelihood of particle decay or interaction is sensitive to atmospheric density for species such as pions [7]. Nevertheless, atmospheric depth remains a convenient unit and is commonly used.

1.2.1.1 Heitler's Model of Electromagnetic Cascades

A simple theory of electromagnetic cascades was proposed by Heitler which, while crude, qualitatively illustrates some basic features of particle showers and serves as a useful introduction to their properties [11]. The model includes electrons, positrons and photons and considers the energy loss mechanisms of bremsstrahlung:

$$e + N \rightarrow e + N + \gamma \tag{1.2}$$

as well as pair production:

$$\gamma + N \rightarrow e^+ + e^- + N \tag{1.3}$$

where N represents a nucleus in the medium in which the cascade develops.

Heitler makes the approximations that these two processes have equal interaction lengths and that interactions occur every time the particles traverse a distance $\Delta t = \ln 2$, where t is measured in units of the radiation length of the particles in the medium. Following each interaction, the two resulting particles (e and γ for bremsstrahlung, or $e^+$and e^- for pair production) are assumed to inherit exactly half of the energy of the interacting particle.

Consider now, with the features of the model as described above, an electron of energy E_0 travelling through some medium. At a depth $t = \ln 2$ radiation lengths, a photon will be created through bremsstrahlung. The energies of the resultant particles are assumed to be equal:

$$E_e = E_\gamma = \frac{1}{2}E_0 \tag{1.4}$$

Following this first interaction, the electron will undergo bremsstrahlung upon reaching a depth $t = 2\ln 2$ radiation lengths, and similarly the photon will undergo pair production at the same depth. These processes continue as the cascade develops, with the number of particles at each interaction point doubling and the energy of the particles following each interaction being half of that of the particles prior to the interaction.

Thus, at a given interaction level $t = k\ln 2$ for an integer k, the number of particles, N_p, in the shower is:

$$N_p = e^t = 2^k \tag{1.5}$$

and the energy of each particle is:

$$E = \frac{E_0}{N_p} = E_0 e^{-t} \tag{1.6}$$

The cascade is assumed to continue until the particle energy drops below a value E_c, at which point collisional energy losses dominate radiative losses. Below such a threshold, it is assumed that multiplication ceases and the particles quickly dissipate their energy through ionisation. In air, this critical energy is approximately 80 MeV [10].

Thus, the maximum number of energetic particles in the shower is reached when $E_c = E_0 e^{-t}$, which gives the depth at which the shower reaches its maximum size:

$$t_{max} = \ln \frac{E_0}{E_c} \tag{1.7}$$

which can be given in terms of atmospheric depth by multiplying both sides by the assumed radiation length, λ_r, in $\mathrm{g\,cm^{-2}}$ (which, for electrons, has a value of $\lambda_r \approx 37\,\mathrm{g\,cm^{-3}}$ in air [10]):

$$X_{max} = \lambda_r \ln \frac{E_0}{E_c} \tag{1.8}$$

Figure 1.1 illustrates the development of a shower using Heitler's model. While details of

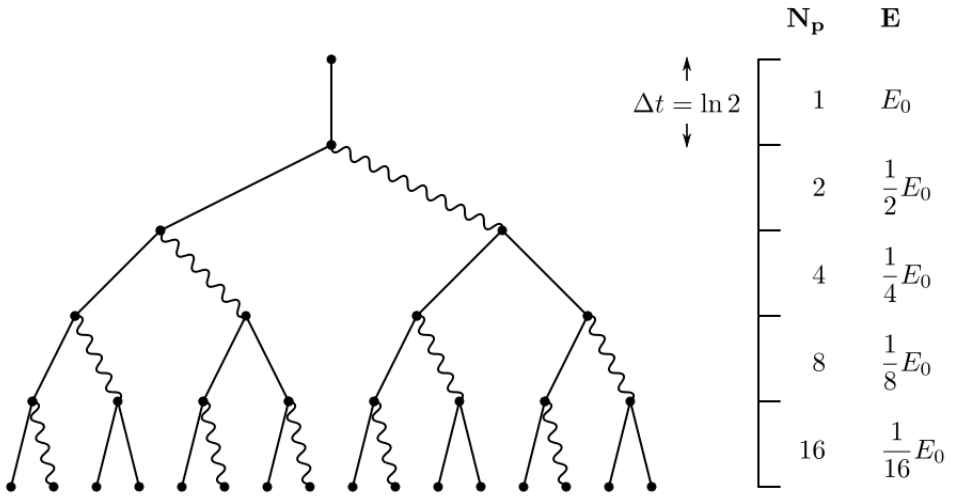


Figure 1.1: An illustration of cascade development in Heitler's toy model. The legend to the right indicates the number of particles N_p, as well as the energy E of each particle, at several stages of shower development. The straight lines represent electrons, while the wavy lines represent photons.

shower development are not well-represented by such a simple model, it correctly describes two important features of cascades: the number of particles at shower maximum is proportional to the primary energy, and the depth of shower maximum is proportional to the logarithm of that energy [7, 11].

1.2.1.2 Hadronic Air Showers

Hadronic showers begin with the interaction of a single nucleon, or a nucleus, with an air molecule, at a depth which follows an exponential distribution. In this interaction approximately half of the primary's energy is converted into mesons, which for simplicity it will be assumed are either charged or neutral pions. Approximately two-thirds of the pions will be charged. The pions will either interact, producing subsequent generations of pions, or decay. Neutral pions, having a mean laboratory lifetime of $\approx 8 \times 10^{-17}\,\mathrm{s}$, decay very quickly unless they are of very high energy, with the main decay mode ($\approx 99\%$ chance) being the production of two photons [12]:

$$\pi^0 \rightarrow \gamma + \gamma \tag{1.9}$$

The resulting photons, if they are of high enough energy, will pair produce and initiate electromagnetic sub-showers. In contrast, charged pions have a much longer mean lifetime of around $3\times10^{-8}\,\mathrm{s}$; the chance of interaction, especially at high energies, is rather high. However, the competition between decay and interaction, especially in an atmosphere of varying density, is complicated and introduces significant fluctuations between the structures of individual hadronic showers [7]. Interaction produces further generations of pions, again in a ratio of

two-thirds to one-third charged and neutral, respectively, and continues until the pion energies reach a threshold after which decay almost always occurs [7]. Decay of the charged pions produces muons with almost 100% probability [12]:

$$\pi^+ \rightarrow \mu^+ + \nu_\mu \tag{1.10}$$

$$\pi^- \rightarrow \mu^- + \bar{\nu}_\mu \tag{1.11}$$

It is through these charged pion decays that hadronic showers develop a significant muon component which, along with large fluctuations in development between individual showers, serves to distinguish air showers of hadronic origin from electromagnetic showers. These muons penetrate deeply into the atmosphere due to slow energy losses ($\sim 2\,\mathrm{MeV}/(\mathrm{g\,cm^{-2}})$) as a result of ionisation of air molecules which leads to much of the muon content surviving at ground level, in contrast to the electromagnetic components which attenuate rapidly after shower maximum [10].

The development of a hadronic EAS can thus be pictured as a number of overlapping independent electromagnetic cascades, initiated at many different points in the atmosphere, with a muon component that contains information on the hadronic interactions occurring within the shower. While a truly hadronic component is present within the shower, it contains far fewer particles than the number of muons, electrons and photons present [7]. Most of the energy of the primary particle is eventually dissipated through the electromagnetic component of the shower, and at 10^{20} eV this fraction is approximately 90% [10, 13].

A parametrization of the average longitudinal development of hadronic air showers, known as the Gaisser-Hillas profile, can be used to describe the size of an EAS as a function of atmospheric depth:

$$N_e\left(X\right) = N_e^{max}\left(\frac{X - X_0}{X_{max} - X_0}\right)^{\frac{X_{max}-X_0}{\lambda}} e^{-\frac{X-X_{max}}{\lambda}} \tag{1.12}$$

where four free parameters are to be fitted; X_0 and λ are shape parameters, while N_e^{max} is the number of electrons reached at the depth of shower maximum X_{max} [14, 15, 16]. Such a profile can be fitted to EAS observed through the fluorescence emission they induce, allowing both the shower energy and X_{max} to be estimated (see Section 1.2.2.2).

As the shower front propagates through the atmosphere, the more massive muons suffer only small deflections from Coulomb scattering and as a result define the shower front's leading edge. By comparison, the electrons undergo considerable scattering and take longer to reach ground level. These effects result in a thin shower front having large particle densities close to the shower core, where the electrons and photons far outnumber the muons, while away from the axis the particle densities quickly decrease and the shower front becomes thick, with the relative muon content becoming much larger due to the muon lateral distribution being flatter than the electromagnetic one. The width of the shower front at core distances of several

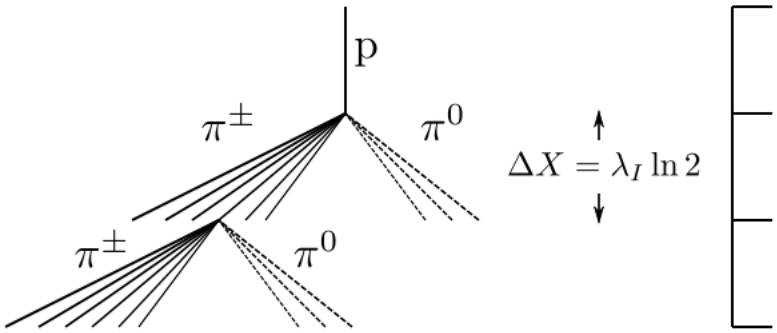


Figure 1.2: An illustration of Matthews' model of hadronic extensive air showers for the first two interaction levels. The dotted lines represent the electromagnetic showers initiated by the decay of neutral pions. Note that, for clarity, the result of only a single one of the second generation interactions is shown.

hundreds of metres means that detector signals can spread out over hundreds of nanoseconds [7, 17]. The scale of lateral spread of the electromagnetic component of the shower can be characterised by the Molière radius, r_M, which gives the spread of low energy particles due to Coulomb scattering - for low energy electrons at sea level it is roughly 80 m [10]. At ground level the total footprint of high energy showers can reach values of many square kilometres.

Matthews has adapted the Heitler model of cascades to include hadronic processes to gain an understanding of shower development from hadronic primary particles [18]. Like Heitler's model of electromagnetic cascades, Matthews' is simplified but its predictions are shown to be in good agreement with both observations and detailed Monte Carlo simulations.

To begin, consider a single proton, having energy E_0, incident upon the Earth's atmosphere. It will be assumed that, by analogy with Heitler's toy model, the interaction points for the hadronic component are spaced apart by $\lambda_I \ln 2$, where λ_I is the interaction length ($\approx 120\,\mathrm{g\,cm^{-2}}$ for pions in air). At each interaction point, N_{ch} charged pions and $\frac{1}{2}N_{ch}$ neutral pions of equal energy are produced, until the energy of the charged pions falls below $E_{\pi,c}$, the critical energy below which they will decay to muons rather than interact - taken to be the energy for which the decay length becomes less than $\lambda_I \ln 2$ (20 GeV is adopted in [18], but it is dependent on the surrounding atmosphere). The neutral pions produced in each interaction immediately decay and the resulting photons produce electromagnetic cascades. Figure 1.2 illustrates the first two generations of a shower under this model.

Approximating the multiplicity N_{ch} of charged particles for proton-proton and pion-proton collisions to be equal, after n interactions there are $(N_{ch})^n$ charged pions in the shower. The total energy contained by the charged pions in the shower is $\left(\frac{2}{3}\right)^n E_0$, giving the energy per charged pion as:

$$E_\pi = \frac{E_0}{\left(\frac{3}{2}N_{ch}\right)^n} \tag{1.13}$$

The remainder of E_0 has been converted into electromagnetic cascades initiated by the neutral pions. We define n_c to be the number of interactions after which $E_\pi \leq E_{\pi,c}$ which, from rearranging Equation 1.13, is found to be:

$$n_c = \frac{\ln\left(\frac{E_0}{E_{\pi,c}}\right)}{\ln\left(\frac{3}{2}N_{ch}\right)} \tag{1.14}$$

In this model, all of the energy is spread between the electromagnetic cascades and the muons. As a result, the primary particle's energy can be expressed as:

$$E_0 \approx 0.85\,\mathrm{GeV}\,(N_e + 24N_\mu) \tag{1.15}$$

where N_e and N_μ are the numbers of electrons and muons at shower maximum, respectively, and the relative weighting between the contributions from the electrons and muons is determined by the ratio of their critical energies. This relationship has been corrected for an overabundance of photons in the electromagnetic component as compared to Heitler's model, which is understood and confirmed by Monte Carlo simulations [18]. From Equation 1.15 it can be seen that, if the number of charged particles at shower maximum is known, an estimate of the primary particle's energy can be made.

From these results it is possible to predict the number of muons in a shower. As it is assumed that all charged pions eventually decay to muons, the number of muons is given by $N_\mu = (N_{ch})^{n_c}$. Use of Equation 1.14 then gives:

$$\ln N_\mu = n_c \ln N_{ch} = \beta \ln\left(\frac{E_0}{E_{\pi,c}}\right) \tag{1.16}$$

or:

$$N_\mu \propto E_0^\beta \tag{1.17}$$

where β is the ratio of the logarithms of the charged and total particle multiplicities in each hadronic interaction:

$$\beta = \frac{\ln N_{ch}}{\ln\left(\frac{3}{2}N_{ch}\right)} \tag{1.18}$$

The value of β varies between different hadronic interaction models, but a value of 0.85 is representative [13, 18].

The energy-dependence of the average value of depth of shower maximum for protons, X^p_{max}, is also possible to predict with this model. To do so, only the first generation of electromagnetic showers is considered - this leads to a systematic underestimation of X^p_{max},

but its behaviour with energy can be approximated [18]. To do so, the depth of shower maximum can be written as the sum of the depth of first interaction, X_1, and the depth of maximum of the electromagnetic showers:

$$X^p_{max} = X_1 + \lambda_r \ln\left(\frac{E_0}{3N_{ch}E_{c,e}}\right) \tag{1.19}$$

where the second term follows from the depth of maximum of an electromagnetic shower in Heitler's model, as initiated by a photon of energy $E_\gamma = \frac{E_0}{3N_{ch}}$. This value of photon energy follows directly from the assumed details of the first interaction. To enable comparison with electromagnetic showers, Equation 1.19 may be written as:

$$X^p_{max} = X^\gamma_{max} + X_1 - \lambda_r \ln\left(3N_{ch}\right) \tag{1.20}$$

where X^γ_{max} is the depth of shower maximum of electromagnetic showers. Thus, it can be seen that showers initiated by proton primaries develop shallower in the atmosphere than purely electromagnetic showers. The energy dependence of X^p_{max} is commonly expressed in terms of the 'elongation rate' Λ^p - the change in average X^p_{max} over a decade of primary energy:

$$\Lambda^p = \frac{dX^p_{max}}{d\log_{10}E_0} = \Lambda^\gamma + \frac{d}{d\log_{10}E_0}\left[X_1 - \lambda_r \ln\left(3N_{ch}\right)\right] \tag{1.21}$$

It is found that X_1 decreases with energy as the proton-proton collision cross-section increases with energy, and N_{ch} increases with energy. This leads to the elongation rate for protons being reduced from Λ^γ, the elongation rate of electromagnetic showers. Matthews' model predicts $\Lambda^p = 58\,\mathrm{g\,cm^{-2}}$ per decade of energy, in good agreement with simulation results [18].

The above results derived for protons may also be applied to heavier nuclei using similar arguments. The 'superposition principle' approximates a nucleus consisting of A nucleons as being A independent protons arriving together, each possessing energy $\frac{E_0}{A}$. Each of these nucleons initiates its own shower which develops independently. Adapting the results above shows that, on average, heavier nuclei produce more muons and their showers reach maximum at a shallower depth than a proton-initiated shower of the same energy:

$$N^A_\mu \propto A\left(\frac{E_0}{A}\right)^{0.85} = A^{0.15}N^p_\mu \tag{1.22}$$

$$X^A_{max} = X_1 + \lambda_r \ln\left(\frac{E_0}{3AN_{ch}E_{c,e}}\right) = X^p_{max} - \lambda_r \ln A \tag{1.23}$$

For example, an iron shower ($A = 56$) would contain roughly 80% more muons and reach shower maximum about $150\,\mathrm{g\,cm^{-2}}$ earlier than a proton shower of equal energy, assuming a radiation length of $\lambda_r = 37\,\mathrm{g\,cm^{-2}}$. In addition, fluctuations in X_{max} will be reduced for nuclei over protons, and the measured depth of maximum distribution at a given energy

will be narrower. This reduction in the fluctuations of X_{max} is due to A individual showers superimposing, although the reduction factor will be less than $\frac{1}{\sqrt{A}}$ due to the interactions of the nucleons not being entirely independent [19]. Finally, if the composition of the CR flux changes with energy (A is energy-dependent) then this will be reflected in the elongation rate where an additional term of the form $\frac{d\ln A}{d\ln E}$ is introduced [20].

Modifications to this model may be made to estimate the effects of varying inelasticity. As the primary particle carries away a fraction of its energy after each interaction, rather than that energy being completely converted into charged and neutral pions as assumed above, the predictions of hadronic shower properties will be altered accordingly. The inelasticity, κ, gives the fraction of the total energy that is converted into charged and neutral pions at each interaction. For values of $\kappa < 1$, some of the energy is retained by the primary particle. Including such an effect alters the predicted value of X_{max}, which increases for decreasing κ, as well as the number of muons produced, through a modification of β in Equation 1.17 [18].

The model described above predicts, with good accuracy, the basic properties of hadronic air showers. However, as previously mentioned, complicated simulations are required to study EAS development in greater detail. More information about such studies can be found in [7, 10, 21] and references therein.

1.2.1.3 Atmospheric Fluorescence and Čerenkov Light Production

Atmospheric nitrogen fluorescence and Čerenkov light are both products of the passage of charged particles through air (as well as water for the latter) and are particularly important for their role in allowing the detection of extensive air showers. Specialised detectors are able to collect such light, allowing analysis and reconstruction of the primary particles' properties.

As the charged particles in an EAS pass through the atmosphere they excite nitrogen molecules. Upon de-excitation of those molecules, light is emitted isotropically in the wavelength range of $\sim 300 - 430\,\text{nm}$. A measured spectrum of such light is shown in Figure 1.3. Importantly, a proportionality exists between the energy deposited by the charged shower particles in the atmosphere and the number of fluorescence photons emitted - for example, in the main emission band of 337 nm the absolute fluorescence yield is approximately 5 photons per MeV of energy deposited by the shower particles at room temperature and atmospheric pressure [22]. Knowledge of such a yield allows the conversion from the number of photons collected by a detector to the energy deposited electromagnetically by shower particles in the atmosphere, which accounts for approximately 90% of the primary particle's energy [23]. This is advantageous as the yield can be accurately measured in laboratory conditions, resulting in cosmic ray energies that are estimated through fluorescence measurements being almost calorimetric. A small correction must be made for the unobserved 'missing energy' - energy that is not fully deposited into the atmosphere, such as that carried by high energy muons and neutrinos which is instead deposited into the ground [22, 24].

Despite the fluorescence yield being quite small, at the highest cosmic ray energies enough

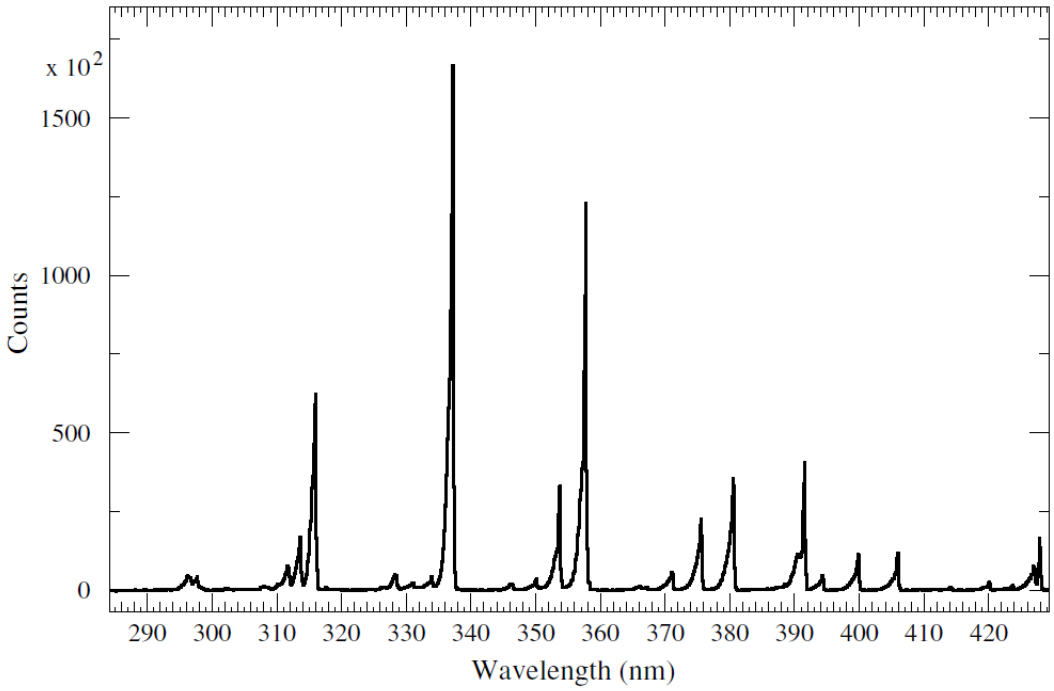


Figure 1.3: The air fluorescence spectrum measured in dry air at 293 K and 800 hPa. From [26].

light is produced to be able to distinguish it from any background sources, given appropriate viewing conditions. The operation of fluorescence detectors is discussed in Section 1.2.2.2.

Presently, lack of knowledge of the fluorescence yield contributes a dominant systematic error to the overall energy scale of fluorescence detector experiments such as the Pierre Auger Observatory, but efforts are ongoing to make much more precise measurements of the fluorescence yield (see, for example, [25]).

Čerenkov light is produced when charged particles move through a dielectric medium at a speed v greater than that of light in the medium, $\frac{c}{n}$ for a refractive index n. The passage of a charged particle polarises surrounding molecules which, upon returning to an unpolarised state, emit radiation. In the case of the charged particles travelling slower than the local speed of light, this radiation is incoherent; however, it becomes coherent and may be observed if $v > \frac{c}{n}$. In air, where the refractive index at sea level is 1.0003, the threshold energy for the production of Čerenkov light is $\approx 21\,\text{MeV}$ for electrons; this value decreases to about 1 MeV in water [7].

This light is emitted in a cone around the axis of the charged particle's motion, with an opening angle θ given by:

$$\cos\theta = \frac{1}{\beta n} \tag{1.24}$$

and intensity given by:

$$I \propto q^2 \left[1 - \frac{1}{(\beta n)^2}\right] \tag{1.25}$$

for particle charge q [2]. In air, the emitted light is predominantly in the visible and UV range [7]. Čerenkov emission has been used in several CR detectors (see Chapter 3), including Yakutsk, Haverah Park and the Pierre Auger Observatory, the latter two employing enclosed pools of water as a detection medium.

1.2.2 Extensive Air Shower Detection

Two common types of detector have been developed for the detection of extensive air showers and the reconstruction of the properties of the primary particle - ground arrays and fluorescence detectors. Both take advantage of the large spatial scales over which EAS develop within the atmosphere to increase their effective aperture to far larger sizes than would be achievable through direct detection methods. The following sections introduce these two methods and describe, in general terms, how the properties of primary particles are reconstructed. Details of several past and current EAS detectors are presented in Chapter 3.

1.2.2.1 Ground Arrays

Although the details of shower measurement and reconstruction vary between ground array detectors, the basic method used remains the same. By employing a number of particle detectors distributed over an area of ground, the array measures signals produced in those detectors at multiple positions as the shower front moves past them. If desired, these signals, and knowledge of the detector response to different particle species, allow the reconstruction of the number of particles crossing each detector. Such signals and the time at which they are measured allow reconstruction of the basic shower properties such as primary energy and arrival direction, while a more complicated analysis may permit studies of properties such as particle composition.

An obvious advantage of ground arrays is their ability to operate constantly (compare with the $\sim 10\%$ duty cycle of fluorescence detectors, see Section 1.2.2.2), giving them an ability to gather very large data sets [27]. Approximately uniform exposure in right ascension - minor corrections are required for detector down-time - is an additional benefit of such a large duty cycle [28].

When designing a ground array, its size, the spacing between individual detectors, and the atmospheric depth at which the array is situated, must all be chosen appropriately for the energy range of interest. For example, above an energy of 10^{19} eV the cosmic ray flux at Earth is of the order one particle per square kilometre per year, and the shower footprint at ground level several square kilometres [28, 29]. To collect reasonable statistics a detector would need to be many square kilometres in size, with a detector spacing no larger than the scale of the shower footprint, to enable accurate EAS reconstruction. In addition, the average depth of

shower maximum at this energy is approximately $750\,\mathrm{g\,cm^{-2}}$; ground arrays at an atmospheric depth greater than $\sim 800\,\mathrm{g\,cm^{-2}}$ are ideally suited to the measurement of these showers, as detection close to shower maximum reduces the effect of fluctuations in shower development [17].

Reconstruction of an event detected by a ground array generally begins with the determination of its arrival direction. This is achieved by fitting the arrival times of the shower in each detector with an adopted model of a shower front, assuming that the shower front propagates at the speed of light, to determine the shower axis. While a planar model may suffice close to the shower core, detectors which can measure signals at large core distances are required to account for shower front curvature [7, 30]. Typically, an angular resolution of several degrees is achieved with ground arrays, although events involving large numbers of detectors may have angular resolutions of less than 1°. Precise arrival direction reconstructions also require good timing resolution in, and synchronisation between, the individual detectors [17, 31].

Following the determination of the shower axis, the core location - the intersection of the shower axis with the ground plane - and energy of the event can be estimated by fitting a lateral distribution function (LDF) to the signals recorded by each detector. Depending on the experiment, the LDF variously describes, as a function of core distance (the perpendicular distance from the shower axis), the shower particle density, signal density produced by the shower particles, or simply the signal recorded by a given detector. This function depends not only on the distribution of particles in the shower but also on the type of detector being employed and, as a result, varies between experiments [7, 17, 32]. For example, the Haverah Park experiment, using water-Čerenkov detectors, described the signal density ($\mathrm{m^{-2}}$) at energies above $2 \times 10^{17}\,\mathrm{eV}$ as:

$$\rho(r) \propto r^{-\left(\eta + \frac{r}{4000}\right)} \tag{1.26}$$

for r in metres and k a normalisation parameter. Zenith angle and energy dependence of the LDF was incorporated into η, given by:

$$\eta = 3.52 - 1.22 \sec\theta + 0.15 \log\left(\rho(500)\right) \tag{1.27}$$

for zenith angle θ and $\rho(500)$ the signal density at 500 m from the shower core [33]. By comparison, the AGASA detector, employing scintillator detectors, adopted the following LDF, where r_M is the Molière radius at the AGASA site:

$$\rho(r) \propto \left(\frac{r}{r_M}\right)^{-1.2} \left(1 + \frac{r}{r_M}\right)^{-(\eta - 1.2)} \left(1 + \left(\frac{r}{1000\,\mathrm{m}}\right)^2\right)^{-\delta} \tag{1.28}$$

Values of $\eta = 3.8$ and $\delta = 0.6$ were found to best describe near-vertical showers with primary energies between 6×10^{17} and $2 \times 10^{18}\,\mathrm{eV}$ [34].

After fitting an appropriate function, a 'ground parameter' is commonly used to relate the fitted LDF to primary energy. Extracted from the fit - or alternatively from interpolation

between detectors with a measured signal - this parameter gives the value of the LDF that would have been observed had a detector been present at a specified distance from the shower core, usually in the range of 500-1 000 m. Next, a correction is required to compensate for the varying attenuation of shower particles at different zenith angles [24]. Finally, a relationship between this ground parameter and primary energy E - obtained from air shower simulations, and typically of the form $E = \beta P_{ground}^{\alpha}$, for the ground parameter P_{ground} used in the experiment and appropriate values of α and β - is used to calculate the energy of the particle that initiated the shower [13, 35, 36].

Use of such a parameter follows from the work of Hillas, in which it was discovered that the signal at these core distances was less prone to large variations resulting from variations in primary particle species, uncertainties in the true LDF, and fluctuations in shower development than estimation of the total number of particles at ground level [37]. This stability is a consequence of the particles observed in this region being produced near shower maximum and hence being more consistent in number for a given energy. Hillas found a value of 500 m was suited for the simulations he considered, but the optimum value will vary with detector spacing, atmospheric depth and energies of interest [24, 29].

Requiring air shower simulations to determine the relationship between a ground parameter and primary energy necessarily introduces a reliance on models of hadronic interactions at energies above which experimental verification is possible [28]. Unless methods independent of these simulations are available for energy estimation (see Section 1.2.2.3) the overall energy scales of ground arrays at extremely high energies are subject to resulting systematic uncertainties which can make comparison of results between detectors difficult [17, 24, 38].

Of the three basic properties of the CR flux to be measured - arrival direction distribution, energy spectrum and composition - the latter is perhaps the most difficult to determine with ground array detectors. Unlike fluorescence detectors, which can be used to reconstruct the longitudinal development profile of an EAS, ground arrays only sample the shower front at a single altitude. This does not allow X_{max} to be directly determined and, as a result, other indicators of primary composition must be sought. Being sensitive to the primary particle species makes the study of the muon content a candidate for the discrimination of particle species. Examples of such studies include the measurement of muon densities at ground level, which for a given energy contains some information about the primary species, and calculating the 'rise-time' - the time taken for the integrated signal to increase from 10% to 50% of its final value - of the detector signals, which is also an indication of the muon-electromagnetic component ratio. In addition, the amount of atmosphere traversed by shower particles varies as a function of angle around the shower axis for inclined showers - effectively offering measurements of shower development at several different stages of development. As the longitudinal development of a shower is affected by composition, such 'asymmetry' studies can be used to reveal information on primary particle species [17, 39, 40, 41].

Scintillators and water-Čerenkov tanks are two commonly used detector types in ground

arrays. As their name suggests, scintillator detectors emit pulses of light as charged particles pass through them. The emitted light is usually collected by a photomultiplier tube. In effect, scintillators count the number of charged particles incident upon them, and are not able to easily distinguish muons from electrons. In order to do so, shielding is generally employed to absorb the electromagnetic component of the shower front before it reaches the detector [10]. Photons are not efficiently absorbed by scintillator detectors [17].

Water-Čerenkov tanks detect the Čerenkov emission which occurs as shower particles pass through a pool of water enclosed within the detector. Advantages over scintillators include a larger zenith angle acceptance, thus giving the array a larger aperture than an equal-size scintillator array, as well as the ability to cheaply construct deep detectors - each of the Pierre Auger Observatory's roughly 1600 ground array detectors (see Chapter 3) contains a pool of water 1.2 m deep [42]. This depth means the water-Čerenkov detectors present several radiation lengths to incoming photons, allowing them to undergo pair-production and the resulting electron-positron pair to be seen by the detector. The result is to ensure consistent sampling of all the dominant shower components and to reduce reconstruction uncertainties due to statistical fluctuations in the number of detected particles [32]. The signals produced by muons in such tanks are easily distinguished from those of the electromagnetic component owing to their distinct size and shape [7, 42].

It is important to note that the reconstruction methods described above generally only apply within a certain range of zenith angles, outside which the degree of attenuation of shower particles may become large enough that the adopted LDF is no longer a good description of the particle distribution at ground level, and zenith angle corrections of the ground parameter break down. Consequently, the reconstruction of highly-inclined showers requires different methods to those at smaller zenith angles to account for such differences [36].

1.2.2.2 Fluorescence Detection

In contrast to ground arrays, which sample the shower front at a single altitude using detectors spread out over a large area, fluorescence detectors (FDs) measure the nitrogen emission induced by an EAS over a range of shower ages, typically from a single location. While the duty cycle of fluorescence detectors is less than that of ground arrays as a result of practical requirements, they offer the ability to study longitudinal shower development in detail and offer some unique benefits for cosmic ray studies.

A basic instrument design for fluorescence detection which has been used several times (see Chapter 3) with great success is to image the atmospheric nitrogen emission using a number of pixels that each view a separate region of sky. Optics appropriate for the detector design direct light to the sensitive surface of each pixel which, in practice, is the photocathode of a photomultiplier tube (PMT). A distant shower traversing the atmosphere is seen as a small spot of light that moves across this photomultiplier camera as the shower approaches ground level. By measuring the sequence of pulse times from each pixel, the shower geometry can be

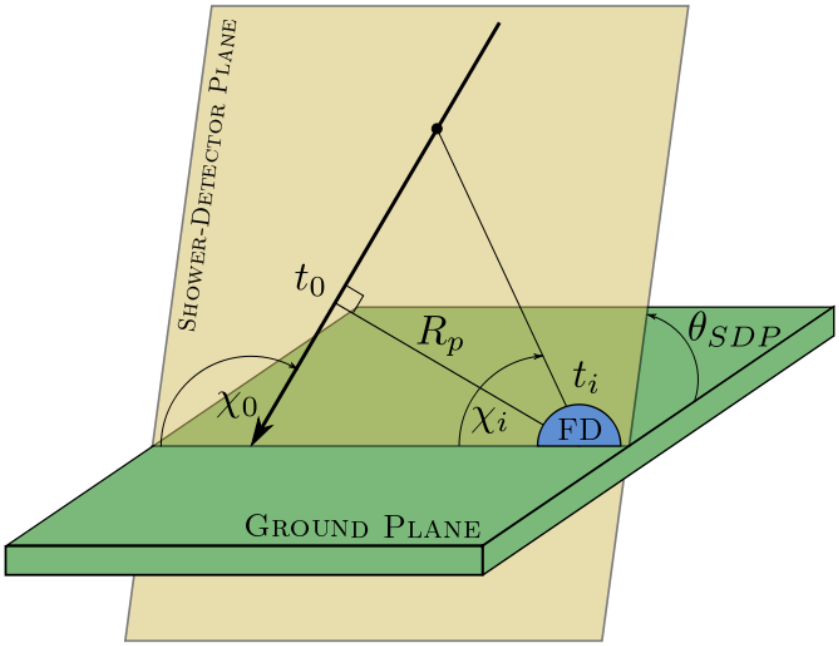


Figure 1.4: This diagram indicates the parameters used to define the geometry of an EAS observed by a fluorescence detector, as described in the text.

reconstructed. The primary particle's energy is estimated from the integral of these pulses. These concepts are discussed below.

As the reconstruction of event energy relies on relating the flux of photons observed at the detector to that which is emitted from within the shower, the geometry must first be defined; that is, the axis along which the EAS developed must be found. The standard method for describing such an axis is to define the 'shower-detector plane' (SDP), the angle that the shower axis (which lies within the SDP) makes with the ground (χ_0), and the distance of closest approach to the detector. Together these three quantities fully describe the shower geometry. Figure 1.4 illustrates a fluorescence event and the quantities used to describe its geometry.

Determination of the shower-detector plane is the first step in geometry reconstruction and is simply a fit to the pointing directions of each pixel which registered a signal, weighted by the size of that signal [27]. This gives a plane that intersects the detector and contains the axis of shower propagation. The accuracy of this fit relies on the rejection of pixels which have triggered accidentally, and is also affected by the angular size of each pixel. The angle θ_{SDP} is defined as that between the ground plane and the SDP.

Following the SDP fit, timing information is used from each pixel to find the shower axis. If the impact parameter, or distance of closest approach of the shower to the detector, is defined

as R_p, the time at which the shower crosses this point as t_0, and the angle that the axis makes with the horizontal as χ_0, then the expected time of arrival of the light in the i^{th} pixel is given by:

$$t_i = t_0 + \frac{R_p}{c} \tan\left(\frac{\chi_0 - \chi_i}{2}\right) \tag{1.29}$$

where c is the speed of light and χ_i is the angle of elevation of the i^{th} pixel within the SDP [23]. For a fluorescence detector working in isolation (referred to as 'monocular mode') the axis is found by taking t_0, R_p and χ_0 as the parameters to be found and minimizing the sum:

$$\chi^2 = \sum_{i=1}^{N} \frac{(t_i - t_{i,obs})^2}{\sigma_i^2} \tag{1.30}$$

where $t_{i,obs}$ is the observed signal time in the i^{th} pixel and σ_i the uncertainty in its value. Problems can occur when little variation is present in the angular velocity $\frac{d\chi}{dt}$ of the shower as seen by the FD, causing degeneracy between the fit parameters R_p and χ_0. In these cases the shower geometry is not uniquely determined. However, the use of extra timing information from a ground array can serve to break this degeneracy and recover the shower axis (see Section 1.2.2.3).

Once the shower's geometry is established, calculation of the primary particle's energy can begin. To do this, the light collected at the PMT photocathodes must be converted to the energy deposited by the shower into the atmosphere as a function of atmospheric depth. While simple in concept, significant challenges arise in the conversion from collected light to energy deposit profile. Altitude profiles of atmospheric density must be well understood, as well as air pressure, temperature and humidity which all affect the yield of nitrogen fluorescence emission [43]. Account must also be taken of the attenuation of the fluorescence light as it travels from its point of emission to the detector, requiring accurate and detailed knowledge of contemporaneous atmospheric conditions; in addition, clouds and aerosols can obstruct and scatter the light, further altering the flux at the detector, meaning knowledge of their presence is highly important [28]. In addition, all sources contributing to the light flux must be accounted for. Čerenkov light from the shower may reach the detector directly or be scattered into its optics, background light from stars and terrestrial sources must be subtracted, and even some fluorescence photons that would not otherwise have reached the detector may be scattered into it, increasing the amount of light being gathered and affecting the energy estimation if not accounted for, as well as altering the shape of the profile and affecting the determination of X_{max} [44].

Following the calculation of the energy deposit profile, a Gaisser-Hillas function may be fitted to it. This process returns the estimated value of X_{max} for the shower. Subsequent integration of the fitted function returns the energy of the event, apart from the small correction which must be made to account for missing energy. At high energies such corrections are minor

and are found through air shower simulations [7]. The proportionality between the number of charged particles in the shower and the number of fluorescence photons emitted means that energy estimation in this manner is almost calorimetric, in contrast to the heavy reliance on air shower simulations required to derive energies with ground arrays. Systematic uncertainties in energy estimation with a fluorescence detector are different to those from ground arrays. For experiments such as the Pierre Auger Observatory, the dominant uncertainties arise from current knowledge of the fluorescence yield and detector calibration [23].

The ability to find X_{max} directly means that fluorescence detectors are ideally-suited for composition studies. Measurements of X_{max} with uncertainties of roughly $20\,\mathrm{g\,cm^{-2}}$ are achievable, permitting useful studies of primary mass. Methods used to study composition with fluorescence detectors include measuring the elongation rate, discussed in Section 1.2.1.2, and finding the width of measured X_{max} distributions as a function of energy. In both cases, results can be compared with air shower simulations. Such studies can reveal how 'heavy' the CR flux is at a given energy, commonly with reference to expectations from pure proton or iron primaries [17, 19, 45].

A single fluorescence detector's aperture can, in favourable atmospheric conditions, reach very large values - thousands of $\mathrm{km^2\,sr}$ are achievable. In addition, due to the proportionality between primary energy and the amount of fluorescence light produced, this aperture grows with energy [46]. However, in contrast to ground arrays, which are capable of operating with a 100% duty cycle, fluorescence detectors average around 10%. Detection of atmospheric fluorescence is possible only on clear nights with low levels of background illumination, prohibiting data collection during daytime, bad weather or when significant moonlight is present. An additional consequence of such requirements is a somewhat complicated detector exposure that requires careful monitoring of uptime for its calculation [47]. Operation in conjunction with a ground array, discussed below, can offer a solution to such a low duty cycle.

As an alternative to operation with a ground array, the quality of FD reconstruction can be improved with an experiment consisting of multiple fluorescence detectors, spaced apart and operating independently. Such a configuration increases the aperture of the experiment, but also allows some events to be detected by multiple FDs. For example, 'stereo' events - those seen by two different FDs - may have their axis reconstructed simply by finding the intersection of the SDPs calculated using the data from each detector, thereby avoiding the need for a timing fit [17]. In addition, multiple FDs allow statistical reconstruction uncertainties to be evaluated directly from data, through fitting stereo (or higher-multiplicity) events independently with each detector and then comparing the parameters found through each reconstruction [38, 45].

1.2.2.3 'Hybrid' Detection

While many detectors employing a single measurement technique have operated successfully (see Chapter 3), the complementary nature of ground arrays and fluorescence detection nat-

urally leads to a design which incorporates both in a single detector. Such a detector is commonly, and hereafter, referred to as a 'hybrid' [7, 17, 27].

Many of the difficulties encountered by ground arrays or fluorescence detectors can be alleviated or avoided entirely in a hybrid detector. In general terms, the ground array is used for the majority of data collection owing to this component's large duty cycle. The fluorescence detector, operating only a fraction of the time, also observes a subset of events - referred to as hybrid events - of those detected by the ground array. Combining information from the two types of detector allows an event to be reconstructed with greater accuracy than would otherwise be possible with the individual detectors [38, 48]. In addition, hybrid events may also be used for deriving the conversion between the signal an event produces at ground level and its energy as measured by the FD, giving a self-consistent energy scale for the detector. These unique advantages of a hybrid detector are discussed in further detail below.

Perhaps the most straightforward benefit of a hybrid system is the ability to incorporate timing information from the ground array into the track seen by the fluorescence detector [23]. For showers where the angular speed $\frac{d\chi}{dt}$, as observed by a single fluorescence detector, displays little departure from linearity, the fit parameters R_p and χ_0 can display a high degree of degeneracy. When this occurs, a highly correlated, but large, range of possible (R_p, χ_0) solutions satisfy the data points and the shower geometry is not well constrained. Typically, this is the case when only a small track segment is observed. If, however, the shower also triggers ground-based detectors, the time of ground impact, t_{ground}, is known. Owing to the difference in elevations between the data points from the two detectors, the addition of timing information from the ground array serves to establish the curvature in $\frac{d\chi}{dt}$ and consequently the fit is constrained. The time of arrival in even a single detector of the ground array can be sufficient to resolve the (R_p, χ_0) degeneracy and hence determine shower geometry accurately [27]. This minimal requirement also enables the detection of events which would not have triggered the ground array only and allows good-quality reconstruction at low energies [23]. It is important to note that any timing offset between the two detectors must be well-understood to enable hybrid reconstruction [38, 49]. Figure 1.5 shows an example of core location uncertainties of an event reconstructed with both monocular and hybrid methods using the Pierre Auger Observatory (see Chapter 3).

Well-constrained geometries reduce the statistical uncertainty in the estimation of the primary particle's energy through integration of its reconstructed energy deposit profile [48]. Hence, hybrid events allow a comparison between the ground parameter used for energy estimation from the ground array and a good-quality calorimetric estimate of the primary energy from the fluorescence detector, E_{FD}. Using this information from many events allows a 'calibration curve' to be produced which parameterises such a relationship as a function of E_{FD}, circumventing much of the reliance upon air shower simulations which introduce large systematic uncertainties in ground array energy estimation; the only need for air shower simulations is the estimation of the missing energy described in Section 1.2.2.2 [7, 36, 50]. Once evaluated,

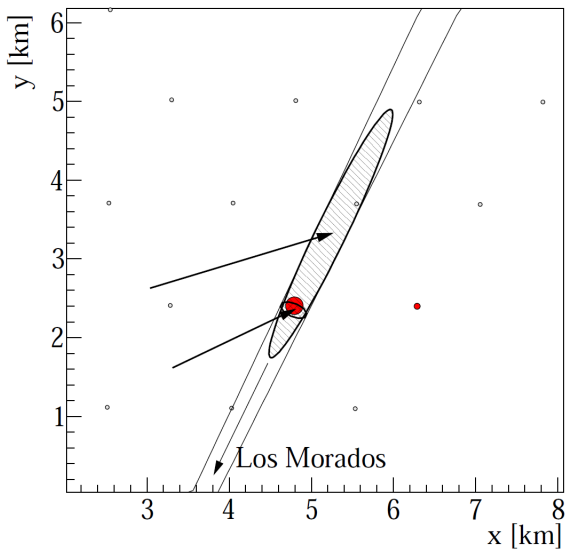


Figure 1.5: The extra precision allowed by hybrid reconstruction is well-illustrated in this image. A single event measured by the Pierre Auger Observatory has been reconstructed using both monocular and hybrid methods, and the resulting uncertainties in the core location are shown by the two ellipses - the larger showing the result of monocular reconstruction, and the smaller that of hybrid. Arrows pointing to the centres of these ellipses show the reconstructed shower axis in each case. The circles represent detectors of the ground array, with open being untriggered and filled being triggered, with the latter's size proportional to the logarithm of the measured signal. Uncertainty in the intersection of the SDP with the ground is illustrated by two solid lines. Finally, an arrow pointing to the bottom of the diagram indicates the direction to Los Morados, the name of the fluorescence detector which observed the event. Image from [48].

this calibration curve allows an internally-consistent energy estimation to be made for any event detected by a hybrid detector, regardless of the method of detection. Importantly, however, the energy scale of the ground array subsequently adopts any systematic uncertainties present in the FD energy scale.

As mentioned in Section 1.2.2.1, composition measurements are difficult with ground arrays. In a similar manner to the calibration of the ground array with calorimetric energy, hybrid detectors introduce the opportunity for deriving a relationship between composition estimators gathered from ground array data and direct measurements of X_{max} by the FD. Again, this avoids the requirement of air shower simulations to interpret the ground array results, and presents the opportunity to take advantage of the large data gathering power afforded by ground arrays [40].

Hybrid data sets, while limited in statistics when compared with those gathered by ground arrays, are well-suited to studies requiring very high quality reconstructions. Some examples of these studies are discussed in Chapter 2.

Chapter 2

Experimental Observations of the Cosmic Ray Flux

Most observations of the cosmic ray flux at Earth can be classified under one of three main areas - its energy spectrum, composition or anisotropy. Information about the sources of cosmic rays, as well as processes they undergo during propagation from those sources to Earth, may be revealed through the study of each of these areas. To produce a consistent model of CR production and propagation, however, it is necessary to consider all three categories together. It is for this reason that cosmic ray observatories will attempt to measure each of these properties at the energy range of interest to gain a comprehensive understanding of CR astrophysics.

The discussion in this chapter will focus on recent measurements of the energy spectrum, composition and anisotropy of ultra-high energy cosmic rays (UHECRs), taken here to mean those at EeV (10^{18} eV) energies and above. More detailed information of past observations and those at lower energies can be found in [7, 13, 17, 51] and references therein. Details of many of the detectors mentioned below may be found in Chapter 3.

2.1 The Energy Spectrum of Cosmic Rays

The cosmic ray energy spectrum describes the number of particles arriving at Earth as a function of energy. It can be expressed in the form:

$$\frac{dN}{dE} \propto E^{-\gamma} \quad \mathrm{m}^{-2}\,\mathrm{s}^{-1}\,\mathrm{sr}^{-1}\,\mathrm{eV}^{-1} \tag{2.1}$$

where E is the particle energy and N the number of particles arriving at Earth. Remarkably, a single value of $\gamma \sim 3$ describes the spectrum quite well across more than 10 decades of energy, from $\sim 10^{10}$ eV to $\sim 10^{20}$ eV, indicating a non-thermal origin. Such a steeply-falling spectrum introduces significant experimental challenges, as only about one particle per square kilometer per year arrives at Earth above an energy of 10^{19} eV [28]. Indirect detection is necessary at

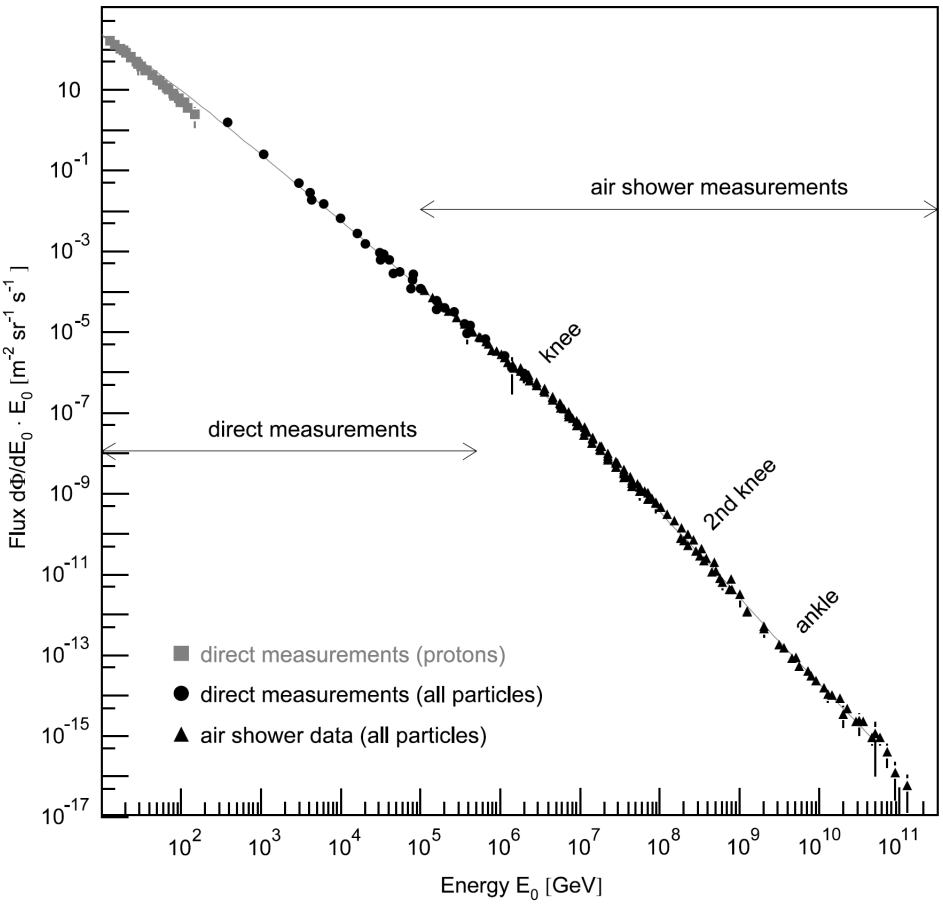


Figure 2.1: The measured energy spectrum of cosmic rays, compiled from a variety of sources. This graph clearly illustrates the absence of a major deviation from a single power law over a large range of energies. However those deviations which are present are of physical significance. Also indicated are the positions of the knee, second knee and ankle. Note that the ordinate has been multiplied by energy. Image from [13].

energies above around 10^{14}-10^{15} eV as described in Chapter 1.

Observed changes in the value of γ are generally considered meaningful and indicative of phenomena such as energy-dependent propagation effects, changing composition of the primary CRs or different CR sources becoming dominant. Three spectral breaks stand out in particular - the 'knee' and the 'ankle' - so named for the spectrum's resemblance to a human leg that these features introduce - and a sharp cut-off at the very highest end of the spectrum. Figure 2.1 shows the all-particle energy spectrum above $\sim 10^{12}$ eV, as well as the proton spectrum below this energy.

The first major feature in the high-energy spectrum appears at an energy of $\sim 10^{15.5}$ eV, and is known as the knee. At this point the spectrum steepens with γ changing from ≈ 2.7 to ≈ 3.1 [13]. This steepening has been found to be a result of the spectra of individual species falling off successively, with the cut-off energies proportional to the charge of each

species (see Section 2.2). One interpretation of this observation is that the knee is the point past which Galactic sources - commonly thought to be sites of diffusive shock acceleration such as supernova remnants [7, 52] - can no longer accelerate protons due to the particles escaping the acceleration region [53]. As particle rigidity is inversely proportional to charge at a given energy (see Chapter 4), nuclei can be accelerated to higher energies than protons before they escape the source region. However, if it is assumed that the entire knee-to-ankle (see below) region is dominated by cosmic rays of Galactic origin, a difficulty is encountered with current Galactic acceleration models being capable of accelerating particles to such high energies [9, 53, 54]. Several alternatives to this model exist, including the explanation that the knee is caused by the inability of the Milky Way to confine higher-rigidity particles within its magnetic field rather than being due to a source acceleration limit, or a combination of several effects [55].

A feature known as the 'second knee' has been observed in the decade between 10^{17} eV and 10^{18} eV, where a further steepening of the spectrum occurs following that of the knee [13, 51]. The exact energy at which the steepening occurs is not well-measured, and the reason for such a feature is not clear; it may be associated with the Galactic acceleration limit of nuclei heavier than iron [56] or, alternatively, a signature of the transition between Galactic and extragalactic CRs [55, 57]. Determination of the composition at these energies will hopefully distinguish between the models - the former predicts a heavy composition around the second knee, while the latter predicts a light one.

The next major feature of the spectrum is the ankle, which is observed at an energy of $\approx 4 \times 10^{18}$ eV [36, 58]. At this point, the spectrum becomes harder with γ taking a value of ≈ 2.7; values of γ of 2.81 ± 0.03, 2.68 ± 0.04 and 2.68 ± 0.01 have been reported from data gathered at HiRes [58], Telescope Array [59] and the Pierre Auger Observatory [36], respectively. The ankle has traditionally been seen as the point at which an extragalactic flux of CRs begins to dominate over the Galactic flux [47]. An alternative theory, assuming an extragalactic flux of protons, predicts this transition to occur at a lower energy in the range 10^{17} eV to 10^{18} eV. This model explains the second knee and ankle as resulting from the energy spectrum flattening on either side of a dip - which has been observed in spectra measured by several detectors - caused by pair production of extragalactic protons interacting with the cosmic microwave background (CMB) [51, 55, 57]. Clearly, identification of the true nature of the ankle will rely on further studies of the preceding decade of energy, which are being performed by experiments such as KASCADE-Grande [60]. A review and discussion of these different models from below the knee to above the ankle can be found in [13].

The dominant feature expected above the ankle is a suppression of the flux above an energy of approximately 6×10^{19} eV. The presence of a suppression above this energy was predicted by Greisen [61], and independently by Zatsepin and Kuz'min [62], following the discovery of the CMB by Penzias and Wilson in 1965 [63]. This so-called GZK (Greisen, Zatsepin & Kuz'min) cut-off is predicted to result from the interactions between protons and photons of the CMB

background at energies close to 10^{20} eV, where the centre-of-mass energy is large enough for pion photoproduction to occur:

$$p + \gamma \rightarrow p + \pi^0 \tag{2.2}$$

$$\rightarrow n + \pi^+ \tag{2.3}$$

These interactions are particularly important in the region of the $\Delta(1232)$ resonance that lies just above the GZK threshold energy [7]. Each time a proton undergoes pion photoproduction it loses $\approx 20\%$ of its energy. The mean free path for this interaction is ≈ 5 Mpc above 3×10^{20} eV, although both the mean free path and inelasticity are energy-dependent. Such interactions serve to limit the horizon within which sources of significant numbers of trans-GZK CRs can lie to the order of 100 Mpc, leading to the expectation of a cut-off in the case of uniformly-distributed sources in the local universe [17, 28, 51].

Nuclei are also expected to interact with photons - CMB as well as infra-red, optical and ultraviolet - as they propagate through intergalactic space, undergoing photodisintegration and pair production. Such an energy loss process occurs at a higher energy threshold than for protons, but the energy loss length is shorter [7]. However, spectra similar to that of protons subject to the GZK effect may be obtained for nuclei following propagation from their source to Earth [17, 64].

Due to the low statistics at the end of the cosmic ray spectrum, it was not until AGASA was operating that enough EAS were observed above the GZK threshold to make a statement about the existence of a cut-off. The AGASA collaboration published this result in 1998 [65]. A single power law was found to adequately describe the energy spectrum above a break energy of $10^{19.01}$ eV, at which point the spectral index took a value of $\gamma = 2.78^{+0.25}_{-0.33}$. Such a result was inconsistent with an expected GZK cut-off from a uniform source distribution, as six events above 10^{20} eV were observed compared to the expectation of less than one from such a model. This result was revisited in 2003 with a new study of the impact of the assumed LDF and energy calibration on the assigned primary energies. This new analysis did not significantly alter the 1998 result, and still the data were not seen to support a GZK suppression [39, 66]. The spectrum published by the AGASA collaboration is shown in Figure 2.2. Note that a later analysis revealed that an overestimation of energies at AGASA had occurred which, when corrected for, reduced the number of events at high energies and rendered the existence of a continuation of the spectrum uncertain [51, 67].

In contrast to the AGASA result, in 2005 the HiRes Collaboration, using monocular fluorescence detector observations, published an observation of a spectral break occurring at an energy of $\log_{10}(E/\text{eV}) = 19.79 \pm 0.09$. Above this energy, the spectral index was found to be $\gamma = 5.2 \pm 1.3$ [68]. A description of the data with a single power law above the ankle was tested and found to be incompatible, indicating the presence of a suppression above this break energy. In 2008, another analysis of HiRes data confirmed the existence of a spectral break

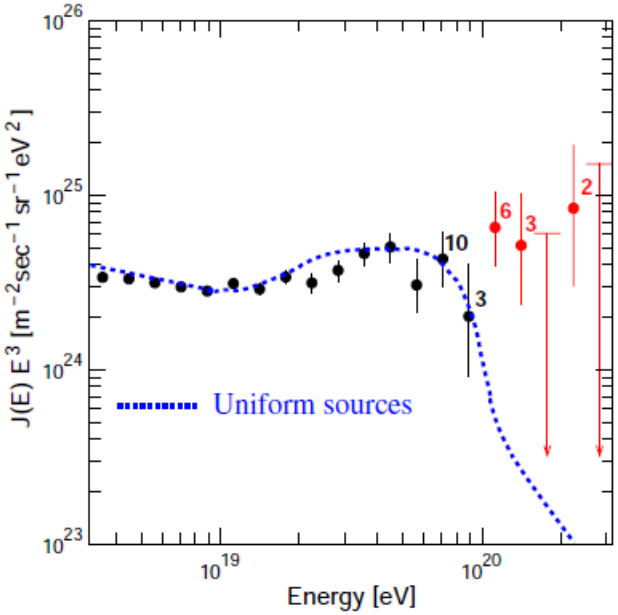


Figure 2.2: The upper end of the CR energy spectrum as measured by AGASA. The dotted line indicates the spectrum expected for the GZK suppression of CRs from uniformly-distributed sources. The data points indicate the measurements of AGASA with Poisson bounds at a 68% confidence level, and upper limits given at a 90% confidence level. The numbers indicate how many events are in each of the highest-energy bins. Image from [39].

occurring at $\log_{10}(E/\text{eV}) = 19.75 \pm 0.04$ after which the spectral index was $\gamma = 5.1 \pm 0.7$ - both values compatible with previously-reported values. Comparing these results with theoretical modelling, the HiRes Collaboration interpreted their spectrum as supporting the existence of a GZK cut-off with a significance of 5 standard deviations [69]. A spectrum was also published in 2009 using stereo data which, while a smaller data set, possessed superior energy resolution. It was found to be consistent with the monocular data [45, 51, 58].

In 2008, the Pierre Auger Collaboration also reported an observation of a flux suppression at high energies. Using ground array data with an energy scale calibrated with fluorescence detectors, the hypothesis of a single power law describing the spectrum above an energy of 4×10^{18} eV was rejected with a significance higher than 6 standard deviations [70]. Following a break energy of 4×10^{19} eV, the spectral index was found to change to $\gamma = 4.2 \pm 0.4 \pm 0.06$, where the quoted errors are statistical and systematic, respectively. This break energy is compatible with that reported by the HiRes Collaboration within the systematic uncertainties of 22% and 17% in the overall energy scales of the Pierre Auger Observatory and HiRes data, respectively.

This measurement was updated in 2010 with larger statistics and a data set combining both ground array and fluorescence detector data. A power law above the spectral break at $\log_{10}(E/\text{eV}) = 19.46 \pm 0.03$ was fitted with an index of 4.3 ± 0.2, agreeing with the previous

Figure 2.3: The upper end of the CR energy spectrum as reported by the Pierre Auger Collaboration in 2011. Note that the ordinate has been multiplied by E^3 to more easily display the ankle and high-energy suppression. Statistical uncertainties are indicated on the data points, while the entire spectrum is subject to a systematic uncertainty of 22%. Also plotted are two sets of fits to the data (see text for details). Image from [36].

measurement within uncertainties. The flux suppression in the larger data set was assigned a significance of greater than 20σ [47].

More recent measurements from the Telescope Array and the Pierre Auger Observatory have continued to support the presence of a spectral cut-off above the GZK threshold [36, 59]. Figure 2.3 shows the energy spectrum measured by the Pierre Auger Observatory as reported in 2011. Also shown are two sets of fits to the data points: the first is constructed from three power laws, and the second from a power law plus a smooth function above the ankle given by:

$$J(E; E > E_{ankle}) \propto E^{-\gamma_2} \frac{1}{1 + \exp\left(\frac{\log_{10} E - \log_{10} E_{1/2}}{\log_{10} W_c}\right)} \tag{2.4}$$

where values of $\log_{10}(E_{ankle}/\text{eV}) = 18.61 \pm 0.01$, $\log_{10}(E_{1/2}/\text{eV}) = 19.63 \pm 0.02$, $\log_{10}(W_c/\text{eV}) = 0.15 \pm 0.02$ and $\gamma_2 = 2.63 \pm 0.02$ are found to best describe the measured spectrum.

An important point to note is that while the observed cut-off is compatible with expectations from GZK interactions, it may also be indicative of other effects such as a change in the source spectrum [47].

2.2 Composition of Cosmic Rays

Composition measurements of the cosmic ray flux aim to determine the type and proportion of different species present as a function of energy. Such knowledge is critical for the interpretation of energy spectrum and anisotropy measurements. In contrast to lower energies where detectors are capable of determining composition directly, the measurement of composition-sensitive parameters of an EAS must be compared with predictions from hadronic interaction models to infer the nature of the primary particle. Thus, it is important to recognise that the interpretation of composition results is somewhat ambiguous due to the reliance upon these models at energies higher than can currently be verified in particle accelerator experiments. A recent review of such issues may be found in [71]. In addition, efforts to improve knowledge of hadronic interactions using air shower measurements are ongoing (see, for example, [19]).

A general expectation for the composition as a function of energy can be formed from simple arguments requiring the magnetic confinement of a cosmic ray being accelerated in a source region, assuming a 'bottom-up' model for CR production (see Section 2.3.1 for discussion). It may then be expected that the maximum attainable energy of a particle of charge Z is approximately Z times the maximum energy of a proton. This leads to a qualitative expectation of the average mass of the cosmic ray flux becoming heavier with increasing energy as the maximum energy of each species is sequentially passed, until an extragalactic component of the flux begins to dominate and a transition back towards lighter masses is observed (assuming that a significant fraction of the extragalactic flux consists of protons) [13, 53]. As previously mentioned, models such as that of the pair production dip predict such a transition to occur in the 10^{17} eV to 10^{18} eV decade [55, 57].

At energies where an extragalactic flux may be dominant it is expected that, unless sources are close by, the primary flux will be composed of some mix of mostly protons and iron. Intermediate-mass particles are expected to disintegrate rather quickly during propagation and do not arrive at Earth in large quantities [13]. Modelling of cosmic ray propagation shows that a variety of source compositions are capable of reproducing the observed energy spectrum at the highest energies [64].

In the knee-to-ankle region, composition measurements show that the abundance of light elements decreases with increasing energy. Measurements of spectra for individual species have shown that the cut-off energies are proportional to the particle charge, consistent with a rigidity dependence of either source acceleration limits or confinement of CRs within the Galactic magnetic field (GMF) [52, 72].

In 1993 it was reported that observations of the X_{max} distributions of events observed by the Fly's Eye detector appeared to indicate a transition from a heavy composition below the ankle, at $10^{17.5}$ eV, to a lighter composition at energies up to 10^{19} eV [73]. This was interpreted as possible evidence, along with Fly's Eye observations of anisotropy and energy spectrum, for a transition from Galactic to extragalactic dominance of the cosmic ray flux at the ankle.

Analysis of shower muon densities at AGASA returned similar results to the Fly's Eye

Figure 2.4: Measurements of $\langle X_{max} \rangle$ from data collected by the HiRes detector. Shown for comparison are expectations from several hadronic interaction models for pure proton and iron compositions. The number of events in each bin is indicated. From [76].

composition measurement [39, 74]. Between $10^{17.5}$ eV and 10^{19} eV the average primary mass was found to decrease with a dominant light component at the upper end of this energy range. An upper limit of the iron fraction above 10^{19} eV of 40% was found at a confidence level of 90%. It was recognized that a strong dependence on the hadronic interaction model used was present for this study.

Measurements of $\langle X_{max} \rangle$ by the HiRes detector were published in 2005 for energies between 10^{18} eV and $10^{19.4}$ eV. In this energy range an elongation rate of $54.5 \pm 6.5 \pm 4.5\,\mathrm{g\,cm^{-2}}$/decade was measured, with the quoted errors arising from statistical and systematic sources, respectively. Such an elongation rate is consistent with a constant or slowly-changing composition, and the values of $\langle X_{max} \rangle$ appear consistent with a light composition. Together with an earlier result from the HiRes-MIA detector, which observed an elongation rate of approximately $90\,\mathrm{g\,cm^{-2}}$/decade between 10^{17} eV and 10^{18} eV, this was interpreted as possible evidence for a heavy end to the Galactic spectrum giving way to a lighter extragalactic flux [16, 75]. It was determined that these results were, within errors, in good agreement with those of the Fly's Eye [51]. An updated measurement of the elongation rate in 2010 was found to be consistent with a constant elongation rate above 1.6 EeV of $47.9 \pm 6.0\,(\mathrm{stat}) \pm 3.2\,(\mathrm{syst})\,\mathrm{g\,cm^{-2}}$/decade and, by comparison with air shower simulations, close to a pure proton composition [76]. These measurements are shown in Figure 2.4. Preliminary results from the Telescope Array are also consistent with a light composition [77].

The Pierre Auger Collaboration published its most recent measurements of $\langle X_{max} \rangle$ and RMS(X_{max}) - the root mean square of the measured X_{max} distribution within an energy bin, which is a measure of the degree of shower-to-shower fluctuations of the depth of maximum - at energies above 10^{18} eV in 2011, using data gathered between December 2004 and September

NOTE:
This figure is included on page 30 of the print copy of the thesis held in the University of Adelaide Library.

Figure 2.5: The values of $\langle X_{max} \rangle$ and RMS(X_{max}) measured by the Pierre Auger Observatory as reported in 2011, with the number of events in each bin indicated. Plotted for comparison are the values of these parameters predicted by several hadronic interaction models in the case of a pure proton and pure iron composition (red and blue lines, respectively). Also indicated on the plot of $\langle X_{max} \rangle$ are the two linear fits described in the text. Systematic errors in both cases are indicated by the shaded regions. From [19].

2010 via fluorescence detection [19]. These results are updated from those presented in [20]. A fit to the data of two different slopes was found to be required rather than a single linear one. Below $\log_{10}(E/\text{eV}) = 18.38^{+0.07}_{-0.17}$, an elongation rate of $\Lambda = 82^{+48}_{-8}\,\text{g}\,\text{cm}^{-2}$/decade was found, and above that energy $\Lambda = 27^{+3}_{-8}\,\text{g}\,\text{cm}^{-2}$/decade. The RMS values of the X_{max} distributions were found to decrease with increasing energy, changing from approximately $55\,\text{g}\,\text{cm}^{-2}$ to $26\,\text{g}\,\text{cm}^{-2}$ between energies of about $10^{18}\,\text{eV}$ and $10^{19.5}\,\text{eV}$. The highest energy event in the data set was $7.2 \times 10^{19}\,\text{eV}$. These results are displayed in Figure 2.5.

Both the $\langle X_{max} \rangle$ and RMS(X_{max}) measurements appear to indicate a light composition changing to an increasingly heavier one in the energy range studied, when compared to predictions of current hadronic interaction models. These results appear to differ from those reported by HiRes, the reason for which is not currently clear. However, both measurements seem to indicate a light composition at energies below the ankle, consistent with expectations from the pair production dip model.

2.3 Cosmic Ray Sources and Anisotropies

Determining the origin of cosmic rays is a major challenge for CR detectors. The detection of a departure from a uniform distribution of CR arrival directions - known as an anisotropy - is important for its ability to reveal information about such sources. Anisotropies are searched for in the data of all CR observatories and may, when considered together with energy spectrum and composition results, form the last major component required to develop a consistent picture of cosmic ray origins. However, the interpretation of anisotropy results is somewhat

complicated due to the charged nature of cosmic rays, which introduces deflections in their trajectories as they travel from their source to Earth through magnetic fields that are not currently well-understood (see Chapter 4).

Many different methods may be used to search for anisotropies in CR data sets and will be dictated by such considerations as physical expectations from models of CR production and propagation, the number of events available or the detector resolution. Some example methods include autocorrelation analysis, the simplest of which is a two-point version but it may be extended to higher dimensions [78], harmonic analysis or the correlation between the CR data and another data set such as a possible source distribution catalogue.

In any study using statistical methods to ascertain the presence of a signal it is important to recognise the impact of performing many statistical trials on a data set. In searching for an analysis that appears to indicate the presence of a significant pattern, many parameters may be varied such as an energy threshold, an angular cut for correlation studies, or even the catalogue against which the CR data are compared. The result of this can be to perform a very large number of trials on the data that increase the possibility of a positive detection for one or more of those trials due to a statistical fluctuation rather than the presence of a true signal. If such a result is obtained it may be very difficult, or even impossible, to estimate its true statistical significance. Consequently, the results of such an *a posteriori* analysis can be difficult to interpret - in the cases where a large number of trials have been performed to extract a result, an identical *a priori* analysis should be performed on an independent data set to confirm the presence of a signal and its statistical significance [28, 79].

Prior to discussion of particular anisotropy results, it is appropriate to briefly discuss possible sources of UHECR.

2.3.1 Cosmic Ray Sources

Many theories have been suggested to explain the origin of cosmic rays, particularly those of ultra high energies, due to the difficulty of modelling the acceleration of such energetic particles. Broadly, these models may be classed as one of two types - 'bottom-up' or 'top-down' models. The former refers to the acceleration of a population of charged particles from low to high energies, while the latter models predict UHECR to originate from exotic particles.

Such top-down models presume UHECRs to be the decay products of extremely massive particles that may be remnants of the early stages of the universe. Under many models, the decay of these so-called X particles is predicted to produce a large fraction of photons at high energies. An alternative, called the Z-burst model, also predicts a large number of photon primaries at high energies from the interactions of ultra high energy neutrinos with relic neutrinos. However, recent results such as those from the Pierre Auger Observatory place strong limits on the photon fraction across a large range of energies, disfavouring such models [80, 81, 82].

Bottom-up models describe the production of UHECRs as the acceleration of a population

Figure 2.6: The 'Hillas plot', displaying the dimensions and magnetic fields of various astrophysical objects compared to those required for the containment of a 10^{20} eV proton and iron nucleus, respectively. Objects in the shaded region are considered possible source candidates for these particles. From [83].

of particles within a source region until they reach such an energy that they can escape. Galactic CRs are commonly considered to be accelerated through the process of diffusive shock acceleration [54] in sites such as supernova remnants. This model requires the confinement of the cosmic rays to their acceleration region during the acceleration process, with the maximum energy attainable (E_{max}) by a particle of charge Z being given by:

$$E_{max} \approx 10^{18}\,\mathrm{eV}\, Z\beta_s \left(\frac{R}{\mathrm{kpc}}\right)\left(\frac{B}{\mu\mathrm{G}}\right) \tag{2.5}$$

for an acceleration region of size R, magnetic field strength B and shock velocity β_S given in units of c [13]. It should be noted that this is a necessary, but not sufficient, requirement for the stochastic acceleration of CRs within a source. For example, diffusive shock acceleration is a rather slow process, requiring that sources are capable of the acceleration of particles to very high energies for extended periods of time [83].

In the context of the acceleration of UHECRs of extragalactic origin, the containment requirement of Equation 2.5 led Hillas to produce the diagram reproduced in Figure 2.6, which displays the dimensions and magnetic field strengths of various astrophysical objects compared to those values satisfying Equation 2.5 for 10^{20} eV particles.

From this argument it is apparent that few known astrophysical objects may be capable of producing particles at the highest observed energies. Active galactic nuclei (AGN), radio galaxies and galactic clusters are commonly considered potential sources. However, the mechanism through which acceleration occurs is not well-understood - for example, it is currently

unclear how CRs may be accelerated to the highest observed energies when energy losses at those energies, such as those due to interactions with ambient photon fields in the source, are expected to be significant. Further discussion of such issues may be found in [7, 84].

2.3.2 Recent Anisotropy Results at EeV Energies

Many anisotropy results have been reported in recent years in a variety of different types of studies. Discussed here are a selection of those results at energies $\approx 10^{18}$ eV and above to illustrate the range of analyses that have been performed and highlight several important results. A recent review of anisotropy results at lower energies may be found in [13], and analyses at EeV energies further to those discussed here may be found in [13], [17] and [28].

The Galactic Centre (GC) is an attractive target for CR anisotropy studies owing to the presence of a supermassive black hole, which may be capable of accelerating cosmic rays to high energies, and gamma ray observations that may indicate the presence of high energy cosmic rays interacting with ambient matter (for example [85]). In addition, the decay length of neutrons becomes large enough for them to reach Earth from the GC at EeV energies, which could produce a point-like source [28]. The AGASA collaboration reported an excess of arrival directions from near the GC in 1999 [86]. Applying the method of harmonic analysis to events of energies between $10^{17.9}$ eV and $10^{18.3}$ eV, a first harmonic of amplitude $\approx 4\%$ was fitted to the data, with the peak being attributed to excess numbers of events near the Galactic Centre and Cygnus regions - note that the GC was located slightly south of the field of view of AGASA. Following this harmonic analysis, smoothing was applied to compare the expected versus observed numbers of events using a circle of radius 20°. A significance of 4σ was assigned to the observed excess compared to expectations from isotropy, although a value of 20° for the smoothing radius was chosen from several possibilities to maximise this significance.

A separate analysis was performed in 2001 using data from the SUGAR array, which contained the GC well within its field of view [87]. To avoid statistical penalties from applying a large number of trials, parameters for the analysis were chosen *a priori*, based on those used to obtain the AGASA result. At energies between $10^{17.9}$ eV and $10^{18.5}$ eV - the energy range was enlarged to account for the poorer energy resolution of SUGAR - an excess of events compared to expectations from isotropy was observed approximately 6° from the peak of that of the AGASA data. This excess was 7.5° from the GC and consistent with a point source, with a probability of the signal occurring by chance of 0.005.

The Pierre Auger Observatory performed a subsequent analysis in 2007 to look for the existence of a GC excess using a larger data set than either AGASA or SUGAR [88]. Repeating the analyses reported in [86, 87], no significant excesses were found to be present even if possible energy shifts between experiments were taken into account. In addition, evidence of a point-source that could be a result of a neutron flux was searched for without a detection.

A large-scale anisotropy study was performed on Pierre Auger Observatory data at energies

above 2.5×10^{17} eV in 2011 [89]. The amplitudes and phases of the first harmonics fitted to data in six energy bins were calculated to search for evidence of dipole-like anisotropies. Several scenarios may produce such an anisotropy; CRs of Galactic origin attaining an energy with which they are able to escape the Galactic magnetic field may produce a dipole-like effect, or it may result from the Compton-Getting effect [90] due to the motion of the Earth with respect to the reference frame in which extragalactic cosmic rays are isotropic. With statistics allowing sensitivity to amplitudes as low as 1%, and careful management of possible exposure-related effects, no statistically-significant amplitudes were found, leading to the placement of upper limits on dipole anisotropies as a function of energy. In particular, these limits were found to be inconsistent with the 4% amplitude variation observed by AGASA. Despite this, the phase of the first harmonic appeared to be correlated between adjacent, independent energy bins, giving a smoothly-changing phase as a function of energy - a result that could be suggestive of a real anisotropy at a level below current detector sensitivity. However, a larger data set is required to draw definite conclusions.

While finding no evidence of a large-scale anisotropy above 10^{19} eV, the AGASA collaboration reported the presence of 'clusters' of events in their data, in which multiple events arrived from similar directions on the sky above 4×10^{19} eV [91]. Three groups of two events and one group of three events were observed to arrive within windows of 2.5° radius, which was calculated to have a less than 1% chance of occurring in an isotropic data set. It should be noted, however, that such a result was obtained by *a posteriori* analysis, and may well be much less significant than originally thought [28]. Two later analyses with different data sets - an autocorrelation analysis using HiRes data [92] and a likelihood analysis searching for point sources using HiRes data combined with the AGASA data [93] - did not show any significant results that would indicate the presence of point sources.

Correlations between UHECR data sets and catalogues of possible sources have been performed several times. For example, following the claim of correlations between data from AGASA and Yakutsk and the directions of BL Lacertae (BL Lac) objects (e.g. [94]) - found through *a posteriori* analysis to maximise the correlation signal - a maximum likelihood analysis was used to search for a similar effect in HiRes data [95]. No significant correlation was found for the subset of BL Lac objects which produced the maximum correlation for the AGASA and Yakutsk data, although a possible correlation with the HiRes data was found for a different subset of BL Lac objects. No statistical significance was claimed as it was noted that the analysis using an independent data set would be required to do so. Application of this analysis to Pierre Auger Observatory data did not indicate the presence of such a correlation [96], although the difference in known BL Lac populations between the northern and southern skies may complicate the interpretation of this result [28].

However, in 2007 the Pierre Auger Collaboration reported the detection of a correlation between UHECR arrival directions and the 12^{th} edition of the Véron-Cetty and Véron (VCV) catalogue of active galactic nuclei (AGN) [97, 98, 99]. This result was arrived at by using an

initial data set, collected between January 2004 and May 2006, to perform a scan of parameter space to find the set of parameters which maximised the correlation signal. The scan parameters were the minimum energy threshold E_{th}, the maximum redshift of galaxies in the VCV catalogue z_{max} and the maximum angular separation between the UHECR arrival directions and the AGN directions ψ, which were scanned over to minimise the cumulative binomial probability of the observed correlation occurring from an isotropic distribution. This scan of the exploratory data set yielded values that minimised the chance probability of $z_{max} = 0.018$ (corresponding to a distance of 75 Mpc), $E_{th} = 56$ EeV and $\psi = 3.1°$, giving the probability of a single event from an isotropic distribution correlating of $p = 0.21$. Following this analysis, these parameters were used as an *a priori* test to apply to an independent data set as a running prescription. This prescription was applied until the hypothesis of an isotropic arrival direction distribution could be rejected at the 99% level, which occurred after 6 out of 8 detected events correlated with the VCV catalogue.

Following the passing of the prescription, which established the presence of a correlation, a further parameter scan was performed on the two data sets together to maximise the correlation signal of the entire set. This yielded parameters of $z_{max} = 0.017$, $\psi = 3.2°$ and $E_{th} = 57$ EeV, with 20 of 27 events correlating with AGN from the VCV catalogue. It is recognised that the interpretation of this result is not entirely clear - for example, it does not necessarily imply that AGN are the source of UHECRs, as the VCV catalogue traces the distribution of the local large-scale matter distribution - but it lends support to the presence of a GZK suppression due to E_{th} being above the GZK threshold, and z_{max} limiting the catalogue to AGN within the GZK horizon. The events involved in this analysis are shown in Figure 2.7.

In 2010 this result was updated with a larger data set [100]. The degree of correlation was found to have decreased, although the most likely value of $0.38^{+.07}_{-.06}$ was still greater than the 0.21 expected for isotropy. However, as an *a priori* prescription was not formed before the analysis of this larger data set a confidence level for determining the data to be anisotropic could not be defined. The chance of observing such a correlation in an isotropic data set is 3×10^{-3}.

In light of the composition measurements of the Pierre Auger Observatory, in which it appears that the CR flux tends toward an intermediate to heavy mass at high energies (although such results do not currently extend to trans-GZK energies) a correlation on such small scales is somewhat puzzling. Further data, both composition and anisotropy, will be required to resolve such an apparent discrepancy.

In addition, the results of several further analyses of Pierre Auger Observatory data were reported in [100], although these were made *a posteriori* with statistical penalty factors that cannot be easily quantified. The data were found to be consistent with models of source distributions constructed from the Swift-BAT and 2MRS galaxy catalogues, although the parameters giving such agreement were not strongly constrained. An autocorrelation analysis was also performed which did not indicate the presence of event clustering greater than expec-

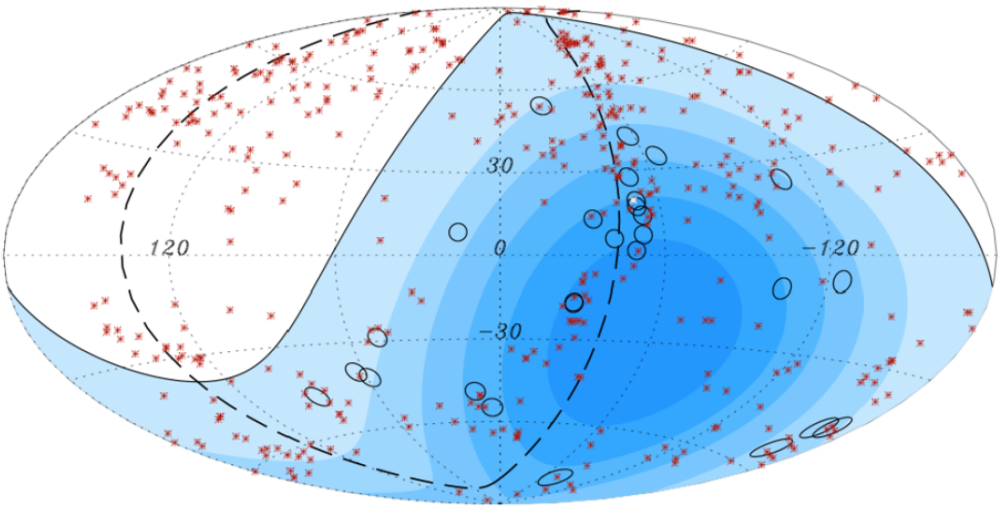


Figure 2.7: This skymap, in Galactic co-ordinates, displays the results of the original AGN correlation study performed with data from the Pierre Auger Observatory. Red crosses indicate the positions of the AGN in the VCV catalogue within a redshift of 0.017; black circles of radius 3.2° are drawn around the arrival directions of the 27 CRs having energies greater than 57 EeV. Shades of blue indicate the relative exposure of the detector, while a dashed line shows the position of the supergalactic plane. The position of Centaurus A is marked by a white star. From [99].

tations from isotropy at any angle, although it was noted that the presence of a large number of events (18.8% of the data set, compared to 4.7% expected from isotropy) within 18° of the nearby radio galaxy Centaurus A was possibly of interest. Such an excess had previously been pointed out, for example in [101]. No obvious conclusions were able to be drawn from these studies and more data are required to allow a better interpretation of these results.

Prompted by the Pierre Auger Collaboration's original publication of the correlation between UHECR arrival directions and local AGN, data from the HiRes detector were analysed for evidence of a corresponding anisotropy [102]. Using the parameters found to maximise the correlation signal of the Pierre Auger Observatory's exploratory data set, the HiRes data found no evidence for a statistically-significant correlation; 2 events out of a total of 13 were found to correlate with AGN of the VCV catalogue. A search to maximise the signal of the HiRes data was also performed, the results of which were again consistent with isotropic arrival directions. However, it has been noted that the same degree of correlation is not necessarily expected in both hemispheres - the distribution of AGN is not uniform, and the VCV catalogue exhibits different amounts of completeness between the two hemispheres [100]. The results of an autocorrelation analysis of HiRes data above 56 EeV were also reported in [102], in which no evidence of significant small-scale clustering was found.

Finally, several studies have been performed to look for evidence of energy ordering in arrival direction data. Such ordering may be expected for events of equal charge passing through a magnetic field in which the induced deflections are not too large and the number of

bright sources small, resulting in groups of events of common origin with angular deflections from their source direction inversely proportional to their energy (see Chapter 6 for more detailed discussion of energy ordering, as well as the papers introduced below).

The HiRes Collaboration searched for arrival directions forming arcs in HiRes data at energies above 10^{19} eV, but found no significant evidence for the presence of such structures [103]. Data from the Pierre Auger Observatory at energies above 20 EeV were analysed for similar features but, again, the results were consistent with expectations from isotropic arrival directions [104].

A different method to search for coherent deflections of CRs through cosmic magnetic fields, termed the 'energy-energy correlation' method, was also applied to the Pierre Auger Observatory data for energies above 5 EeV [82]. This analysis also yielded results consistent with expectations from isotropic arrival directions.

The last analysis of energy ordering in UHECR arrival directions to be discussed here was reported in [105]. The arrival directions of events detected by the Pierre Auger Observatory above 55 EeV, which were published in [100], were analysed using a method adapted from [106]. While possible evidence of energy ordering of the arrival directions of a subset of the data was found, which was consistent with a source direction close to the Virgo cluster of galaxies, this analysis was *a posteriori* and as such a statistical significance to the result cannot be unambiguously determined.

2.4 Interpretation of Energy Spectrum, Composition and Anisotropy Results

Upon consideration of the results discussed here, an overall picture of the cosmic ray flux may be formed. Composition and energy spectrum measurements at energies around the knee are consistent with a flux of cosmic rays of Galactic origin falling off as the particles are either no longer contained by the Galactic magnetic field or their sources reach the maximum acceleration energies for different rigidities. At an energy before or around the ankle, an extragalactic flux may become dominant where the steep spectrum of heavy Galactic particles gives way to a flatter spectrum of possibly lighter composition.

Above the ankle a suppression is well established that is consistent with the GZK effect attenuating incoming protons or nuclei, although this suppression could also be produced by a change in the source spectra. Composition measurements from different detectors do not exhibit as much agreement at such energies, and the resolution of this discrepancy will be an important step in the understanding of UHECR origins. Finally, at such energies a number of anisotropy results have been reported, although it is important to recognise the *a posteriori* nature of many of them in which search parameters were optimised to maximise the significance of the signal. Without knowledge of the number of statistical trials undertaken to produce such results it is difficult to verify their statistical significance, and in many cases

such observations have not been reproduced in the analysis of data from other detectors [28]. Perhaps the most promising result is the identification of a correlation between UHECR arrival directions and AGN within 75 Mpc, a result which was verified by the *a priori* analysis of data collected by the Pierre Auger Observatory. While the HiRes data did not display the degree of correlation observed in the initial Pierre Auger Observatory data set, the updated estimate of a correlating fraction of ≈ 0.38 may be consistent with the HiRes result [100]. The presence of such a correlation may indicate that the sources of UHECR are local (within ~ 100 Mpc) and follow the distribution of large scale structure.

With the current operation of detectors such as the Pierre Auger Observatory and the Telescope Array, it is hoped that many more clues as to the origins of UHECR will soon be found.

Chapter 3

Ultra High Energy Cosmic Ray Detectors

3.1 Previous UHECR Detectors

Since Hess' discovery of cosmic rays, many efforts have been made to construct detectors with the aim of measuring the properties of the CR flux at Earth. However, it was not until the construction of the first giant air shower array, the Volcano Ranch detector, that the ability to study CRs at the highest energies was realised. Since that time many more detectors have operated around the world to detect extensive air showers. Aided by technological advances and the development of new detection techniques, cosmic ray detectors have grown larger and become more accurate with each generation.

This section will discuss the history of these giant air shower arrays - detectors which covered areas of several square kilometres or more - and fluorescence detectors. A brief summary of detector types and configurations, as well as their major results, will serve to highlight some of the major contributions to the field and also illuminate some of the motivations for the construction of the largest CR detector to date, the Pierre Auger Observatory. Further discussion of some of the observations mentioned here may be found in Chapter 2.

3.1.1 Volcano Ranch

The cosmic ray detector at Volcano Ranch, near Albuquerque, New Mexico (35.2° N) [107], was the first of the giant air shower arrays [7, 17]. Volcano Ranch was built by the group led by Bruno Rossi from the Massachusetts Institute of Technology to investigate cosmic rays above 10^{17} eV and operated between 1959 and 1963 [17].

Nineteen detectors located on a triangular grid made up the array, with an additional 20^{th} detector being used in several different positions throughout Volcano Ranch's lifetime. For some of the operating period, this 20^{th} detector was shielded by 10 cm of lead to investigate the penetrating (muonic) component of EAS [108]. Each detector was a 3.3 m^2 plastic scintillator

Figure 3.1: The original layout of the Volcano Ranch array, with a detector spacing of 442 m. Unlabelled open circles indicate the positions of the nineteen detectors which made up the standard array. Positions labelled 1, 2 and 3 indicate the various positions taken by the 20^{th} detector while the array was in this configuration. Taken from [108].

[1, 107] viewed by a 5 in photomultiplier tube [17]. The 19 detector component of the array operated in two configurations, the first of which used a detector spacing of 442 m and enclosed an area of 2 km^2 (see Figure 3.1). This was later changed to a spacing of 884 m with which the array enclosed approximately 8 km^2 [17, 108].

Operation of Volcano Ranch yielded several important results pertaining to the CR flux, examples of which include the earliest measurements at the highest energies of the arrival direction distribution and energy spectrum, the latter of which included possible evidence for the ankle [109]. In addition, much work was done towards the description of the lateral distribution of EAS at high energies [7]. With around 1 000 showers detected above 10^{18} eV, Volcano Ranch contributed important early knowledge leading to our current understanding of cosmic rays [1].

Perhaps the most startling discovery, however, was the detection of an event with an energy estimated to be 1.0×10^{20} eV [110]. At the time, this was the most energetic particle ever observed [1]. The energy was later revised to 1.4×10^{20} eV and remains one of the highest energy cosmic rays ever detected [7, 17]. The detection of an event of such a high energy is particularly surprising in light of the relatively small exposure of Volcano Ranch and extremely low flux of CRs above 10^{20} eV.

3.1.2 Haverah Park

The 12 km^2 ground array located at Haverah Park, England (54° N), was the result of a collaboration between the Universities of Durham, Leeds, Nottingham and Imperial College and

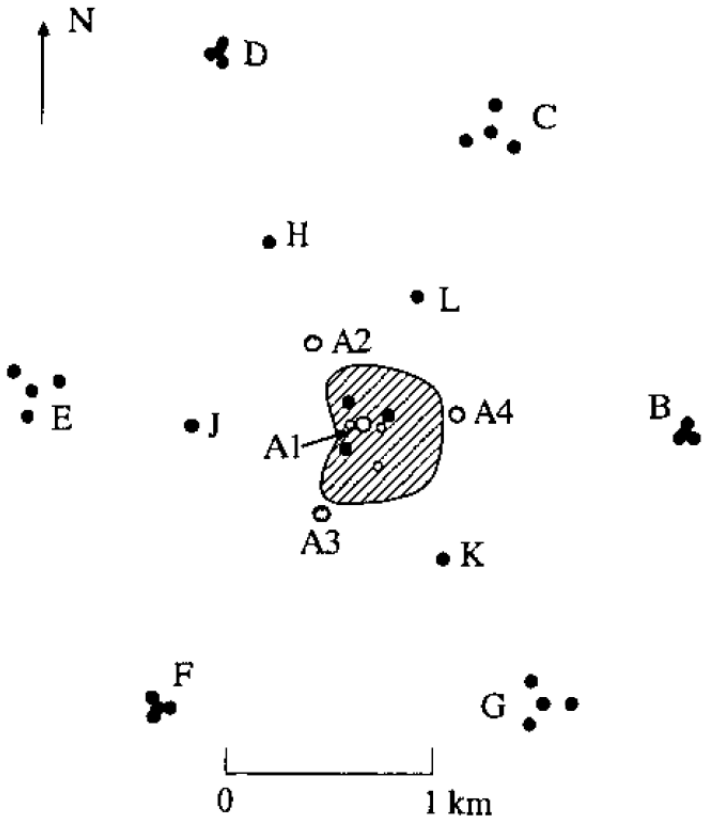


Figure 3.2: The layout of the Haverah Park array. Circles indicate detector positions, with large open circles indicating 34 m^2 detectors, and filled circles the smaller detectors. The shaded area outlines the array of 1 m^2 tanks, and three small open circles denote muon detectors which were also present. See text for details of the individual detectors. From [35].

took data between 1968 and 1987 [7, 24, 35, 111, 112]. In a departure from contemporaneous experiments, the detectors of Haverah Park consisted of water-Čerenkov tanks - a design which proved successful and would later be incorporated into the Pierre Auger Observatory. The somewhat irregular design of the array (see Figure 3.2) was a consequence of limited land access, which dictated the placement of detectors [17].

The water-Čerenkov tanks which made up each detector were galvanized steel containers of 2.29 m^2 in area, containing a pool of water 1.2 m deep which was lined with diffusing plastic. A single 5 in PMT viewed the water from above. An individual detector contained one or more of these tanks; for larger detectors, they were housed in a hut with a thin (4 $g\,cm^{-2}$) roof. The detectors labelled A1 to A4 in Figure 3.2 each totalled 34 m^2 in area (15 tanks); three 9 m^2 detectors were deployed at three locations centred on A1; B to G were each a small sub-array of 4 huts of 13.5 m^2 (6 tanks); H was a single hut of 13.5 m^2; finally, J, K and L were single tanks [17, 35, 111].

Between 1976 and 1981, an extra 30 water-Čerenkov tanks of $1\,\mathrm{m}^2$ area and 1.2 m depth were arranged in a lattice of approximately 150 m spacing at the centre of the array (indicated by the shaded area in Figure 3.2). These were used to measure the lateral distribution of shower particles in finer detail than allowed by the existing array [17, 35, 113]. The detection of an event of energy $10^{20}\,\mathrm{eV}$ which triggered 50 detectors emphasised the quality of data which could be measured with this array.

Data storage occurred at a single location for the detectors in the centre of the array, while the outlying sub-arrays recorded data independently for later collection. Triggering of the array was performed by the central detectors - if A1, plus any of A2, A3 or A4 detected a coincidence within a $4\,\mu$s window of greater than 10 VEM, a signal was sent by microwave link to initiate data recording at the other stations. While this limited the aperture of the array, it allowed highly accurate measurements of the exposure to be made due to careful monitoring of the triggering detectors [35].

Near the end of Haverah Park's operation, eight scintillators were added near the centre of the array. These were employed to cross-calibrate results with those from other experiments such as Volcano Ranch and Yakutsk in order to permit easier comparison of results between detectors [17, 114, 115].

The demonstration of the viability of water-Čerenkov detectors was an important one, as the Pierre Auger Observatory would later use a similar detector design for its ground array. In addition, the ability to study in fine detail particle densities of showers triggering the array - particularly close to the shower core - enabled a good description of lateral distribution functions of showers to be made [33, 35].

3.1.3 SUGAR

SUGAR (the Sydney University Giant Airshower Recorder) took data between 1968 and 1979 [24, 87, 116]. Located in the Pilliga State Forest, New South Wales, Australia (30.5° S) at an altitude of 250 m, SUGAR was the only giant air shower array to have operated in the Southern Hemisphere prior to the Pierre Auger Observatory [17, 24]. A detector of advanced design, SUGAR incorporated many of the elements which would appear in future generations of UHECR detectors. However, unfortunate limitations of the technology used in the detector electronics led to poor estimates of shower energies, limiting the usefulness of data collected by the array [6, 7].

At its peak of operation, SUGAR consisted of 47 detectors enclosing an area of approximately $70\,\mathrm{km}^2$ [1, 117]. Most of these were located on a square grid of 1 600 m spacing but some detectors were located on smaller grids of 800 m and 400 m spacings to gather data on smaller showers [117, 118]. Each detector consisted of two liquid-scintillators separated by 50 m in a North-South direction and buried approximately 1.5 m underground, giving a muon threshold energy of $0.75\sec\theta\,\mathrm{GeV}$ (for the zenith angle θ) [117, 118]. The scintillators had an effective collecting area of $6\,\mathrm{m}^2$ and were designed such that muons at a given angle of

incidence would produce the same flux of light at the photomultiplier tube, regardless of the entry point into the scintillator.

The large size of SUGAR coupled with the terrain of the Pilliga State Forest necessitated the development of a new detector design. Previous ground arrays had linked the detectors to a central data acquisition site using cables, which was not a practical solution for SUGAR; instead, each detector was autonomous. Each provided its own power, and data were recorded to audio tape. Triggering of the detectors, which required a pulse in both scintillators to occur within 350 ns of each other, was also independent, the timing for which was taken relative to a pulse emitted from a clock in the proximity of the array. The data tapes from each station were collected every 7-10 days and collated to allow event reconstruction [1, 117]. The development of an array of autonomous detectors was a step forward in array design, one which would be required to allow the construction of extremely large detectors such as the Pierre Auger Observatory [6].

Unfortunately, the sparsity of the array and low sensitivity to the electromagnetic component of extensive air showers resulted in poor measurements of shower LDFs, and consequently energy estimation [17]. Exacerbating this problem was the issue of photomultiplier afterpulsing which, when coupled with the logarithmic pulse height-to-time converters used in SUGAR, led to energy overestimates - the effect of which was not realised for several years [118, 119, 120]. Consequently, the focus of SUGAR data analysis was on arrival direction studies (e.g. [116, 121]), including an analysis of arrival directions around the Galactic Centre [87] - a unique opportunity at the time due to it being a region of sky which is not easily accessible for many Northern Hemisphere detectors.

3.1.4 Yakutsk

In operation since 1970, the array near Yakutsk, Russia (61.7° N), has undergone several reconfigurations throughout its lifetime. An expansion in 1974 achieved a collecting area of approximately 18 km^2, but it was subsequently contracted in 1995 to cover about 10 km^2. A somewhat unique feature of Yakutsk is the presence of upward-facing PMTs in addition to scintillation detectors. These PMTs detect the Čerenkov light produced by EAS in the atmosphere, from which a calorimetric estimation of primary energy may be made. The ground parameter obtained with the scintillator array is calibrated against the energies estimated with the PMT array, thus allowing the energy scale of Yakutsk to be defined without heavy reliance upon air shower simulations [17, 122, 123].

The current layout of Yakutsk is shown in Figure 3.3. The present configuration is designed to detect air showers from primary particles of energies between 10^{15} eV and 10^{20} eV, a range which is achieved through the use of different detector spacings within the array. 58 scintillator stations - each consisting of two 2 m^2 scintillator detectors - are distributed over the 10 km^2 area with a further 6 scintillator stations buried for the detection of muons. In addition to this, 48 PMTs are arranged in two arrays , the first of which has a spacing of roughly 500 m

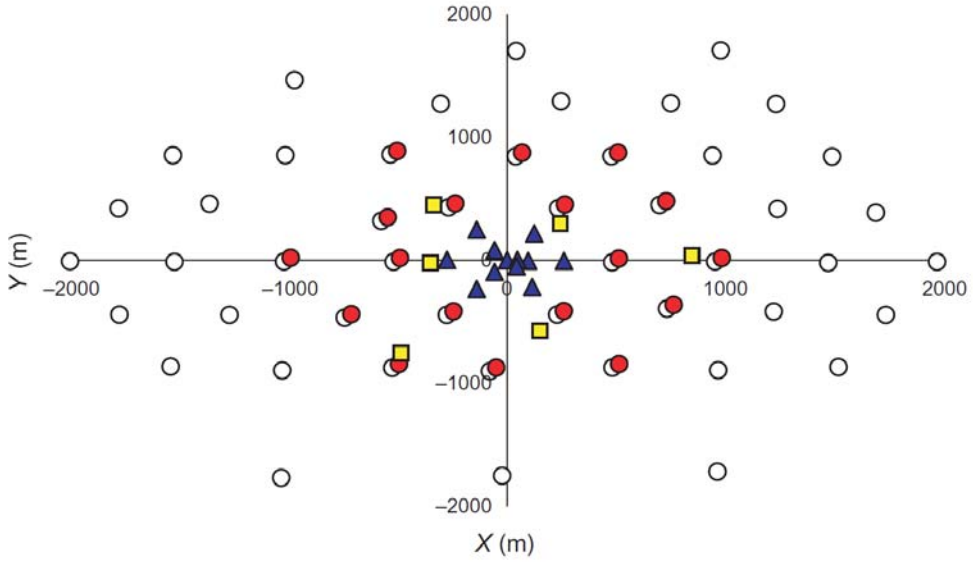


Figure 3.3: The layout of the Yakutsk array. The positions of the scintillator detectors are indicated by open circles; those of the muon detectors by filled squares; the C_1 and C_2 sub-arrays of Čerenkov detectors by filled circles and filled triangles, respectively. See the text for further details of these detectors. From [122].

(the C_1 array). The second PMT array, C_2, was added in 1995 and is arranged at the centre of the site with spacings ranging from 50 m to 200 m. Triggering for the C_1 array is provided by the scintillator detectors, while C_2 triggers independently of the rest of the array [122].

Development of the Yakutsk array is ongoing, and the addition of imaging atmospheric Čerenkov telescopes to the detector is currently planned. This enhancement will allow the longitudinal development profile of showers to be measured and significantly increase the ability of Yakutsk to study the composition of the primary CR flux [123].

3.1.5 Fly's Eye

The Fly's Eye detector is the first of those discussed here to utilise the air fluorescence technique. While it was not the first detector built around this method - early fluorescence detectors were built at Ithaca, New York [124] and the Dōdaira Observatory, Japan, the latter of which reported a detection of EAS through atmospheric fluorescence emission in 1970 [125] - Fly's Eye became the first to successfully detect large quantities of events. Following the success of a prototype fluorescence detector built at Volcano Ranch, in which coincident detections were made by both fluorescence and scintillator detectors - indicating that the detection by the former was of a real event - the first of the two Fly's Eye detectors was built at Dugway Proving Grounds, Utah, approximately 160 km from Salt Lake City at a latitude of 40° N [17, 24, 126]. Fly's Eye I (FE I), as it was known, began taking data in 1981. A second detector (Fly's Eye II, or FEII) was built 3.3 km away from Fly's Eye I [127] and began operation

in 1986 [17, 24]. Fly's Eye operated in this configuration until the end of the experiment in 1992 [7, 17].

Both Fly's Eye I and II consisted of PMTs to detect the fluorescent light emitted by atmospheric nitrogen. 67 detectors were located at Fly's Eye I, each housed in a steel pipe section 2.13 m long and 2.44 m in diameter. A 1.6 m diameter spherical mirror section was mounted inside each section of pipe. At the focus of the mirror were located groups of either 12 or 14 Winston light collectors combined with PMTs and necessary electronics. In total, Fly's Eye I was comprised of 880 PMTs, arranged such that each was viewing a unique section of sky. Fly's Eye II was similar to Fly's Eye I but was of reduced size - 36 detectors viewed half of the sky, oriented in the direction of Fly's Eye I [17, 24]. Each detector operated independently, but the triggering of Fly's Eye I caused a signal to be sent to Fly's Eye II to record a marker in the case of the latter also triggering. Thus, a subset of events observed simultaneously by both detectors could be reconstructed in stereo (see Section 1.2.2.2) [128].

Fly's Eye was crucial for demonstrating that the fluorescence technique was a viable means of CR detection. The development and demonstrated success of a detection technique completely independent of ground arrays was an important advance in the field of CR astronomy. Results published by Fly's Eye include the measurement of the ankle in the CR energy spectrum at around 3×10^{18} eV and a change in composition from heavy to light primaries in the interval of 0.1 EeV to 15 EeV [128]. In addition, it detected the highest energy CR to date - the energy of which was estimated to be 3×10^{20} eV [129]. The shower profile of this event appears in Figure 3.4.

3.1.6 AGASA

Prior to the construction of the Pierre Auger Observatory, AGASA (Akeno Giant Air Shower Array) was the largest air shower array in the world [7, 17, 65]. It was located approximately 120 km west of Tokyo, Japan, at a latitude of 38.8° N [130, 131]. Originally divided into four separate branches, AGASA was unified in 1995 to operate as a single detector [7, 131]. Construction of the main array began in 1987, with smaller arrays located at the site having been constructed earlier [130]. Data collection began in February 1990 and continued until January 2004 [39].

AGASA covered approximately 100 km^2 with 111 2.2 m^2 plastic scintillator detectors spaced approximately 1 km apart. Muon detectors, consisting of proportional counters located underneath an absorber giving a vertical muon energy threshold of 0.5 GeV, were built at 27 of the scintillator sites. Optical fibres connected the detectors sequentially with one of four branch centres which processed and stored the data from detectors within its branch [130, 131]. Prior to the four branches of AGASA being unified, triggering and shower reconstruction were handled independently by each branch; linking the branches together allowed the reconstruction of showers which triggered detectors close to or across branch boundaries, increasing the effective area of the array by a factor of approximately 1.7 [131].

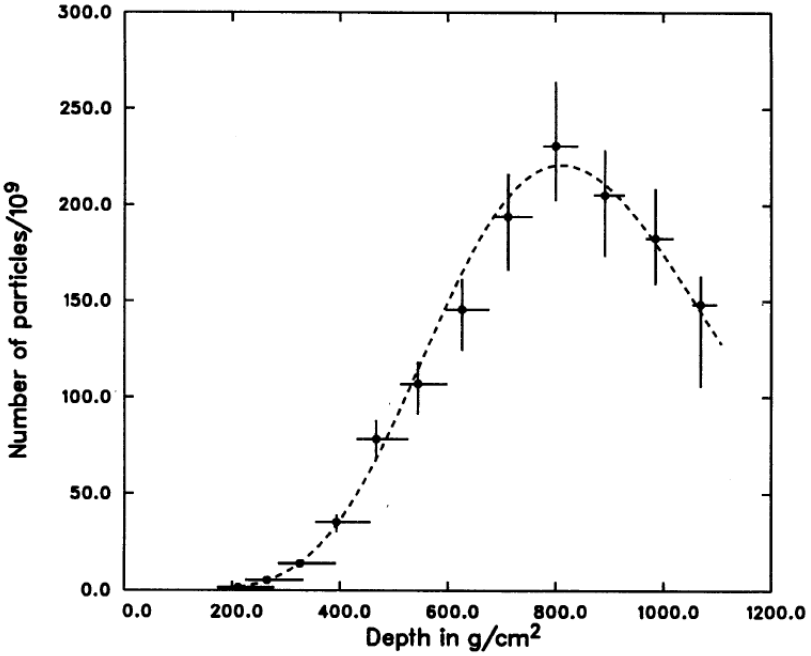


Figure 3.4: The shower profile of the highest energy cosmic ray recorded, as measured by Fly's Eye. The event's energy is estimated to be 3×10^{20} eV. Plotted as a dotted line is the best-fit Gaisser-Hillas function. From [129].

The AGASA collaboration published results regarding the CR composition, anisotropy and energy spectrum. Of particular curiosity was the latter - significant numbers of super-GZK events were observed at the array, including one possessing an estimated energy of between 170 and 260 EeV which arrived soon after operation began [132]. In 1998 the AGASA collaboration published their UHECR spectrum and claimed to see no spectral cut-off compatible with the GZK effect [65]. The subsequent measurement by HiRes of the existence of such a suppression appeared to be at odds with the AGASA measurement, a mystery which was not resolved for several years (for further discussion, see Section 2.1).

The excellent angular resolution - approximately 3° at 10^{18} eV and 1.5° at 10^{19} eV [133] - achieved at AGASA allowed for both large and small-scale studies of anisotropy at high energies, two major detections of which were reported by the AGASA collaboration. The first was a large scale anisotropy at energies around 10^{18} eV. Using the harmonic analysis method, a first harmonic of amplitude 4% was found with the peak directed towards the Galactic Centre. By performing smoothing on the data, the scale of the observed excess was found to be approximately 20° [133]. In addition, at energies above 4×10^{19} eV a small-scale clustering of arrival directions was observed. One cluster of three events, plus three clusters of two events were found in 2.5° windows, the chance coincidence of which was reported to be less than 1% [91]. However, the true statistical significance of these results is unclear (see Chapter 2).

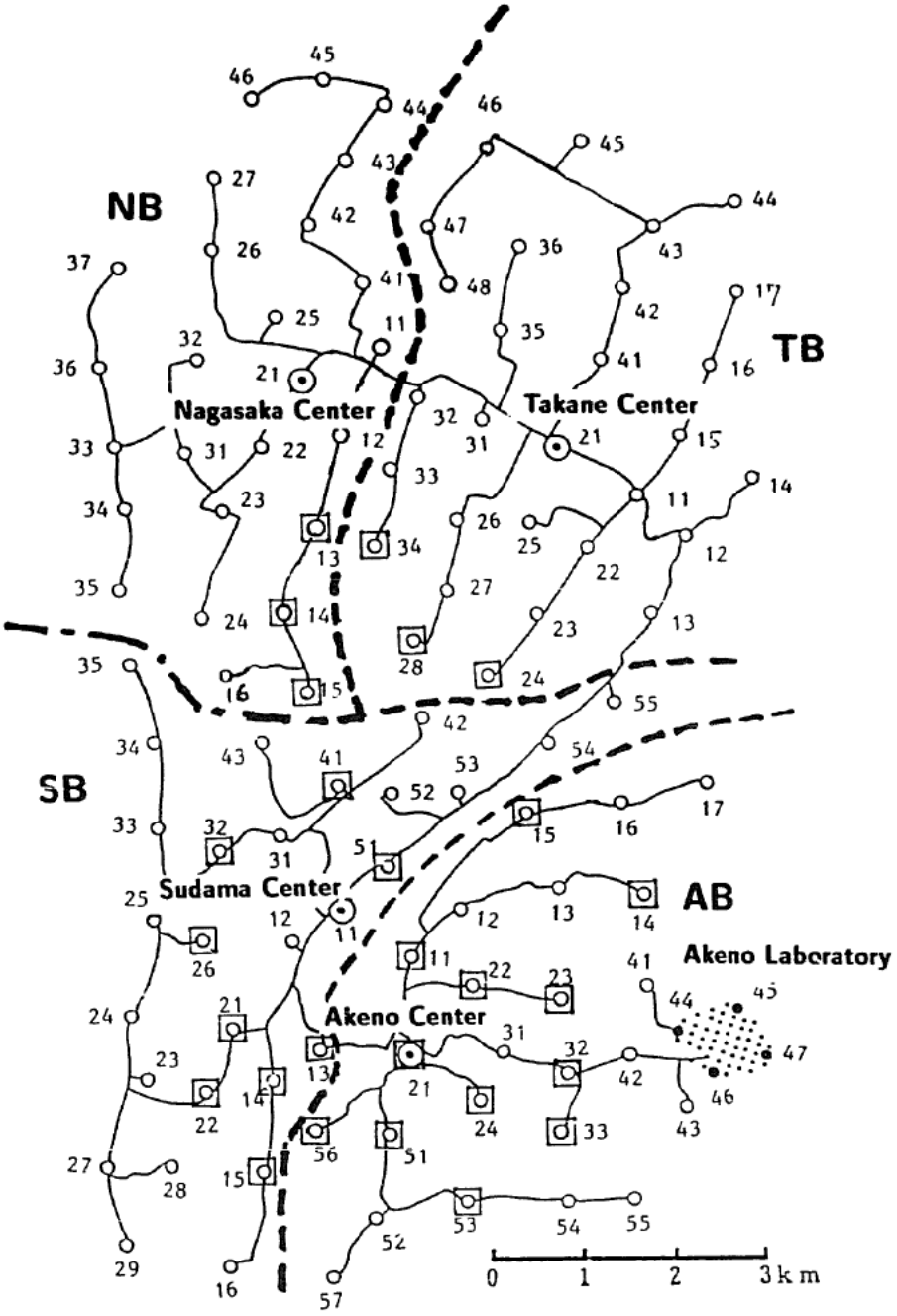


Figure 3.5: The layout of the AGASA detector, as shown in [131]. The positions of the above- and below-ground scintillators are indicated by open circles and squares, respectively. The solid lines display the optical fibre routes which connect the detectors to data processing centres, and the dashed lines show the boundaries of the four branches of the array.

3.1.7 HiRes

The High Resolution Fly's Eye observatory, or HiRes, was, as its full name suggests, the successor to the Fly's Eye experiment. Like its predecessor, HiRes was located at the Dugway Proving Grounds in Utah. Two detectors named HiRes I and HiRes II comprised the experiment and began operation in 1997 and 1999, respectively [7, 134]. Operation of both HiRes detectors ceased in 2006 [45].

A distance of 12.6 km separated HiRes I and II [69]. Each detector consisted of a number of individual telescope modules which were arranged to cover 360° in azimuth. An individual module was made up of a roughly $4\,\mathrm{m}^2$ spherical mirror, at the focus of which was located a 16×16 array of PMTs each viewing a 1° diameter cone on the sky. HiRes I contained a single ring of 22 of these modules, viewing elevations between 3° and 17°. The number of modules in HiRes II was increased to 42 which consequently increased the visible range of elevations to between 3° and 31° [45, 69]. This hardware provided HiRes with several improvements over Fly's Eye including greater sensitivity due to larger mirrors as well as better spatial (and hence angular) resolution and an improved signal-to-noise ratio due to the reduced pixel size [135]. In addition, the photomultiplier read-out electronics of HiRes II employed an FADC system operating at 10 MHz which afforded several improvements over the $5.6\,\mu$s sample-and-hold integration of HiRes I, including definition of pulse shapes as well as sampling of background noise before and after a signal pulse [126, 134].

HiRes I and II were triggered independently, allowing the reconstruction of monocular data. A subset of these events, seen by both detectors, were also reconstructed in stereo, permitting a higher quality estimate of the shower properties. As an example, the angular resolution was approximately 0.8° for stereo events [102] compared to several degrees for monocular events [92]. These data sets allowed a variety of studies to be performed, with the monocular data being used for studies in which uncertainties on shower parameters were not critical but large statistics required, while the stereo data were used in studies such as composition and anisotropy measurements, where precise measurements were necessary [45].

The HiRes data were used to make many measurements of properties of the CR flux, including composition and anisotropy [45]. A particularly notable result was the observation of a flux suppression above an energy of $\approx 6 \times 10^{19}\,\mathrm{eV}$, which was compatible with predictions of the GZK effect [69]. This did not seem compatible with the AGASA spectrum, which - with approximately half of the exposure of HiRes above the GZK threshold - showed a lack of a spectral steepening above this energy (see Section 2.1).

It should also be noted that the HiRes prototype served as the fluorescence detector component of an early hybrid detector, for which the Michigan muon array took ground-level measurements. HiRes/MIA was operational for several years and successfully measured several thousand showers, permitting composition measurements between 10^{17} and $10^{18}\,\mathrm{eV}$ [75]. A successor experiment, the Telescope Array (TA), is a hybrid experiment using one of the HiRes fluorescence telescopes as well as two new telescopes in addition to a newly-installed

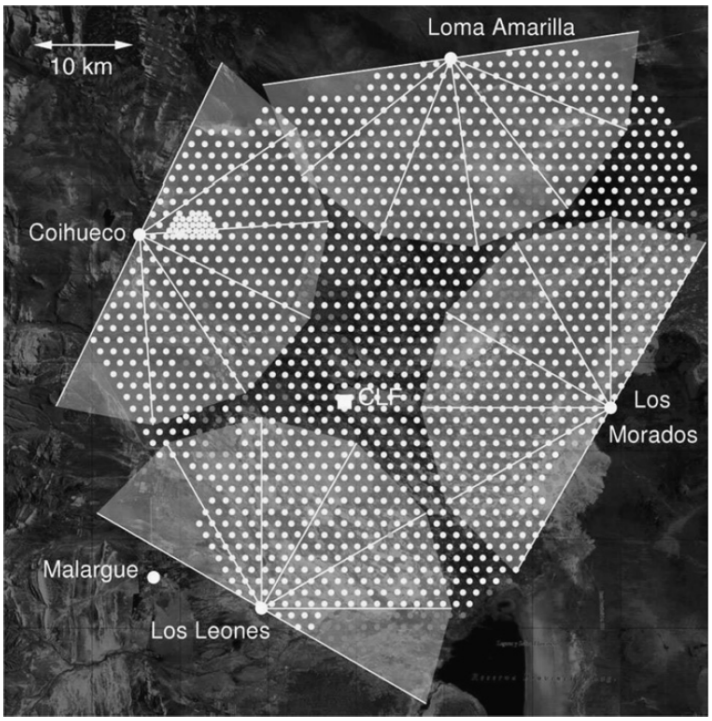


Figure 3.6: Overhead view of Auger South as of March 2009. Each dot corresponds to a vertex of the SD array, with light grey indicating the positions of deployed stations and dark grey showing unoccupied positions. The positions of the four FD sites are illustrated and their respective fields of view shaded in light grey; white lines separate the fields of view of the 24 telescopes. The position of the central laser facility (CLF) is also shown. From [23].

scintillator ground array. Construction was completed and full operation underway as of 2009, with preliminary results having been reported shortly thereafter [136, 137].

3.2 The Pierre Auger Observatory

Taking advantage of the knowledge gained from several decades of cosmic ray experiments, in 2004 data collection began at the world's largest cosmic ray detector - the Pierre Auger Observatory. Planned as a pair of hybrid detectors (see Section 1.2.2.3) - one in each of the Northern and Southern Hemispheres, called Auger North and Auger South, respectively - the Pierre Auger Observatory is designed to investigate the cosmic ray flux at the highest energies, gathering an unprecedented number of events from EeV up to trans-GZK energies [27]. An operational lifetime of each site of 20 years is called for in the design of the Pierre Auger Observatory [42].

The site chosen for Auger South was Pampa Amarilla, near the town of Malargüe in the Province of Mendoza, Argentina. It covers the region of 35.0°-35.3 S, 69.0°-69.4° W with an average altitude of 1420 m above sea level; the altitudes of the detectors range between 1340

and 1610 m above sea level. Factors contributing to the choice of this site were its relative flatness which permits wireless communications as well as ease of deployment and maintenance, its proximity to existing infrastructure in the form of Malargüe, and clear night skies with low levels of light pollution which allow good measurements of nitrogen fluorescence emission [27, 42].

In what follows, Sections 3.2.2 to 3.2.5 refer to the various components of Auger South, while Section 3.2.6 discusses its northern counterpart.

3.2.1 Communications and CDAS

Communications at Auger South are based on a wireless network owing to the large size of, and the number of detectors in, the array (see Figure 3.6). Data transmission from the stations of the surface detector (SD, which is the ground array component of Auger South) to their base stations (see below) is handled by a wireless local area network (LAN) operating in the 902-928 MHz band, while a microwave network in the 7 GHz band transmits data back to the Observatory Campus in Malargüe, where it is forwarded to the Central Data Acquisition System (CDAS) [27].

The wireless LAN is managed by dividing the SD array into many sectors, to each of which a base station is assigned to handle communications for the stations belonging to that sector. Four communications towers, one at each of the fluorescence detector (FD) sites, each hold eight base station units. Dividing the array up in such a manner efficiently distributes the processing load required to handle array-wide communications, and also serves to prevent data loss from the entire array in the event of base station failure.

Each of the four communications towers transmits data from its base stations as well as the local FD site (see Section 3.2.3 for a description of these sites). Data are transmitted directly to the CDAS from Coihueco and Los Leones, while that from Loma Amarilla and Los Leones are relayed via the former and latter, respectively.

The Central Data Acquisition System serves several purposes. It controls the stations of the SD and allows access to SD monitoring data and configuration tools, and manages the storage of data. Its primary purpose, however, is to process triggers for SD-only and hybrid events and initiate data collection. Triggering of the FD occurs independently of the CDAS, but notification is transmitted to the latter on the detection of an event of interest for the purpose of hybrid triggering (see Section 3.2.3.2). Triggering of the SD is discussed in more detail in Section 3.2.2.3.

3.2.2 Surface Detector

Extending over an area of some 3 000 km^2, the surface detector of Auger South consists of over 1 600 water-Čerenkov detectors arranged on a triangular grid with a spacing of 1.5 km. This choice of spacing is a compromise between energy threshold and cost, with an overlap of the bottom of the energy range accessible by the Pierre Auger Observatory with that of

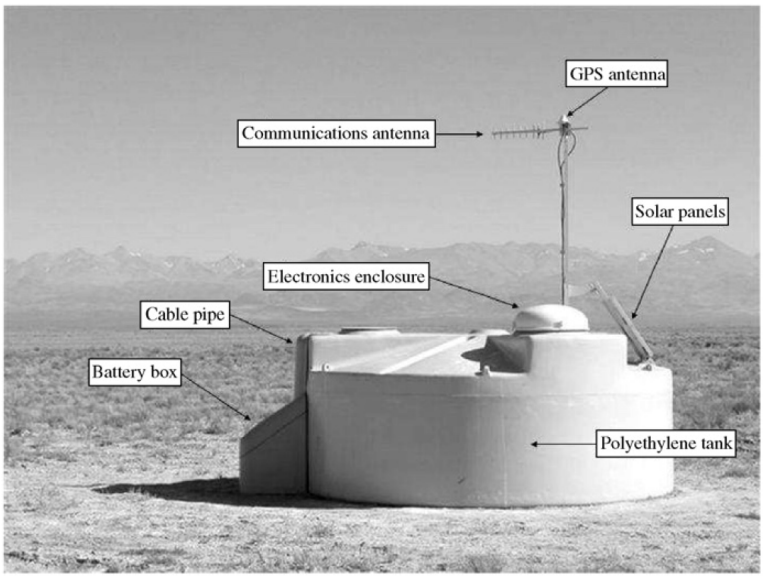


Figure 3.7: One of the SD stations of Auger South. Labelled are the main components of the station. Also of note is the flatness of the terrain surrounding the station, which is typical of the site of Auger South. From [42].

previous cosmic ray experiments being desirable as it allows comparison of results [42]. With this spacing, the surface detector becomes fully efficient - meaning that the trigger probability for the 3ToT trigger (see Section 3.2.2.3) reaches 100% - at an energy of 3×10^{18} eV for zenith angles less than 60° [138] and a shower of energy 10^{19} eV will trigger a minimum of five stations [42].

Construction of the SD began in 1999 and was completed in 2008, with stable data collection beginning in January 2004 [139]. The choice of water-Čerenkov detectors was made as a result of their low cost, ability to withstand harsh conditions and their almost uniform exposure to particles incident at zenith angles out to 60° [42]. In addition, their depth corresponds to approximately 3 photon radiation lengths which achieves efficient conversion of the photons in an EAS to charged particles and reduces the statistical fluctuations in the sampling of the electromagnetic component of EAS [32].

3.2.2.1 Station Design

Each station of the SD is a self-contained unit that operates independently of the rest of the array. One such station is shown in Figure 3.7. A polyethylene tank encloses most of the components and protects them from the elements, leaving only communications antennae and solar panels to be mounted on its exterior.

A cylindrical volume of 3.6 m diameter and 1.2 m depth is enclosed by the tank and contains 12 000 L of ultra-pure water that is required to maximise the transmission of Čerenkov light as well as avoid the growth of microbes, which could cause the detector performance to

become unstable. Surrounding this water is a strong, reflective plastic liner which prevents contamination of the water and offers an additional barrier against external light. Three windows of UV-transparent polyethylene are located in the top surface of the liner at a distance of 1.20 m from its centre, through each of which a 9 in diameter PMT views the water [42, 140].

The signals from the PMTs are processed by the electronics package mounted in an enclosure at the top of the station. The large dynamic range required for measuring signals both close to, and at large distances from, a shower core is achieved by reading the output from both the anode and the last-stage dynode of the PMTs, the latter of which is amplified. This range enables the measurement of signals from several to roughly 10^5 photoelectrons. An FADC operating at 40 MHz filters and digitises these signals for triggering (see Section 3.2.2.3). The timing for these signals is provided by a 100 MHz clock which is synchronised by a GPS receiver at each station. Also included in the electronics package are a processor which performs the tasks of local data acquisition, detector monitoring and software triggering, as well as memory for the storage of data.

Power is provided to the station through a solar power system, providing an average of 10 W. A pair of solar panels, together with two lead acid batteries, power each station for an estimated 99% of the time. Accounting for possible degradation of the power system over the lifetime of the SD, an uptime fraction of about 98% is still expected [42].

3.2.2.2 Calibration

Accurate reconstruction of shower energies and stable triggering of the SD relies on knowledge of the particle response of every individual station involved in an event, requiring constant calibration of each station. Owing to the number of stations of the SD and large distances involved, this calibration is performed autonomously and is achieved through the continuous monitoring of signals produced by the atmospheric muon flux [141].

The calibration of each station is performed with data recorded immediately prior to an event trigger. Along with the event data, upon such a trigger, histograms of the charges and pulses produced at the PMT outputs by atmospheric muons in the preceding minute (corresponding to approximately 150 000 muons) are also transmitted to the CDAS. Examples of these histograms are shown in Figure 3.8. This allows the accurate conversion of a measured signal, which is in units of FADC channels and is dependent upon the status of individual detectors, into units which are station-independent and allow straightforward comparison of signals.

This reference unit is known as a VEM. One VEM (also denoted as Q_{VEM}), or vertical equivalent muon, is defined as the average measured charge produced at the PMT output by a muon travelling vertically through the centre of a station. Knowing the value of 1 VEM in integrated FADC channels allows the conversion of any other signal into these units. While an SD station is unable to pick out vertical, central muons only, the value of 1 VEM can be determined from a charge histogram produced by background muons. Through an independent

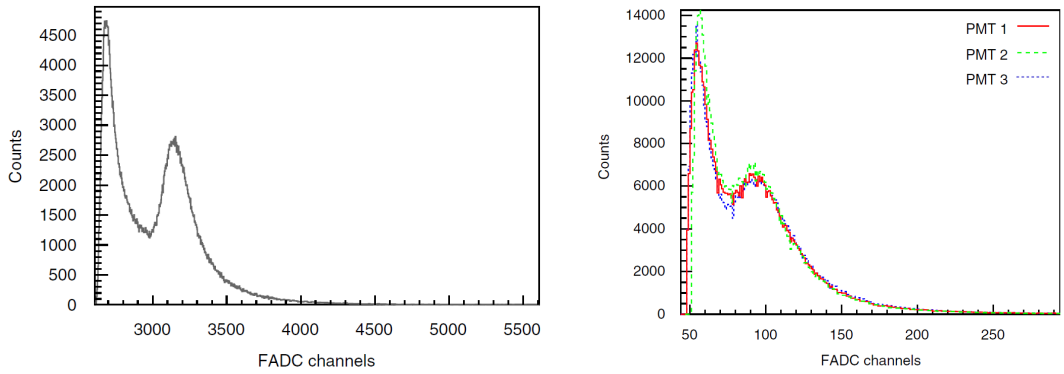


Figure 3.8: Examples of histograms returned to the CDAS along with information about a triggered event. The histogram in (a) shows the summed charges for the 3 PMTs over the preceding minute while (b) shows the pulse height histogram for each of the 3 PMTs. In both histograms, the peak produced by atmospheric muons and used for calibration is that on the right; the left-hand peak is produced by low-energy particles and is not of interest for calibration purposes. From [141].

experiment, the position of the second peak in a histogram of the sum of the charges from all 3 PMTs in a station is known to correspond to 1.09 VEM. Thus, when the data are analysed, the position of this peak is found by a polynomial fit and its value used to convert measured signals to VEM accordingly. The accuracy of this conversion is around 3%.

Use of the atmospheric muon flux is also made for the purposes of setting the triggering levels of the SD (see below). Thresholds are set with respect to a value derived from the current produced by the PMT as a result of the atmospheric muon flux, so again the relationship between this value and FADC channels must be known. The reference unit in this case is I_{VEM}^{peak}, which is the value of the peak of a histogram of pulse heights measured by the station. However, directly measuring this value would result in unacceptable downtime for a station, and as a result an estimate of I_{VEM}^{peak} is used that is denoted as I_{VEM}^{est}. Using I_{VEM}^{est} allows for stable triggering of the SD while performing continuous calibration. A convergence algorithm is performed to maintain the rate of a calibration trigger at a pre-defined value. A rate of approximately 70 Hz (determined from a reference station) is required to be measured for signals above a value of $2.5I_{VEM}^{est}$ in a given PMT and above $1.75I_{VEM}^{est}$ in all three PMTs. If this is not the case, the value of I_{VEM}^{est} is adjusted until these conditions are satisfied. Once convergence is achieved, the value of I_{VEM}^{peak} is known to an accuracy of 6% which gives stable triggering conditions.

3.2.2.3 Trigger System

The SD employs a hierarchical trigger system, starting at the low-level triggers processed independently by each station - known as the 'local' triggers - and leading up to the high-

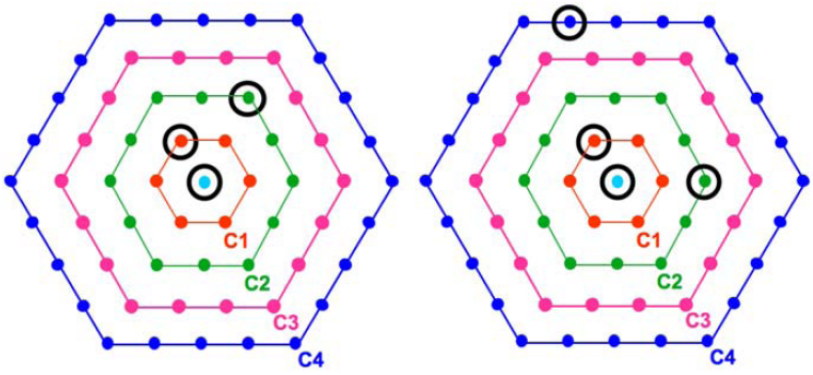


Figure 3.9: Two examples of events satisfying the spatial criteria of the T3 trigger. The example on the left is a $ToT2C_1\&3C_2$, while that on the right is a $2C_1\&3C_2\&4C_4$. See the text for further details of these configurations. Originally appeared in [139].

level triggers which are processed off-line and select physical events for reconstruction. Local triggering is required for efficient use of the communications bandwidth at the site, as it filters out background signals and selects only events of interest to be transmitted to the CDAS [139, 142].

The lowest level trigger, known as T1, occurs when either of two conditions is satisfied. The first is a time over threshold trigger (denoted as ToT-T1), which requires 13 FADC bins in a 120 bin window to be above a level of 0.2 I^{est}_{VEM} for two PMTs in coincidence. The ToT-T1 is designed to select low-amplitude, dispersed signals such as those from distant or low energy showers. The rate of this form of the trigger is about 1.6 Hz. The alternate condition for a T1 (TH-T1) is a threshold value of 1.75 I^{est}_{VEM}, required to be exceeded in coincidence with all three PMTs. TH-T1 performs the task of selecting the quick signal expected from the muonic component of highly inclined showers, at the cost of increased background noise - a rate of approximately 100 Hz is seen due to this trigger.

Following a T1 trigger, a station's processor checks the signal against the T2 criteria. This is designed to reject background noise and select signals which are likely to result from showers. All ToT-T1 triggers are upgraded to a T2, while the TH-T1 triggers are required to pass a threshold of 3.2 I^{est}_{VEM} in coincidence with all three PMTs in the station. The results of the T2 trigger are subsequently sent to the CDAS for processing, which occurs for a given station at a rate of roughly 20 Hz.

Triggering of the array occurs at the T3 level and is assessed by the CDAS using coincidences of T2 triggers between separate stations. Once the CDAS recognises a T3 trigger, FADC trace data are collected from every station exhibiting either a T2 trigger or a T1 trigger within 30 μs of the T3. Two sets of conditions will trigger a T3 and examples are shown in Figure 3.9: the first requires a coincidence of ToT triggers in 3 stations in addition to exhibiting a specified level of compactness, and is referred to as "$ToT2C_13C_2$". In this notation, C_n refers to the n^{th} successive hexagon surrounding a given tank. This trigger then requires that

one of the stations has at least one of its nearest neighbours triggered (giving two within the smallest hexagon around the station of interest, or $2C_1$) as well as at least one of its second nearest neighbours (giving 3 stations within the second hexagon, or $3C_2$). An additional time criterion ensures that the T2 triggers occur within $(6 + 5C_n)\,\mu$s of the first. The T3 trigger under these conditions occurs around 1600 times per day, 90% of which correspond to real events.

A more relaxed set of spatial criteria give the second configuration which will trigger a T3. It is designed for the detection of highly inclined showers in which the electromagnetic component is highly attenuated, and the remaining muons produce signals of short duration but spread over a large area of the array. This second type of T3 is known as a "$2C_1\&3C_2\&4C_4$", with the timing criterion the same as for the $ToT2C_1\&3C_2$. A rate of about 1200 triggers per day is observed as a result of the relaxed configuration, approximately 10% of which are real showers.

Next in the trigger sequence is the T4, the first of two off-line trigger levels. This is a 'physical' trigger which is designed to accept real showers and reject noise triggers remaining after the T3 level. The first of the two criteria (3ToT) is a requirement of three stations in a triangular cell to pass the ToT trigger requirement with their arrival times consistent with a plane shower front moving at the speed of light. At fully-efficient energies and zenith angles below 60°, more than 98% of real showers satisfy these requirements. The remaining ∼2% trigger the alternative T4 criteria (4C1), which requires 4 stations in a hexagon to trigger any type of T2 with the same timing requirements as the 3ToT. Following this selection of T4 triggers, the removal of stations which have triggered accidentally is performed by imposing timing criteria to stations outside the triangle of stations which measured the highest signal in the event. Of the events which reach this stage, 99.9% are reconstructed fully to determine their arrival direction, core position and value of $S(1000)$ (see Section 3.2.2.4).

Finally, the T5 trigger is designed to select events which are reconstructed with well-understood accuracies in both energy and geometry. Its main purpose is to avoid incorrect shower reconstructions arising from shower cores falling outside the array boundaries. The adopted trigger requires that the core of the shower is located within a triangle of working stations, and that the station registering the highest signal (the 'hottest' station) is surrounded by 6 working nearest neighbours. Application of this trigger also allows ease of aperture calculation for the SD, as the aperture at any given time for energies above 3×10^{18} eV is calculated by the sum of the number of hexagons composed entirely of working stations, multiplied by the effective aperture of a single hexagon. For purposes such as anisotropy studies, where the calculation of aperture is not critical, a relaxed T5 condition (T5 ICRC2005) can be used which requires only 5 of 6 neighbours of the hottest station to be working; this increases the available statistics while still allowing accurate reconstruction of event properties [142].

The implementation of this trigger chain reduces the single station signal rate from roughly

3 kHz, due mainly to atmospheric muons, to about 3 triggers per day, produced by real showers.

3.2.2.4 Event Reconstruction and Performance

Following the identification of a physical event by the CDAS, reconstruction can occur in order to estimate its properties. The methods discussed below pertain to the reconstruction of 'vertical' showers, which are those incident upon the array with zenith angles smaller than 60°. At zenith angles larger than this the attenuation of electromagnetic component of the shower is very large, and as a result the standard analysis as described below is no longer used. Discussion of the reconstruction and analysis of these 'horizontal air showers' (HAS) can be found in [36].

Determination of the arrival direction by the SD is performed through fits to the time of arrival of particles in each station involved in the event [143]. A parabolic shower geometry with variable curvature is used to represent the shower front for fitting purposes, although the particular functional form of the adopted shape does not introduce significant uncertainty into the analysis. Both the precision of the clocks at each station and natural fluctuations in the arrival time of the first particle at each station contribute to the uncertainty in arrival direction. These effects are incorporated into a time variance model which describes the probability distribution of arrival times at each station; the presence of 'doublets' in the array (pairs of tanks separated by 11 m) allows experimental verification of the model used through the repeated reconstruction of events involving a doublet with one of the two tanks excluded. The validity of this model allows the angular resolution to be determined on an event-by-event basis. For vertical showers involving only 3 stations an angular resolution of 2.2° is typical, a value which improves with increasing station multiplicity. Above an energy of 10 EeV, the angular resolution at all zenith angles is better than 1°. An independent check of these values is afforded by the subset of hybrid events which can also be fully reconstructed by the SD alone, and the resolution estimated through this method is found to be in reasonable agreement with that of the time variance model [31].

Energies of events are estimated by the SD through the fitting of a lateral distribution function to the signals recorded by each station involved in the event. From this LDF, $S(1000)$ - the signal that would have been produced in a station at a core distance of 1000 m - is calculated. To adjust for the zenith-angle dependent attenuation of signal, $S(1000)$ is converted to S_{38°, which is the $S(1000)$ that the shower would have produced had it arrived at the median observed zenith angle of 38°. This conversion is performed through the use of the Constant Intensity Cut (CIC) method [144]. Finally, the conversion from S_{38° (in units of VEM) to primary energy is made by using a calibration curve that is derived from hybrid events. This process gives statistical uncertainties of approximately 19%, where again the doublets allow experimental verification of the uncertainty in the signal measured by the SD [32]. Systematic uncertainties due to the calibration procedure increase from about 7% at 10^{19} eV to 15% at 10^{20} eV, while an overall systematic uncertainty of 22% arises from the use of the FD energy

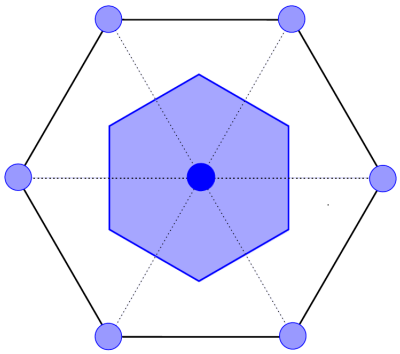


Figure 3.10: An illustration of the basic hexagonal cell which is used for exposure calculations. The blue shaded area indicates the hexagon corresponding to the station in the centre, which is surrounded by 6 working stations. From [139].

for calibration (see Section 3.2.3) [70].

Composition measurements using the SD are currently in development, with the viability of several different observables being studied [40]. The first of these is the 'rise-time' of the FADC signal, which is the time taken for the integral signal to increase from 10% to 50% of its total value. This parameter is sensitive to the ratio of the muonic and electromagnetic components in the shower front. Another observable is the slant depth $XAsymMax$ at which, for a given energy, the distribution of rise-times as a function of azimuthal angle reaches its maximum asymmetry. This parameter correlates with the amount of longitudinal development undergone by the shower, which itself is related to the primary species [41]. In addition to these methods, several more are under consideration - testament to the difficulty of precise composition measurements with ground arrays.

Knowledge of the state of each SD station at any given time allows the accurate determination of the instantaneous (and hence, integrated) exposure of the array. Above 3 EeV, the full efficiency of the array reduces the aperture calculation to a simple geometrical one. The individual element of the array for exposure calculations is taken as a hexagon defined by a central detector surrounded by 6 working neighbours, an example of which is shown in Figure 3.10. The area of this hexagon is $1.95\,\mathrm{km}^2$, giving an aperture for zenith angles within 60° of $a_{cell} \approx 4.6\,\mathrm{km}^2\,\mathrm{sr}$. Monitoring of the number of cells, $N_{cell}(t)$, is performed on a second-by-second basis which allows very good determination of the total array aperture $N_{cell}(t) \times a_{cell}$. Integration of this quantity over time gives the exposure of the array as a whole. The uncertainty in $N_{cell}(t)$ is approximately 1.5% and arises from communication problems between stations and the CDAS, while the uncertainty in the live time of the array is approximately 1% and is due to problems in communications across the array or local data storage at individual stations. The total uncertainty on integrated exposure is less than 3%. With the full array having been deployed, the SD exposure above $3 \times 10^{18}\,\mathrm{eV}$ increases by approximately $500\,\mathrm{km}^2\,\mathrm{sr\,yr}$ every month [139].

3.2.3 Fluorescence Detector

The fluorescence detector of the Pierre Auger Observatory comprises 24 telescopes (in addition to those of HEAT - see Section 3.2.5.2) - configured as 4 groups of 6 telescopes each - overlooking the SD from elevated positions around the edge of the array. This configuration provides an FD triggering efficiency of 100% across the entire ground array for energies above 10^{19} eV and also allows for stereo and higher multiplicity events. However, the latter feature was not as important a requirement as in previous fluorescence-only detectors due to the excellent reconstruction of shower geometry permitted by hybrid detections [23]. Each telescope views 28.1° in elevation and 30° in azimuth, giving each site a 180° azimuthal field of view. The four sites are named Los Leones, Coihueco, Loma Amarilla and Los Morados and are located on the southern, western, northern and eastern edges of the SD, respectively.

The four buildings of the FD are climate-controlled to minimise the effects of temperature variations on the system. Retractable shutters cover the windows to each telescope bay and protect them from the elements, only being opened when the FD is being operated. The requirement of clear nights with minimal moonlight gives an average duty cycle for the FD of about 13% [23].

3.2.3.1 Telescope Design

The telescopes of the FD use Schmidt optics for imaging EAS which achieves an angular spot diameter of no greater than 0.5° (meaning that 90% of the light from a distant object falls within this region). A camera, consisting of 440 hexagonal PMT pixels each viewing a 1.5° diameter region of sky, is situated in the focal surface of a $\sim 13\,\mathrm{m}^2$ spherical mirror. Spaces between the pixels are covered with a light collector constructed of reflective 'Mercedes stars' to maximize light collection and minimize the spreading of light between pixels. The efficiency of light collection averaged over the camera is 93%, compared to a value of 70% when the Mercedes stars are not in place [23]. The accuracy of the pixel alignment is excellent, with laser shots and tracking of bright stars across the field of view enabling the absolute pointing direction of each pixel to be determined to within 0.1° [23, 38].

The mirror viewed by the camera is made from segmented pieces, and is one of two designs due to the use of two different manufacturers. Los Leones and Los Morados have anodised aluminium mirrors comprising 36 rectangular pieces, while Coihueco and Loma Amarilla use mirrors made of 60 hexagonal glass segments layered with a reflecting aluminium coating, producing reflectivities at a wavelength of 370 nm in excess of 90%. The scale of irregularities in both types of mirrors is below 10 nm.

To provide a larger effective aperture for the telescopes, an annular corrector ring is utilised to negate the spherical aberration of the mirror. This ring has inner and outer radii of 0.85 and 1.1 m, respectively, and is constructed from 24 glass segments. The presence of the corrector ring doubles the optical aperture of the FD telescopes, while maintaining a maximum spot diameter of 0.5°.

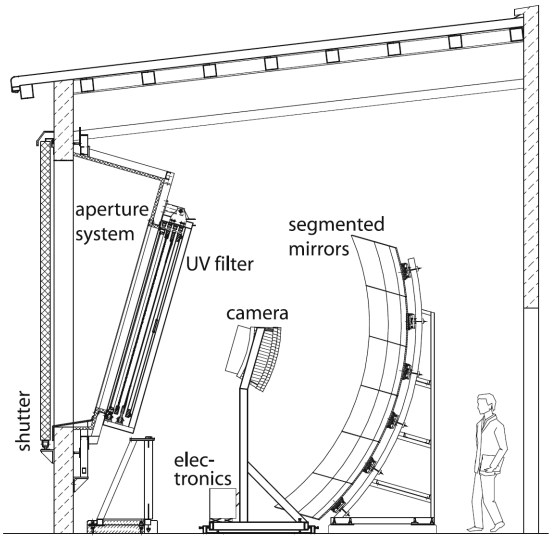


Figure 3.11: Schematic view of an individual FD telescope, with the main components labelled. Six such telescopes make up each of the four FD sites. Image from [23].

Finally, the aperture of each telescope is spanned by an optical filter for the removal of background light. The filter transmits photons having wavelengths between approximately 290 nm and 410 nm, which covers most of the nitrogen fluorescence spectrum. In the absence of such a filter, the low flux of photons produced by EAS would be indistinguishable from the background. The layout of one such telescope is illustrated in Figure 3.11.

3.2.3.2 Electronics and Triggering

As with the stations of the SD, triggering of the FD comprises hardware triggers which are applied by the telescope electronics on-line to filter out background events efficiently, as well as software triggers which are evaluated off-line and used to make the final selection of real events.

The electronics of the FD are required to perform well under a variety of background light conditions, detecting both low and high energy events over a large volume surrounding each site. A large dynamic range is achieved through a compression technique described in [23] while local intelligence is used to adapt to changing light conditions by adjusting pixel thresholds to maintain a trigger rate of 100 Hz. The latter is performed by the First Level Trigger (FLT) boards, which also generate pixel triggers when the threshold is exceeded.

The output of the FLT boards is processed by the Second Level Trigger (SLT) boards, which process the data and identify triggered pixel combinations compatible with straight tracks. Examples of the five fundamental patterns that the SLT boards search for are shown in Figure 3.12. To allow for untriggered pixels, only four of the five pixels in each combination are required to trigger. Following the identification of a straight track segment, which occurs at a rate of 0.1-10 Hz per telescope, the data from the camera are processed by the Third Level

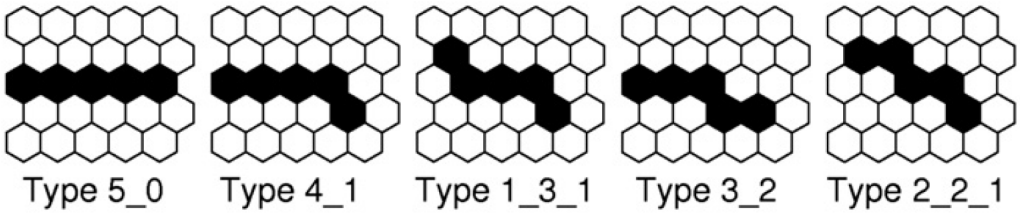


Figure 3.12: The five fundamental trigger patterns used in the track identification algorithm. From [23].

Trigger (TLT) algorithm.

Rejection of triggers caused by lightning is the main purpose of the TLT. It is the first software trigger for FD data and is essential to prevent lightning strikes - which can trigger hundreds of pixels in a single camera - from causing large amounts of dead time in the FD. This trigger is very efficient and rejects 99% of all lightning events at the cost of only 0.7% of real events. After an event passes the TLT conditions, which happens at a rate of about 0.01 Hz per telescope, it is passed to the final trigger level.

Events which have passed the TLT are handled by a further algorithm which collects data from all telescopes of the site and also initiates a T3, or hybrid, trigger. The purpose of the T3 trigger is to initiate data collection from the SD for the purposes of forming a hybrid event. A simple reconstruction of the FD data is performed to estimate the location and time of impact at the ground for a shower seen by the FD. The FD data, along with the T3 trigger, are sent to the CDAS which subsequently requests signals from the SD that are in the vicinity of the relevant eye and compatible with the ground impact time. This serves to retain SD data for events below 3×10^{18} eV, which may have triggered one or two stations but not produced a T3. The resulting FD and SD signals are stored, and later merged as a hybrid event during off-line analysis (see Section 3.2.4).

3.2.3.3 Calibration and Atmospheric Monitoring

Calibration of the FD entails determining the response of each pixel to an incident photon flux in order to enable the conversion from a value in ADC counts to a number of photons, and hence the ability to accurately estimate shower energy. Both absolute and relative calibrations are performed, allowing the efficient tracking of any change in telescope response.

Absolute calibration is performed periodically and is relatively time consuming - it takes approximately 3 days to calibrate each site [23]. An end-to-end calibration is used to calculate the detector response to light incident upon the aperture which implicitly accounts for the effects of optics and electronics. This removes the need for the detailed study of each component of each telescope individually. This absolute calibration is performed by mounting a drum which provides uniform illumination of known intensity at the aperture of each telescope. Measuring the response of each camera to pulses of light from this drum allows a

pixel-by-pixel calculation of the camera response, with measurements at several wavelengths possible to evaluate the spectral response curve of each camera. Absolute calibration in this manner gives an uncertainty of 9% in the pixel response.

Due to the low frequency of absolute calibration of the FD, relative calibration is performed to track the detector response on a nightly basis during operation. Fibre optics connect common light sources to one of three positions - A, B or C - in each telescope to illuminate different components of the optical system. Position A provides diffuse illumination of the camera face at a wavelength of 470 nm, Position B illuminates the mirror through a filter which approximates the wavelength acceptance of the telescopes, and Position C illuminates diffuse reflectors mounted on the inside of the telescope bay shutters, from which the light reflects and enters the telescope aperture. This system is used before and after each night of data collection to allow regular updates to the calibration constants of every pixel in the FD.

In addition to absolute calibration using the drum, an independent check is provided by a portable 337 nm nitrogen laser. This laser emits light of a known intensity and, with the laser positioned relative to the FD such that the effects of aerosol scattering are minimized, a known fraction of light is scattered to the telescope. Then, by comparison with the calibration constants calculated from the drum calibration, the response of the telescope can be cross-checked between the two absolute calibration methods.

Calibration of the FD allows the number of photons reaching the telescope aperture to be calculated but the effects of photon propagation from their point of emission in the EAS, through the atmosphere and to the camera must also be accounted for. As a result, an extensive campaign is in place to monitor atmospheric conditions at Auger South [43]. These measurements are made through various instruments including weather stations, regular balloon launches, infra-red cameras and lasers. Together, this monitoring network provides regular measurements of atmospheric conditions such as pressure, temperature and humidity, aerosol scattering properties, and cloud coverage of the site. All of these factors are incorporated into the reconstruction of FD events to reduce the effect of atmospheric fluctuations.

3.2.3.4 Performance

Reconstruction of events triggering the FD begins with the calculation of their geometry [23]. A view of one such shower as displayed by analysis software is shown in Figure 3.13. The SDP is determined first by fitting a plane to the line of triggered pixels in the camera. This fit can be performed with an accuracy of the order of 0.1°, depending on the observed track length. Following this, a fit is performed to the pulse times in each pixel to estimate the axis within the SDP along which the shower developed. While this may be achieved using FD data alone, the uncertainty in shower geometry when doing so can be quite large due to degeneracy in the fit parameters (see Section 1.2.2.2). However, the almost 100% duty cycle of the SD means that most events triggering the FD are observed as hybrids, in which case the shower geometry is reconstructed with excellent resolution (see Section 3.2.4). Laser shots from the

Figure 3.13: An example EAS seen by the FD, as displayed in analysis software. The left-hand image displays the pixels of one camera, with pixels triggered by the event appearing in colour. The gradation of colours represents the triggering times. To the right are shown the individual signal traces as a function of time from each of the four pixels marked by black dots. Clearly visible is the time ordering of the pulses in subsequent pixels. From [23].

Central Laser Facility (CLF), which is located near the centre of the SD, allow a comparison of reconstructed geometries with known ones for validation of these uncertainties.

Following the reconstruction of a shower's geometry, estimation of its energy can commence. The light flux at the telescope aperture needs to be converted to energy deposited in the atmosphere as a function of depth, which requires the calculation of contributions from both direct and scattered fluorescence and Čerenkov light. Following this, the profile of energy deposited as a function of slant depth can be reconstructed, and integration of this profile returns the shower energy. This process gives a statistical energy uncertainty of 10% for events in which the geometry is well-established. An overall systematic uncertainty on the energy scale of 22% is also present, with the factors contributing most to this value being the uncertainties on fluorescence yield (14%), the uncertainty arising from reconstruction of the energy deposit profile (10%), and the uncertainty in the absolute calibration of the FD (9%). Due to the SD energy scale being calibrated with respect to that of the FD, this 22% systematic uncertainty propagates through to the former. The fitting of the profile also returns the depth of shower maximum X_{max}, which is determined to an accuracy of around $20\,\mathrm{g\,cm^{-2}}$ [38].

The aperture of the FD, unlike that of the SD, does not saturate at high energies. The amount of light produced increases with shower energy, leading to a dependence on energy of the viewable distance. This distance is affected by background light and atmospheric conditions, but simulations using average values for these environmental factors give an aperture of about $900\,\mathrm{km^2\,sr}$ at $10^{17.5}\,\mathrm{eV}$, increasing to about $7400\,\mathrm{km^2\,sr}$ at $10^{19}\,\mathrm{eV}$. However, for the purposes of studies such as energy spectrum calculations, the hybrid exposure is of greater importance and consequently is calculated very accurately. This is discussed below.

3.2.4 Hybrid Operation

Hybrid operation of Auger South, while only available for the periods during which the FD operates, yields excellent reconstructions of shower properties. The hybrid data are the highest quality measured by the Pierre Auger Observatory and allow several studies requiring accurate and precise data, such as calculation of the CR energy spectrum with hybrid data [47] and composition studies [20].

Accurate hybrid reconstructions require the trigger times between the FD and SD to be well understood, so that any offsets can be corrected for. Two techniques are used to investigate this, and give knowledge of the timing offsets to within 100 ns [38, 49]. The first of these is to use 'golden hybrid' events, which are those that can be reconstructed either as hybrids or by the SD alone. The reconstructed core location is sensitive to time offsets between the SD and FD, allowing the identification of such an offset if systematic shifts are observed. In addition to this, the CLF is used to fire a vertical laser at the same time as a pulse from the same laser is sent to a nearby tank through an optical fibre. Comparison between the reconstructed laser firing times for the two systems can also be used to reveal any time offset present.

Following the collection of a hybrid trigger, as discussed in Section 3.2.3.2, reconstruction of a hybrid event occurs. For the Pierre Auger Observatory, this uses the FD data combined with the pulse time found in the SD station measuring the highest signal [47]. Shower geometry can be reconstructed extremely well for hybrid events, due to the longitudinal timing information from the FD being combined with a precise time of arrival at ground level. Core location and arrival direction have resolutions of 50 m and 0.6°, respectively. In a similar manner to monocular measurements, the CLF is used to validate these uncertainties using laser shots with known geometries [23].

Measurement of hybrid events with the FD means that the process of energy reconstruction, and the associated uncertainties, are the same as for standard FD operation. However, events which can be reconstructed either as hybrid or SD-only events allow the direct comparison between reconstructed energies of the two modes of operation, through which the SD energy scale is derived [50].

Knowledge of the hybrid exposure is important for studies such as the energy spectrum, and as a result it is carefully calculated [48]. Through the use of extensive Monte Carlo simulations which incorporate the time-dependent state of the entire detector, the aperture of Auger South is calculated taking into account status of detectors, weather effects and other relevant factors. The hybrid exposure is determined to within 10% at 10^{18} eV and 6% above 10^{19} eV.

3.2.5 Auger South Enhancements

Following the success of the baseline design of the Pierre Auger Observatory, several enhancements have been added to the detector. Of particular interest is the lowering of the minimum energy threshold of Auger South to better measure the properties of showers in the knee-to-

Figure 3.14: An example of the timing fit to a hybrid event. The red line represents the fit to monocular data, while the blue line indicates that of the hybrid fit. The coloured circles represent the FD data, while the squares represent that of the SD, with the filled square being the station with the highest signal. The obvious discrepancy between the two fits illustrates the improvement offered by hybrid analysis. From [23].

ankle transition region, an important region for the discrimination of astrophysical models [145]. Two projects - AMIGA and HEAT - have been devised to achieve this. In addition, an array of radio antennae has been added to the Pierre Auger Observatory to detect EAS at radio frequencies and investigate this form of emission. These enhancements are briefly described below.

3.2.5.1 AMIGA

As noted in Section 3.2.2, the main SD array becomes fully efficient at an energy of 3×10^{18} eV. AMIGA - Auger Muons and Infill for the Ground Array [145] - is designed to lower this threshold to 10^{17} eV by forming two sub-arrays of smaller spacing than the 1.5 km of the main array. These sub-arrays are constructed by adding densely-clustered detectors to a section of the already existing SD. The first array has a spacing of 750 m and will cover an area of 23.5 km^2 with 61 new detectors when completed. Full efficiency of this array will occur at an energy of $10^{17.5}$ eV. The second sub-array, which will be fully efficient at 10^{17} eV, will have a spacing of 433 m and cover an area of 5.9 km^2 using 24 new detectors. These small areas are viable due to the rapidly increasing flux of cosmic rays with decreasing energy.

Each station of AMIGA consists of two components: a water-Čerenkov tank of the same design as the rest of the SD, and a buried 30 m^2 scintillator muon detector. The latter is designed as a method of composition measurement complementary to the detection of the position of shower maximum by the FD. The scintillators will be buried at a depth of approximately 2.3 m. The water-Čerenkov tanks function in a similar way to the remainder of the SD, with the addition of the ability to provide a triggering signal to their associated scintillators

to initiate data recording. The output from the scintillators is in the form of a logical bit which indicates when the measured signal exceeds an adjustable threshold; the sampling for this output occurs at 320 MHz.

As of April 2011, the construction of the water-Čerenkov component of the 750 m array is nearing completion with 53 out of 61 detectors installed. Associated scintillators have been installed at three of those locations. Construction of the 433 m array is yet to begin. Figure 3.15a displays the layout of these installed components.

3.2.5.2 HEAT

HEAT (High Elevation Auger Telescopes) comprises a group of three fluorescence telescopes located 180 m north-east of the Coihueco site and oriented toward AMIGA, enabling hybrid measurements to be made at energies around 10^{17} eV [140, 145]. Its design is similar to that of the standard FD telescopes, except that it incorporates a tilting mechanism which can pivot the telescopes upward by 29°. When used in conjunction with the telescopes of Coihueco, elevations between 3° and 58° are viewable .

This expanded field of view is important for making unbiased measurements of the depth of shower maximum at $\approx 10^{17}$ eV; as showers at these energies develop high in the atmosphere and are also more likely to be detected close to a fluorescence telescope due to the low flux of light they produce, shower maximum is often outside the field of view of the standard FD telescopes. Consequently a measurement of the average depth of shower maximum will be biased toward deeper showers and particular geometries. By extending the field of view to higher elevations, an unbiased measurement can be made at low energies.

HEAT operates as an independent detector in the manner of the original FD sites and event data are merged with other detectors off-line. Stable operation of HEAT began in September 2009 and an example shower that was measured by both Coihueco and HEAT is shown in Figure 3.15b. In addition to extending the field of view of the FD at Auger South, the electronics of HEAT also serve as a prototype design for the telescopes of the FD at Auger North [140] (see Section 3.2.6).

3.2.5.3 AERA

Beginning in March 2011, the Auger Engineering Radio Array (AERA) has been recording data at Auger South [145]. AERA is the first component of what will be a 160 detector radio array covering about 20 km^2, located within the bounds of the AMIGA enhancement and within the field of view of HEAT. The goal of the full radio array is to measure the radio-frequency emission, the dominant source of which is the acceleration of charged particles in an EAS by the geomagnetic field, to allow detailed study of emission mechanisms. As of mid-2011, the 21 stations of AERA have been deployed on a triangular grid of 150 m spacing, with future detectors to be arranged on 250 m and 375 m grids. The energy threshold of AERA is approximately 10^{17} eV.

Each of the 21 stations consists of an antenna sensitive to frequencies between 27 MHz and 84 MHz, an electronics package, power source and wireless communications for the transmission of data to a central data acquisition system. This system forms triggers and initiates data readout from the radio array. Calibration of signals is autonomous, and a beacon located at Coihueco is used for adjusting time stamps at each station.

AERA has been constructed to prove the viability of the station design and prove the radio detection technique prior to the completion of the full array. In April 2011, the first radio/SD hybrid events were detected, with the first 'super-hybrid' - an event seen by the SD, FD and AERA - following in late April. Following the demonstrated success of the method and detector design, completion of the full radio array is expected within several years.

3.2.6 Auger North

With the bulk of construction finished for Auger South, planning is underway for its counterpart in the Northern Hemisphere [46]. Data collected at Auger South, particularly at trans-GZK energies, have hinted at several interesting properties of the UHECR flux including a cut-off consistent with the GZK effect, an apparent change in composition above 10^{18} eV, and a correlation with local AGN (see Chapter 2 for more details). However, the rate of trans-GZK particles detected at Auger South is insufficient to permit the rigorous study of such aspects as source identification and spectra measurement, particle composition, the measurement of cosmic magnetic fields, and high energy hadronic interaction cross-sections at these energies. As a result, the size of Auger North will be greatly expanded from Auger South to facilitate a high rate of data collection above ~ 60 EeV; a rate of about 180 events collected each year above this threshold is expected for the combined Pierre Auger Observatory [140].

A site in the South-East of Colorado, USA (38° N), has been selected for the location of Auger North. The proposed design for the detector is shown in Figure 3.16. The main ground array will cover approximately 20 000 km^2 with one of approximately 4 000 detectors installed on every second corner of a square 1 mile (1.6 km) grid, giving a detector spacing of 2.3 km. This is a natural choice of spacing due to the pre-existing road network in the area that defines a 1 mile grid, offering ease of tank deployment and maintenance. Full efficiency is expected for this array at 80 EeV. In addition, approximately 10% of the array will be infilled with 400 extra detectors occupying the empty corners of the 1 mile grid, achieving full efficiency at 10 EeV. The detectors will be of similar design to those of Auger South with several changes made: thermal insulation will be added to endure the winters of Colorado during which temperatures can reach $-35°$ C, one PMT will be used instead of three to reduce costs, and data will be passed between stations on the way to collection points due to the topography of the site not allowing reliable long-range wireless communication for transmission directly to the latter.

Monte Carlo simulations predict the core location resolution to be 130 m for the SD, with angular resolutions of 2.2°, 1.2° and less than 1° for zenith angles of 25°, 45° and 60°, respectively, for an energy of 50 EeV. A similar energy resolution to that of Auger South is expected,

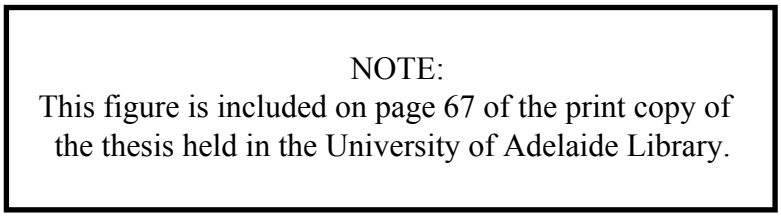


Figure 3.15: (a) The status of AMIGA in April 2011. Seen are the stations of the regular SD grid together with the infill stations. Also shown are the 3 deployed muon detectors out of the 7 that will complete the indicated hexagon and form the prototype muon detector. (b) An example shower that was detected by both HEAT and Coihueco, the pixels of which are shown in the top and bottom halves of the figure, respectively. The increase in observable track length offered by the presence of HEAT is clearly visible. (c) A polar diagram, with array zenith at centre, showing reconstructed arrival directions of radio/SD hybrid events. Crosses indicate the arrival directions of the SD reconstructions, while contours indicate the arrivals directions that triggered three or more stations of AERA. Figures are from [145].

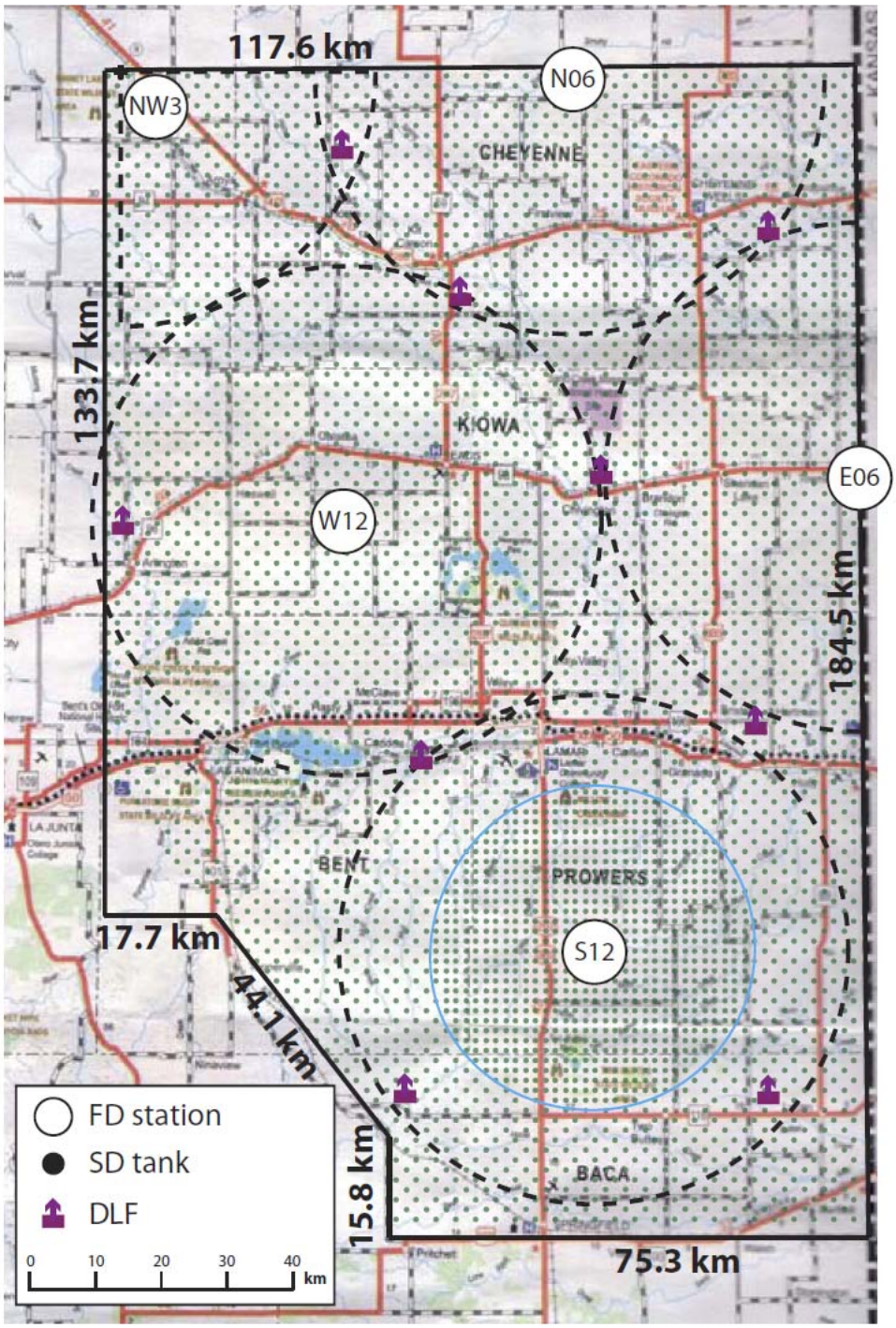


Figure 3.16: The proposed design of Auger North. The locations of the FD sites are indicated by circles, while each green dot indicates the position of an SD station. The 1 mile infill is centred on the S12 FD site. Also indicated are the positions of the distant laser facilities (DLF, see text). From [46].

however, the change in array spacing means the signal at 1 500 m from the core, $S(1500)$, will be used instead of $S(1000)$.

Fluorescence detection at Auger North will be accomplished with 39 telescopes spread amongst five sites which, together, will provide coverage of almost all of the SD, assuming a viewing distance of 40 km for the site. The optics of these telescopes will be similar to those of the Auger South FD, although an updated electronics package will be used which reflects the availability of components and improves performance. A variety of monitoring and calibration systems, similar to those of Auger South, will be employed at Auger North, including the permanent installation of UV laser facilities (DLFs, or distant laser facilities) to replace the somewhat unwieldy mobile laser used at Auger South.

The acceptance of the FD will be extremely large, with expected values above energies of 100 EeV of around 40 000 $\mathrm{km}^2\,\mathrm{sr}$ (varying slightly with season) [46]. While only a small fraction of events will be seen by multiple sites, above ~ 30 EeV the detector will have almost complete hybrid coverage [140].

Construction of Auger North is planned to begin in 2011 and be completed in 2016. The greatly increased collecting power provided by this detector will help shed light on many fundamental questions in cosmic ray research and provide important data for multi-messenger studies, in conjunction with current and future high energy astrophysical observatories.

Chapter 4

Cosmic Magnetic Fields and Cosmic Rays

Cosmic rays, being charged particles, are deflected by magnetic fields present both inside and outside the Milky Way galaxy as they propagate from their source to Earth. These deflections introduce significant challenges for the study of CR sources, as the arrival direction of a CR at Earth may be very different from the direction to its point of origin. Without knowledge of the intervening magnetic fields, the interpretation of CR arrival directions is not clear.

For this reason, it is highly desirable to understand both the strength and structure of cosmic magnetic fields. However, their measurement is a challenging task, leading to only a very general understanding of the properties of magnetic fields within the Milky Way, and little knowledge of those outside. In this chapter the propagation of CRs through magnetic fields is introduced, followed by a brief discussion of the methods by which magnetic fields are measured and a description of the Galactic magnetic field (GMF) models used in this work.

4.1 Cosmic Rays in Magnetic Fields

Fundamentally, the propagation of CRs is governed by the Lorentz force equation for a particle of atomic number Z having velocity vector $\boldsymbol{v}$ moving through electric and magnetic fields $\boldsymbol{E}$ and $\boldsymbol{B}$, respectively:

$$\boldsymbol{F} = Ze\left(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}\right) \tag{4.1}$$

where e is the elementary charge. Recognising that force is equal to the time derivative of momentum, $\frac{d\boldsymbol{p}}{dt}$, this becomes

$$\boldsymbol{F} = \frac{d\boldsymbol{p}}{dt} = Ze\left(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}\right) \tag{4.2}$$

Integration of Equation 4.2 will then define the particle's momentum at a given time. However, the first term can be neglected in the case of propagation through interstellar and

intergalactic space. The very high conductivity of these regions hinders their ability to maintain charge separation, meaning that electrostatic fields are negligible for the study of CR propagation [146, 147]. An alternative source of electric fields has been studied in [148], where the rotation of the galaxy (and consequently its magnetic fields) with respect to the reference frame in which the Solar System is at rest gives rise to electric fields in that frame. The resulting changes in energy of the CRs were found to be very small and do not need to be considered for this work. Finally, the magnetic fields in the models discussed in Section 4.2.2 are assumed to be unchanging, meaning no electric field results from time-varying magnetic fields.

Disregarding $\boldsymbol{E}$, then, reduces Equation 4.2 to:

$$\frac{d\boldsymbol{p}}{dt} = Ze\boldsymbol{v} \times \boldsymbol{B} \tag{4.3}$$

which, when integrated with respect to time, allows calculation of the trajectory of the CR. Chapter 5 discusses the numerical implementation of this in further detail.

Several basic features of CR propagation can be derived from these equations. Firstly, the gyroradius of a particle r_g - which is the radius of the arc along which a CR will move in a uniform magnetic field - is given by equating the Lorentz Force to the centripetal force acting on the particle:

$$|\boldsymbol{F}_L| = |\boldsymbol{F}_C| \tag{4.4}$$

$$\Rightarrow Ze\,|\boldsymbol{v} \times \boldsymbol{B}| = \frac{\gamma m\,|\boldsymbol{v}_\perp|^2}{r_g} \tag{4.5}$$

$$\Rightarrow Ze\,|\boldsymbol{v}_\perp|\,|\boldsymbol{B}| = \frac{\gamma m\,|\boldsymbol{v}_\perp|^2}{r_g} \tag{4.6}$$

for a CR having rest mass m and Lorentz factor γ and component of the velocity perpendicular to the magnetic field $\boldsymbol{v}_\perp$. Rearranging for r_g gives:

$$r_g = \frac{\gamma m\,|\boldsymbol{v}_\perp|}{Ze\,|\boldsymbol{B}|} \tag{4.7}$$

If the angle between $\boldsymbol{v}$ and $\boldsymbol{B}$ is δ, then $|\boldsymbol{v}_\perp| = |\boldsymbol{v}| \sin\delta$. This gives:

$$r_g = \frac{\gamma m\,|\boldsymbol{v}|}{Ze\,|\boldsymbol{B}|} \sin\delta \tag{4.8}$$

Recognising that $\gamma m\,|\boldsymbol{v}|$ is the magnitude of the CR's relativistic momentum, which can be approximated as $|\boldsymbol{p}| \approx \frac{E}{c}$ for very high energies where the rest mass can be neglected, one comes to:

$$r_g \approx \frac{E}{Zec\,|\boldsymbol{B}|} \sin\delta \tag{4.9}$$

which can be expressed in units useful for propagation through the Galaxy as:

$$r_g = 1.08 \left(\frac{E}{\mathrm{EeV}} \right) \frac{1}{Z \cdot (|\boldsymbol{B}| / \mu\mathrm{G})} \,\mathrm{kpc} \tag{4.10}$$

for a particle travelling in a direction perpendicular to the magnetic field. Units of μG are convenient for describing the GMF, where $1\,\mathrm{G} = 10^{-4}\,\mathrm{T}$.

From Equation 4.9 it can be seen that the gyroradius is directly proportional to the energy of the CR, while being inversely proportional to its charge and the strength of the magnetic field. The quantity $\frac{E}{Zec}$ is termed the 'rigidity' of a particle and is useful for the discussion of propagation of different species of CRs having different charges, as particles of equal rigidity will follow the same trajectory in a magnetic field.

At low rigidities, where the gyroradius of a CR is less than or comparable to the scale of irregularities in the magnetic field through which they are propagating, diffusive propagation occurs [149]. A recent review of cosmic ray diffusion may be found in [146]. During diffusive propagation, CRs are efficiently scattered by irregularities in the magnetic field and their velocities are quickly randomised. Consequently, their arrival directions will hold little information about their point of origin. For this reason, CRs at ultra high energies are considered for the present work, where rigidities are large enough that diffusive propagation is not expected within the GMF.

At these higher rigidities the deflections suffered are much smaller, and the 'rectilinear' regime of propagation is entered. This occurs when $r_g \gg L_c$, where L_c is the coherence length of the turbulent field. In this regime, the deflection of a CR as it propagates through the GMF is primarily due to the 'regular' magnetic fields which contain structure on kiloparsec scales. The turbulent field, which is random and exhibits variations on much smaller scales introduces deflections which are, in general, much smaller [149, 150].

The deflection of a CR after travelling a distance L perpendicular to a uniform magnetic field $\boldsymbol{B}$ can be expressed as:

$$\Delta\theta \approx 13^\circ \frac{40\,\mathrm{EeV}}{E/Z} \frac{L}{2\,\mathrm{kpc}} \frac{|\boldsymbol{B}|}{5\,\mu\mathrm{G}} \tag{4.11}$$

in the limit of small deflections. By comparison, the RMS deviation during rectilinear propagation of a CR traversing a purely turbulent field of coherence length L_c and RMS strength B_{rms} is [150]:

$$\delta_{rms} \approx 1.5^\circ \frac{40\,\mathrm{EeV}}{E/Z} \frac{B_{rms}}{5\,\mu\mathrm{G}} \sqrt{\frac{L}{2\,\mathrm{kpc}}} \sqrt{\frac{L_c}{50\,\mathrm{pc}}} \tag{4.12}$$

in the case of $L \gg L_c$, meaning the particle has traversed many 'cells' of the turbulent magnetic field within each of which the field is approximately uniform. It is expected that L_c will be of the order of tens of parsecs in the case of the GMF, and it can thus be seen that at multi-EeV energies the dominant deflections are those due to the regular field, with a small amount of

additional scattering resulting from the particle crossing many cells of the turbulent field that each impart a small deflection in a random direction.

4.2 Galactic Magnetic Fields

Knowledge of the strength and structure of cosmic magnetic fields, both Galactic and extragalactic, is highly desirable as it will allow a clearer understanding of many areas of cosmic ray astrophysics. For example, the interpretation of arrival directions relies heavily upon the assumed characteristics of the magnetic fields through which CRs have passed and their effect on those particles' trajectories. The measurement of cosmic magnetic fields is challenging, however, which leads to few constraints on their properties being currently available [151].

4.2.1 Measurement Methods

Cosmic magnetic fields may be measured in one of several ways. These methods include the observation of the polarisation of light from stars or molecular clouds, the Zeeman splitting of spectral lines, measurement of synchrotron emission from free electrons being accelerated in the magnetic fields, or measurements of the Faraday rotation experienced by the propagation of linearly-polarised electromagnetic radiation through free electrons [151, 152]. Of these methods, the latter two are the best-suited for studies of the large scale GMF and are, in practice, performed at radio frequencies.

Synchrotron radiation is produced by energetic electrons undergoing acceleration in the GMF. The high degree of linear polarisation of synchrotron radiation ($\approx 70\%$ in a uniform field) allows the strength and structure of the magnetic field in the plane of the sky to be estimated, if the density of the illuminating electrons is known [152]. However, depolarisation occurs due to changes in the magnetic field orientation and emission from background sources along the line of sight, complicating such measurements [153]. In addition, energetic electrons are required for the emission of synchrotron radiation at detectable frequencies, limiting measurements to regions where significant populations are present [154].

A complementary measurement is known as the rotation measure (RM), which may be used to estimate the integrated line-of-sight components of the GMF between Earth and a source. The sources are required to emit light with a large degree of linear polarisation and are often Galactic pulsars or extragalactic sources such as radio galaxies [155]. The latter are more numerous and allow finer spacing between measurements, while distances to the former can be estimated to make three-dimensional maps of the GMF [151]. The RM is given by the integral of the product of the number density of free electrons n_e and the line-of-sight component of the magnetic field $\boldsymbol{B}$ between the source and the observer:

$$\mathrm{RM} \approx 0.81 \int_0^d n_e(l) \boldsymbol{B}(s) \cdot d\boldsymbol{l} \,\mathrm{rad\ m^{-2}} \tag{4.13}$$

for n_e in cm^{-3}, $\boldsymbol{B}$ in μG and l in pc [156, 157]. It is derived from measuring the polarisation angle ψ of electromagnetic radiation at the telescope, which is related its polarisation angle at the source ψ_0 and wavelength λ by:

$$\psi - \psi_0 = \mathrm{RM}\lambda^2 \tag{4.14}$$

Consequently, measuring ψ from a single source over a range of wavelengths allows the estimation of the RM for that line-of-sight. This still does not yield the integrated line-of-sight magnetic field, however, as the RM is a measurement of the magnetic field weighted by the local electron density. Estimates of n_e are then required for extraction of the magnetic field. For example, the dispersion measure DM $= \int_0^d n_e dl$, which describes the spread in arrival times of different frequency components of a short pulse, may be obtained from the observation of pulsars. Division of the RM by the DM then allows an estimate of the magnetic field. However, difficulties in the interpretation of magnetic field estimates arise in cases where there is significant variation in $\boldsymbol{B}$ along the line of sight, or the distributions of n_e and $\boldsymbol{B}$ are not independent [153]. Such issues result in significant ambiguity in the interpretation of magnetic field measurements and make a consistent description of the GMF currently very difficult [151]. It is hoped that new radio telescopes such as the Low Frequency Array and the Square Kilometer Array will enable a greater understanding of magnetic fields in our galaxy in the coming years [154, 155].

4.2.2 Magnetic Field Models

Due to the complexity inherent in the interpretation of RM surveys and measurements of synchrotron radiation, the structure of the GMF is not currently well-understood. Summaries of recent measurements may be found in [151, 154, 155] and references therein.

As a result of such ambiguity, a variety of different models have been proposed to describe the large-scale features of the GMF, some examples of which may be found in [158, 159, 160] and [161]. Studies considering many of these proposed models have not found any which are fully compatible with observational data such as RM surveys and radio polarisation measurements, leading to difficulties in excluding or permitting particular models [162, 163, 164].

Consequently, while recognising that there are a number of different GMF models available in the literature, a small number of representative ones have been chosen for the work presented here. These models fall into three categories: spiral field models, which describe the field in the disk of the galaxy, halo models, which describe the field at larger distances from the galactic plane, and the turbulent field, which is present throughout. The first two are considered to comprise the regular component of the GMF model. The simulated Galactic volume is a sphere of radius 20 kpc from the Galactic Centre (GC).

Extragalactic magnetic fields are not included in this work. Many difficulties arise in the measurements of these fields and as a result very little is known about their strength and structure, apart from estimates of upper and lower limits which can vary by several orders

of magnitude [165]. As a result of such uncertainty their influence has not been considered, although future extensions of the work presented here could incorporate extragalactic magnetic fields as well as a larger variety of plausible GMF models.

4.2.2.1 Spiral Field

Two descriptions of the spiral field structure have been adopted. These models were adapted from those proposed in [159], which were modifications of those originally proposed in [158] in order to smooth out discontinuities of the field. The two models are differentiated by the behaviour of the magnetic field both between the spiral arms and above and below the mid-point of the Galactic plane.

Cylindrical polar co-ordinates are used to describe the spiral field models, giving $\boldsymbol{B} = \boldsymbol{B}(\rho, \phi, z)$. If the position of Earth in Cartesian co-ordinates is $(x, y, z) = (8.5, 0, 0)\,\mathrm{kpc}$ then the polar angle is given by $\phi = \tan^{-1}\left(\frac{-y}{x}\right)$. No component of the magnetic field perpendicular to the Galactic plane is present in the spiral field models. The models as described in [159] have radial (B_ρ) and azimuthal (B_ϕ) components given by:

$$B_\rho = B_{sp}(\rho, \phi) B_{A/S}(z) \sin p \tag{4.15}$$

$$B_\phi = B_{sp}(\rho, \phi) B_{A/S}(z) \cos p \tag{4.16}$$

for the relevant value of $B_{sp}(\rho, \phi)$ and $B_{A/S}(z)$ (where the subscript A or S is chosen as required) as described below. The angle p is known as the 'pitch angle', and indicates the deviation of the magnetic field lines from being purely azimuthal. It is taken to be $p = -10°$, meaning that the spiral field at Earth points towards $l = 80°$.

The first model is a bisymmetric spiral (BSS) in which the magnetic field orientation is reversed from one spiral arm to the other. The model is also taken to be symmetric across the Galactic plane (a BSS-S configuration) such that $\boldsymbol{B}_{BSS}(\rho, \phi, -z) = \boldsymbol{B}_{BSS}(\rho, \phi, z)$. The value of B_{sp} in this case is:

$$B_{sp}(\rho, \phi) = B_{BSS}(\rho, \phi) = B_{norm} \frac{8.5}{\rho} \tanh^3\left(\frac{\rho}{\rho_1}\right) \cos\left(\phi - \beta \ln\left(\frac{\rho}{\xi_0}\right)\right) \tag{4.17}$$

for an appropriate normalisation factor B_{norm} and other parameters as described below. The factor $B_S(z)$ introduces a z-dependence of B_{BSS-S} through:

$$B_S(z) = \left(\frac{1}{2\cosh(z/z_1)} + \frac{1}{2\cosh(z/z_2)}\right) \tag{4.18}$$

The second model is an asymmetric spiral (ASS) with asymmetry across the Galactic plane (ASS-A) such that $\boldsymbol{B}_{ASS-A}(\rho, \phi, -z) = -\boldsymbol{B}_{ASS-A}(\rho, \phi, z)$. The value of B_{sp} in this model is given by:

$$B_{sp}(\rho,\phi) = B_{ASS}(\rho,\phi) = B_{norm}\frac{8.5}{\rho}\tanh^3\left(\frac{\rho}{\rho_1}\right)\cos^2\left(\phi - \beta\ln\left(\frac{\rho}{\xi_0}\right)\right) \tag{4.19}$$

and the field reversal across the plane is introduced through the factor:

$$B_A(z) = B_S(z)\tanh(z/z_3) \tag{4.20}$$

where $z_3 = 0.02\,\text{kpc}$ is taken to give a smooth but rapid reversal in the magnetic field direction across the Galactic plane.

The other parameters describe the various features of the magnetic field structure and were not changed between models. Following [159], the parameters were taken as $\rho_1 = 2\,\text{kpc}$, $\beta = \frac{1}{\tan p} = -5.67$, $\xi_0 = 10.55\,\text{kpc}$, and z_1 and z_2, which describe two scale heights for the decay of the field away from the plane, given values of $0.3\,\text{kpc}$ and $4\,\text{kpc}$, respectively.

The spiral field models as described above do not obey $\boldsymbol{\nabla}\cdot\boldsymbol{B} = 0$. As a result, a minor correction was made to satisfy this condition, which was achieved by modifying the azimuthal component of the magnetic field vector such that the total field became divergence-free. The strength and structure of the spiral field were not significantly affected in this process. When this correction is made, Gaussian fits to the distributions of the ratio of the magnitude of the corrected magnetic field to that of the uncorrected magnetic field, for both the ASS and BSS fields, have means of 1 and standard deviations of less than 0.1%. In a similar manner, Gaussian fits to the distributions of the deviation of the pitch angle of the corrected field from the original value of $-10°$ are centred at 0 and have standard deviations of less than 0.1° for both spiral field models.

Finally, for each of the ASS and BSS fields, two variations were included by changing the overall normalisation of the fields. The value of B_{norm} was adjusted to give a value of B_{sp} of either $2\,\mu\text{G}$ or $5\,\mu\text{G}$ at the position of Earth, which is assumed to lie at a distance of $8.5\,\text{kpc}$ from the GC. Figure 4.1 displays the structure of each of these models.

4.2.2.2 Halo Fields

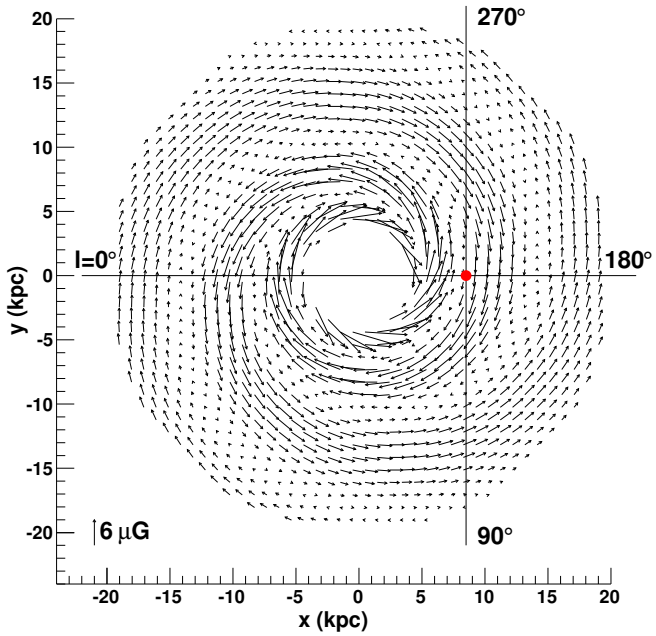
Halo fields describe components of the magnetic fields away from the Galactic plane in addition to the vertical extension of the spiral field models as described above. A two-component model was used consisting of a dipole field added to a toroidal field, similar to those described in [160]. Such a configuration has been suggested to explain observed features of the GMF [151].

The dipole field is given by the following in spherical polar co-ordinates:

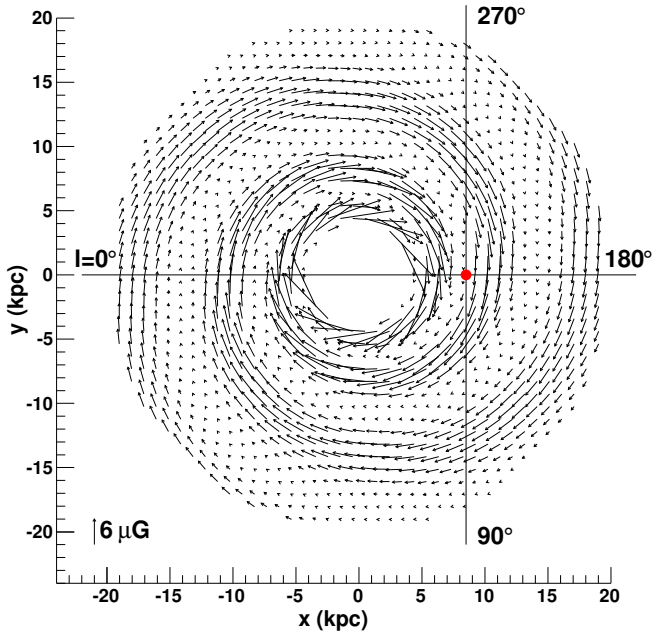
$$\boldsymbol{B}_D(\rho,\theta) = \frac{2D\cos\theta}{\rho^3}\hat{\boldsymbol{\rho}} + \frac{D\sin\theta}{\rho^3}\hat{\boldsymbol{\theta}} \tag{4.21}$$

where D is the dipole moment. The axis of symmetry is the same as the central axis of the spiral field and no azimuthal component is present in the dipole field.

The toroidal field is purely azimuthal, and is given in cylindrical polar co-ordinates by:



(a)



(b)

Figure 4.1: Plots showing the spiral field models used in this work, as described in the text, looking from the positive z-axis towards the negative z-axis. (a) displays the BSS-S model at $z = 0$ kpc, while (b) displays the ASS-A model at $z = 0.01$ kpc as the field vanishes at $z = 0$ kpc. In both cases the arrows indicate the direction and relative strength of the field at each point, while the overall normalisation is indicated on the plots. For clarity the magnetic field within a radius of 4 kpc of the GC is omitted. The position of Earth is marked by a red circle, and the orientation of Galactic longitude is shown.

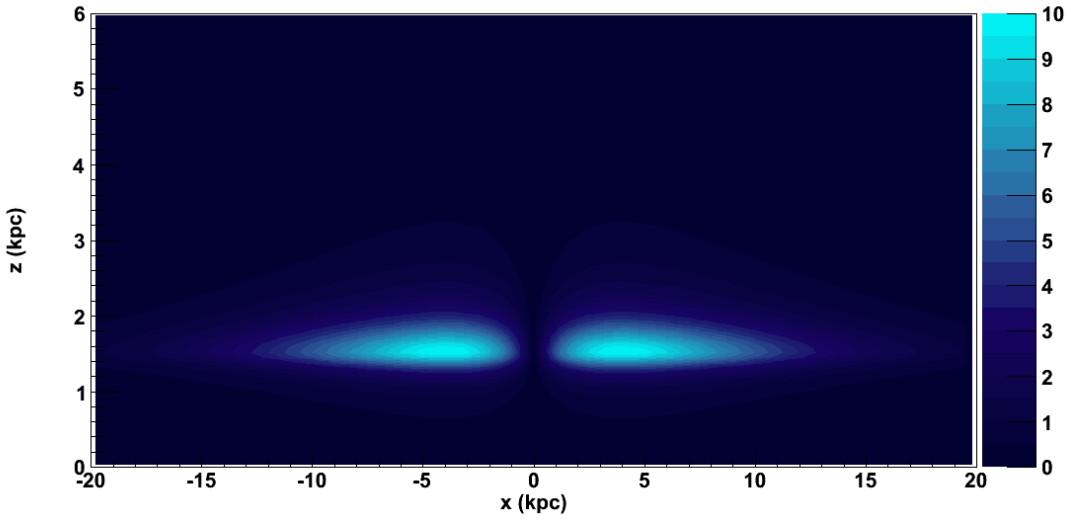


Figure 4.2: A contour plot of the strength of the toroidal component of the halo field, in the case of $B_T^0 = 10\,\mu$G. Only the values for $z > 0$ are shown; the strengths are identical for $z < 0$, with only the direction of the field reversed. The legend to the right indicates the strength of the field in μG.

$$\boldsymbol{B}_T(\rho, z \geq 0) = B_T^0 \frac{1}{1 + \left(\frac{|z| - z_0^T}{z_1^T}\right)^2} \frac{\rho}{\rho_0} \exp\left(-\frac{\rho - \rho_0}{\rho_0}\right) \hat{\boldsymbol{\phi}} \tag{4.22}$$

for $z_0^T = 1.5\,\mathrm{kpc}$, $z_1^T = 0.2\,\mathrm{kpc}$ for $|z| < z_0^T$ or $z_1^T = 0.4\,\mathrm{kpc}$ otherwise, and $\rho_0 = 4\,\mathrm{kpc}$. For $z < 0$, the direction of the field is reversed to give an antisymmetric halo which has been suggested in [166]. This choice of parameters follows [162]. Figure 4.2 illustrates the strength of the toroidal field in the $x - z$ plane of the Galaxy.

Two cases were considered, the first in which the halo field vanished ($D = B_T^0 = 0$). In the second case, $D \approx -123\,\mu\mathrm{G\ kpc}^3$, giving a vertical component of the magnetic field at Earth of $0.2\,\mu$G pointing towards Galactic north, and $B_T^0 = 10\,\mu$G. These values were chosen following [151] and [162]. However, the value of B_T^0 may be significantly smaller than $10\,\mu$G [167]. The two configurations chosen here are representative of the two differing cases of a very small halo field component or a very strong one.

4.2.2.3 Turbulent Fields

The third and final component of the GMF to be modelled is a turbulent field. The method of implementation of this field follows that of [168] and is described in further detail in Chapter 5. The scale sizes of this turbulence follow a power law given by a Kolmogorov spectrum, which has been reported to describe well the fluctuations of interstellar plasma density over a large range of scales [169]. As magnetic fields are 'frozen in' to astrophysical plasma due to its very high conductivity [2], the form of the turbulence in the magnetic field is expected to be related

to that of the density fluctuations of the plasma.

The maximum scale size of the magnetic turbulence was taken to be 100 pc, a value that is expected for the injection of irregular magnetic fields by supernovae. The RMS strength of the turbulence is equal to 1.5 times the normalisation of B_{sp} at the position of Earth (as described in Section 4.2.2.1) which is within the range of values expected for the relative strengths of the regular and turbulent components [150, 153]. The turbulent field is approximated to be homogeneous throughout the simulated volume of the Galaxy. Future work could incorporate different strengths and scale sizes for different regions of the galaxy, as such variations are consistent with observations [170].

Chapter 5

Numerical Simulations of Particle Propagation

The calculation of individual CR trajectories through models of the Galactic magnetic field may be accomplished with the numerical integration of Equation 4.3. Software was developed for the purposes of this work to perform these calculations with good accuracy while keeping the required computational time to a practical level, and also to incorporate the ability to adapt to a large range of particle rigidities and magnetic field models. The algorithms used in this software, and considerations that led to those choices of algorithms, are described in this chapter.

5.1 Numerical Integration

The complicated field structures of the models described in Chapter 4 mean that an analytical solution to Equation 4.3 is not generally possible, so numerical methods must be employed for the calculation of individual CR trajectories. This is achieved through an iterative procedure that calculates the CR particle's trajectory with a large but finite number of steps. A first-order method of integration was chosen here for its speed, as a large number of CRs are required to be propagated (see Chapter 7). In addition, as discussed in Chapter 4, the true form of the Galactic magnetic field is subject to great uncertainty, and the models used in this work are considered representative only. As a result, the increased accuracy provided by the use of more sophisticated, but slower, integration methods is not of critical importance for the current studies.

Consider a particle of energy E and charge Ze. At the i^{th} step of this particle's trajectory, we denote its position and momentum vectors as $\boldsymbol{r}_i$ and $\boldsymbol{p}_i$, respectively. For ultra-relativistic particles, the velocity is approximated as $\boldsymbol{v}_i \approx c\hat{\boldsymbol{v}}_i$, for the unit velocity vector $\hat{\boldsymbol{v}}_i$. At position $\boldsymbol{r}_i$, the magnetic field vector is $\boldsymbol{B}(\boldsymbol{r}_i) = \boldsymbol{B}_i$. The momentum vector, in the case of an infinitesimally small change in time dt, at the $(i+1)^{\text{th}}$ step is given by:

$$\boldsymbol{p}_{i+1} = \boldsymbol{p}_i + \frac{d\boldsymbol{p}_i}{dt} dt \tag{5.1}$$

For the purposes of the numerical integration, dt must be approximated as a finite step Δt_i, where the subscript is introduced to signify that this time step varies along the trajectory. Discussion of this adaptive step size may be found in Section 5.2. This then gives:

$$\frac{d\boldsymbol{p}_i}{dt} dt \approx \frac{d\boldsymbol{p}_i}{dt} \Delta t_i \tag{5.2}$$

leading to:

$$\begin{aligned} \boldsymbol{p}^*_{i+1} &= \boldsymbol{p}_i + \frac{d\boldsymbol{p}_i}{dt} \Delta t_i & (5.3) \\ &= \boldsymbol{p}_i + Ze\boldsymbol{v}_i \times \boldsymbol{B}_i \Delta t_i & (5.4) \end{aligned}$$

which follows from Equation 4.3. The asterisk indicates that the magnitude of the new momentum vector is too large due to the addition of the finite second term. As the force on the particle due to the magnetic field is perpendicular to the particle's velocity, no work is done and the energy (and, hence, the magnitude of the momentum vector) of the particle should not change. In the approximation of finite Δt_i, this energy conservation does not hold and consequently the particle momentum must be rescaled to avoid an effective increase in E. Consequently, renormalisation of $\boldsymbol{p}^*_{i+1}$ must take place to give $\boldsymbol{p}_{i+1}$, which has magnitude $\left|\boldsymbol{p}_{i+1}\right| = \frac{E}{c}$ (again, an approximation made for ultra-relativistic particles).

This renormalisation must be performed with some care to avoid deformation of the particle's trajectory. A simple rescaling of all three components of the momentum vector through:

$$\boldsymbol{p}_{i+1} = \frac{E}{c} \frac{\boldsymbol{p}^*_{i+1}}{\left|\boldsymbol{p}^*_{i+1}\right|} \tag{5.5}$$

modifies the component of the momentum vector parallel to the local magnetic field direction, $\boldsymbol{p}_{\parallel}$, which should not occur. The effect of such a rescaling can be clearly seen in Figure 5.1, which shows the trajectory of a particle spiralling through a uniform magnetic field calculated using Equation 5.5. The particle's motion parallel to the magnetic field direction is suppressed, and its trajectory transitions from being helical to circular in the limit of an infinite path length. Such an effect is of greatest importance in the cases of r_g being small compared to the local scale of irregularities of the magnetic field (such as during propagation within a regular field). In these circumstances, the direction of $\boldsymbol{p}_{\parallel}$ is approximately constant resulting in significant deformation of the particle's trajectory.

To avoid this effect, a rescaling scheme is used that preserves the component of the particle momentum parallel to the local magnetic field, and modifies only those components perpendicular to it. To do so, $\boldsymbol{p}^*_{i+1}$ (in the unprimed co-ordinate system) is found in the primed co-ordinate system through the basis transformation $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}}) \rightarrow \left(\hat{\boldsymbol{x}}'_i, \hat{\boldsymbol{y}}'_i, \hat{\boldsymbol{z}}'_i\right)$, where

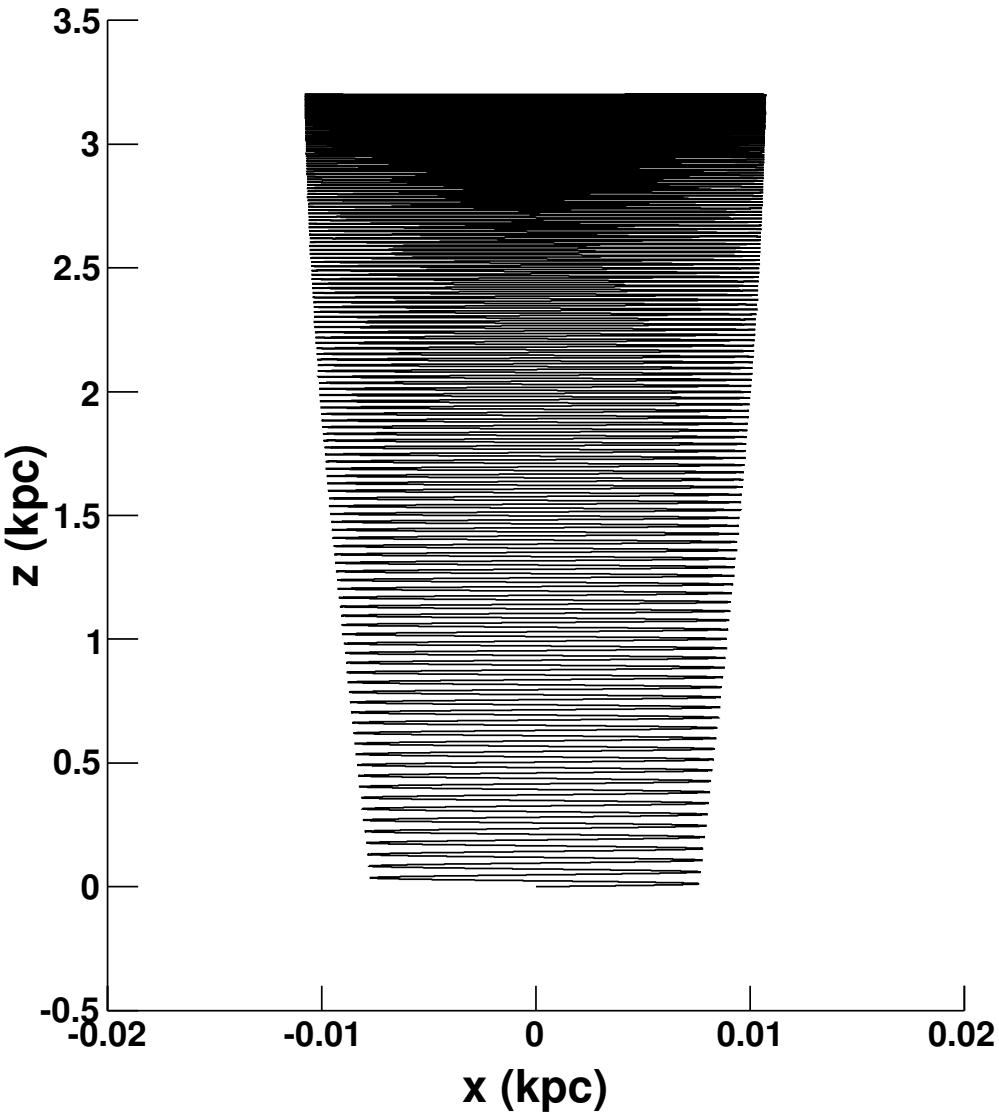


Figure 5.1: The trajectory of a 10^{16} eV proton in a uniform magnetic field, when its momentum is rescaled according to Equation 5.5, projected onto the $x-z$ plane. The magnetic field has strength $1\,\mu$G and is parallel to the z-axis. The transition from helical to almost circular motion can be clearly seen as the particle travels from the origin toward positive z.

the primed co-ordinate system is chosen to give $\boldsymbol{B}_i' = |\boldsymbol{B}_i|\,\hat{\boldsymbol{z}}_i'$. In this primed frame, the unscaled momentum vector is given by:

$$\boldsymbol{p}_{i+1}^{*'} = p_{(i+1),x}^{*'}\hat{\boldsymbol{x}}_i^{'} + p_{(i+1),y}^{*'}\hat{\boldsymbol{y}}_i^{'} + p_{(i+1),z}^{*'}\hat{\boldsymbol{z}}_i^{'} \tag{5.6}$$

With these definitions, the rescaled momentum vector is found by modifying its components through:

$$p_{(i+1),x}^{'} = kp_{(i+1),x}^{*'} \tag{5.7}$$
$$p_{(i+1),y}^{'} = kp_{(i+1),y}^{*'} \tag{5.8}$$
$$p_{(i+1),z}^{'} = p_{(i+1),z}^{*'} \tag{5.9}$$

for

$$k = \sqrt{\frac{\frac{E^2}{c^2} - \left(p_{(i+1),z}^{*'}\right)^2}{\left(p_{(i+1),x}^{*'}\right)^2 + \left(p_{(i+1),y}^{*'}\right)^2}} \tag{5.10}$$

and rotating back to the unprimed co-ordinate frame to give $\boldsymbol{p}_{i+1}$, which has magnitude $\frac{E}{c}$. Finally, the position of the particle is updated by stepping it in the direction of $\boldsymbol{p}_{i+1}$:

$$\boldsymbol{r}_{i+1} = \boldsymbol{r}_i + c\Delta t_i \hat{\boldsymbol{p}}_{i+1} \tag{5.11}$$

In this way, the particle trajectory can be followed through the magnetic field, which at each point is calculated as the sum of regular and turbulent components. The spiral and halo field models, as described in Chapter 4.2.2, form the regular component of the field $\boldsymbol{B}_{REG}(\boldsymbol{r}_i)$ to which the turbulent component $\delta\boldsymbol{B}(\boldsymbol{r}_i)$ - calculated as described in Section 5.3 - is added:

$$\boldsymbol{B}_i = \boldsymbol{B}_{REG}(\boldsymbol{r}_i) + \delta\boldsymbol{B}(\boldsymbol{r}_i) \tag{5.12}$$

Only propagation within a sphere of radius 20 kpc from the Galactic Centre was considered for the work presented here. At the CR energies considered in Chapter 7, the path lengths of the cosmic ray trajectories within this volume are of the order of $10 - 100$ kpc, which is much smaller than the mean free paths of the dominant energy loss mechanisms of pair production, pion photoproduction and photodisintegration (in the case of nuclei) [17, 64]. Consequently, energy losses were not incorporated into the software as they are considered negligible.

5.1.1 Cosmic Ray Back-tracking

The back-tracking of cosmic rays is a useful technique for the study of cosmic ray trajectories. From Equation 4.3 it can be seen that the force due to magnetic fields acting upon a CR of

charge Ze and velocity $\boldsymbol{v}$ at position $\boldsymbol{r}$ is the same as that acting upon a CR of charge $-Ze$ and velocity $-\boldsymbol{v}$ at the same point. As a result, the trajectory of a CR propagating from position $\boldsymbol{r}_0$ to $\boldsymbol{r}_1$, at which point it has velocity $\boldsymbol{v}_1$, will be the same as that of a 'back-tracked' particle, of opposite charge, that begins at $\boldsymbol{r}_1$ with velocity $-\boldsymbol{v}_1$.

It is important to note that back-tracking works only in the case of negligible or no stochastic energy loss processes. If the forward-tracked particle is subject to random interactions causing energy losses, then its trajectory will not be traced out by the back-tracked particle.

5.2 Step Size and Accuracy

The simulation of particle trajectories through complicated magnetic field models is very computationally intensive, and requires the careful choice of step size Δt_i in Equation 5.11. Very small values of Δt_i result in good accuracy in the determination of CR trajectories, but the length of time taken to perform the simulations may become prohibitive. On the other hand, increasing Δt_i results in faster simulations at the cost of accuracy. A compromise between these two extremes must be found.

For the simulations considered here, an adaptive step size is used which varies Δt_i along the particle trajectory. In a similar manner to [171, 172], the step size scales with the particle gyroradius at each point along the trajectory but is limited to a maximum value derived from the scale sizes present in the turbulent field.

To begin with, at each step in the particle trajectory a trial value of Δt_i is calculated through:

$$\Delta t_i = \frac{6 \times 10^9 \,\mathrm{s}}{Z} \frac{E}{\mathrm{EeV}} \left(\frac{|\boldsymbol{B}_i|}{\mu\mathrm{G}} \right)^{-1} \tag{5.13}$$

which corresponds to a step size of $\approx 60\,\mathrm{pc}$ for a $1\,\mathrm{EeV}$ proton in a $1\,\mu\mathrm{G}$ magnetic field and is approximately $r_g/20$. The accuracy of this step size was estimated by tracking a CR through circular motion within a uniform magnetic field and finding its distance from its origin after the simulated path length was equal to an integer number of revolutions. The relative error in the particle position - given by the distance from the origin divided by the total path length travelled - was found to be better than 0.1%.

Following the calculation of Δt_i according to Equation 5.13, its value is compared with $L_{min}/10$ (where L_{min} is the smallest scale size of the turbulent field - see below) and the minimum of the two values taken. This serves to avoid resonant effects that may be introduced through the accidental sampling of the turbulent field at spatial intervals comparable to the wavelength of a mode of turbulence.

5.3 Calculation of the Turbulent Field

The turbulent component of the magnetic field was generated using the method presented in [168], which is formed by summing, at each point in space, a large number of plane waves each having a random orientation and phase. Before discussing its implementation, however, it should be noted that a common alternative method of generating turbulence is through a fast-Fourier transform (FFT) [171, 173]. In this method, a three-dimensional lattice of wave number vectors $\boldsymbol{k}$ is defined, for each of which a corresponding magnetic field vector $\boldsymbol{B_k}$ is chosen with random phase and direction, and an amplitude defined by the desired power spectrum. Taking the FFT of this lattice gives the magnetic field in co-ordinate space defined on a regular lattice within a cubic cell of side length D. This cell is pre-computed, and the turbulent field throughout the simulation volume is constructed by repeating the same cell contiguously. Within a single cell, the magnetic field along the cosmic ray's trajectory may be calculated by interpolation between lattice points.

While the FFT method is computationally faster than the plane wave method, it is somewhat limited in the dynamic range of CR gyroradii with which it can be used. For a given cell size, the spacings of the grid points within it are limited by available computer memory. The FFT method is not suitable for the propagation of particles with gyroradii larger than D or smaller than the spacing between lattice points, as those particles will not fully sample all components of the turbulent field. This could introduce anomalies in the particle trajectories due to periodic effects (in the case of the former) or the absence of irregularities on the scale of r_g (in the case of the latter). Thus, the dynamic range of the FFT method is restricted for practical reasons unless different turbulent field configurations are used for different particle rigidities.

A large dynamic range was desired for the software described here, which led to the choice of the plane wave method. Following [168], then, the turbulent magnetic field is formed by taking the real component of a sum of N_m plane wave modes:

$$\delta\boldsymbol{B}(x,y,z) = \sum_{n=1}^{N_m} A(k_n)\hat{\boldsymbol{\xi}}_{\boldsymbol{n}} e^{i(k_n z'_n + \beta_n)} \tag{5.14}$$

where $\hat{\boldsymbol{\xi}}_{\boldsymbol{n}}$ is given by:

$$\hat{\boldsymbol{\xi}}_{\boldsymbol{n}} = \cos\alpha_n \hat{\boldsymbol{x}}'_{\boldsymbol{n}} + i\sin\alpha_n \hat{\boldsymbol{y}}'_{\boldsymbol{n}} \tag{5.15}$$

which ensures $\boldsymbol{\nabla}\cdot\delta\boldsymbol{B} = 0$. The primed Cartesian co-ordinates are related to the unprimed ones by:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\theta_n\cos\phi_n & \cos\theta_n\sin\phi_n & -\sin\theta_n \\ -\sin\phi_n & \cos\phi_n & 0 \\ \sin\theta_n\cos\phi_n & \sin\theta_n\sin\phi_n & \cos\theta_n \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{5.16}$$

where α_n, β_n, θ_n and ϕ_n are randomly chosen and determine the orientation and phase of the

n^{th} plane wave. Finally, the amplitude $A(k_n)$ of a mode n, having wavenumber k_n (related to the scale size L_n of the n^{th} mode through $k_n = 2\pi/L_n$), is defined through:

$$A^2(k_n) = \sigma^2 G\left(k_n\right) \left[\sum_{n=1}^{N_m} G\left(k_n\right)\right]^{-1} \tag{5.17}$$

with $G(k_n)$ calculated by:

$$G(k_n) = \frac{\Delta V_n}{1 + (k_n L_c)^{\gamma}} \tag{5.18}$$

where L_c is the correlation length of the turbulent field and ΔV_n and γ are defined below. The value of σ^2 defines the variance of the turbulent field amplitude such that $\sqrt{\left\langle |\delta \boldsymbol{B}|^2 \right\rangle} = \sigma$. Each of the three field components is distributed according to a Gaussian distribution with mean 0 and standard deviation $\frac{1}{\sqrt{3}}\sigma$. It should be noted, however, that the variance of the magnitude of the real component of the turbulent field is given by $\frac{1}{2}\sigma^2$, with the variance of the imaginary component also equal to this. As only the real component of $\delta \boldsymbol{B}$ is considered, this must be taken into consideration when defining the desired strength of the turbulent field. Hereafter, the variance of the real component of the turbulent field is denoted as σ_r^2; the value of σ^2 is larger by a factor of two.

For three-dimensional turbulence as considered here, $\Delta V_n = 4\pi k_n^2 \Delta k_n$ is the volume element corresponding to Δk_n. The spacing Δk_n is the distance between consecutive wave numbers which, following [168], is taken to be logarithmic to give a constant relative spacing $\Delta k_n / k_n$ between modes. The value of γ defines the spectrum according to which the amplitude of each mode is given.

For the simulations of propagation within the Galactic magnetic field discussed in Chapter 7, a Kolmogorov spectrum is chosen for the turbulent field which requires $\gamma = \frac{11}{3}$. The maximum scale size is taken as $L_{max} = 0.1\,\text{kpc}$, which is representative of the expected scale of magnetic irregularities introduced by supernova remnants [150]. The correlation length L_c is set to be 20 pc, consistent with the expectation for Kolmogorov-like turbulence where $L_c \approx L_{max}/5$ [150, 174].

The choice of N_m and L_{min} must be made according to practical considerations, the latter of which is discussed below in Section 5.3.1. The summation of a large number of plane waves at every point in a particle's trajectory is time consuming, and so N_m should not be too large. A value of 100 modes per decade of wave number was found by [174] to give comparable results to larger values while keeping the required computational time to an acceptable level, and was adopted in these simulations.

5.3.1 Choice of Minimum Turbulent Scale Size

While the choice of the minimum wave number to use when modelling the turbulent magnetic field of the Galaxy is physically-motivated, the choice of the maximum wave number (corre-

sponding to minimum scale size L_{min}) is subject to practical considerations. The scales at which magnetic field irregularities are present in the Galaxy may be inferred from the measured scales of density fluctuations in interstellar plasma, as the two quantities are expected to be coupled [175]. Such fluctuations in the plasma density have been observed to extend to scales as small as $\sim 10^6$ m [169], far smaller than is practical to simulate. Consequently, it is necessary to choose an appropriate value of L_{min} for the simulation under consideration to achieve accurate results while keeping the required computational time to a practical level. The relationship between the rigidity of the particles being simulated and the effect of varying L_{min} is studied in this section.

Despite the study of cosmic ray diffusion not being the focus of this work, a brief study of CR transport in purely-turbulent fields enables a suitable choice of L_{min}. As cosmic rays propagate through turbulent magnetic fields, they are expected to be efficiently scattered by irregularities in the field that are of a scale comparable to the CR's gyroradius [149]. Irregularities on scales $L \gg r_g$ and $L \ll r_g$ are not expected to cause significant scattering. Consequently, an unsuitable choice of scale sizes can significantly alter the mode of propagation of simulated particles. The importance of this choice is illustrated in Figure 5.2 where a $10^{15.5}$ eV proton has been propagated through two purely-turbulent magnetic fields which both have values of $\sigma_r = 10\,\mathrm{nG}$ and $L_{max} = 1\,\mathrm{Mpc}$, but for which L_{min} is taken as either 100 kpc or 0.01 kpc. In a field of this magnitude, $r_g \approx 0.3\,\mathrm{kpc}$, although this value will obviously vary due to the turbulent nature of the magnetic field. The lack of scattering in the case of $L_{min} \gg r_g$ can be clearly seen when compared to the case of $L_{min} \ll r_g$, reflecting the importance of choosing an appropriate value of L_{min}.

To find a suitable value of L_{min}, cosmic rays were propagated through several configurations of a purely-turbulent field of equal σ. However, in each successive configuration the value of L_{min} was decreased by a decade. Following [174, 176], the average 'instantaneous effective diffusion coefficient' (IEDC) - defined as $\frac{\Delta r^2}{6\Delta t}$ for Δr the distance from the particle's point of origin and Δt the time since propagation began - was calculated for an ensemble of particles propagating through each turbulent field configuration. This allowed the value of L_{min} to be identified past which the addition of smaller scales would have a negligible effect upon the IEDC. To enable comparison with the results of [174, 176], the values of $L_{max} = 1\,\mathrm{Mpc}$ and $\sigma_r = 10\,\mathrm{nG}$ as used in those studies were also adopted here.

For each of the turbulent field configurations a number of particles was injected at a random position with a randomly-chosen velocity vector. Changing the origin of each particle in this way avoids any systematic effects that could arise from the local properties of the turbulent field near a common origin. These particles were subsequently tracked through the magnetic field until $\Delta t = 10^3\,\mathrm{Myr}$.

The results of these simulations for particles of energy $10^{16}\,\mathrm{eV}$ and $10^{17}\,\mathrm{eV}$ are shown in Figure 5.3, at which energies $r_g \approx 1\,\mathrm{kpc}$ and $10\,\mathrm{kpc}$, respectively. While $L_{min} > r_g$, varying its value changes the results quite dramatically as L_{min} approaches the value of r_g. However,

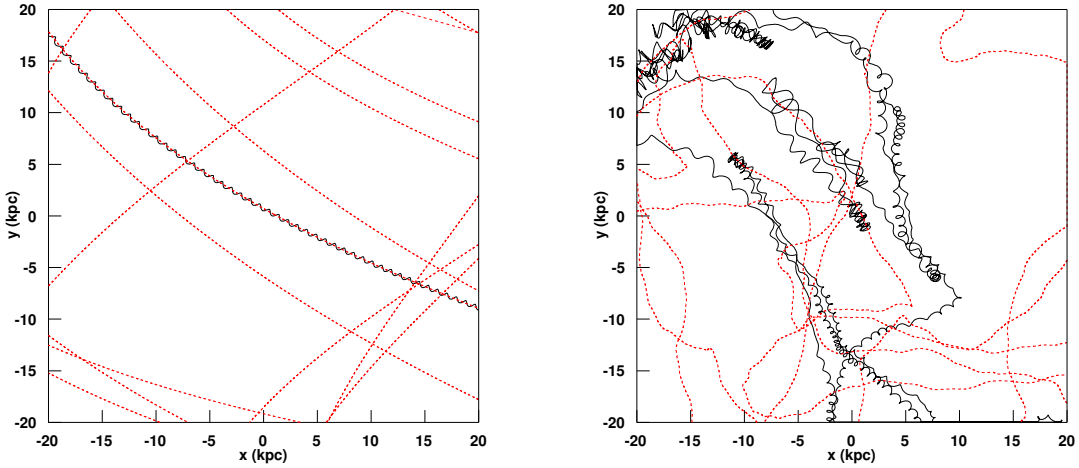


Figure 5.2: Examples of a $10^{15.5}$ eV proton in two different turbulent magnetic field configurations (see text for details). Sub-figure (a) illustrates the case of the smallest turbulent scale $L_{min} = 100\,\mathrm{kpc}$ being much larger than r_g while (b) illustrates the case of $L_{min} = 0.01\,\mathrm{kpc} \ll r_g$. In both figures, the particle trajectory is indicated by the black line, while sample magnetic field lines are indicated by dashed red lines. The lack of scattering can be clearly seen in (a) when compared to (b). All lines are projected onto the x-y plane.

once L_{min} becomes comparable with r_g, decreasing it further does not significantly alter the calculated IEDC curve, and the results from doing so begin to converge. The fluctuations in each curve result from a sample size limited to ≈ 150 particle trajectories, although this number is sufficient for the purposes of defining L_{min}. Following these results, taking L_{min} to be a factor of 10 lower than the gyroradius of the lowest-energy CR being propagated is considered a suitable compromise between computational speed and accuracy.

From these results, a value of L_{min} may be chosen for propagation within the Galaxy. The Galactic magnetic field models used reach maximum values of the order $10\,\mu$G, while the lowest-rigidity CRs considered are iron nuclei of energy 55 EeV - leading to a minimum value of r_g of the order 0.1 kpc. Thus, a value of $L_{min} = 0.01\,\mathrm{kpc}$ is adopted for the simulations discussed in Chapter 7.

As an additional cross-check of these results, the IEDC curves for a larger range of energies - now taking L_{min} to be 0.1 kpc - are plotted in Figure 5.4, along with the results of [176]. The value of L_{min} used in the simulations of the latter is not quoted, but the results appear to agree well. Again, the fluctuations in the curves produced here are due to low statistics.

The general shape of these curves can be understood as the transition from rectilinear motion to diffusive motion, which occurs on a time scale that is dependent upon particle rigidity. Immediately after a particle is injected into the turbulent field, it begins propagating

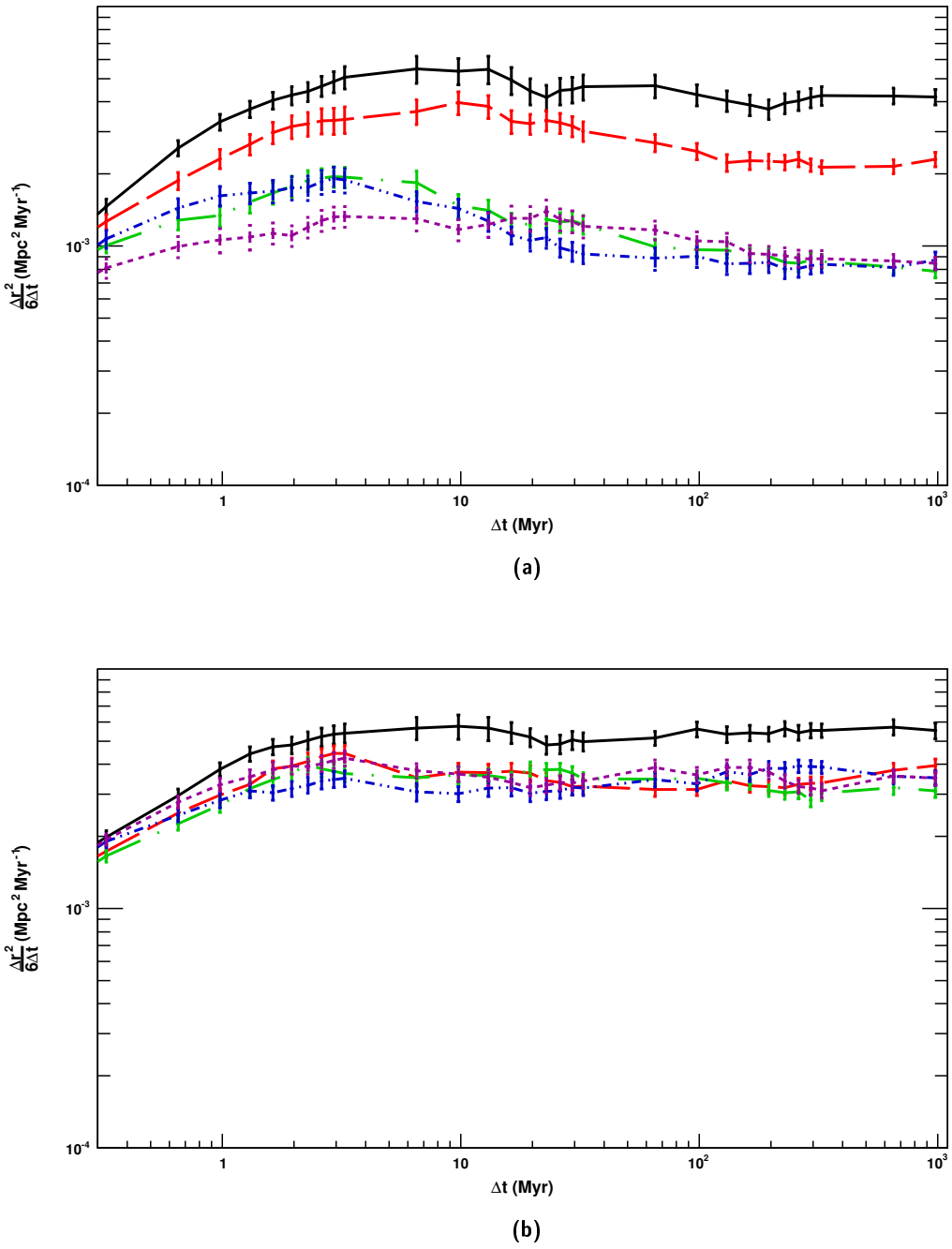


Figure 5.3: The instantaneous effective diffusion coefficient calculated for (a) 10^{16} eV protons and (b) 10^{17} eV protons in a turbulent field of $\sigma_r = 10$ nG and maximum scale size of $L_{max} = 1$ Mpc. Five configurations have been simulated in which the minimum scale size is successively decreased. They are: $L_{min} = 100$ kpc (solid black line), 10 kpc (long-dashed red line), 1 kpc (long-dash-dot green line), 0.1 kpc (short dash-double-dot blue line) and 0.01 kpc (short dash purple line). Note that the error bars on each point are the one standard deviation spread of values. As the points on a single curve are not independent of the preceding ones, these error bars do not represent the uncertainty of each value.

in the rectilinear regime where $\Delta r \sim \Delta t$. This gives a slope of 1 when the IEDC is plotted on logarithmic axes as in Figure 5.4.

Subsequent deflections due to the turbulent magnetic field serve to randomise the velocity vectors of the particles, inducing a transition to diffusive motion. During diffusive propagation, the average value of Δr for an ensemble of particles of a given rigidity is proportional to $\sqrt{\Delta t}$. This results in a slope of zero when the IEDC is plotted on logarithmic axes, and the value of $\frac{\Delta r^2}{6\Delta t}$ at which each curve becomes flat defines the diffusion coefficient for particles of the relevant rigidity (again, see Figure 5.4).

In the case of particles with gyroradii much larger than L_c, each crossing of a coherent cell introduces only a small deflection in a random direction. Thus, many such crossings are required before the directions of the velocity vectors of particles with $r_g \gg L_c$ are no longer correlated with their original velocities. This is in contrast to the cases of $r_g \approx L_c$ and $r_g \ll L_c$, where large deflections are experienced upon crossing a correlated cell (and even within a single cell for the latter), causing rapid decorrelation of the particle velocities, and hence a quick transition to diffusive motion. Quantitatively, for the energy-dependent diffusion coefficient $D(E)$ (for simplicity, assume a constant particle charge) defined through the relationship $\Delta r = \sqrt{6D(E)\Delta t}$ for diffusive propagation, the time scale for the transition to diffusive propagation can be expressed as [176]:

$$\tau_{diff}(E) = \frac{6D(E)}{c^2} \tag{5.19}$$

In [174], the energy-dependence of D was found to be well-described by:

$$D(E) = D_* \left[\left(\frac{E}{E_*} \right)^{\frac{1}{3}} + \left(\frac{E}{E_*} \right) + \left(\frac{E}{E_*} \right)^2 \right] \tag{5.20}$$

where the value of E_* is defined through $r_g(E_*) = L_c/5$ and $D_* = \frac{1}{4} c r_g(E_*)$ (again, assuming a constant particle charge). The first term in Equation 5.20 describes the low-energy dependence of the diffusion coefficient, where the scattering of particles is efficient, the second term describes Bohm-like diffusion where the mean free path of scattering (which is the length over which the particle velocity becomes decorrelated) is equal to r_g, and the third term describes the high-energy dependence where r_g is much greater than L_c and scattering is inefficient. Using Equation 5.20 to estimate τ_{diff} for 55 EeV protons in the Galaxy ($L_c = 20\,\mathrm{pc}$, $\sigma_r \approx 5\,\mu\mathrm{G}$) gives a diffusion length of $c\tau_{diff} \approx 50\,\mathrm{Mpc}$. This value is much larger than the path lengths of these protons within the GMF, so it may be clearly seen that, at the particle energies considered in Chapter 7, propagation in the rectilinear regime is expected.

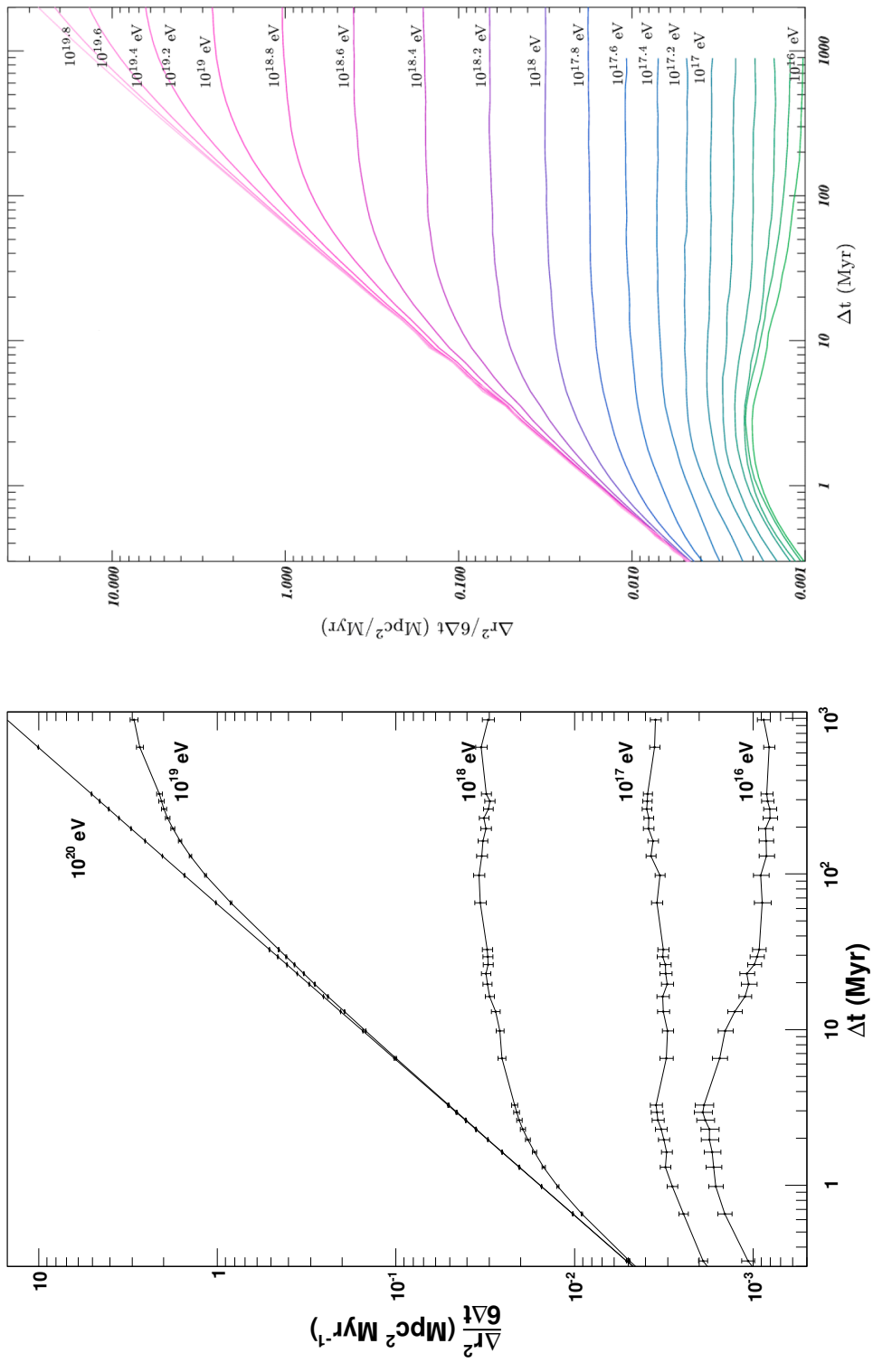


Figure 5.4: (a) The IEDC calculated for a range of energies in a purely turbulent field, with $L_{max} = 1\,\mathrm{Mpc}$, $L_{min} = 0.1\,\mathrm{kpc}$ and $\sigma_r = 10\,\mathrm{nG}$. Marked on each of the curves is the particle energy it corresponds to. The error bars are calculated in the same way as for Figure 5.3. (b) For comparison, the results published in [176]. Note the slight differences between the axes of the two plots.

Chapter 6

Energy Ordering in UHECR Arrival Directions

The unambiguous identification of cosmic ray sources is a challenging task, and, to date, has not been achieved. Many factors contribute to its high difficulty. As discussed in Chapter 4, the charged nature of cosmic rays means that they suffer deflections as they propagate from their source to Earth through cosmic magnetic fields about which little is currently known. Consequently, the relationship between a given event's arrival direction and the direction to its source is difficult to determine.

To mitigate this effect, one may study the arrival directions at very high energies. Above approximately 40 EeV the GZK effect is expected to limit the source horizon to about 100 Mpc, thus serving to restrict the path lengths of those CRs in extragalactic magnetic fields and reducing its influence on their trajectories. In addition, the deflections suffered by such energetic CRs as they propagate through the GMF may be small enough that their arrival directions retain information about their point of origin. However, this picture is complicated if the composition of the cosmic ray flux is energy-dependent. For example, in an energy range where the CR composition transitions from being light to heavy, the average rigidity may deviate from being proportional to primary energy, with the dependence relying on the details of such a transition. Currently, the composition of the cosmic ray flux has not been measured at trans-GZK energies, and below this threshold recent measurements are subject to some ambiguity (see Chapter 2).

Historically, this problem has been compounded by the extremely small flux of events at trans-GZK energies, meaning that the gathering of sufficient statistics to identify individual cosmic ray sources requires detectors with extremely large apertures, of the order of several thousands of $\mathrm{km}^2\,\mathrm{sr}$. It is only recently that detectors such as the Pierre Auger Observatory and Telescope Array that possess such statistical gathering power have been constructed. As a result, directional cosmic ray astronomy - including the identification of individual cosmic ray sources - may be possible in the near future.

This chapter discusses the use of 'energy ordering' in cosmic ray arrival directions for the purposes of identification of the directions to the sources of those events. Section 6.1 discusses the expectations of energy ordering as well as several methods which have been proposed to search for evidence of it within UHECR data sets. Following this, Sections 6.2-6.4 detail the analysis method that has been developed as part of this work and its application to data.

6.1 UHECR Energy Ordering and Analyses

The search for energy ordering in UHECR data sets is motivated by the ability to detect unprecedented numbers of high energy events with the current generation of UHECR detectors. At energies around 40 EeV, the gyroradii of protons in the GMF may become comparable with the radius of the Galaxy. As a result, the deflections suffered by many observed protons at these energies are expected to be small. In addition, the regular component of the GMF is currently understood to vary on kiloparsec scales. This means that, if the deflections due to the extragalactic magnetic field are not too large, extragalactic protons of a common source will traverse similar regions of the GMF. This introduces coherent deflections due to the regular field, with a small amount of scattering introduced by the turbulent field (see Chapter 4). In this case, the arrival directions of those protons at Earth will be deflected from the direction to their source by an angle that is approximately proportional to the inverse of their energy. An example of several sets of arrival directions after propagation through the ASS-A model of the GMF is shown in Figure 6.1, in which such structure can be clearly seen.

It is important to note that, in the case of heavy nuclei or very strong magnetic fields, such an approximation may no longer hold and the pattern of arrival directions upon reaching Earth becomes more complicated with multiple images of the source appearing [150]. In these circumstances, the presence of energy ordering is not expected. For this reason, the work presented here focuses on proton sources, although the influence of mixed composition is briefly investigated in Chapter 7.

Several types of analyses have been proposed to identify such energy ordering within UHECR data sets. Although the details of each method vary, the basic aim is to identify groups of events which have angular deviations from a particular point in the sky that are inversely proportional to the particle energy. If found, that point would be presumed to be approximately the direction to the source of those events.

A method of 'energy-energy correlations' (EEC), presented in [177], has been developed for the analysis of cosmic ray data above energies of 5 EeV. This method defines one or more 'regions of interest' - areas on the sky subtending a cone of opening angle $\approx 11.5^{\circ}$ - through an iterative procedure that begins by centering on events of energy greater than 60 EeV. This region is further refined by weighting each event within it by the event's energy and recalculating the centre of mass, a process that is performed three times in total. The EEC is defined as:

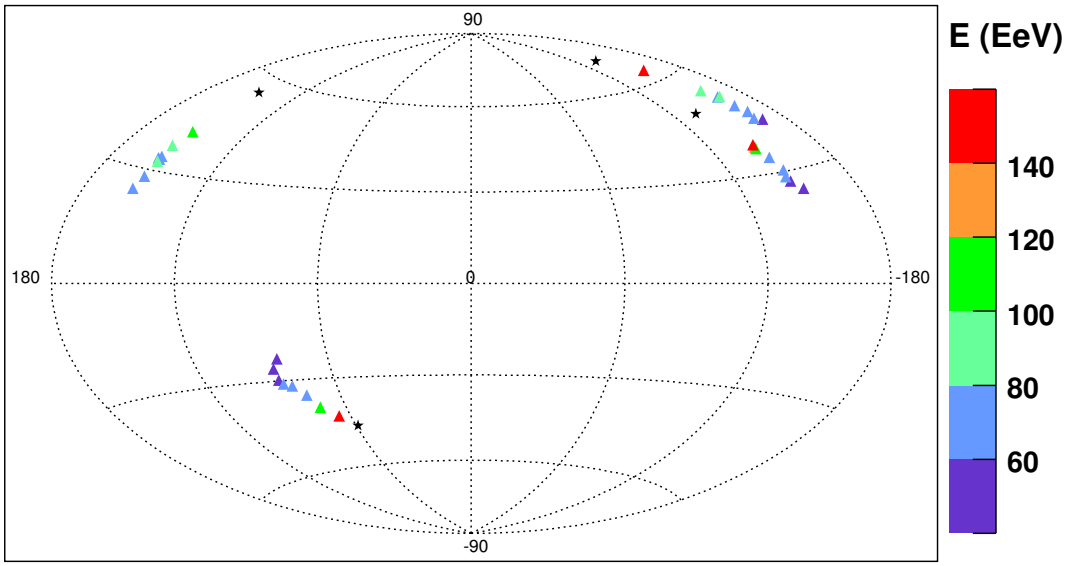


Figure 6.1: A skymap, in Galactic co-ordinates, of four sets of arrival directions of protons at Earth after having propagated through an ASS-A model of the GMF. The black stars indicate each of the four source directions. The energies of the events are indicated by coloured triangles, with the energy decreasing from red ($\approx 150\,\mathrm{EeV}$) to violet ($\approx 50\,\mathrm{EeV}$). The approximately linear structure, and energy ordering, in the arrival directions may be clearly seen.

$$\Omega_{ij} = \frac{(E_i(\alpha_i) - \langle E_i(\alpha_i)\rangle) \cdot (E_j(\alpha_j) - \langle E_j(\alpha_j)\rangle)}{E_i(\alpha_i) \cdot E_j(\alpha_j)} \tag{6.1}$$

where E_i is the energy of the i^{th} event, α_i is the angle between the i^{th} event and the centre of the region of interest and $\langle E_i(\alpha_i)\rangle$ is the average energy of events within the ring defined by α_i (see Figure 6.2). The indices i and j are taken to scan over all possible pairs of events within the region. Thus, data containing isotropic background plus signal data (which have experienced coherent deflections through the GMF) will be expected to produce a larger value of Ω_{ij} than data containing only the former. The distribution of average Ω_{ij} values (denoted simply as Ω) as a function of opening angle α provides information on the angular scale of any correlation present.

Although the EEC method is sensitive to energy-dependent deflections, it does not assume a particular structure to be present within the arrival directions of the events; only an energy-dependence of the angular deviation from the source is required. Applying the method to Monte Carlo data sets, which were produced by generating arrival directions from 10 sources and accounting for random deflections by extragalactic magnetic fields and coherent deflections by the GMF (chosen to be a uniform field), allowed an estimate of the EEC method's ability to reconstruct source positions and discriminate between the parameters of magnetic field strength and the number of sources in the sky. The angular resolution, given by the 68%

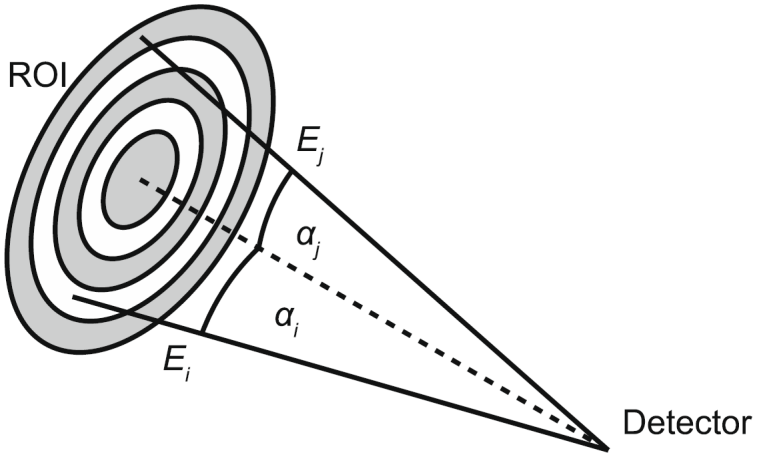


Figure 6.2: A diagram illustrating the quantities used in the calculation of the energy-energy correlation, as described in the text. From [177].

quantile of the angular separation between a source and the nearest region of interest, was found to be $\approx 7^\circ$ through such studies.

Through the comparison of Ω distributions - calculated from an ensemble of data sets having various numbers of sources and field strengths - with that of the signal data, the EEC method was shown to constrain the possible parameter space to a small region near the true parameters. The choice of a uniform GMF model was made for its ability to qualitatively reproduce the effects of energy ordering, although the performance of this method is not clear if more complicated field models are assumed.

The application of the EEC method to data recorded by the Pierre Auger Observatory was reported in [82]. The distribution of Ω gained from doing so was found to be compatible with that expected from isotropic arrival directions, leading to the conclusion that no evidence of energy ordering within the Pierre Auger Observatory data above 5 EeV was found using energy-energy correlations.

HiRes data were analysed for evidence of arc structures in the arrival directions of events above 10^{19} eV using Hough transforms in [103]. This method divided the sky into bins of longitude and latitude. For each event in the data set, a great circle was constructed that was perpendicular to that event's arrival direction. This was repeated for all events in the data set, and finally the value of each bin was set to be equal to the number of great circles which passed through it. If the arrival directions of a subset of events lay on an arc, their corresponding great circles would all intersect within the bounds of two bins (see Figure 6.3), leading to higher values for those bins compared to the rest of the sky. The technique did not require the existence of energy ordering within such an arc, although its presence could be determined if an arc was found.

This analysis was applied to the real HiRes data as well as 2×10^4 mock data sets, produced according to the detector exposure to an isotropic flux. Following this analysis, the significance

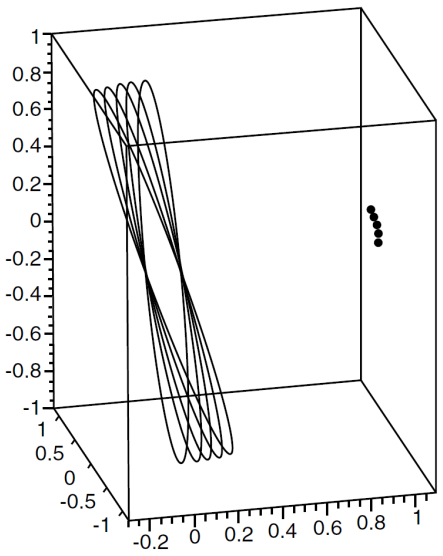


Figure 6.3: This diagram illustrates the use of a Hough transform to search for cosmic ray arrival directions forming an arc, as performed on HiRes data. The arrival directions of the cosmic rays are indicated by black dots, and the great circles perpendicular to those arrival directions shown to the left. As can be seen here, if the arrival directions form an arc, the intersections of their respective great circles all lie at the same two points. This leads to a large value of the bins to which the points of intersections belong (see text). From [103].

of each bin from the analysis of the real data was calculated through finding the fraction of mock data sets in which the corresponding bin had a value equal to or higher than that of the real data. Subsequent analysis of the mock data sets found that about 41% contained a bin of equal or higher significance than the most significant from the real data. Consequently, it was concluded that this analysis found no evidence for significant arc-like structures in the HiRes data.

An analysis of Pierre Auger Observatory data to search for 'multiplets' above an energy of 20 EeV was reported in [104]. A multiplet is defined as a group of events which have arrival directions that form approximately linear structures on the sky, within which energy ordering is present. For a given set of events, an axis was found that maximised the correlation between the arrival directions and inverse of the energies of the events. If the maximum transverse deviation of the arrival directions from this axis was less than 1.5°, the correlation coefficient between the arrival directions and inverse energy greater than 0.9, and the angular separation between the highest energy event (required to be at least 45 EeV) and all other events less than 20°, then that set of events was considered a multiplet. These values of maximum transverse deviation and correlation coefficient were chosen by simulating proton arrival directions after propagation through a model of the GMF and maximising the significance of those multiplets compared to the results from applying the analysis to isotropic data sets.

Upon application to data from the Pierre Auger Observatory data, two multiplets contain-

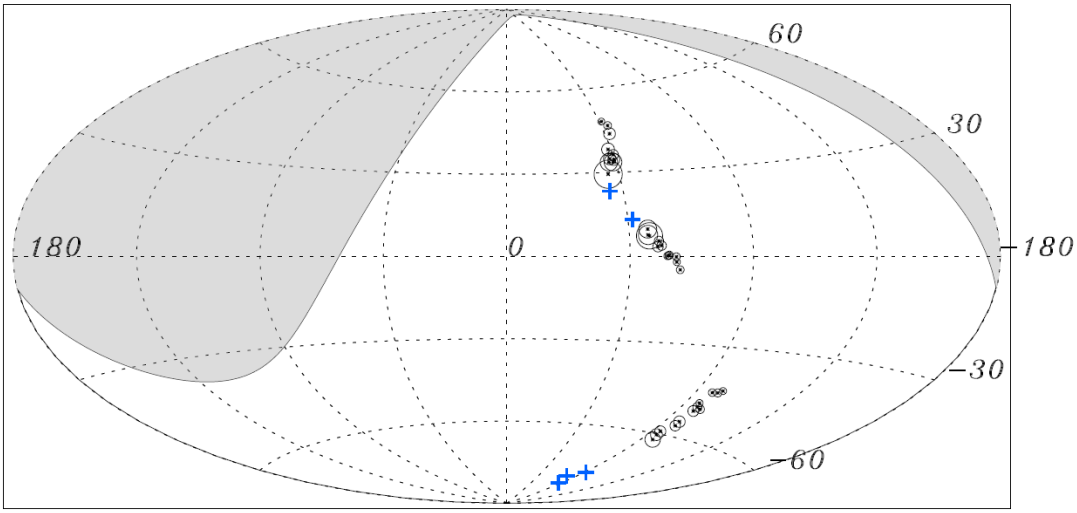


Figure 6.4: In galactic co-ordinates, the multiplets found through the analysis of Pierre Auger Observatory data as published in [104]. The black circles are centred on the arrival direction of the events in each multiplet, with a radius proportional to each event's energy, while the blue crosses indicate the reconstructed source direction from each multiplet. Three crosses are shown near $b = -90°$ which are the reconstructed sources for each of the three decuplets contained in the multiplet of multiplicity 12. The grey shaded area indicates the region of sky not visible to the Pierre Auger Observatory for zenith angles less than $60°$.

ing 10 events (decuplets) and one multiplet containing 12 events (which contained 3 decuplets, although they were not independent) were found to satisfy the above criteria. These are shown in Figure 6.4. A comparison with isotropic mock data sets found that 6% contained at least one multiplet with 12 or more events and 20% contained at least three multiplets of 10 or more events. Thus, it was concluded that no evidence for statistically-significant energy-ordering was found through this analysis.

It is worth remarking that the energy threshold of 20 EeV required in this analysis is below the GZK threshold, and the requirement of a 45 EeV event is just above that threshold. While adopting a lower threshold increases the number of events available for the analysis, it also means that the source horizon for those events is ≈ 1 Gpc [104]. Thus the interpretation of any statistically-significant multiplet, if it were observed, may be somewhat ambiguous due to both the large volume in which sources could be present and the degree of deflection that may be suffered during propagation through extragalactic magnetic fields over such a large distance.

Another method, presented in [106], also searched for approximately linear structures, in this case produced by bright sources of protons at energies above $\approx 10^{19}$ eV. Beginning with the identification of at least two events above 60 EeV in close proximity, a sector of sky was defined (see Figure 6.5) that was subsequently searched to find lower-energy events that maximised the correlation between the arrival direction and inverse energy of each event in

Figure 6.5: The definition of a sector in which energy ordering of arrival directions is searched for in the methods presented in [105] and [106]. The events labelled 1 and 2 are the highest and second-highest energy events in the sector, respectively. The sector (outlined by blue lines) is defined with a central axis formed from the arc that passes through the arrival directions of events 1 and 2 and has opening angle Θ. The axis X' is defined by the centre of mass of the arrival directions of events in the sector, and does not necessarily coincide with the central axis of the sector. The angular extent of the sector is given by R (defined in the text). The true source region is indicated by the black circle S, while the red cross labelled S' indicates the location of the reconstructed source. From [105].

a manner similar to the multiplet analysis described above. A sector shape is chosen from expectations of deflections through both turbulent and regular magnetic fields; as the coherent deflection (along the central axis of the sector) increases with decreasing rigidity, so too does the random deflection (which has components with respect to the sector's central axis both longitudinal and transverse). The angular extent of the sector R is chosen from expectations of the deflections suffered by CRs in the energy range of interest through the GMF.

The parameters of the analysis (such as the angle between the initial events, the opening angle and extent of the sector and the cut on correlation coefficient to be accepted as a source) were optimised using Monte Carlo data, which was produced by the propagation of protons through a representative model of the GMF. In the case of a source being identified, its position could be reconstructed with a resolution of 1° and the deflection power calculated with a resolution of 25%. However, such accuracy required an accurate knowledge of the deflection power for each direction in the sky, which is not well-known for the GMF. A source reconstruction was possible for $\approx 50\%$ of the simulated data sets, although this value varied depending on the simulated parameters used such as relative source luminosity, magnetic field strengths and the ratio of deflections resulting from the turbulent and regular magnetic fields.

This method was later extended to a search for nuclei sources [105], where the possibility of energy ordering even for highly-charged CRs was considered for favourable magnetic field models and source positions, at CR energies above $\approx 55\,\mathrm{EeV}$. Application of this method to

data from the Pierre Auger Observatory resulted in possible energy ordering within the events close to the direction of Centaurus A being found, from which a reconstructed source was located near the direction of the Virgo cluster. While the statistical significance of this result was estimated to be roughly 3×10^{-5} after taking into account a statistical penalty factor for the trials performed, this analysis was *a posteriori,* with scans being performed in parameter space to optimise the significance of the signal. As a result, the true significance of this result is unknown without application to an independent data set.

6.2 Introduction to the Magnetic Spectrometer Method

A new method of searching for energy ordering in UHECR data sets has been developed for this work and is detailed here - the 'magnetic spectrometer' analysis, so named because, in the circumstances outlined at the beginning of Section 6.1, the arrival directions of CRs from a common source will be separated according to their rigidities. While the analysis may be applied to any set of arrival directions and energies, it is applied in this work to data sets at trans-GZK energies. The minimum energy threshold considered here is 55 EeV, which is the threshold at which data from the Pierre Auger Observatory exhibits the most significant correlation with local large scale structure [100]. This choice has the benefit of using events that may undergo small deflections in cosmic magnetic fields (particularly if they are protons or light nuclei) and the sources of such events are expected to be relatively local if the observed spectral cut-off at high energies is due to GZK interactions.

Advantages of the magnetic spectrometer analysis method include the ability to reconstruct cosmic ray source directions using only the properties of the UHECR data, with minimal assumptions about magnetic field models, and an ability to directly account for experimental uncertainties in reconstruction of source positions. The analysis method, and these points, are expanded upon below.

The basic configuration that leads to the development of the magnetic spectrometer method is illustrated in Figure 6.6. This is the ideal case of cosmic rays travelling from an (unknown) source to a detector, located a distance D away, through a plane perpendicular to the direction of a uniform magnetic field $\boldsymbol{B}$. At the detector, the angle between the direction from which a cosmic ray arrives and the direction to the source, θ, is given by:

$$\sin\theta = \left(\frac{ceD\,|\boldsymbol{B}|}{2}\right)\frac{Z}{E} \tag{6.2}$$

where the constant of proportionality $\left(\frac{ceD|\boldsymbol{B}|}{2}\right)$ may be derived from simple geometrical arguments. The factor of 2 arises from θ being the angle between the CR's arrival direction and the source direction rather than between its arrival direction and the direction in which it leaves the source. If two or more cosmic rays from the source are detected, it is not necessary to know $|\boldsymbol{B}|$ or D to calculate the direction to the source. Consider two events, denoted Event 1 and Event 2 and having energies and atomic numbers E_1, Z_1 and E_2, Z_2, respectively, where

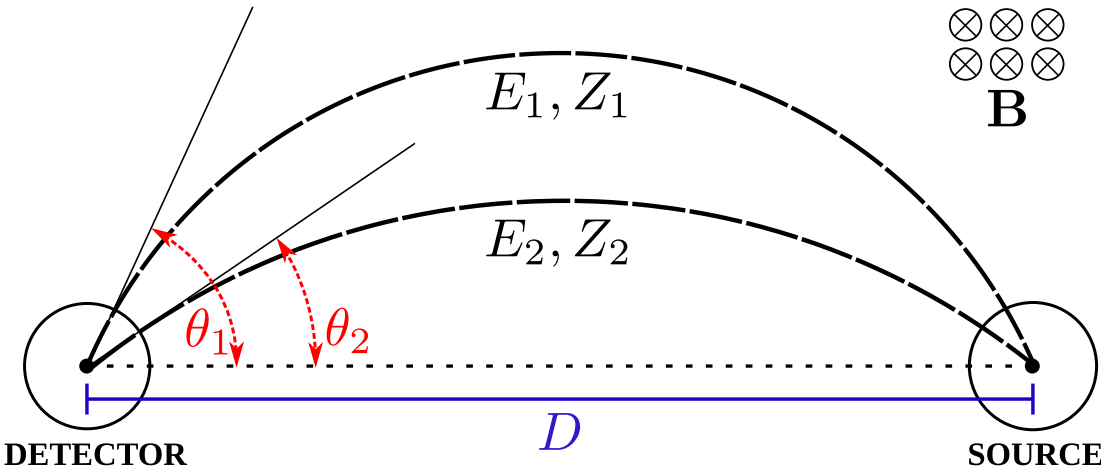


Figure 6.6: The geometry of the 'ideal' magnetic spectrometer, where cosmic rays from a common source arrive at a detector - located a distance D away - after propagating through a uniform magnetic field $\boldsymbol{B}$ perpendicular to their trajectory. The angular separation between the arrival direction of a cosmic ray and the direction to its source is a function of the energy of the cosmic ray. In this image, $E_2/Z_2 > E_1/Z_1$.

$\frac{E_2}{Z_2} > \frac{E_1}{Z_1}$ such that $\theta_2 < \theta_1$. Taking the ratio of Equation 6.2 for the two events gives:

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{Z_1 E_2}{Z_2 E_1} \tag{6.3}$$

which gives:

$$\theta_1 = \sin^{-1}\left(\frac{Z_1 E_2}{Z_2 E_1}\sin\theta_2\right) \tag{6.4}$$

Now, as the angle between the arrival directions of Events 1 and 2, $\Delta\theta$, is known, one can solve for θ_2:

$$\Delta\theta = \theta_1 - \theta_2 \tag{6.5}$$

$$= \sin^{-1}\left(\frac{Z_1 E_2}{Z_2 E_1}\sin\theta_2\right) - \theta_2 \tag{6.6}$$

$$\Rightarrow \theta_2 = \tan^{-1}\left(\frac{\sin\Delta\theta}{\frac{Z_1 E_2}{Z_2 E_1} - \cos\Delta\theta}\right) \tag{6.7}$$

which may be used to define the direction to the source of the two events. It can therefore be seen that, in this ideal case, the source direction may be found using only the properties of those cosmic rays, without reference to any other information.

The magnetic spectrometer analysis presented here makes the assumption that Equation 6.3 is approximately true for cosmic rays of sufficient energy, following their propagation from their source to Earth. Thus, for the analysis method to be valid, the following

assumptions are implicit:

- In the cosmic ray data set under consideration, there exists at least one subset of events which have a common origin. These are referred to hereafter as 'signal' events. The presence of a cut-off in the UHECR energy spectrum that is consistent with the GZK effect, and the correlation with local large scale structure observed in Pierre Auger Observatory data, together hint at extragalactic sources of trans-GZK events being located in the local universe. If this is true, it is reasonable to accept the possibility of a small number of discrete sources contributing multiple events to the total UHECR flux observed at Earth.

- The deflections experienced by those subsets of signal events are small enough such that Equation 6.3 approximately holds following propagation of the events through the GMF. If a significant component of protons or light nuclei is present in the cosmic ray flux at trans-GZK energies then, with magnetic field strengths at μG levels within the Galaxy, this assumption is expected to be valid, at least for some regions of the sky.

- The deflections caused by the turbulent component of the GMF are smaller than the coherent deflections caused by the regular component. Again, this is expected to be valid within the Galaxy from the current knowledge of the GMF (see Chapter 4).

- Deflections introduced by the extragalactic magnetic field do not significantly randomise the UHECR trajectories prior to entry into the GMF so as to destroy the energy ordering of their arrival directions. As previously discussed, the strength and structure of the extragalactic magnetic field are not well-constrained, so for this work it is assumed that its influence is negligible. It should be noted that most analyses designed to look for energy ordering must make a similar assumption.

In addition, owing to fluctuations in EAS development resulting in the inability to determine the primary particle species on a shower-by-shower basis, it is assumed that a majority of the measured events are of the same species. The lack of composition measurements at trans-GZK energies, and the ambiguity of current measurements above the ankle, mean that some assumption about composition must be made. Consequently, $Z_1 = Z_2$ is considered true, at least for a subset of the events in the data set, and Equation 6.3 becomes:

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{E_2}{E_1} \tag{6.8}$$

with Equation 6.7 reducing to:

$$\theta_2 = \tan^{-1}\left(\frac{\sin\Delta\theta}{\frac{E_2}{E_1} - \cos\Delta\theta}\right) \tag{6.9}$$

where $E_2 > E_1$ (a convention that is adopted for the remainder of this section).

Following these assumptions, then, the method may be applied to data sets consisting solely of event arrival directions and energies in an effort to find evidence of energy ordering within those arrival directions, and reconstruct a source direction. To do so, the scenario illustrated in Figure 6.6 is extended to three dimensions. The arrival directions and energies of the events in the data set are paired up (described further in Section 6.4) and, for each pair, a set of Cartesian axes defined. If the unit vectors corresponding to the arrival directions of events 1 and 2 of the pair are $\hat{\boldsymbol{r}}_1$ and $\hat{\boldsymbol{r}}_2$, respectively, then the Cartesian basis vectors are defined as:

$$\hat{\boldsymbol{z}} = \hat{\boldsymbol{r}}_2 \tag{6.10}$$

$$\hat{\boldsymbol{y}} = \frac{\hat{\boldsymbol{r}}_2 \times \hat{\boldsymbol{r}}_1}{|\hat{\boldsymbol{r}}_2 \times \hat{\boldsymbol{r}}_1|} \tag{6.11}$$

$$\hat{\boldsymbol{x}} = \hat{\boldsymbol{y}} \times \hat{\boldsymbol{z}} \tag{6.12}$$

With these definitions, the reconstructed source direction for the pair is taken to lie in the $x - z$ plane. The angles θ_1 and θ_2 are found through Equations 6.5 and 6.9. If $\theta_1 \leq 90^\circ$ then the source for the pair is accepted, otherwise it is rejected; Equation 6.8 does not hold for larger deflections. While this cut could be made more restrictive, doing so implicitly constrains the strength of the magnetic field through which the pair of CRs have propagated. On the other hand, a more restrictive cut also has the effect of increasing background rejection. For this work a 90° cut was acceptable, although the effect of varying it could be investigated in the future.

In the case of the source being accepted, the unit vector corresponding to its position, $\hat{\boldsymbol{s}}$, may then be uniquely defined by requiring that $\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{r}}_1 = \cos\theta_1$ and $\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{r}}_2 = \cos\theta_2$. Figure 6.7 illustrates the definition of these quantities for a single pair of events.

Repeating this process for all possible pairs within the data set (see Section 6.4) results in multiple 'sources' being reconstructed. If the magnetic spectrometer assumptions listed above are valid, then the source directions that are reconstructed from pairs of signal events will be clustered (in the ideal case, they will be exactly the same). On the other hand, source directions reconstructed from pairing either two background events or a background and a signal event will be randomly located on the sky and will display a lower degree of clustering. Figure 6.8 sketches these different cases. In a similar way, if the assumption that $Z_1 = Z_2$ is incorrect, the reconstructed source for that pair will also be incorrect and effectively contribute to the background. A brief investigation of this can be found in Chapter 7.

Thus, when applying the magnetic spectrometer method to a UHECR data set, a significant clustering of reconstructed source directions is the signature of energy ordering that is to be searched for. If such a direction clustering is found, it may be indicative of multiple events having a common origin that is located in a direction near that cluster.

Advantages of the magnetic spectrometer analysis include its ability to be applied to

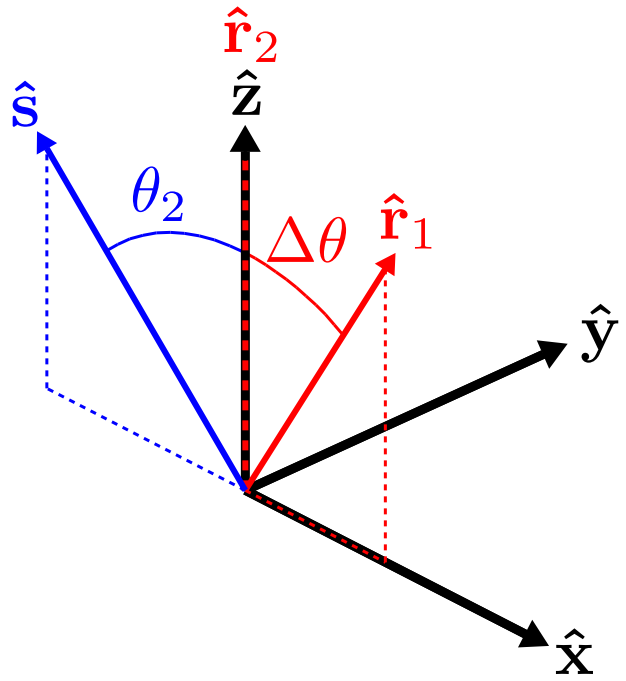


Figure 6.7: This diagram shows the geometry that is defined for each pair of events in the magnetic spectrometer analysis. The arrival direction unit vectors of events 1 and 2, $\hat{\boldsymbol{r}}_1$ and $\hat{\boldsymbol{r}}_2$, lie in the x-z plane with $\hat{\boldsymbol{r}}_2$ coinciding with $\hat{\boldsymbol{z}}$. The direction to the reconstructed source, defined by $\hat{\boldsymbol{s}}$, also lies in the x-z plane at an angle θ_2 from $\hat{\boldsymbol{z}}$.

UHECR data while making minimal assumptions regarding the strength and structure of the GMF. Only the measured properties of event energy and arrival direction are required to perform the analysis, with no reference to other information required. In addition, the method has little sensitivity to overall systematic shifts in the energy scale of a detector due to the ratio of the event energies being used, with such a shift affecting only the cut on a data set based on a minimum energy threshold. Finally, no cuts based on the overall geometry of the arrival directions (such as their deviation from linearity) are required to be made, as the signature from signal events to be searched for is a clustering of reconstructed source positions rather than a specific pattern of their arrival directions.

On the other hand, the magnetic spectrometer analysis as described here has several disadvantages when compared to some other energy ordering analyses. While in some cases it is an advantage, making minimal assumptions about magnetic fields can mean that this method is less sensitive than others when cosmic magnetic fields are known. However, adaptations of the method (such as altering the cut on θ_1 or restricting the angular range within which two events of given energies are paired up) in line with physical expectations of deflections through the GMF could improve its sensitivity in that case. The consideration of events only at trans-GZK energies means that data sets are limited in size when compared to other methods that

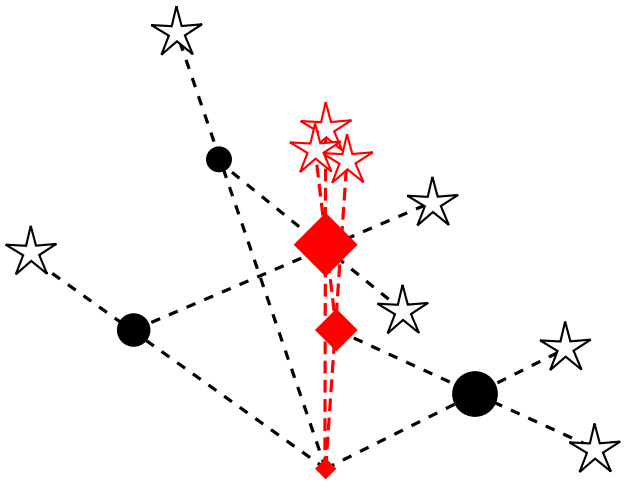


Figure 6.8: This sketch illustrates the reconstruction of sources from a data set containing both signal and background events. The arrival directions on the sky of the signal and background events are depicted by red diamonds and black circles, respectively, and the size of the symbols proportional to energy. Dashed lines connect each pair of events to their reconstructed source position, represented by a star, with the sources reconstructed from pairs of signal events in red. When pairing up either two background events, or one signal event and one background event, the reconstructed source is isolated. However, the pairing of signal events produces a small cluster of reconstructed source positions, indicating the presence of energy-ordering within the data. Note that, for clarity, not all possible pairs of events are shown.

include a lower-energy threshold - the collection of enough independent data to verify any observed signal may take a long period of time. Finally, as source positions are calculated for each event pair independently of the other pairs, consistent energy ordering across all events contributing to a cluster of source positions is not enforced. However, if a source position is identified, the arrival directions of the events associated with that source may be checked for such consistency. This is discussed in Chapter 7.

6.3 Incorporating Experimental Uncertainties Into Source Position Reconstruction

When applying the magnetic spectrometer analysis method to real UHECR data sets, it is necessary to take into account experimental uncertainties in the energies of the events and their arrival directions. The effect of these uncertainties can be incorporated directly, as described here. In doing so, the reconstructed source position for a pair of events is described by a probability distribution rather than the single direction calculated using the method in Section 6.2.

The process of calculating the source direction probability distribution is the same for each

pair and obtains the function $\frac{dN}{d\Omega}(\vartheta, |\varphi|)$ that describes (in terms of the spherical polar angles ϑ and φ in the co-ordinate frame defined in Figure 6.7) the differential probability of the reconstructed source lying at (ϑ,φ) under the magnetic spectrometer assumption. Calculation of $\frac{dN}{d\Omega}(\vartheta, |\varphi|)$ for every pair in the data set allows the production of a smoothed skymap, as described in Section 6.4, which gives the integral number of sources within an angle ζ_{smth} of each point of the sky. From this map, possible source directions may be identified.

As the process of calculating a source direction probability distribution is time-consuming, the fitting of $\frac{dN}{d\Omega}(\vartheta, |\varphi|)$ with simple functions, as described below, permits a simple description of its shape. Thus, the description of $\frac{dN}{d\Omega}(\vartheta, |\varphi|)$ for given values of $\frac{E_2}{E_1}$ and $\Delta\theta$ is generated when required, and stored for later use. This effectively creates a look-up table, which has chosen binning in $\frac{E_2}{E_1}$ of 0.01 and in $\Delta\theta$ of 0.1°. Subsequent analyses may access these stored values, thereby greatly reducing the time taken to perform an analysis in which previously encountered pairs of the energy ratio and angular separation are found.

The distribution for an event pair is calculated through a Monte Carlo method in which the energies and arrival directions of each event in the pair are resampled within their respective uncertainties. As the magnetic spectrometer method will be applied to data from the Pierre Auger Observatory, its experimental uncertainties are adopted here. Above an energy of 10 EeV, the relative statistical energy uncertainty is approximately 12% [36] and the uncertainty in arrival directions measured with the surface detector is approximately 1° [143]. Consequently, each event's energy is drawn from a Gaussian distribution having a mean of the event's nominal energy E and standard deviation $0.12E$. The event's arrival direction is generated such that the space angle η between the event's nominal and resampled arrival directions follows a distribution given by [31]:

$$dn \propto \exp\left(-\frac{\eta^2}{2\sigma^2}\right) d(\cos\eta) \tag{6.13}$$

such that 68% of the resampled arrival directions lie within 1° of the nominal arrival direction. This distribution is plotted in Figure 6.9. The azimuthal distribution of the resampled arrival directions is uniform.

To allow an accurate description of the source direction distribution, 10^7 instances of these resampled values are created. In every case, a reconstructed source direction is calculated as described in Section 6.2. The values of ϑ and φ of this reconstructed source position are calculated according to the Cartesian axes defined by the nominal arrival directions of the event pair and are subsequently binned into a two-dimensional distribution. As this distribution is symmetric with respect to the x-z plane of the Cartesian axes, the absolute value of φ is binned instead - this approximately doubles the statistics of each bin, permitting more accurate fitting the distribution's values (see below). An example distribution is shown in Figure 6.10, which has been calculated for $\frac{E_2}{E_1} = 1.2$ and $\Delta\theta = 7.3°$.

Following the generation of a distribution of $(\vartheta, |\varphi|)$ pairs, each of its rows - for which ϑ is constant - is considered separately. Every bin in the row is assigned a Poissonian uncertainty

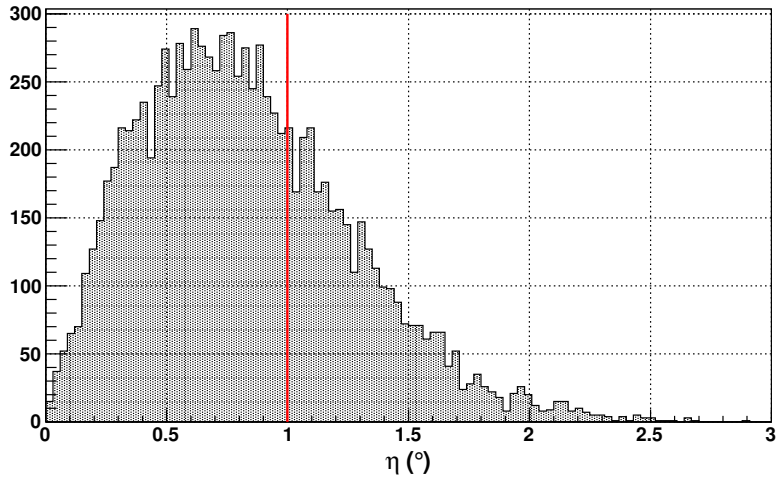


Figure 6.9: The distribution of angle η, the space angle between an event's nominal arrival direction and its resampled arrival directions. The red line indicates the 68% quantile at $\eta = 1°$.

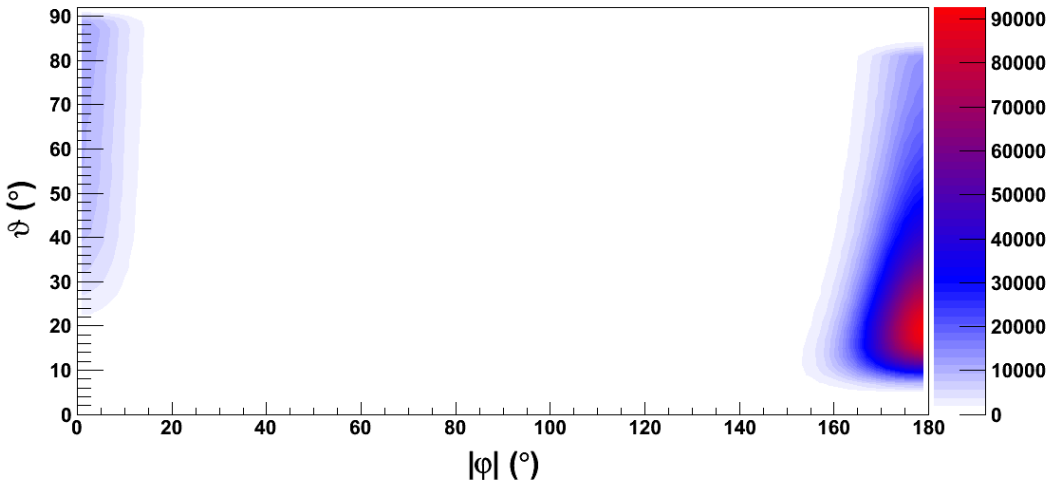


Figure 6.10: An example distribution of the $(\vartheta, |\varphi|)$ pairs calculated for an event pair having parameters $\frac{E_2}{E_1} = 1.2$ and $\Delta\theta = 7.3°$. The legend to the right gives the colour scale as a function of the number of reconstructed sources per bin. Note that there are two peaks present, one at $|\varphi| = 0$ and a larger one at $|\varphi| = \pi$; the latter is due to the instances where the resampled energies give $E_2 > E_1$, while the former is a result of the energy ordering being reversed ($E_1 > E_2$).

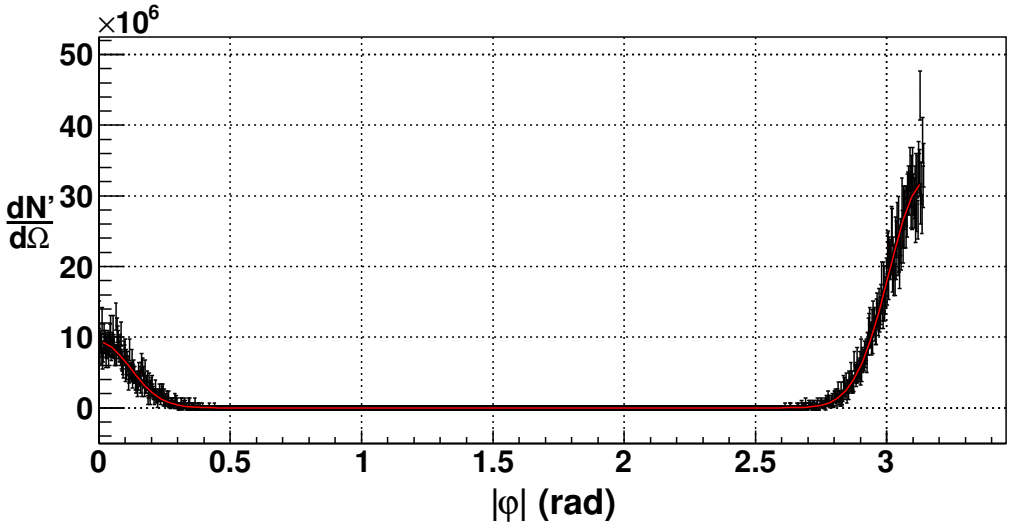


Figure 6.11: An example fit of $\frac{dN'}{d\Omega}$ to the histogram of reconstructed source positions. This fit corresponds to the row for which $\vartheta = 50°$ in Figure 6.10. The values of each bin are shown by the data points with their corresponding Poissonian errors, and the fitted function is shown by the red line.

and is divided by the solid angle $d\Omega$ that the bin subtends. This solid angle correction is the same for every bin in a single row, but varies between rows owing to its ϑ-dependence. Finally, a fit to the adjusted bin values is performed which describes the number of sources per steradian as a function of $|\varphi|$ for a given value of ϑ. A sum of two Gaussian functions was found to describe the data well, and is given by:

$$\frac{dN'}{d\Omega}(\vartheta, |\varphi|) = A_1(\vartheta) \exp\left(-\frac{|\varphi|^2}{2\sigma_1^2(\vartheta)}\right) + A_2(\vartheta) \exp\left(-\frac{(|\varphi| - \pi)^2}{2\sigma_2^2(\vartheta)}\right) \tag{6.14}$$

with the prime indicating that the function still remains to be normalised (see below).

By repeating this fit for all values of ϑ, the distributions of $A_1(\vartheta)$, $\sigma_1(\vartheta)$, $A_2(\vartheta)$ and $\sigma_2(\vartheta)$ may be found for a particular pair of $\left(\frac{E_2}{E_1}, \Delta\theta\right)$ values. In the case of varying energy and arrival direction uncertainties, $\frac{dN}{d\Omega}$ would display a dependence on those as well. However, since these uncertainties are considered constant within the energy range of interest, only the energy ratio and angular separation dependencies need to be considered. A single fit to the histogram of Figure 6.10 for the row corresponding to $\vartheta = 50°$ is shown in Figure 6.11. The distributions of $A_1(\vartheta)$, $\sigma_1(\vartheta)$, $A_2(\vartheta)$ and $\sigma_2(\vartheta)$ for the entire histogram are shown in Figure 6.12.

Near the edges of the distribution, large fluctuations are evident in the fitted values of σ_1 and σ_2, which result from difficulty in fitting bins with very small values. Such fluctuations are not important due to the corresponding fitted amplitudes also being very small, meaning the overall shape of the fitted distribution is not significantly affected.

Finally, $\frac{dN'}{d\Omega}$ is normalised by multiplying by a normalisation constant:

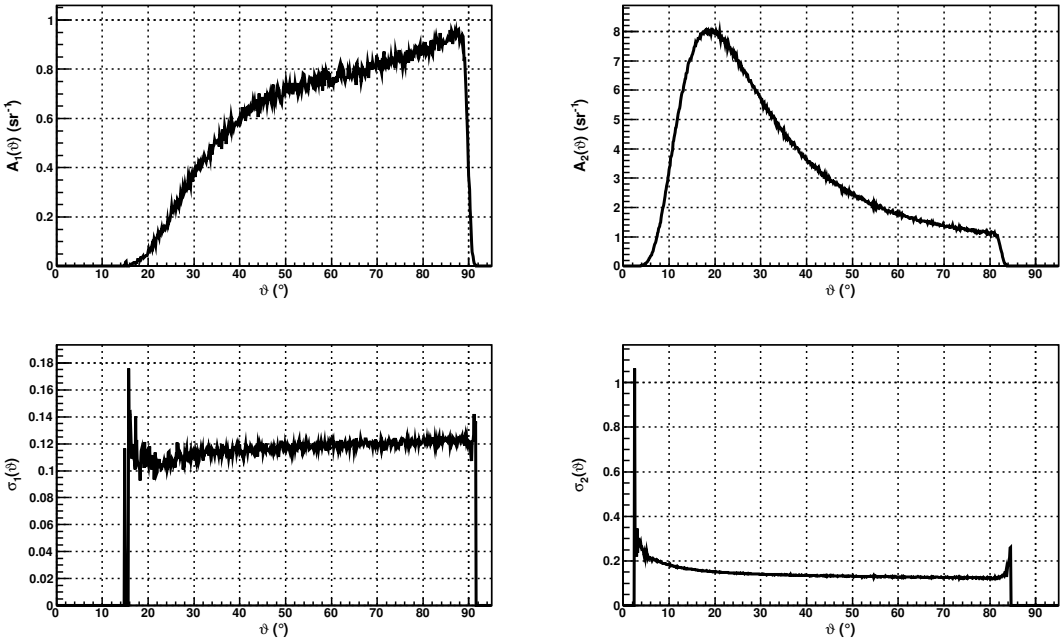


Figure 6.12: The values of $A_1(\vartheta)$, $A_2(\vartheta)$, $\sigma_1(\vartheta)$ and $\sigma_2(\vartheta)$ for the values $\frac{E_2}{E_1} = 1.2$ and $\Delta\theta = 7.3°$. Note that the amplitudes that correspond to large fluctuations in σ_1 and σ_2 are extremely small, so those fluctuations do not significantly distort the distribution. The truncations of the distributions near $\vartheta = 90°$ arise from the cut placed on θ_1. The 'dip' that is evident in the plot of $A_1(\vartheta)$ near $\vartheta = 70°$ is due to convolution with solid angle effects.

$$\frac{dN}{d\Omega} = C_{norm}\frac{dN'}{d\Omega} \tag{6.15}$$

such that the integral of $\frac{dN}{d\Omega}$ across the sky is no greater than 1. This represents the probability that the source for a given pair exists under the magnetic spectrometer assumption, as not all of the 10^7 resampled pairs used to generate the source distribution may produce a reconstructed source due to the cut placed on the value of θ_1. If the number of reconstructed sources is denoted N_{rec}, the value of C_{norm} is defined through:

$$\frac{N_{rec}}{10^7} = C_{norm}\int \frac{dN}{d\Omega}d\Omega \tag{6.16}$$

$$= C_{norm}\sqrt{\frac{\pi}{2}} \times \int \left(A_1(\vartheta)\sigma_1(\vartheta)\,\mathrm{erf}\left(\frac{\pi}{\sqrt{2}\sigma_1(\vartheta)}\right) + A_2(\vartheta)\sigma_2(\vartheta)\,\mathrm{erf}\left(\frac{\pi}{\sqrt{2}\sigma_2(\vartheta)}\right)\right) d\vartheta \tag{6.17}$$

where the error function is used to account for the truncation of the first Gaussian at $\varphi = \pi$ and the second at $\varphi = 0$.

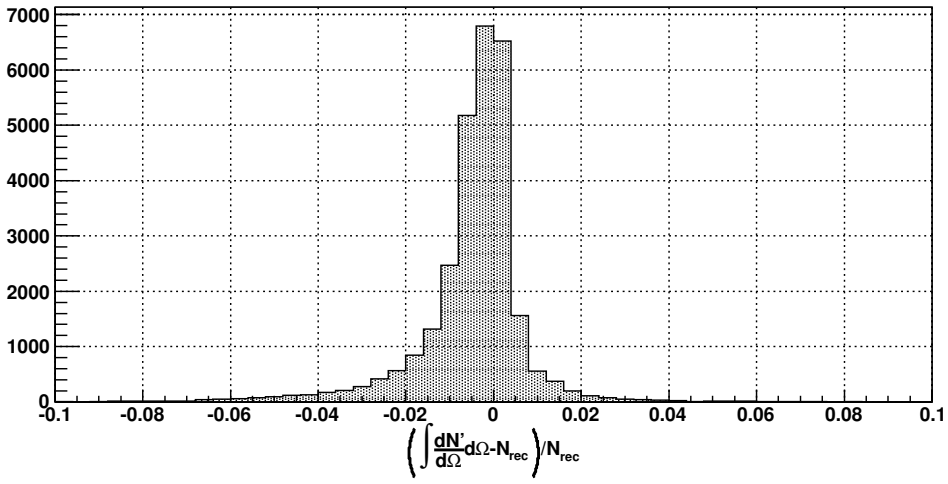


Figure 6.13: Histogram of the relative differences between $\int \frac{dN'}{d\Omega} d\Omega$ and N_{rec} for the pair distributions calculated in this work. A slight bias towards negative values is apparent, but the effect is negligible.

The accuracy of describing the probability distributions with the fitted functions $\frac{dN}{d\Omega}$ was found to be very good. A comparison of the integral of $\frac{dN'}{d\Omega}$ over the sky with N_{rec} resulted in good agreement - in all cases the difference was less than 10%, while in most cases it was of the order 1%. A histogram of the relative errors is shown in Figure 6.13.

6.4 Application to Cosmic Ray Data Sets

With the calculation of source position probability distributions, as described above, taking into account the experimental uncertainties in event energies and arrival directions, the magnetic spectrometer method may be applied to real UHECR data sets.

Firstly, the arrival directions of events in a UHECR data set above an energy threshold E_{th} are taken. As mentioned previously, for this work $E_{th} = 55\,\mathrm{EeV}$ is adopted. When maximising the significance of any signal seen in a data set, E_{th} could be one of the parameters varied. For this work, the choice was motivated by both physical considerations and results of previous analyses, with the effects of varying other parameters focussed upon instead.

Secondly, all events which have arrival directions separated by less than an angle $\Delta\theta_{cut}$ are paired. However, a pair is rejected if $\frac{E_2}{E_1} < 1.02$ or $\Delta\theta < 2^\circ$, as little information is expected to be obtained regarding a source position for pairs of events with similar energies or closely-separated arrival directions (for example, the random scattering in arrival directions expected from turbulent fields is comparable to 2°). The value of $\Delta\theta_{cut}$ reflects the maximum separation between events that may be expected for the energy range and magnetic field under consideration. In the case of well-known magnetic fields, a dependence on $\frac{E_2}{E_1}$ could be incorporated into the value of $\Delta\theta_{cut}$ used, which could also be extended to considering pairs

only within some $\Delta\theta$ range. However, since it is desirable, in the absence of well-understood magnetic fields, to decouple the method from *a priori* assumptions about the GMF that may be prone to error, only a maximum separation is used. For this work, the effect of varying $\Delta\theta_{cut}$ is investigated using Monte Carlo data in Chapter 7, and $\Delta\theta_{cut}$ is considered a parameter to be optimised in Chapter 8.

Once the pairs are defined, each of their source direction distributions are calculated as described in Section 6.3. As these are differential distributions, integration over a chosen solid angle is performed to return the integral number of sources within a local region of each point on the sky. Top-hat smoothing is used here, with angular radius ζ_{smth} being a second variable parameter in the magnetic spectrometer method. The optimum value of ζ_{smth} reflects the size of the individual source direction distributions, as well as the degree to which the sources reconstructed from signal events are clustered. If the energy ordering of the signal events is disrupted following their propagation from their source to Earth then the value of ζ_{smth} would need to be larger than in the case of energy ordering being well preserved, in order for the angle of smoothing to include the majority of the signal source distributions.

In order to perform this smoothing, the sky is divided into a large number of pixels of approximately equal solid angle. These are arrayed in rows with a spacing in latitude of $\Delta(\text{lat}) = 0.25°$ and longitude $\Delta(\text{lon}) \approx 0.25°/\cos(\text{lat})$ (the spacing in longitude is adjusted to keep an integer number of equal-sized pixels in each row). This gives pixels which subtend a solid angle that varies across the sky by only $\approx 1\%$, apart from at the poles where a single pixel at each pole subtends a somewhat smaller solid angle. Such a deviation for only two pixels can be ignored. This consistency between solid angles is required for the peak-finding algorithm used in Chapter 7.

For each of the (i,j) pixels, its $(\vartheta, |\varphi|)$ position is found relative to the co-ordinate system of each of the n_p pairs found in the data. Then the differential number of sources at the location of that pixel is given by:

$$\frac{dS_{ij}}{d\Omega} = \sum_{p=1}^{n_p} \frac{1}{2} \frac{dN_p}{d\Omega}(\vartheta_p, |\varphi|_p) \tag{6.18}$$

where the p subscript denotes which pair the function $\frac{dN_p}{d\Omega}$ is calculated for. The factor of $\frac{1}{2}$ is required to take into account the fitting of $\frac{dN_p}{d\Omega}$ to $|\varphi|$ rather than φ, as each bin in the distribution of $(\vartheta,|\varphi|)$ pairs contains an equal contribution from $+\varphi$ and $-\varphi$. Finally, the map of $\frac{dS_{ij}}{d\Omega}$ is smoothed through:

$$S_{ij} = \sum_k \sum_l \frac{dS_{kl}}{d\Omega} \Delta\Omega_{kl} H(\zeta_{smth} - \delta_{ij,kl}) \tag{6.19}$$

where $\Delta\Omega_{kl}$ is the solid angle subtended by the $(k,l)^{\text{th}}$ pixel, $\delta_{ij,kl}$ is the angle between the centre of the $(i,j)^{\text{th}}$ and $(k,l)^{\text{th}}$ pixels and $H(x)$ is the Heaviside step function.

Note that, due to the arrangement of the pixels used to calculate maps of S_{ij} (described

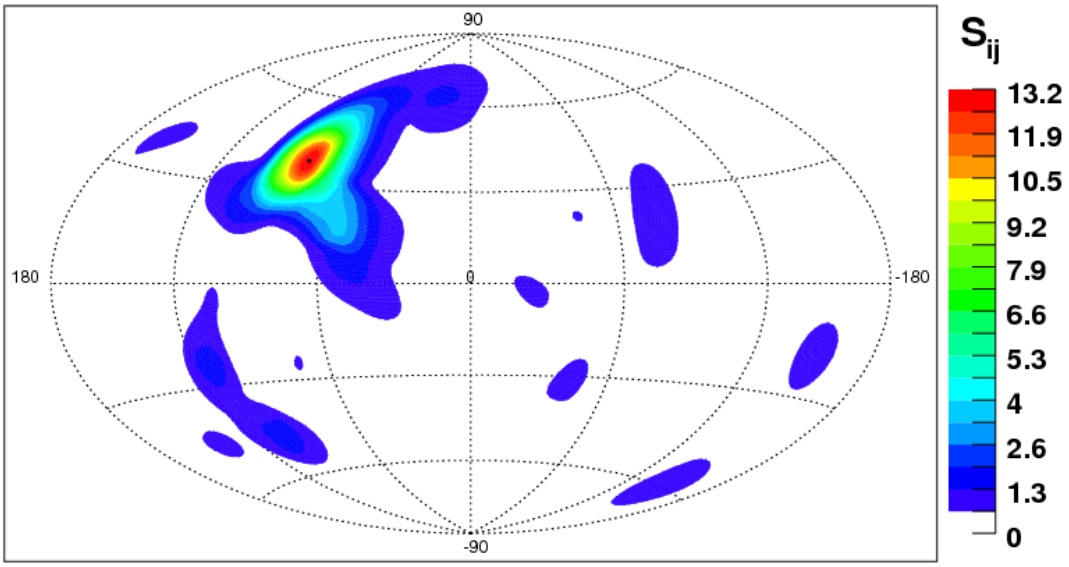


Figure 6.14: An example of a skymap following the application of the magnetic spectrometer method to a simulated data set containing both signal events, which have been propagated from a common source through a model of the GMF, and isotropic background events. The map is coloured according to the value of S_{ij} (see text). A smoothing angle of $\zeta_{smth} = 10^\circ$ has been used here. The reconstructed source direction is indicated by a black dot, and is approximately 1.6° from the simulated source direction.

above), the nominal smoothing angle must be increased by half of the pixel size (0.125°) to give a constant value of S_{ij} in the case of a uniform value of $\frac{dS_{ij}}{d\Omega}$. Thus, if ζ_{smth} is reported to be 10°, for example, the actual value used is 10.125°. The additional 0.125° is omitted hereafter in the text for neatness.

These steps produce a skymap in which the value of each pixel, S_{ij}, describes the integral number of sources within an angle ζ_{smth} of that pixel. An example of such a skymap, produced from a data set containing an isotropic background with a group of signal events generated through a Monte Carlo simulation (see Chapter 7), is shown in Figure 6.14. The peak of the resulting S_{ij} map, which is formed by multiple overlapping source direction distributions, lies close to the direction of the simulated source. Further analysis of these simulated data sets may be found in the next chapter.

Chapter 7

Magnetic Spectrometer Analysis of Simulated UHECR Data

To validate the magnetic spectrometer analysis method, described in Chapter 6, it is applied to simulated UHECR data sets which contain signal events from a known source position. The use of such data sets allows the accuracy of source direction reconstruction to be determined for a range of simulation parameters, such as the GMF configuration and the relative brightness of the source compared to the all-sky UHECR flux.

Section 7.1 describes the production of the data sets to which the magnetic spectrometer analysis method is applied. The remaining sections discuss the results of the analysis of those data to gain an understanding of the ability of the magnetic spectrometer analysis to identify energy ordering and reconstruct the directions of sources of signal events.

7.1 Generation of Simulated Data Sets

The simulated data sets produced for this work contain both signal events and background events. Signal events are generated by propagating particles through a GMF model and recording their energies and directions upon entering a simulated detector, while the arrival directions of background events are randomly generated. This process is described in greater detail below.

In Section 7.3, a scan in parameter space is performed to determine the values of $\Delta\theta_{cut}$ and ζ_{smth} that allow the reconstruction of real source positions while efficiently rejecting those peaks in S_{ij} maps which do not correspond to a true source. In order to avoid biasing those results toward a particular magnetic field model, an equal number of data sets are considered for each GMF configuration (listed in Table 7.1); a total of 8 source directions, providing 8 signal events each, were generated for every model. The computationally intensive nature of the production of such data sets limited the number that could be considered in this work.

The production of a signal data set begins with the definition of a source position, which is a randomly-chosen point at a distance 100 Mpc from the Galactic Centre. Using this source

	Spiral Model	B_{sp} (at Earth,μG)	D (μG kpc^3)	B_T^0 (μG)	σ_r (μG)
1	BSS-S	2	0	0	3
2	BSS-S	2	-122.825	10	3
3	BSS-S	5	0	0	7.5
4	BSS-S	5	-122.825	10	7.5
5	ASS-A	2	0	0	3
6	ASS-A	2	-122.825	10	3
7	ASS-A	5	0	0	7.5
8	ASS-A	5	-122.825	10	7.5

Table 7.1: The magnetic field parameters of each of the eight GMF configurations considered for this work. See Chapter 4 for definitions of each of the quantities listed here.

direction, CRs are propagated through the relevant GMF model until either 8 events (4 events for iron nuclei, as described in Section 7.5) are 'detected' as described below, or 100 000 trajectories have been calculated (200 000 trajectories in the case of iron). The latter condition is required for cases where the source direction is located in a region of sky from which CRs cannot reach the detector, such as directions close to $l = 0°$ for which the Galactic Centre lies between the source and Earth.

The propagation of a particle begins with the selection of its entry point to the Galaxy, which is located on the surface of the 20 kpc sphere taken to be the Galactic volume. This point is randomly-generated but is required to lie within a spherical cap oriented toward the source position. In the case of iron, the edge of this cap is defined by the tangent vectors to the 20 kpc sphere which pass through the source position; for protons, the cap is defined such that the maximum angle between the source-Earth and source-entry point vectors, ω_{max}, is no greater than 75δ, where δ is the angle subtended by the 100 pc radius of the simulated detector (see below) at the source position (see Figure 7.1). This avoids the choice of entry points at positions from which protons are unlikely to reach the detector, which would result in unnecessary calculations and increase the time taken to create signal data sets. In order to determine if restricting the entry points in this manner significantly affects the results, Figure 7.2 shows the values of ω for those protons which reached the detector. Although a cut-off due to the restriction placed on ω is present, it appears in the tail of the distribution. Consequently, increasing the value of the cut placed on ω would not have significantly affected the arrival direction distributions of the particles reaching the detector, but would have decreased the efficiency of the simulations.

Under the assumption of negligible extragalactic magnetic fields, the velocity vector with which the particle enters the Galactic volume is parallel to the source-entry point vector. An energy above 55 EeV is randomly generated according to the energy spectrum reported in [47] and assigned to the event.

The particle is subsequently propagated through the GMF, as described in Chapter 5, until

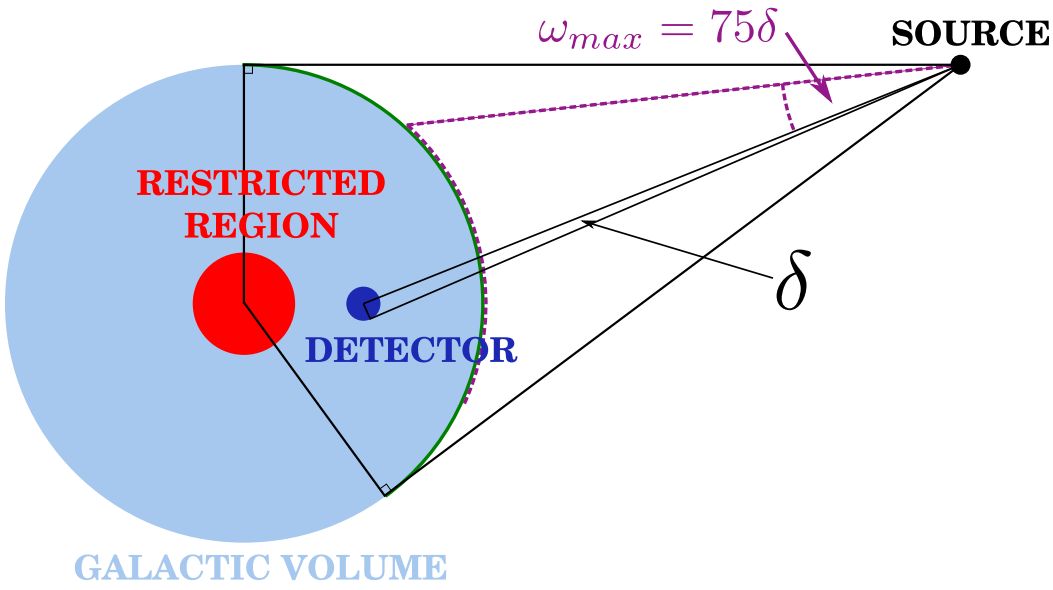


Figure 7.1: This image sketches the setup of the simulations used to generate the signal data sets (not to scale). The Galactic volume is represented by the light blue circle, the detector by a dark blue circle and the sphere of 4 kpc radius at the Galactic Centre, which particles are not allowed to enter, by the red circle. The green arc represents the spherical cap within which the entry points of iron nuclei are allowed, while the shorter purple dashed arc represents that of protons. At the source, the radius of the detector subtends an angle δ, which defines the maximum opening angle between the source-detector and source-entry point vectors of $\omega_{max} = 75\delta$ for protons.

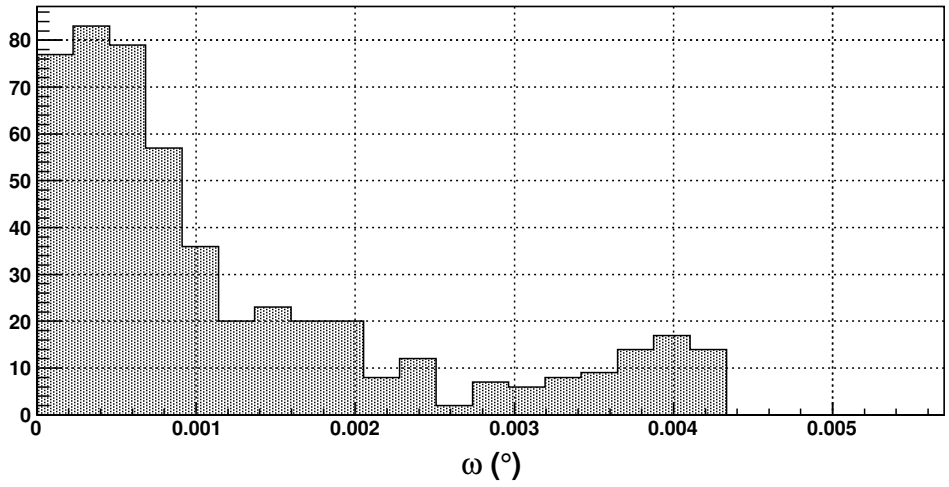


Figure 7.2: A histogram of ω, the angle between the simulated source-Earth vector and source-entry point vectors for the protons that reached the detector. A cut-off appears above $\omega = 0.004°$ due to the restriction placed on the proton entry points to the galaxy but, as this appears in the tail of the distribution, such a restriction will not have significantly affected the results.

one of three conditions is satisfied: the particle exits the Galactic volume, the particle passes within 4 kpc of the Galactic Centre, or the particle enters the volume of the detector, which is a sphere of radius 100 pc centred on the position of Earth. In the latter case, the velocity vector and energy of the particle are recorded as a detected event. The rejection of particles entering the central region of the Galaxy takes place because the GMF models employed here are not considered to be suitable for the description of UHECR trajectories within this region; strong variations in the strength and structure of the GMF are observed near the Galactic Centre which are not reflected in the models used [158, 159, 162].

Finally, to simulate the experimental uncertainties of the Pierre Auger Observatory, the energies and arrival directions as measured at the detector are fluctuated within their respective uncertainties (described in Chapter 6). These fluctuated quantities are taken to comprise the signal data used in the analyses described below.

A background component of the data is constructed by generating events with arrival directions drawn from an isotropic distribution and energies from the same spectrum as the signal data. These events can be considered UHECRs which have suffered large deflections in the GMF (such as heavy nuclei) or which originate at dim sources that contribute very few events to the all-sky UHECR flux at Earth. While the results of [98] indicate that the UHECR arrival direction distribution is likely anisotropic, its exact form is unknown. As discussed in Chapter 2, the fraction of UHECR events that correlate with local AGN is currently estimated to be 0.38 [100], while the fraction expected from isotropy is 0.21. Thus, while an isotropic arrival direction distribution can be considered an approximation of the true UHECR arrival direction distribution, the results presented here may differ to those in which the background data were constructed using an anisotropic distribution. Nevertheless, for the purposes of demonstrating the performance of the magnetic spectrometer analysis an isotropic arrival direction distribution for background events is sufficient.

Note that the generation of data sets in the manner described here assumes that the simulated detector has equal exposure across the sky. However, that of a real detector will, in practice, be non-uniform, effectively leading to greater exposure to particular regions of the Galactic magnetic field than others. The performance of the magnetic spectrometer analysis upon application to data sets drawn from a non-uniform exposure is left for a future study.

7.2 Analysis of Signal Events Only

While the focus of the magnetic spectrometer analysis is on UHECR data sets in which the majority of events are not expected to display energy ordering in their arrival directions, it is instructive to first apply it to data sets containing signal events only. This allows the performance of the analysis to be evaluated in the absence of complicating factors resulting from background events, permitting an estimate of the best possible results.

Thus, each of the 64 sets of 8 signal events were considered separately and the magnetic

spectrometer analysis applied. Results using parameter values of $\Delta\theta_{cut} = 90°$ and $\zeta_{smth} = 5°$ are shown here ($\zeta_{smth} = 10°$ was also applied, and returned similar results). Note that taking $\Delta\theta_{cut}$ to be 90° is equivalent to having no cut at all, owing to the 90° limit that is placed on θ_1 in this work. For each of the smoothed S_{ij} maps, the highest-value pixel was taken to be the reconstructed source direction, and the angle between it and the direction of the simulated source, ϵ_{rec}, the error in the source direction reconstruction. Figure 7.3 displays an example of such a data set (generated using GMF configuration 4), as well as the S_{ij} map produced by applying the magnetic spectrometer analysis to the data. Analysis of this data set gave a source direction error or $\epsilon_{rec} \approx 1.9°$. A histogram of ϵ_{rec} for all 64 of the data sets is shown in Figure 7.4.

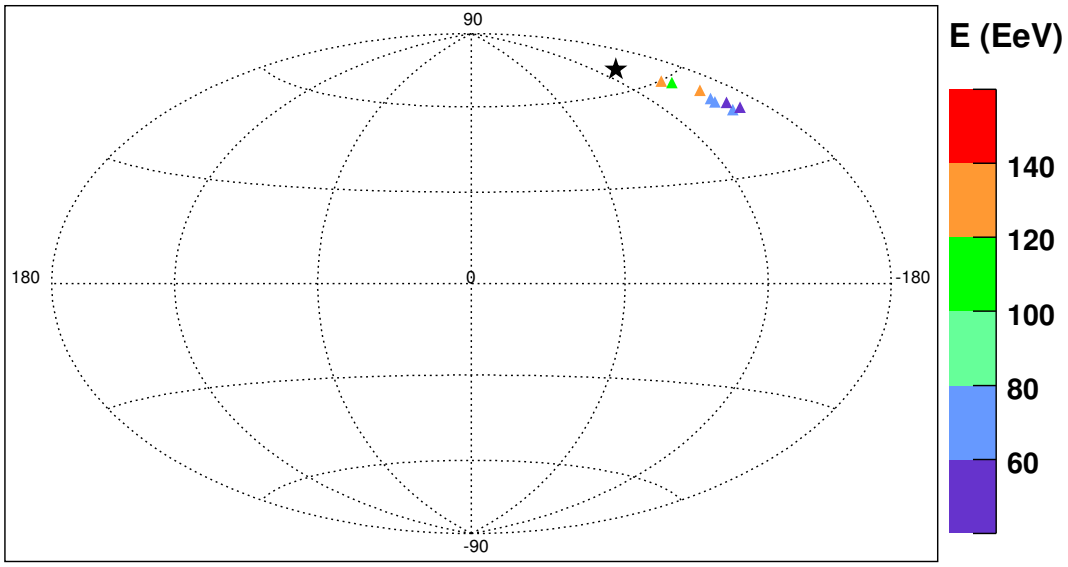
Upon performing the magnetic spectrometer analysis on the 64 signal data sets, in 52 cases ϵ_{rec} is less than 20°. For the purposes of Section 7.3, $\epsilon_{rec} = 20°$ is taken as the limit for a reconstructed source to be considered to correspond to the position of the simulated source, following the apparent separation in the distribution of Figure 7.4 near this value.

The reason for the larger values of ϵ_{rec} in the remaining 12 cases was investigated and found to be a result of the CRs traversing regions of strong magnetic fields. This serves to introduce a combination of very large deflections from the simulated source direction and also disrupt the energy ordering, leading to the assumptions made for the magnetic spectrometer analysis breaking down.

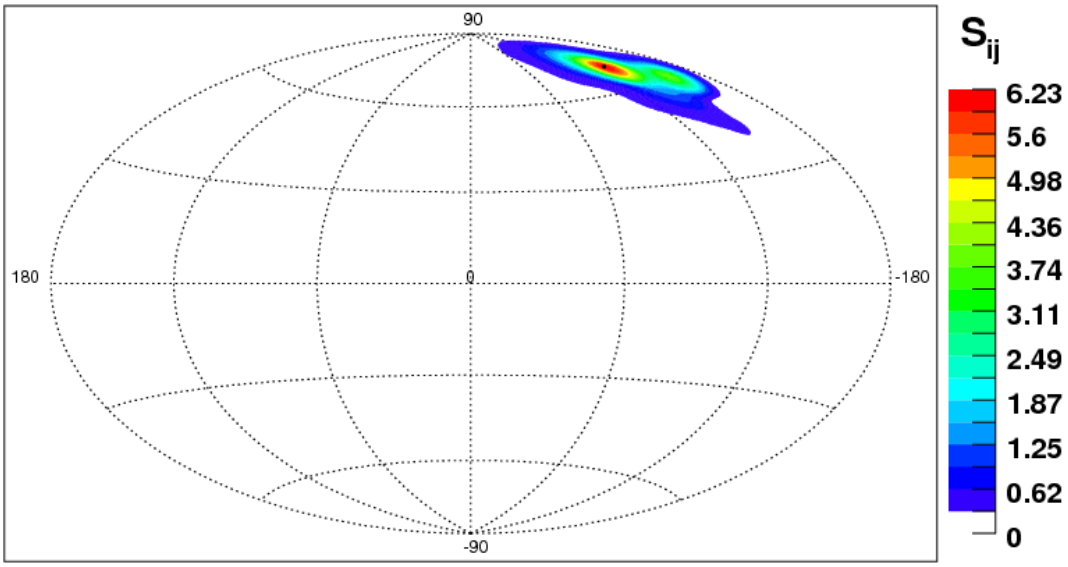
An example of these data sets is shown in Figure 7.5, where skymaps of the arrival directions and the smoothed S_{ij} values following application of the magnetic spectrometer analysis are shown. This data set was produced using GMF configuration 8. Note the scattered nature of the arrival directions, as well as the fact that 2 of the 8 events arrive from very different directions to the remaining events - this is an example of magnetic lensing, as discussed in [159], in which multiple images of an extragalactic source may appear following propagation through strong Galactic magnetic fields. Low values of S_{ij} are evident across the sky, with a maximum value of ≈ 0.49 sources within $\zeta_{smth} = 5°$, indicating little overlap between the source direction distributions calculated for each pair.

Figure 7.6 shows the back-tracked trajectories of the CRs of the example data set shown in Figure 7.5 within the Galactic volume. Strong deflections are experienced by the protons as they pass regions of strong magnetic fields. These deflections serve to shift the arrival directions at Earth by a large angle away from the true origin of the CRs. In addition, the two events which have arrival directions separated from the rest of the data set can be seen to follow very different trajectories within the Galaxy. It is clear from studying these trajectories that the assumptions made for the magnetic spectrometer analysis will not hold for these events.

The remaining data sets for which $\epsilon_{src} > 20°$ were found to display similar properties. Of the 12 data sets exhibiting such large values of ϵ_{rec}, 11 were generated with the ASS-A spiral model. The presence of a squared cosine term in the spiral field component of this



(a)



(b)

Figure 7.3: (a) The arrival directions of the events of an example simulated data set in Galactic co-ordinates. The direction of the simulated source is indicated by a black star, and the arrival directions by triangles coloured according to energy. (b) Smoothed map of S_{ij} gained from applying the magnetic spectrometer analysis to the events shown in (a). The pixel having the largest value is indicated by a black dot, and is located approximately 1.9° from the simulated source direction.

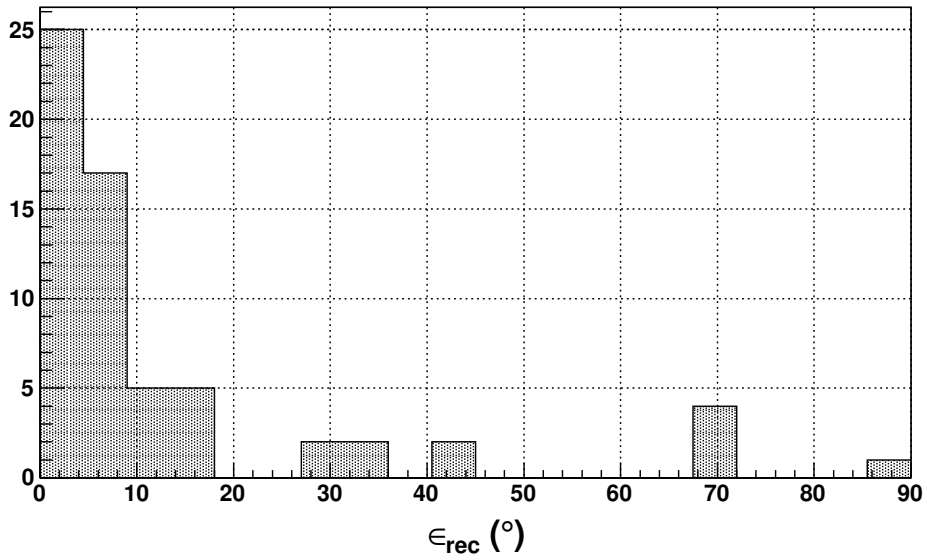
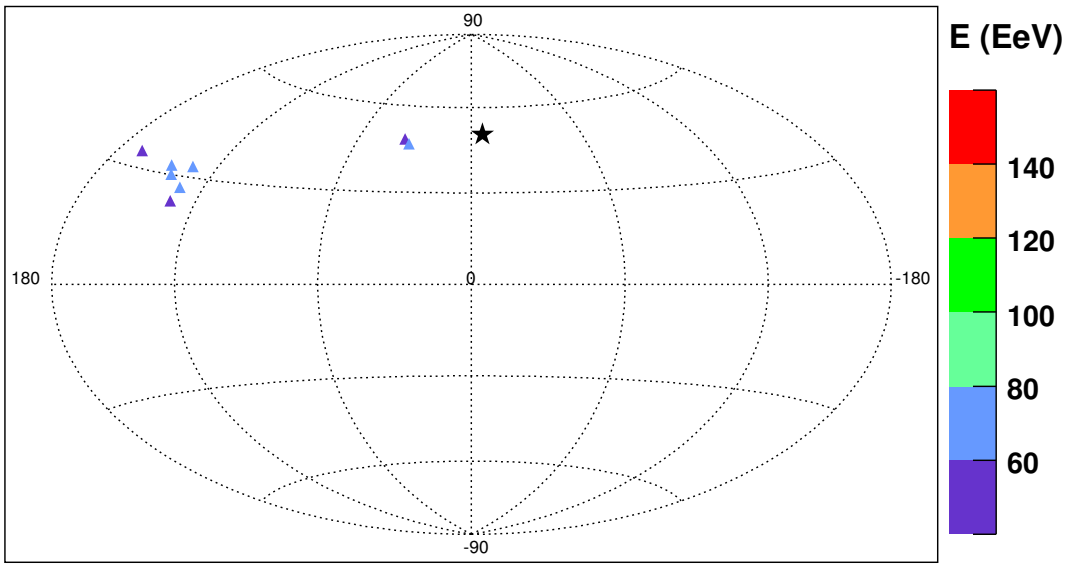


Figure 7.4: A histogram of ϵ_{src}, the angle between the directions of the reconstructed and simulated sources, for signal-only data sets. The largest separation of $\epsilon_{rec} = 142°$ is omitted.

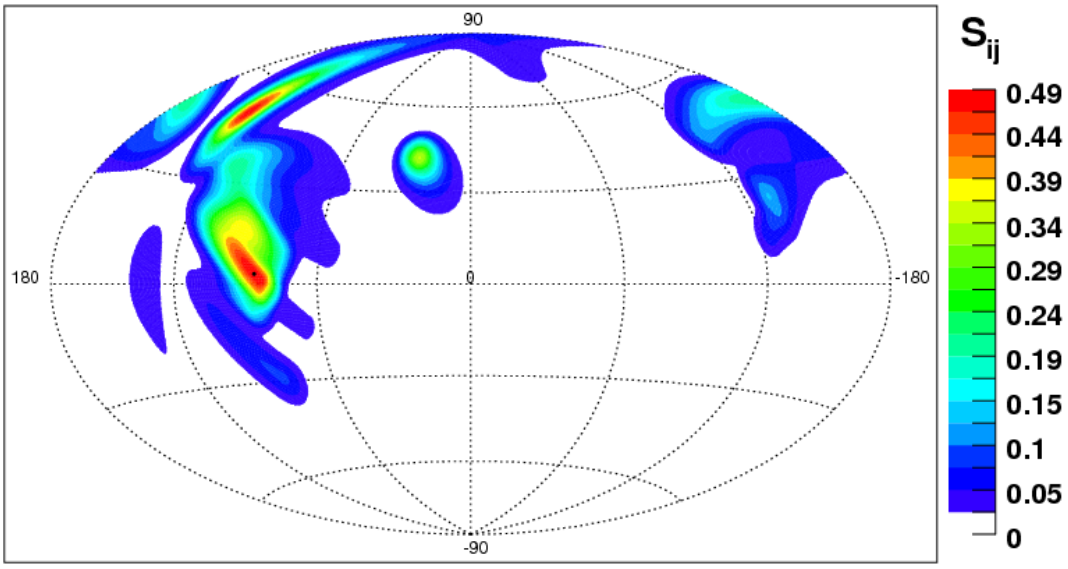
model means that the spiral field strength may reach larger values than the BSS-S field for the same normalisation at the position of Earth, resulting in larger deflections of the particles for particular source directions. The presence of halo fields (7 of the 12 data sets) further complicates the particle trajectories, resulting in regions of sky from which particles may reach the detector, but they do so from a direction that is at a large angle from that of their source.

Although it is known that these 12 data sets contain CRs which have experienced deflections too large to allow an accurate source position reconstruction with the magnetic spectrometer analysis, they are retained in the following analyses. As the magnetic spectrometer assumptions do not apply for these data sets, it is desired that no sources are reconstructed when using the source-finding algorithm described in Section 7.3. Consequently, these data sets are still considered in Sections 7.3 and 7.4 for the purpose of optimising the parameters of the analysis.

Finally, Figure 7.7 shows the angular error in the source position reconstruction as a function of $\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right|$, the magnitude of the integrated magnetic field vector transverse to the line of sight between Earth and each of the 64 source positions. An increase in the average source reconstruction error may be seen with increasing $\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right|$, as expected. Additional factors affecting the value of ϵ_{rec}, such as the maximum event energy in each of the signal data sets, introduce some of the spread between values observed here. However, for $\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right|$ less than approximately $50\,\mu\mathrm{G\,kpc}$, the majority of source directions are reconstructed to within 10°.



(a)



(b)

Figure 7.5: Skymaps of the arrival directions and smoothed S_{ij} map of a signal data set in which the source position is inaccurately reconstructed. The black star indicates the simulated source direction, and the other symbols are as defined for Figure 7.3. (a) Note the large angular separation from the source, and lack of energy ordering, in 6 of the 8 events, as well as the two events which were observed to come from a different region of sky. (b) The highest-value pixel is approximately 89° from the simulated source position. Note the low values of S_{ij} compared to those of Figure 7.3 and the presence of several peaks due to little overlap of reconstructed source position distributions.

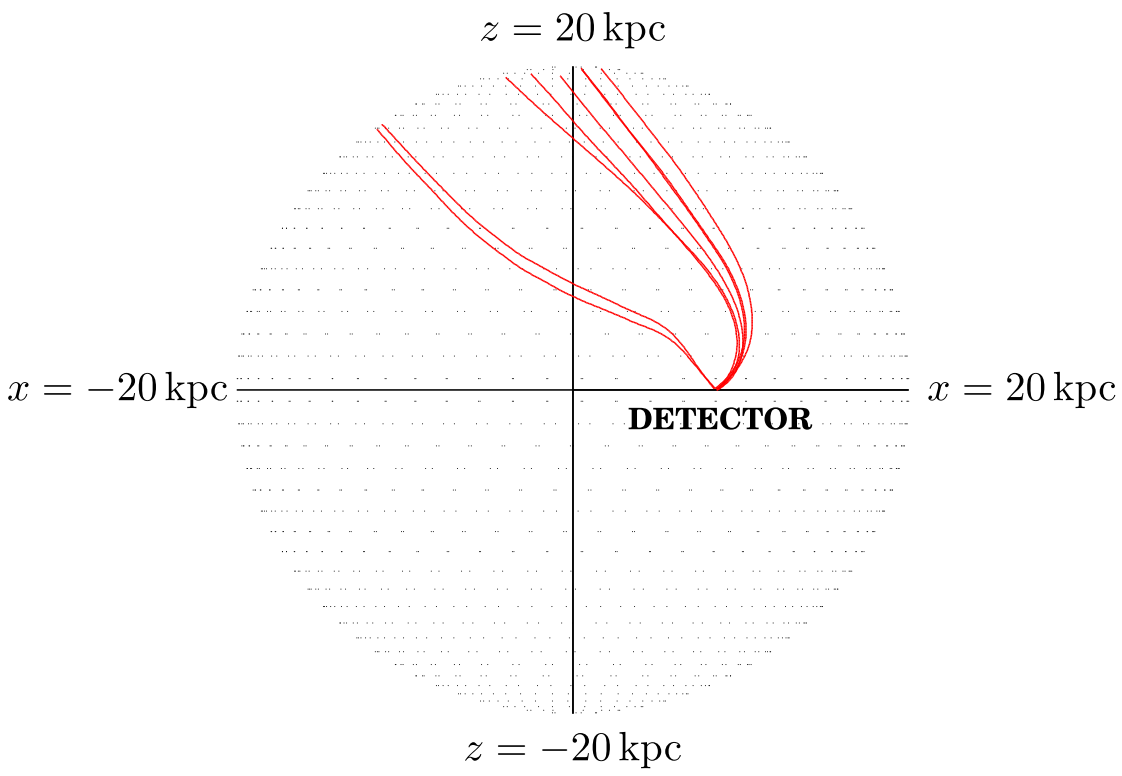


Figure 7.6: The back-tracked trajectories of the events shown in Figure 7.5, projected onto the x-z plane. The red lines indicate the trajectories of each of the 8 protons within the Galactic volume, which is outlined by black dots. The large deflections experienced by the particles may clearly be seen.

7.3 Analysis of Signal and Background Events

Following the analysis of signal-only data sets, the more realistic case of data sets containing both signal and background events is considered. For the analyses presented in this section, a ratio between the number of isotropic background events (n_{iso}) and the number of signal events (n_{sig}) of $n_{iso}/n_{sig} = 6$ is chosen. Such a value corresponds to a bright CR source, such as has been suggested as a possible origin for the cluster of events observed in the Pierre Auger Observatory data (e.g. [105]); approximately 19% of the 69 events published in [100] arrive within an 18° window centred on Centaurus A. In Section 7.4 the effect of varying n_{iso}/n_{sig} is investigated.

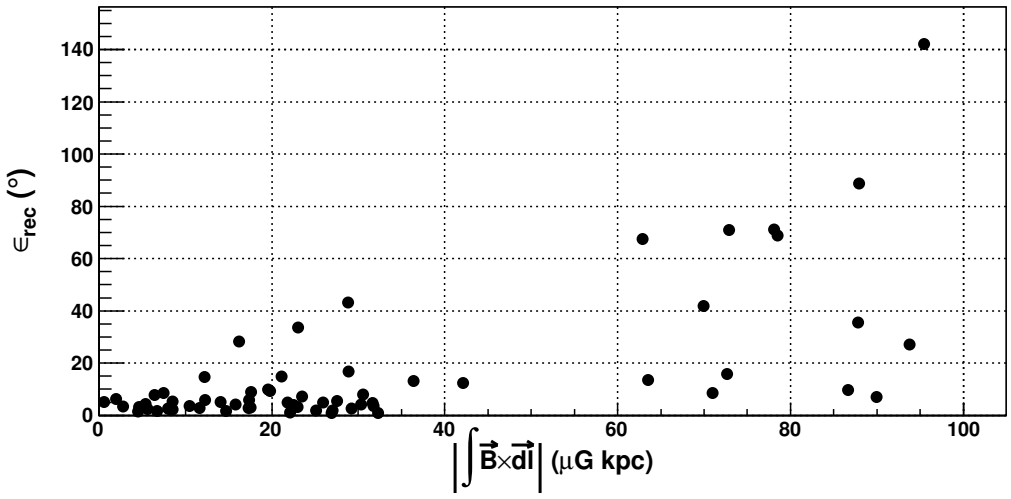


Figure 7.7: Angular error in the reconstructed source direction, ϵ_{rec}, as a function of the magnitude of the integrated magnetic field vector transverse to the line of sight between Earth and each of the 64 source positions. An increase in ϵ_{rec} as $\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right|$ increases may be seen. The gap observed between 40 μG kpc and 60 μG kpc corresponds to lines of sight such as those close to the Galactic Centre and the Galactic poles, which result in geometries from which the collection of a signal data set is difficult.

7.3.1 Peak-Finding Algorithm

As discussed in Chapter 6, energy ordering of UHECR arrival directions is expected to produce a peak in the smoothed map of values S_{ij}. Therefore, an algorithm to search for such peaks is required. While, as previously stated, the number of bright UHECR sources is assumed to be small, the exact number is unknown. To avoid assuming a set number of sources prior to data analysis, an algorithm based on the statistical properties of the S_{ij} map is employed to search for such peaks.

For each smoothed map of S_{ij}, the mean and root mean squared values, denoted as $\langle S \rangle$ and S_{RMS}, respectively, are calculated As each of the n_{pix} pixels of the skymap subtends approximately the same solid angle, $\langle S \rangle$ and S_{RMS} may be calculated directly from the pixel values with no correction required for the pixel size:

$$\langle S \rangle = \frac{1}{n_{pix}} \sum_{ij} S_{ij} \tag{7.1}$$

$$S_{RMS} = \sqrt{\frac{1}{n_{pix}} \sum_{ij} (S_{ij} - \langle S \rangle)^2} \tag{7.2}$$

Subsequently, the value of each pixel may be represented by is deviation from the mean value in terms of S_{RMS}:

$$\alpha_{ij} = \frac{S_{ij} - \langle S \rangle}{S_{RMS}} \tag{7.3}$$

In the studies described in the current section, pixels having a value of α_{ij} greater than a threshold value α_{th} are considered to correspond to a 'source'. Generally, multiple pixels comprising a peak will pass such a threshold cut; in order to avoid many pixels of a single peak being identified as sources, an angular distance of 20° is required between such 'source' pixels. In practice, this value is sufficient to give only a single 'source' pixel per peak.

The values of S_{ij} follow, in general, a distribution that is asymmetric about $\langle S \rangle$. Figure 7.8 contains histograms of the α_{ij} values calculated for two different data sets in which such asymmetry may be clearly seen. As a result, while the interpretation of the statistical significance of α_{ij} is not straightforward, it is used as a convenient parameter for the discrimination of source peaks. As is discussed in Chapter 8, the significance of any peaks observed in real UHECR data may be estimated through comparisons with 'shuffled' data sets.

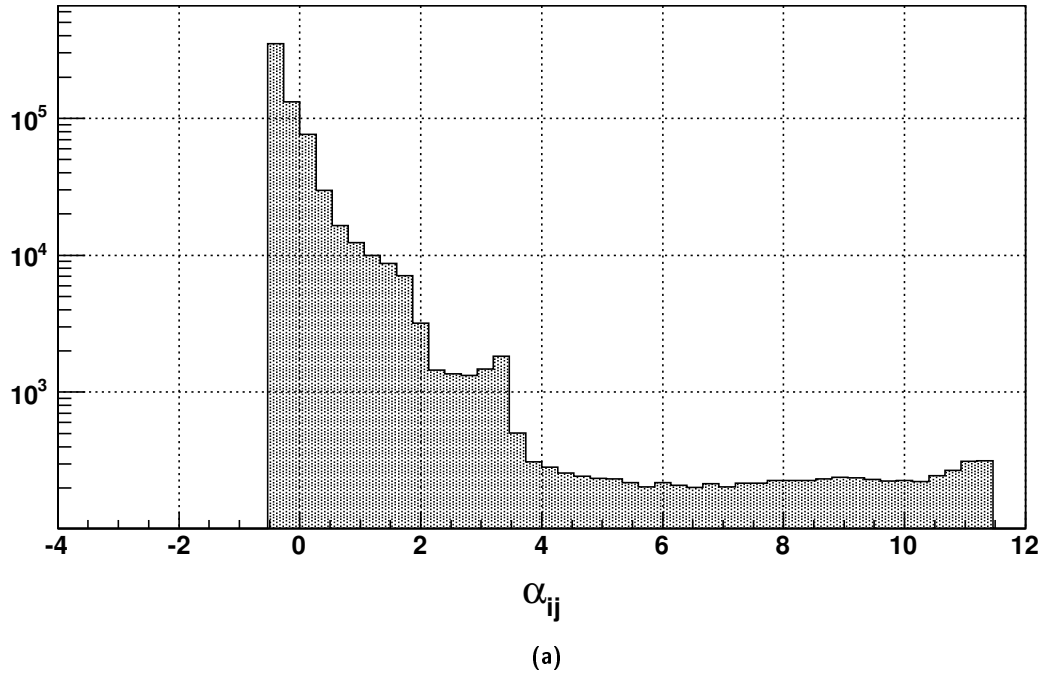
The source-finding algorithm employed here begins with finding the pixel with the maximum α_{ij} value; if it passes the threshold cut, it is classified as a source, otherwise the algorithm ends. In the former case, a 20° window around the source pixel is excluded, and the search continued in this manner until the maximum value found of α_{ij} is less than α_{th}.

Following the results of Section 7.2, a source identified in this manner is considered 'real' - meaning that it is associated with the direction of the simulated source of that data set - if $\epsilon_{rec} < 20°$. Otherwise, it is considered a 'false' source which has been erroneously reconstructed. If multiple sources are reconstructed with $\epsilon_{rec} < 20°$, the closest to the simulated source is taken to be real, and the rest false.

7.3.2 Results of Magnetic Spectrometer Analysis

Due to the possibility of reconstructing false sources in addition to real sources upon the analysis of UHECR data, the effect of varying $\Delta\theta_{cut}$ and ζ_{smth} on the number of each type of reconstructed source is studied in this section. For each combination of $\Delta\theta_{cut}$ and ζ_{smth} considered, all 64 data sets were analysed individually and the total number of reconstructed real and false sources calculated. In addition, two values of α_{th} were considered: $\alpha_{th} = 7S_{RMS}$ and $\alpha_{th} = 8S_{RMS}$. These results are plotted in Figure 7.9. In addition, Figure 7.10 displays the mean values of ϵ_{rec}, $\langle \epsilon_{rec} \rangle$, for the reconstructed real sources for each combination of parameters. It can be seen that $\langle \epsilon_{rec} \rangle$ is not strongly affected by the parameters used for the magnetic spectrometer analysis, reflecting the fact that the distribution of ϵ_{rec} for the reconstructed real sources is similar regardless of the parameters of the analysis.

General trends in the number of real and false reconstructed sources may be observed as the parameters of the analysis are varied. The effect of increasing α_{th} is to decrease the number of reconstructed sources, both real and false. However, the relative decrease in the number of false sources reconstructed is, in all cases, greater than the corresponding relative decrease in the number of real sources. It is of interest to note that no values of $\Delta\theta_{cut}$ and ζ_{smth}



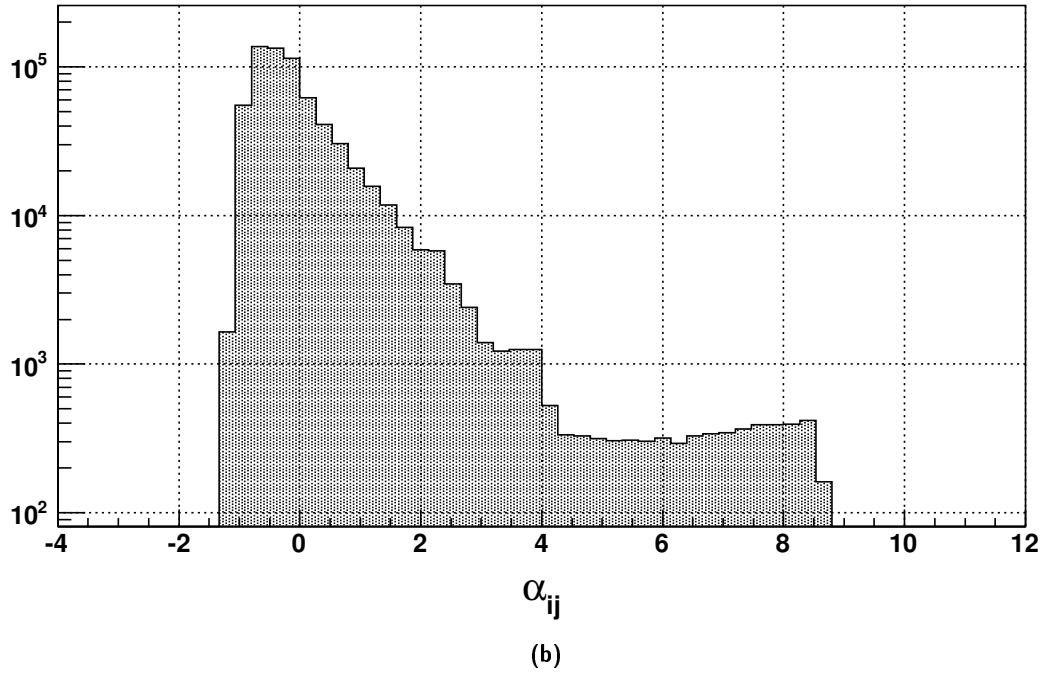


Figure 7.8: Histograms of α_{ij} values following the analysis of (a) data containing 8 signal events and 48 background events, using parameters $\Delta\theta_{cut} = 20^\circ$ and ζ_{smth}=10° and (b) data containing 8 signal events and 160 background events, using the same parameters as (a). Note the asymmetry of both distributions, and the truncation of (a) at $\alpha_{ij} \approx -0.5$ due to the value of S_{ij} being positive by definition.

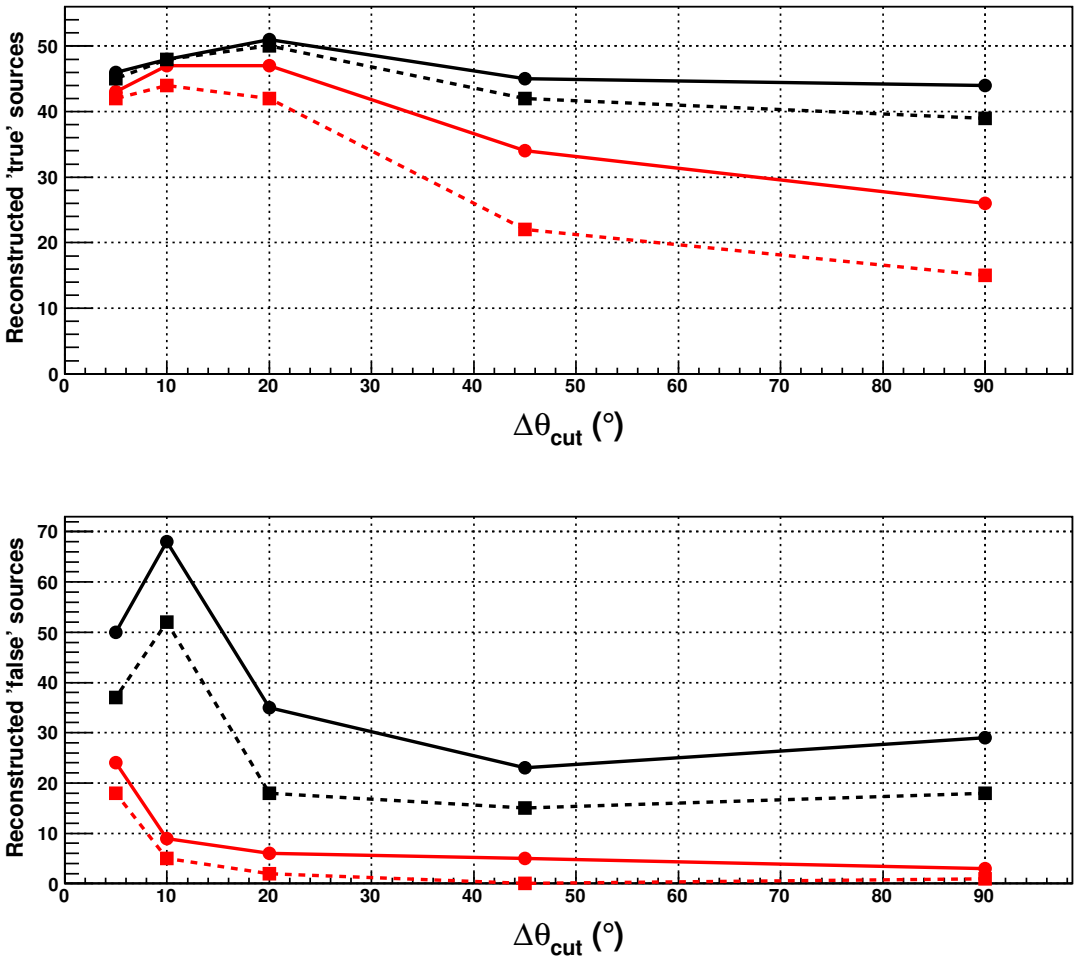


Figure 7.9: Graphs of the number of reconstructed real and false sources (described in the text) after applying the magnetic spectrometer analysis to the data sets for which $n_{iso}/n_{sig} = 6$, as a function of $\Delta\theta_{cut}$. Black and red lines correspond to the results of 5° and 10° smoothing, respectively, and solid and dotted lines to threshold cuts of $7S_{RMS}$ and $8S_{RMS}$, respectively. Following the results of Section 7.2, it is expected that a maximum of 52 real sources may be reconstructed, as the deflections of the cosmic rays from the remaining 12 are too large to satisfy the magnetic spectrometer assumptions.

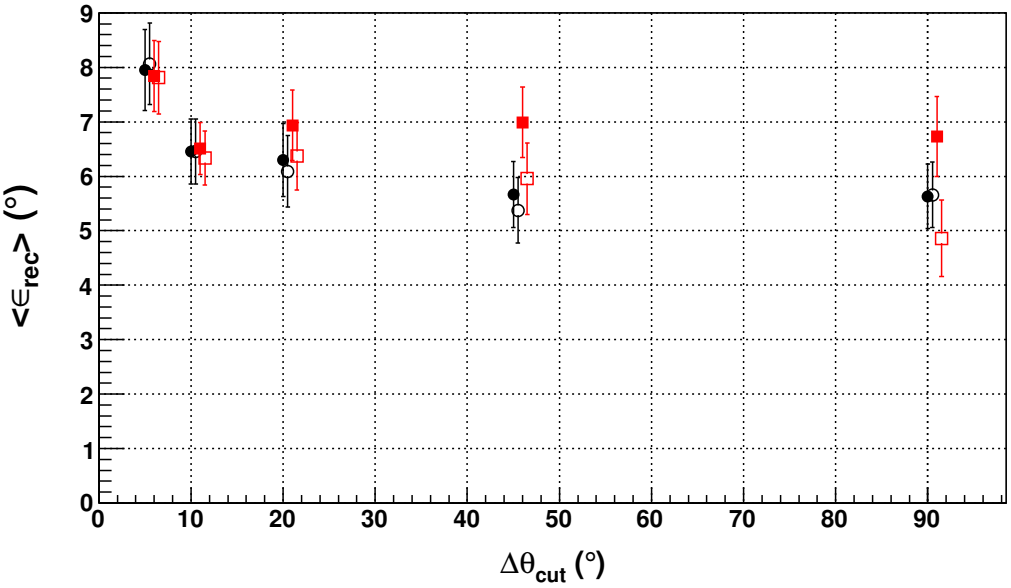


Figure 7.10: The mean values of the reconstructed source error, $\langle \epsilon_{rec} \rangle$, for the reconstructed real sources of the data sets for which $n_{iso}/n_{sig} = 6$. Black circles and red squares correspond to $\zeta_{smth} = 5°$ and $\zeta_{smth} = 10°$, respectively, while filled and hollow symbols correspond to $\alpha_{th} = 7S_{RMS}$ and $\alpha_{th} = 8S_{RMS}$, respectively. Error bars represent the standard error of the mean. For clarity, a small horizontal offset has been introduced between symbols at a given value of $\Delta\theta_{cut}$. The value of $\langle \epsilon_{rec} \rangle$ does not exhibit a strong dependence on the set of parameters used for the analysis.

considered resulted in the reconstruction of 52 real sources, which is the maximum number expected given the results of Section 7.2. With the addition of background events, some signal peaks are not large enough to be identified as sources using the threshold cuts applied here, due to only a minor overlap of source direction distributions.

Increasing ζ_{smth} also has a consistent effect across all values of $\Delta\theta_{cut}$. In all cases, the number of real and false sources reconstructed using $\zeta_{smth} = 10°$ is less than for $\zeta_{smth} = 5°$ for a given value of α_{th}. However, at $\Delta\theta_{cut} = 20°$, both values of ζ_{smth} give a similar number of real sources (for $\alpha_{th} = 7S_{RMS}$, $\zeta_{smth} = 5°$ results in the reconstruction of 51 real sources while $\zeta_{smth} = 10°$ results in 47), while $\zeta_{smth} = 10°$ results in fewer false sources being reconstructed (6 compared to 35 for $\zeta_{smth} = 5°$). Thus, for an appropriate value of $\Delta\theta_{smth}$ (see below), ζ_{smth} may be chosen to reduce the number of false sources reconstructed at only a small cost to the number of real sources reconstructed.

Consider now the behaviour of the curves of the number of reconstructed 'real' events with changing $\Delta\theta_{cut}$. It is expected that the real source positions will be reconstructed effectively when $\Delta\theta_{cut}$ is comparable to the characteristic angular spread of signal events from their point of origin (denoted here as ψ). For the 52 data sets in which the source position can be reconstructed, the maximum observed separation between the 8 arrival directions has a mean

of $\psi \approx 14°$.

If $\Delta\theta_{cut} \gg \psi$, then in addition to all possible pairs of signal events being formed, a large number of background events are also paired together. This gives a correspondingly large number of background source direction distributions, raising the value of $\langle S \rangle$ and decreasing the significance of signal peaks (which remain approximately constant). Decreasing the value of $\Delta\theta_{cut}$, while it remains greater than ψ, reduces the number of background pairs being formed while retaining the size of the signal peaks; consequently, the significance of the latter increases. It is expected that a value of $\Delta\theta_{cut} \sim \psi$ will return the best signal-to-noise ratio, where all of the signal events are paired but few background pairs are formed. This effect causes the increase in the number of 'real' sources being reconstructed that is observed as $\Delta\theta_{cut}$ decreases from 90° to 20° in Figure 7.9.

Decreasing $\Delta\theta_{cut}$ below ψ means that all possible pairs of signal events are no longer being formed. As a result, the size of the signal peak decreases in addition to the number of background pairs becoming fewer. Real source peaks belonging to signal events that are spread over a large angle reduce in value more quickly than those for which the signal events are more tightly-clustered, causing the gradual reduction in the number of real sources observed for $\Delta\theta_{cut} < 20°$.

The decrease in the number of false sources found for $\zeta_{smth} = 5°$ as $\Delta\theta_{cut}$ changes from 10° to 5° is due simply to statistics; a group of 48 isotropically-distributed events on the sky will only have a small number of pairs being formed (with an average of ≈ 2) which serves to limit the possible number of background sources that could be found.

From these results, the combination of parameters $\Delta\theta_{cut} = 20°$, $\zeta_{smth} = 10°$ and α_{th}=$7S_{RMS}$ is taken to be the best of those considered here. This combination results in 47 real sources and 6 false sources being found. While these parameters do not give the highest ratio of the number of real reconstructed sources to the number of false ones, this combination is taken to be a suitable compromise between reconstructing a large number of real sources and keeping the number of reconstructed false sources small. These values are used in the remainder of analyses in this chapter, and are used as a guide for the analysis of real data, as described in Chapter 8.

7.3.3 Estimation of Magnetic Fields

For the data sets in which a source is reconstructed, an estimate of $\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right|$ - the magnitude of the magnetic field vector perpendicular to the line of sight, integrated between the detector and the source - may be made directly from the data. The use of simulated data sets allows the accuracy of such an estimate to be determined through comparison of the estimated and true values.

Equation 6.2, which was derived for the ideal magnetic spectrometer, contains the product $D\,|\boldsymbol{B}|$, where D is the distance to the source and $|\boldsymbol{B}|$ the magnitude of the uniform magnetic field through which the particles passed. In a non-uniform field, the quantity $D\,|\boldsymbol{B}|$ may be

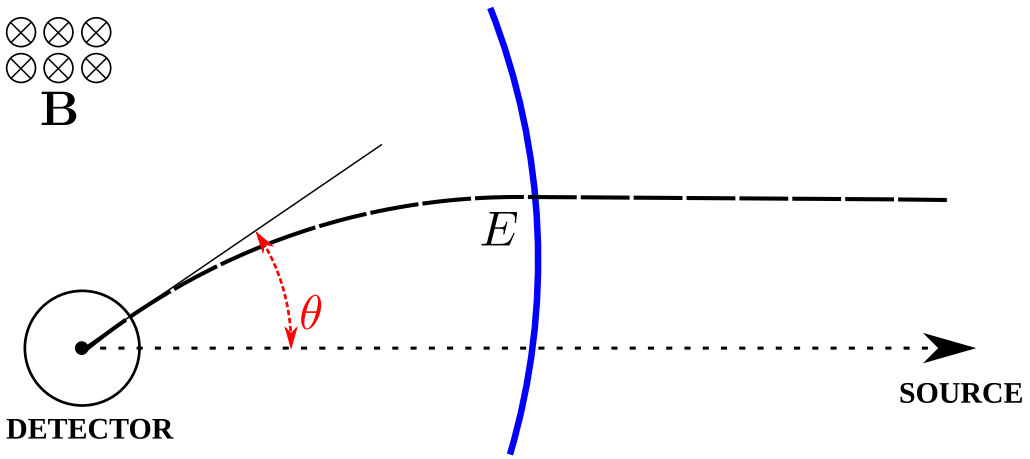


Figure 7.11: Relating the value of $\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right|$ to the fitted value of P requires the consideration of a different geometry to that of the ideal magnetic spectrometer as shown in Figure 6.6. Here, deflections are experienced only within the Galactic volume (the border of which is indicated by a blue arc). Outside the Galaxy, the vector from the source to the detector and from the source to the entry point of the particle into the Galaxy are approximately parallel, owing to the large source distance. Thus, the only contribution to the particle deflection is due to the GMF (approximated to act as a uniform field) which leads to Equation 7.4.

approximated as $D\,|\boldsymbol{B}| \approx \left|\int_0^D \boldsymbol{B} \times d\boldsymbol{l}\right|$ under the magnetic spectrometer assumption. This gives:

$$\sin\theta \approx \left|\int_0^D \boldsymbol{B} \times d\boldsymbol{l}\right| \frac{ceZ}{E} \tag{7.4}$$

Note that the factor of 2 that was present in Equation 6.2 has been removed - the value of $\left|\int_0^D \boldsymbol{B} \times d\boldsymbol{l}\right|$ is no longer twice that which causes the deflection θ as it was for the ideal spectrometer, due to the assumption that only deflections within the Galaxy are significant. Figure 7.11 illustrates the situation considered here.

Thus, if a source is found for a group of events, fitting the function:

$$\sin\theta = \frac{P}{E} \tag{7.5}$$

for the parameter P allows the estimation of $\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right|$ (hereafter the limits are omitted from the integral). In order to do so, the events associated with a reconstructed source peak in a map of S_{ij} must be identified. For this work, a pair is considered to be associated with a reconstructed source (meaning that the source direction distribution derived for that pair has significantly contributed to the peak found in the S_{ij} map) if the nominal source direction - the single reconstructed source direction calculated using the nominal energies and arrival directions of the events in that pair - is located within ζ_{smth} of the peak.

While the signal events are expected to be associated with the reconstructed source direc-

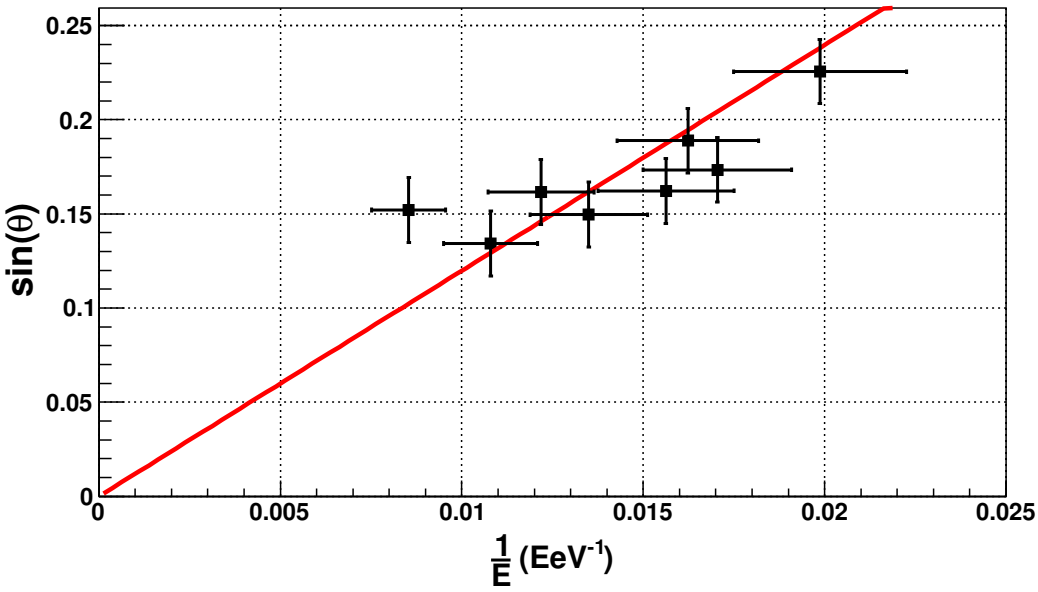


Figure 7.12: An example of a fit of Equation 7.5 to a set of signal events in order to find P. The red line indicates the fitted function.

tion if the magnetic spectrometer assumptions hold, it is also possible that the nominal source direction of a background pair will fall within ζ_{smth} of the reconstructed source by chance. Such events are not expected to display energy ordering that is consistent with that of the signal events, as such consistency is not enforced because of the independent consideration of event pairs. The presence of any such background events will introduce errors in the reconstructed value of P. While not incorporated into the work presented here, the presence of incompatible events could be determined through methods such as least trimmed squares regression, which can identify and remove outliers during a fit to data [178]. An alternative method is rejection of events on the basis of their azimuthal angle around the source position; it may be expected that signal events will lie within a sector of the sky, and events contributing to the source that lie outside that sector considered to belong to the background. Such methods are left for future studies, and the estimation of $\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right|$ considered in this section is performed by allowing only signal events to be considered associated with reconstructed source directions.

Consequently, using the data sets for which a real source direction is reconstructed, a fit is performed to the signal events associated with that source using their angular deviations from its position and the inverse of those events' energies. An example of such a fit is shown in Figure 7.12.

Subsequently, the value of the integrated magnetic field is calculated through (taking $Z = 1$ for protons):

$$\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right| = \frac{P}{ce} \tag{7.6}$$

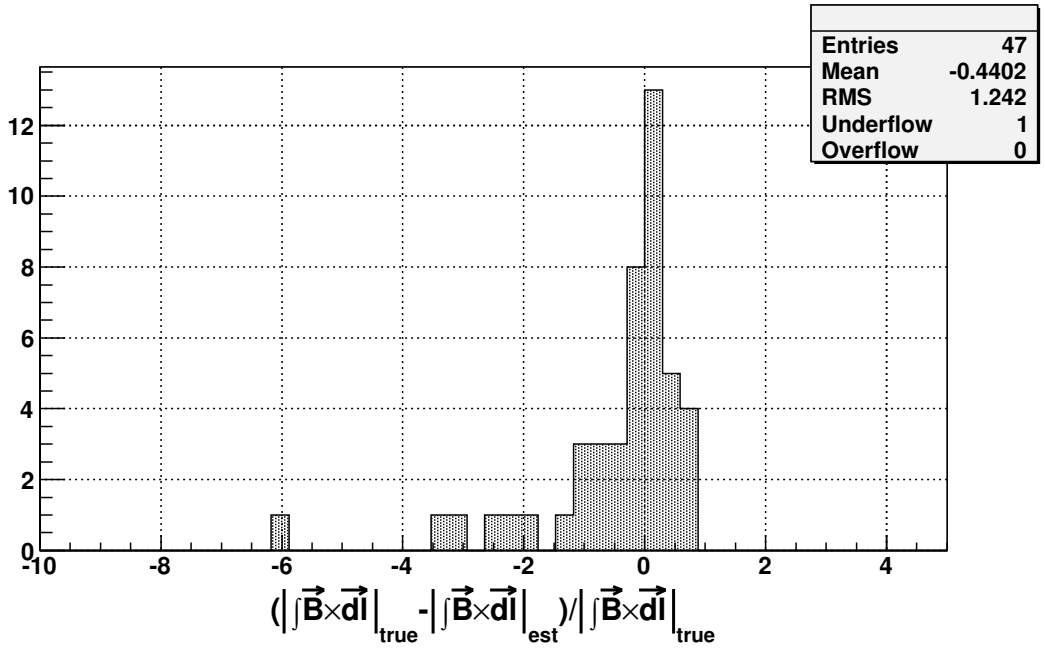


Figure 7.13: The relative error between the true and estimated values of $\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right|$. The estimated values are derived from the fitted value of P in Equation 7.5 as explained in the text. Although the distribution peaks near zero, a tail of negative values is observed.

and is compared with the true value of $\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right|$ calculated from explicitly performing the integral between the detector and the simulated source position for the regular component of the relevant GMF model. The results of this analysis for all 47 reconstructed real sources found using $\Delta\theta_{cut} = 20°$, $\zeta_{smth} = 10°$ and $\alpha_{th} = 7S_{RMS}$ are shown in Figure 7.13.

While the distribution of the relative errors peaks close to zero, it is skewed towards negative values - indicating a tendency to overestimate the value of $\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right|$. Particularly large errors in the estimates of $\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right|$ are found to be due to significant variation in the true value of $\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right|$ along the line of sight to the source. Such variation is a result of large-scale reversals in the regular component of the GMF for some lines of sight. As the CRs do not travel along this line of sight to reach the detector, the value derived from a fit to the event arrival directions may differ greatly from the value found through explicitly calculating $\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right|$.

Restricting the analysis to reconstructed sources for which $\epsilon_{rec} < 5°$, the distribution of relative errors becomes much narrower; Figure 7.14 displays the distribution for the 18 data sets in which such accuracy was achieved. Consequently, the accuracy of the estimation of $\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right|$ is seen to be strongly affected by the accuracy in the reconstruction of the corresponding source position.

An estimate of the limit of the accuracy of this method may be made by fitting for P once more, but using the direction of the simulated source rather than the reconstructed source to

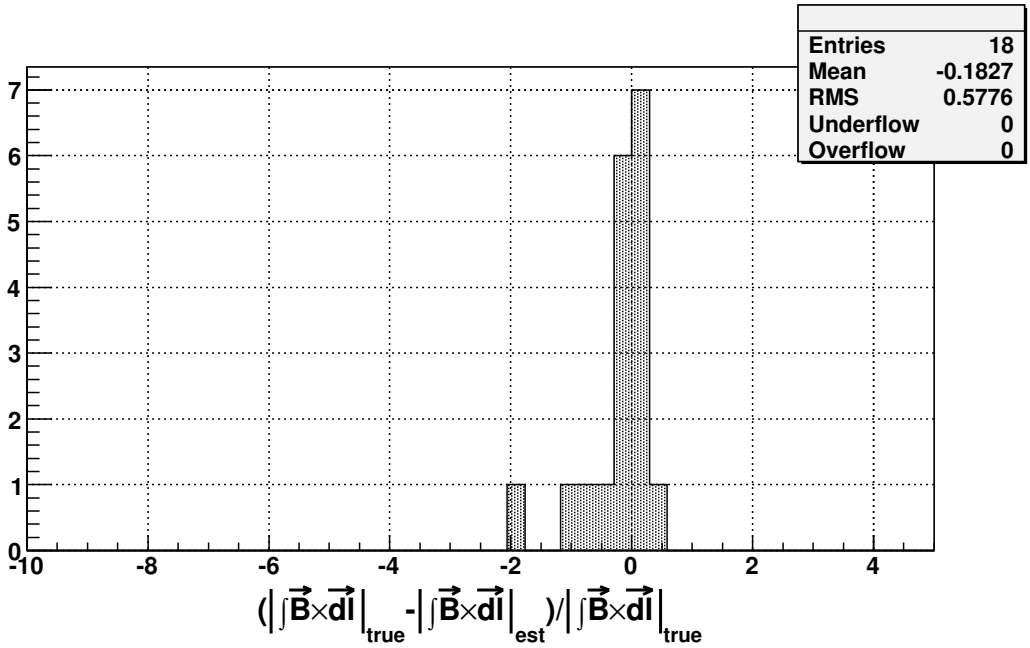


Figure 7.14: Relative errors in the estimation of $\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right|$ for the data sets in which $\epsilon_{rec} < 5^{\circ}$. By comparison to the distribution shown in Figure 7.13, it may be seen that errors in the reconstruction of the source position greatly affect the estimation of $\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right|$.

calculate the values of θ for each event. This serves to remove the error introduced by inaccuracies in the source position reconstruction (effectively taking $\epsilon_{rec} = 0$ for all reconstructions), leaving only the error resulting from the approximations made in the magnetic spectrometer method. These results are displayed in Figure 7.15. The distribution is narrower still than in Figure 7.13, with a width of approximately 30% for the peak around zero. Again, one may see that improving the accuracy of the reconstruction of source directions directly impacts the accuracy of magnetic field estimates.

7.4 Varying Relative Source Brightness

To gather the results presented in Section 7.3, an assumption of the ratio between the numbers of background and source events was required. Here, the effect of varying this ratio is investigated. As discussed in Section 7.3, values of $\Delta\theta_{cut} = 20^{\circ}$, $\zeta_{smth} = 10^{\circ}$ and $\alpha_{th} = 7S_{RMS}$ are used in this section. While such values may not necessarily give the optimal signal-to-noise ratio for different values of n_{iso}/n_{sig}, their use is sufficient to gain an understanding of the sensitivity of the magnetic spectrometer analysis to this ratio.

Additional data sets were produced for this work by simply varying the number of background events that were added to the signal events. The magnetic spectrometer analysis was subsequently applied, and the number of real and false sources calculated as in Section 7.3.

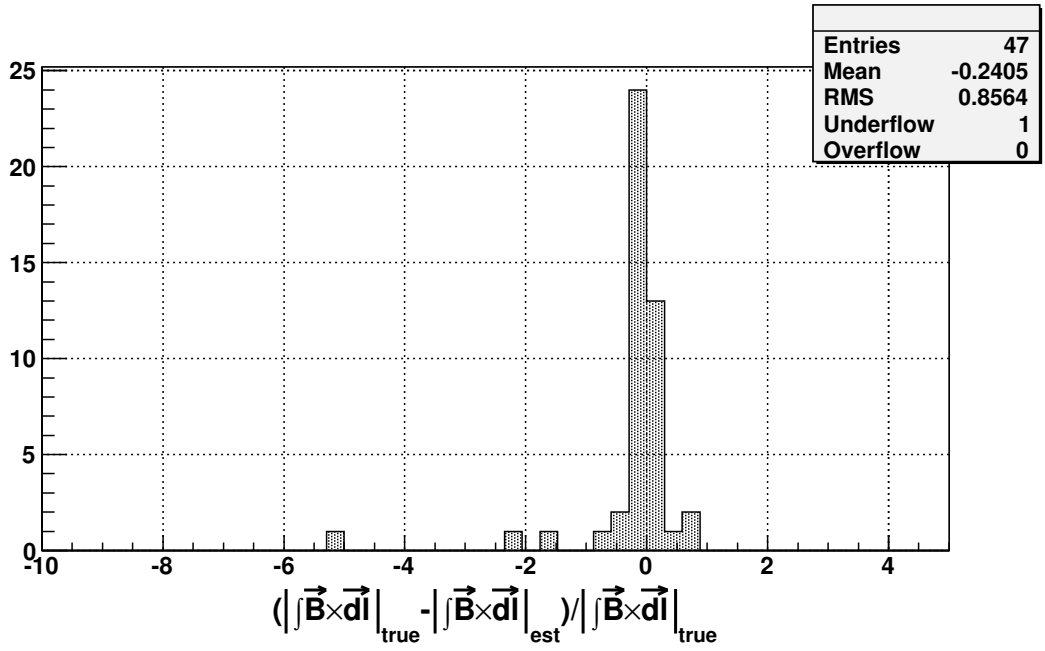


Figure 7.15: Relative errors in the estimation of $\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right|$ when the simulated source direction is used instead of the reconstructed direction. Although a tail of large negative values still exists, in most cases the accuracy is greatly improved over the estimates derived from the latter.

These results are shown in Figure 7.16.

It may be seen that the ability to reconstruct the real source position degrades as the fraction of background events increases. This is a result of both $\langle S \rangle$ and S_{RMS} increasing as n_{iso} increases. The value of the source peak does not increase (unless background pairs produce an overlapping source distribution by chance) making its discrimination using a value of $\alpha_{th} = 7S_{RMS}$ less likely. At the maximum ratio considered ($n_{iso}/n_{sig} = 20$), 28 of the expected maximum of 52 real sources are reconstructed. Although this number may vary for different choices of $\Delta\theta_{cut}$, ζ_{smth} and α_{th}, $n_{iso}/n_{sig} = 20$ (for which the data consisted of 168 events) is indicative of the relative source brightness below which the reconstruction of a real source becomes unlikely for the GMF models considered here.

The number of false sources that are reconstructed also exhibits a decrease as the number of background events increases. Again, this follows from increasing values of $\langle S \rangle$ and S_{RMS} with increasing n_{iso}/n_{sig}; a large number of background pairs are required to produce overlapping source direction distributions to pass the threshold cuts, which does not often occur by chance in the data sets considered.

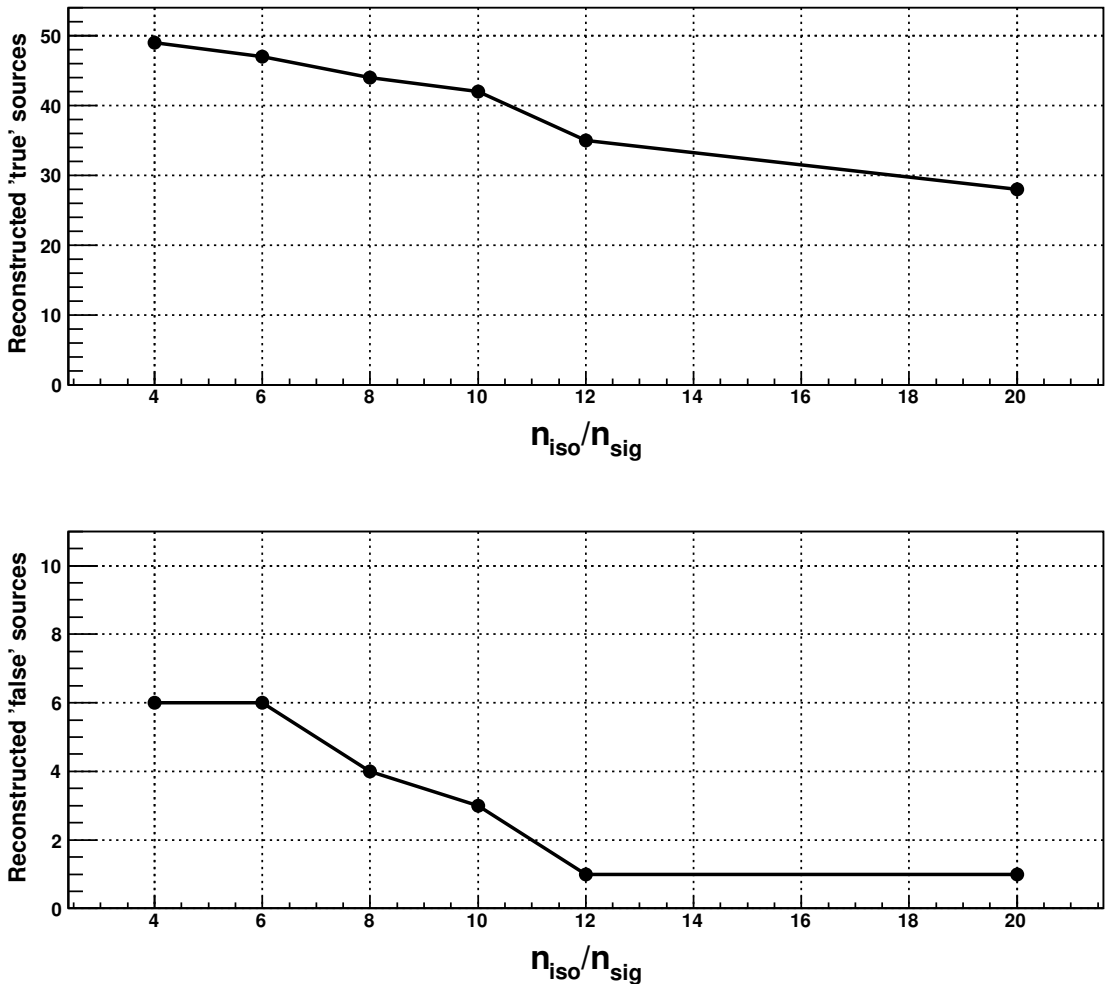


Figure 7.16: The number of reconstructed real and false sources as a function of n_{iso}/n_{sig}. The gradual decrease in the numbers of both types of sources that are reconstructed may be seen. At $n_{iso}/n_{sig} = 20$, only 28 real sources are reconstructed of the maximum of 52 that is expected.

7.5 Mixed Composition Source Flux

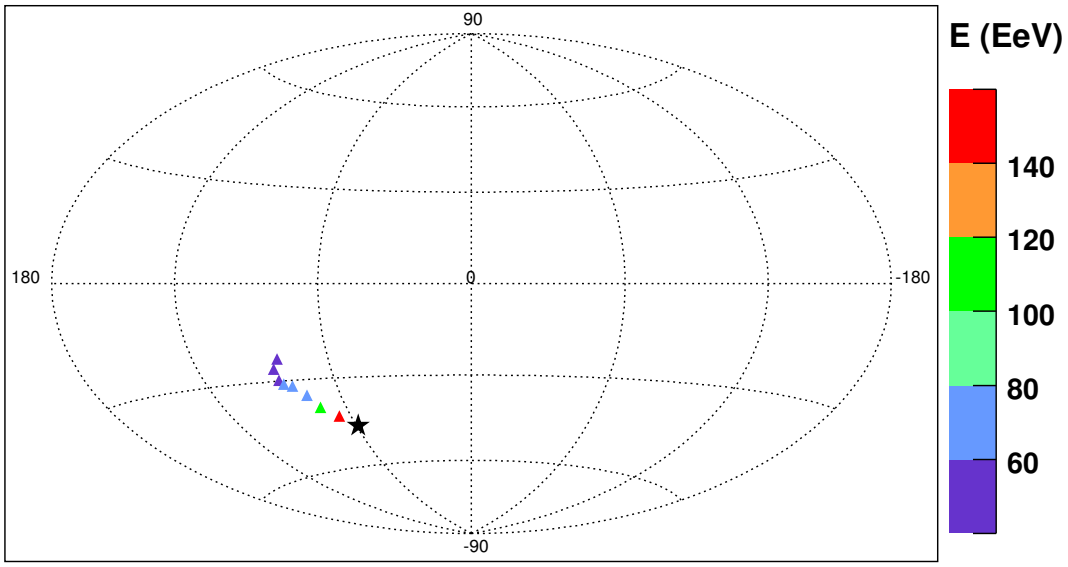
The final study presented here investigates the effect of the signal events being different species. The data sets studied in this section are different to those of the previous sections, but were produced in a similar manner. From each source, 8 protons and 4 iron nuclei were required to be detected. The magnetic spectrometer analysis is subsequently applied to a source flux of either 8 protons, or 4 protons and 4 iron nuclei. This allows a comparison between the analysis of pure and mixed composition data. A total of 22 data sets were produced in this manner. In contrast to the work presented in Section 7.3, an equal number of data sets produced with each magnetic field configuration was not required here, as only the relative change in the number of reconstructed sources in the case of pure and mixed compositions is of interest.

Figure 7.17 displays the arrival directions of the signal events from one of the sources. Energy ordering is clearly visible in the case of all signal events being protons, while the arrival directions of the iron events from the source do not exhibit such structure.

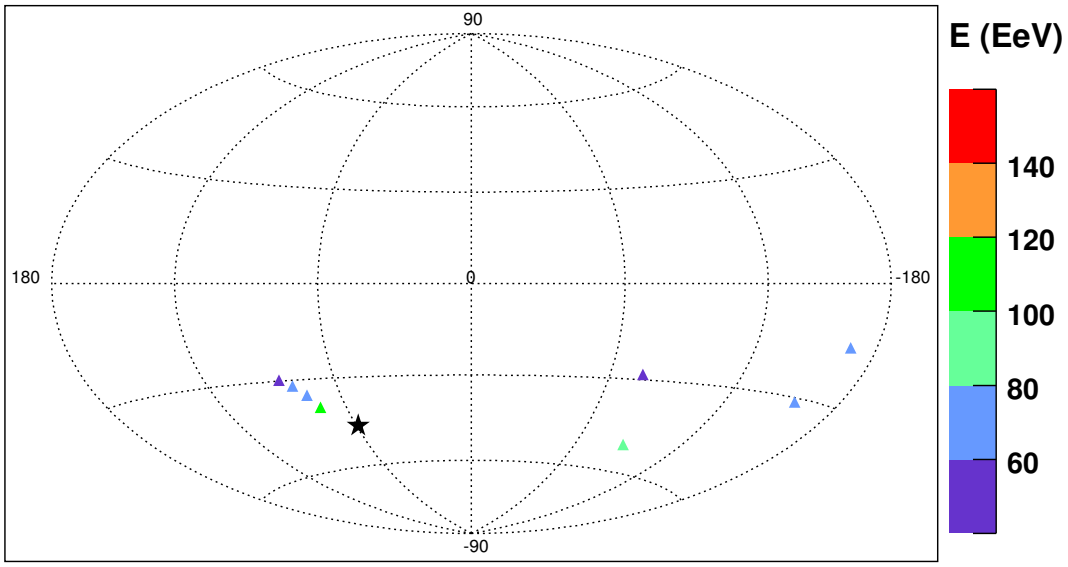
The result of such a lack of energy ordering in the arrival directions of the iron events serves to greatly reduce the size of each signal peak. In the case of n events in an ideal spectrometer, the source will be produced by $\frac{1}{2}n(n-1)$ overlapping distributions. Thus, a flux of 8 protons exhibiting energy ordering in their arrival directions will give 28 overlapping distributions; for the case of 4 protons and 4 iron nuclei, only the former would be expected to contribute to the signal peak which will consequently be composed of 6 such distributions.

This behaviour is reflected in Figure 7.18, which displays the results of applying the magnetic spectrometer analysis, again using the optimal parameters, to the data sets described above. The relative decrease in the number of real sources reconstructed when a mixed composition is present is dependent on the relative source brightness. In the case of $n_{iso}/n_{sig} = 4$, 4 protons are still capable of producing a source peak larger than α_{th} in 10 cases, compared to 16 for the pure proton case, while for $n_{iso}/n_{sig} = 12$, only a single data set contains a reconstructed source peak.

Consequently, while the magnetic spectrometer analysis is capable of reconstructing the positions of UHECR sources if a light component is present in the flux of those sources, for a given source luminosity the sensitivity of the analysis is reduced in the case of a heavy component to that flux. The sensitivity of the magnetic spectrometer analysis will depend on both the relative fraction of the light component in the source flux, as well as the brightness of the source compared to the all-sky UHECR flux. However, as the size of the signal peak will approximately scale as $\frac{1}{2}n(n-1)$ where n is the number of energy-ordered events, a very bright source may still be reconstructed even if a large component of its flux comprises heavy nuclei.



(a)



(b)

Figure 7.17: Skymaps showing the arrival directions of signal events of (a) a pure proton flux and (b) half proton, half iron flux, in both cases originating at the same source. The lack of directionality in the case of iron nuclei may be clearly seen. Symbols are as defined for Figure 7.3.

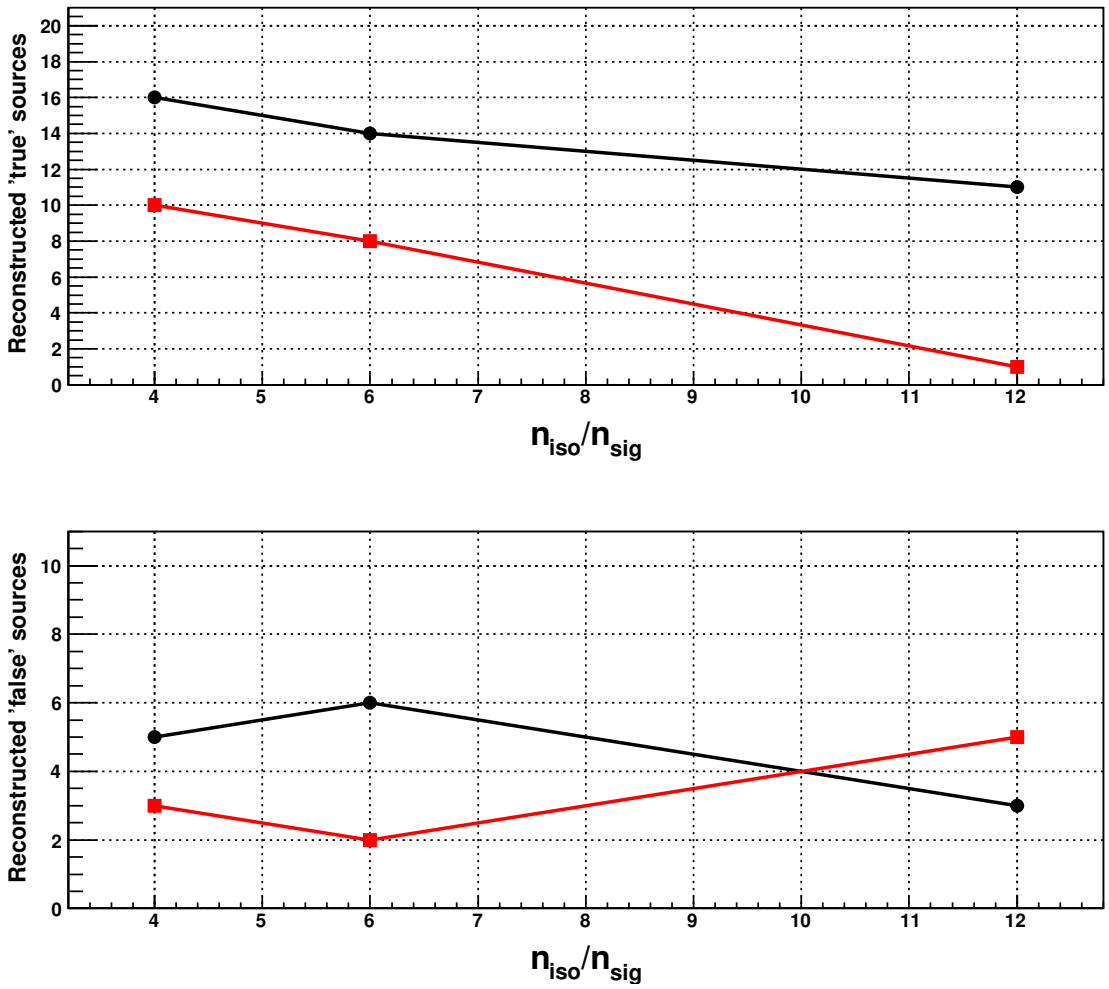


Figure 7.18: The number of real and false reconstructed sources for varying n_{iso}/n_{sig}. The black lines correspond to the data sets containing signal events which are all protons, while the red lines correspond to the data sets in which half the signal events are protons and half are iron nuclei.

Chapter 8

Analysis of Pierre Auger Observatory Data

In Chapter 7, it was demonstrated that, if signal events having energy ordered arrival directions were present in a UHECR data set, the source position for those signal events could be found with good accuracy and reliability. The magnetic spectrometer analysis is now applied to data collected by the Pierre Auger Observatory.

It is important to note that the analyses presented in this chapter are *a posteriori*. Consequently, a statistical significance for any result obtained cannot be claimed. Possible evidence of energy ordering is searched for but only by the application of an identical analysis to an independent data set can the significance of any such result be quantified.

In Section 8.1 the results of applying the magnetic spectrometer analysis to the complete UHECR data set published by the Pierre Auger Collaboration in 2010 [100] are discussed. In Section 8.2, this data set is divided into three smaller sets to investigate the evolution with time of the results of the analysis. The results of both Sections 8.1 and 8.2 are discussed in Section 8.3. Finally, Section 8.4 provides a summary of the main findings of the work presented in this thesis, in addition to possible expansions and enhancements to the magnetic spectrometer analysis which could be incorporated in future studies.

8.1 Analysis of Full Data Set

The analysis of data from the Pierre Auger Observatory begins with the publicly-available data set published in 2010 [100]. This data set consists of 69 events, of energies greater than 55 EeV and zenith angles less than 60°, that were detected between 1$^{\text{st}}$ January 2004 and 31$^{\text{st}}$ December 2009. Each event is required to satisfy the T5 ICRC2005 SD trigger requirement (see Section 3.2.2.3) to ensure a high-quality reconstruction of the properties of the primary particle. A skymap of these data is shown in Figure 8.1.

Although future work could study the effect of varying the energy threshold used to select the data set, a single value of 55 EeV is used here. This is the threshold which was found to

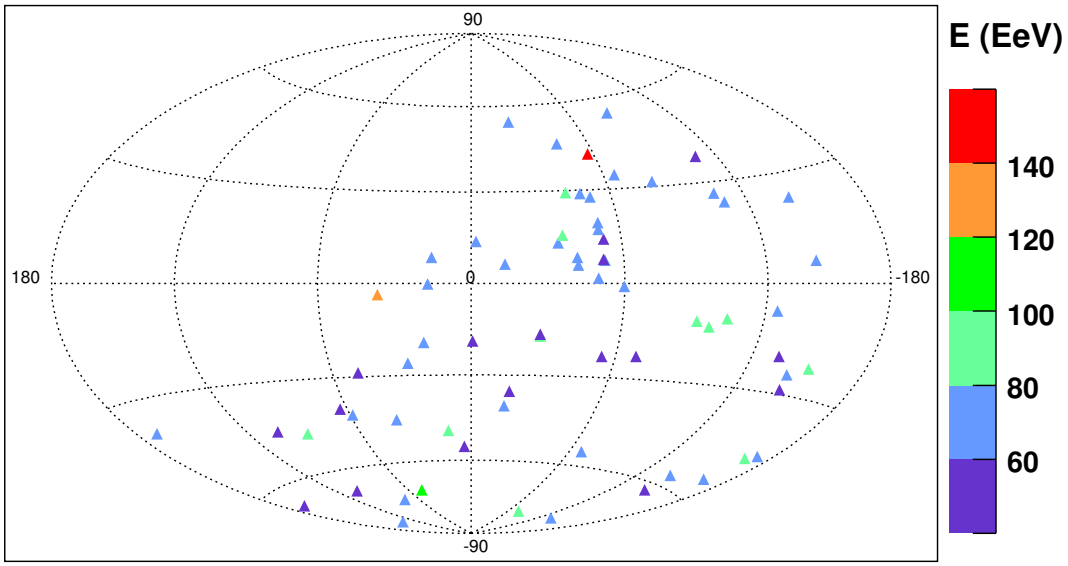


Figure 8.1: Skymap in Galactic co-ordinates of the arrival directions and energies of the 69 events above an energy of 55 EeV detected by the Pierre Auger Observatory between 2004 and 2009. The arrival direction of each event is indicated by a triangle coloured according to the energy of the event.

maximise the correlation between the arrival directions of events detected by the Pierre Auger Observatory and local AGN in [98, 99], after adjusting for a shift in the energy calibration. As this correlation could possibly be interpreted as evidence for a local origin of these UHECR, the use of the data set as described above is reasonable and the assumptions made for the magnetic spectrometer analysis may plausibly hold for at least some of these events.

The magnetic spectrometer analysis was applied to the data set 21 times, each time varying the combination of analysis parameters $\Delta\theta_{cut}$ and ζ_{smth}. It was shown in Chapter 7 that values of $\Delta\theta_{cut} = 20°$ and $\zeta_{smth} = 10°$ enabled good reconstruction of the source directions of simulated data while keeping the number of false sources found small. These values were, thus, used as a starting point for the analysis of the Pierre Auger Observatory data. For the analysis of this section, $\Delta\theta_{cut}$ was varied between 15° and 45° in increments of 5°, while ζ_{smth} took a value of either 5°, 7.5° or 10°.

An example of the results of this analysis is shown in Figure 8.2. In all cases, a prominent peak in the resulting S_{ij} map, which is the map of the integral number of reconstructed sources within the angle ζ_{smth} of each point on the sky, is observed near the Galactic co-ordinates $(l, b) = (-60°, 60°)$. The exact position of the maximum of the peak varies within a window of approximately 10° radius as the analysis parameters are changed. The presence of such a peak, and the consistency in its location, across many different combinations of the analysis parameters may be indicative of energy ordering within the arrival directions of the data set.

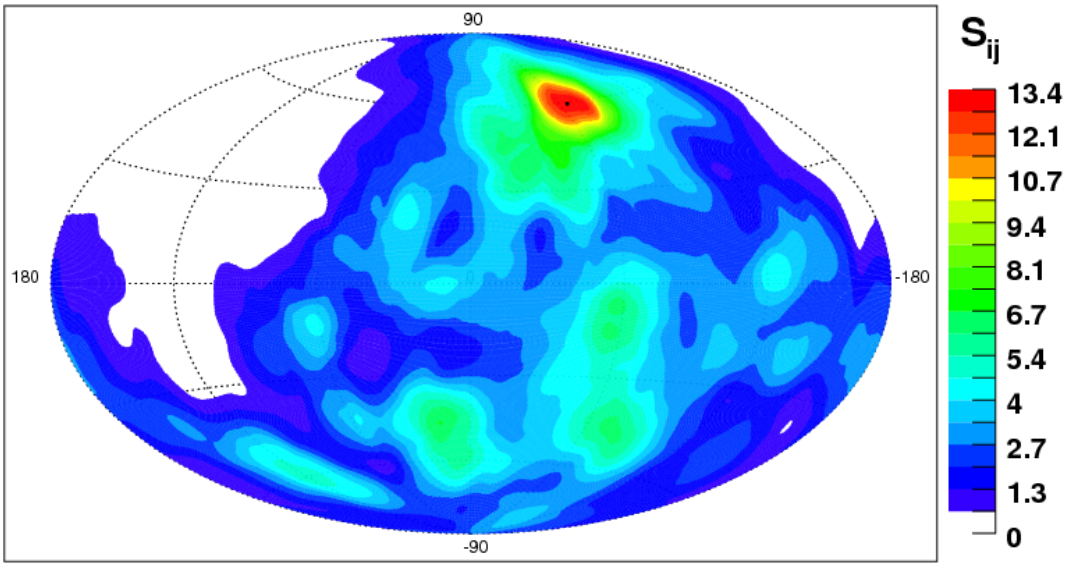


Figure 8.2: Skymap, in Galactic co-ordinates, of S_{ij} values after applying the magnetic spectrometer analysis to the data of Figure 8.1. The parameters used in this analysis are $\Delta\theta_{cut} = 45^{\circ}$ and $\zeta_{smth} = 10^{\circ}$. The location of the highest value of S_{ij} is marked by a black dot.

In order to estimate the likelihood of such a peak appearing by chance, for each set of analysis parameters the peak value of S_{ij} as calculated from the real data is compared to the peak values gained from the analysis of mock data sets, of which 500 were produced through the 'shuffling' method [179]. This method takes the set of measured arrival times of the events of the data set, and randomly reassigns each arrival time to one of the pairs of zenith and azimuth angles measured for the events of that data set. Subsequently, the arrival directions of these shuffled events may be calculated. In doing so, the distributions of arrival times, zenith angles and azimuth angles of events measured by the detector are preserved. Consequently, each shuffled data set represents a possible sample of events that would be detected by the Pierre Auger Observatory in the case of an isotropic flux. In addition to the random reassignment of arrival times, the energies of the real data are also redistributed among the shuffled events to retain an identical energy distribution to that of the real data.

For each pair of $\Delta\theta_{cut}$ and ζ_{smth} values, the magnetic spectrometer analysis is applied to each of the 500 mock data sets. Subsequently, the fraction of mock data sets, F_{mock}, which produce a value of S_{ij}, anywhere on the sky, that is greater than the peak value obtained for the real data using the corresponding parameter values is calculated. This fraction is an estimate of the chance of the signal that is observed in the real data being reproduced by the analysis of an isotropic data set of the same size and possessing an identical energy distribution.

Note that, due to the *a posteriori* nature of the analysis, the value of F_{mock} does not give the statistical significance of the peak in the observed data. In addition, as discussed in

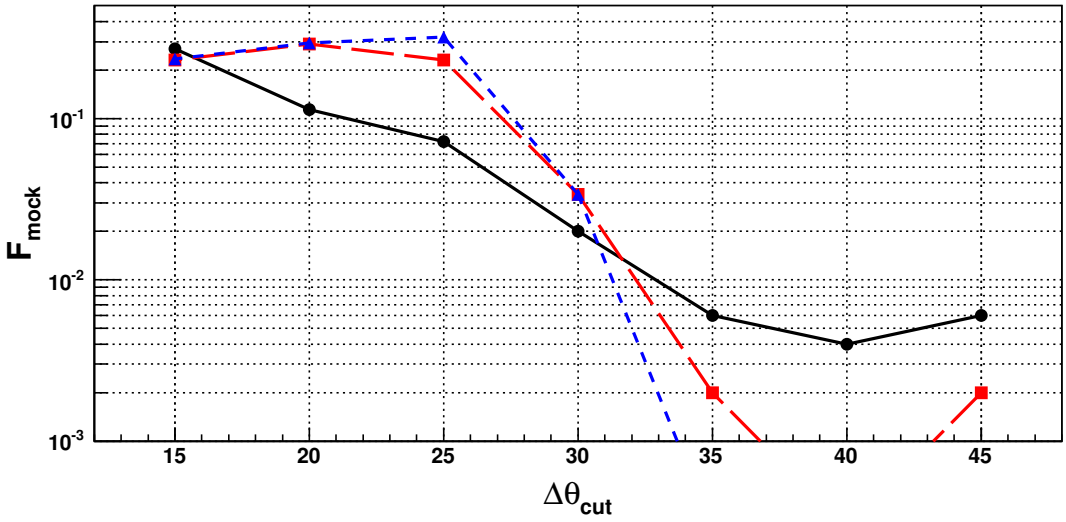


Figure 8.3: Plot of the fraction, F_{mock}, of the 500 mock data sets which produced, anywhere on the sky, a value of S_{ij} greater than the peak value observed from the real data following application of the magnetic spectrometer analysis, as a function of the analysis parameter $\Delta\theta_{cut}$. The solid black, long-dashed red and short-dashed blue lines correspond to values of ζ_{smth} of 10°, 7.5° and 5°, respectively. Note that the points corresponding to $\zeta_{smth} = 7.5^\circ$and $\Delta\theta_{cut} = 40^\circ$, as well as $\Delta\theta_{cut} \geq 35^\circ$ for $\zeta_{smth} = 5^\circ$, are not visible. This is due to the calculated value of F_{mock} for those parameter combinations being zero.

Chapter 7, a subset of the 69 events have been demonstrated to show anisotropy at a confidence level of 99%, and the correlating fraction of the entire 69 events remains above that which is expected for isotropy (although a confidence level for anisotropy cannot be determined for the full data set [100]). It is plausible that energy ordering is more likely to be present in anisotropic than in isotropic UHECR data. As a result, the chance probabilities presented here, which are calculated using isotropic arrival direction distributions, should be considered lower limits to the values of F_{mock} that would be obtained from generating mock data sets according to the true (but currently unknown) arrival direction distribution of UHECRs.

Figure 8.3 displays F_{mock} for each of the combinations of parameters considered. A minimum value of F_{mock} is obtained at $\Delta\theta_{cut} = 40^\circ$ for both $\zeta_{smth} = 10^\circ$ and $\zeta_{smth} = 7.5^\circ$. For a smoothing angle of $\zeta_{smth} = 5^\circ$, $F_{mock} = 0$ for $\Delta\theta_{cut} \geq 35^\circ$. Values of zero are obtained for F_{mock} due to the number of mock data sets being limited to 500, indicating that the true value of F_{mock} is estimated to be less than 0.2% in those cases.

It is of interest to note that the values of $\Delta\theta_{cut}$ and ζ_{smth} which minimise F_{mock} are not the same as were considered optimal in Chapter 7. In particular, $\Delta\theta_{cut} = 40^\circ$, which gives minimum values of F_{mock} for all values of ζ_{smth} considered here, is twice the value of 20° that was considered optimal for the analysis of simulated data sets. See Section 8.3 for discussion of this.

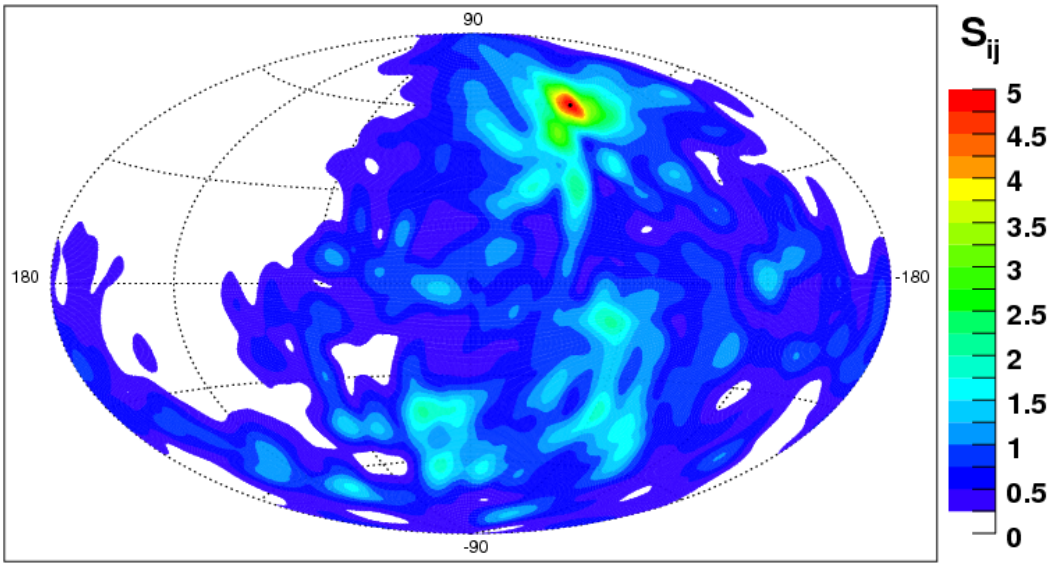


Figure 8.4: Skymap of S_{ij} values following the magnetic spectrometer analysis of the real data set using parameter values of $\Delta\theta_{cut} = 40°$ and $\zeta_{smth} = 5°$, which are the values taken here to be optimal. The peak value, indicated by a black dot, lies at the Galactic co-ordinates $(l, b) = (-66.4°, 58.8°)$. By comparison with Figure 8.2, it may be seen that the maximum value of S_{ij} is lower in this map due to the smaller smoothing angle employed. However, the value of F_{mock} is calculated to be smaller for the analysis parameters used here.

Although several sets of $\Delta\theta_{cut}$ and ζ_{smth} values result in $F_{mock} = 0$, values of $\Delta\theta_{cut} = 40°$ and $\zeta_{smth} = 5°$ will be used here as 'optimal' parameters - $\zeta_{smth} = 5°$ gives a broad minimum of F_{mock} at large values of $\Delta\theta_{cut}$, while $\Delta\theta_{cut} = 40°$ minimised F_{mock} for $\zeta_{smth} = 7.5°$ and $\zeta_{smth} = 10°$. The map of S_{ij} corresponding to these optimal parameters is shown in Figure 8.4.

In a similar manner to Chapter 7, the events which contribute to the reconstructed source peak may be found by calculating the nominal source positions for each pair of events. However, whilst the analysis of the simulated data accepted any pairs of events having a nominal source position within ζ_{smth} of the peak of the calculated S_{ij} map as being associated with that peak, such a simple criterion only results in four events being considered associated with the reconstructed source in the current analysis of real data. This is a result of the degree of clustering of the nominal source positions being less than was present in the simulated data set.

Consequently, a larger angular radius of 15° is used here, although it is recognised that this choice is somewhat arbitrary. Future extensions to the analysis could calculate the integral of the function $\frac{dN}{d\Omega}$, for each pair, within the smoothing angle ζ_{smth} of the peak pixel. A subsequent cut on the value of this integral could then be used to classify event pairs as being associated with a given peak in an S_{ij} map. Nevertheless, Figure 8.5 indicates the events that are considered associated with the source peak using the 15° window adopted here.

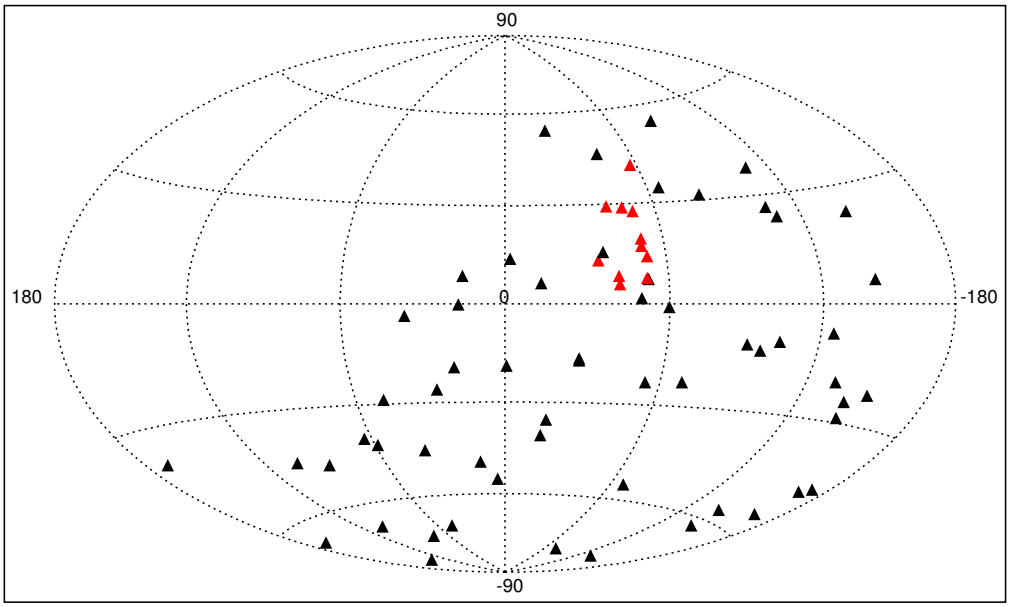


Figure 8.5: Skymap of the arrival directions of the 69 events in the full data set. Marked in red are the arrival directions of the events comprising pairs which, under the magnetic spectrometer assumptions, produce nominal source positions that lie within 15° of the peak of the S_{ij} map shown in Figure 8.4.

As described in Chapter 7, the identification of events associated with the reconstructed source peak allows an estimate of $\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right|$ to be made. The fit to these events is shown in Figure 8.6. It may be seen that a linear function describes the data well, which indicates consistent energy ordering for all 11 events.

From this fit, a value may be derived of $\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right| \approx \frac{50}{Z}\,\mu\mathrm{G\,kpc}$, where the factor of Z follows from Equation 7.4 as the charge of these cosmic rays is unknown. This value is not very sensitive to variations in the size of the angular window within which the nominal source position of a pair is required to lie in order for that pair to be associated with the S_{ij} peak. Taking that angle to be 10° or 20° gives estimated values of $\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right| \approx \frac{47}{Z}\,\mu\mathrm{G\,kpc}$ and $\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right| \approx \frac{49}{Z}\,\mu\mathrm{G\,kpc}$, respectively.

Finally, it is important to recognise that the appearance of such a large source peak relies heavily on the presence of the highest energy event in the data set, which has Galactic co-ordinates $(l, b) = (-57.2°, 41.8°)$ and energy 142 EeV. During the magnetic spectrometer analysis of the data, the pairing of this event with multiple lower-energy events produces many of the overlapping reconstructed source distributions that comprise the observed peak in S_{ij}.

If the magnetic spectrometer analysis is applied once more to the real data set and 500 mock data sets, but the event of energy 142 EeV is removed from every set, F_{mock} takes a value of 0.53 for $\Delta\theta_{cut} = 40°$ and $\zeta_{smth} = 5°$. The S_{ij} map following the analysis of the 68 events remaining after the removal of the highest energy event is shown in Figure 8.7. It may

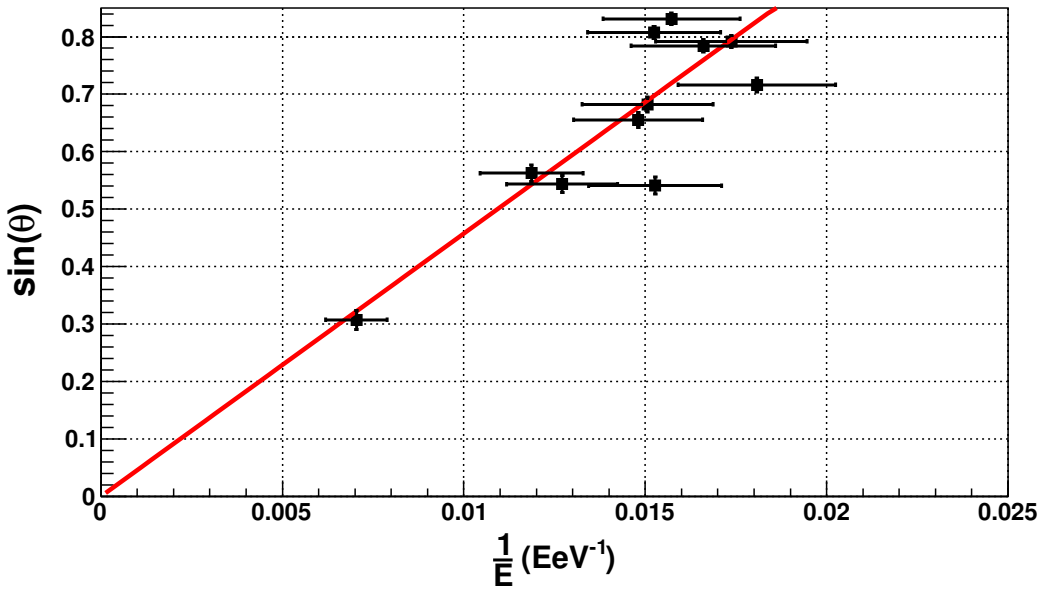


Figure 8.6: A fit to the events, considered to be 'associated' with the S_{ij} peak observed in Figure 8.4, of Equation 7.5. The red line indicates the fitted function.

be seen that the peak is still located near that of Figure 8.4, although the maximum S_{ij} value is approximately half of that found when the 142 EeV event is present. Thus, several pairs of lower energy events still produce overlapping source distributions but their number is fewer than if the 142 EeV event were present.

8.2 Analysis of Epochs

As discussed in [100], the observed degree of correlation between the arrival directions of the highest energy events detected by the Pierre Auger Observatory and the AGN of the 12$^{\text{th}}$ edition of the VCV catalogue appears to have decreased from the value obtained from the earlier analysis reported in [98, 99]. It is of interest, then, to investigate the evolution with time of F_{mock} to establish if a similar effect occurs following the magnetic spectrometer analysis of the data.

To do so, the data set containing all 69 events is split into three smaller sets, each containing 27 events. The first data set (epoch 1) contains the events of [98], the second data set (epoch 2) contains the following 27 events that were detected with respect to time, and the third data set (epoch 3) contains the final 27 events (of which 12 are also present in the second data set). As was performed in Section 8.1, 500 mock data sets were produced for each of the three epochs. Using once more the parameter values $\Delta\theta_{cut} = 40°$ and $\zeta_{smth} = 5°$, the magnetic spectrometer analysis is applied to each of these sets and F_{mock} calculated.

The skymaps of S_{ij} values for each of the real data sets is shown in Figure 8.8. Analysis of

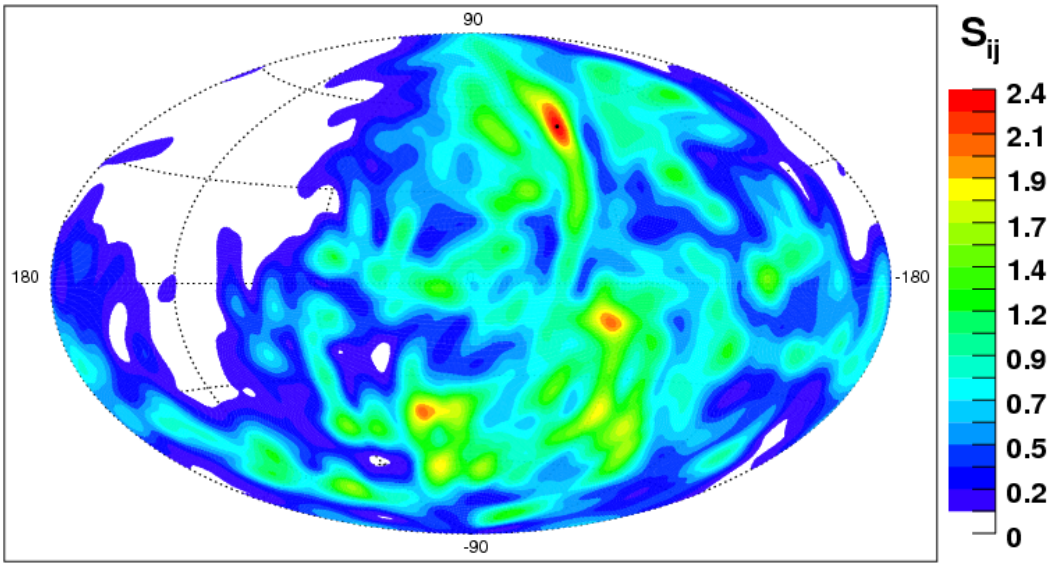


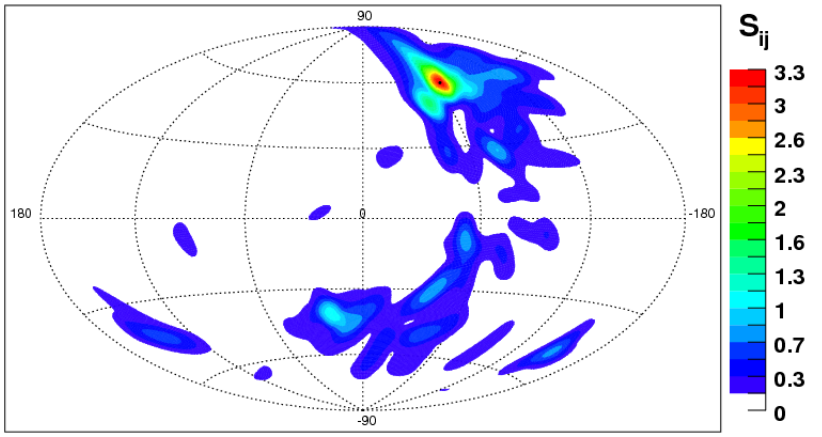
Figure 8.7: Skymap of S_{ij} values following the magnetic spectrometer analysis of the Pierre Auger Observatory data set following the removal of the highest energy event. The analysis parameters are $\Delta\theta_{cut} = 40^\circ$ and $\zeta_{smth} = 5^\circ$. By comparison with Figure 8.4, it may be seen that the peak position is in a similar location but its S_{ij} value has approximately halved upon removal of the 142 EeV event.

epoch 1 returns a peak position at $(l, b) = (-66.0^\circ, 58.3^\circ)$, which is approximately 0.6° from the peak position found through the analysis of all 69 events. However, no prominent peaks are observed in the results of either epoch 2 or 3, indicating little overlap of reconstructed source distributions - indicative of little energy ordering within the arrival directions of those data sets.

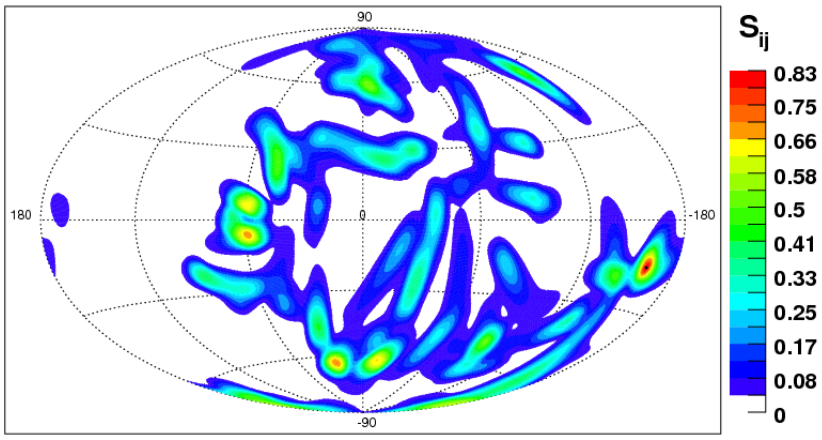
The values of F_{mock} are calculated as in Section 8.1 and are found to be 0, 0.91 and 0.75 for epochs 1, 2 and 3, respectively. Thus, while the analysis of the first set of 27 events returns an S_{ij} peak that is greater in size than that found in any of the 500 isotropic mock data sets considered, the second and third data sets return values which are consistent with those produced by isotropic data sets. It may be seen that the peak found through the magnetic spectrometer analysis of all 69 events arises mainly as a result of a subset of the first 27 events to be detected.

8.3 Interpretation of Results

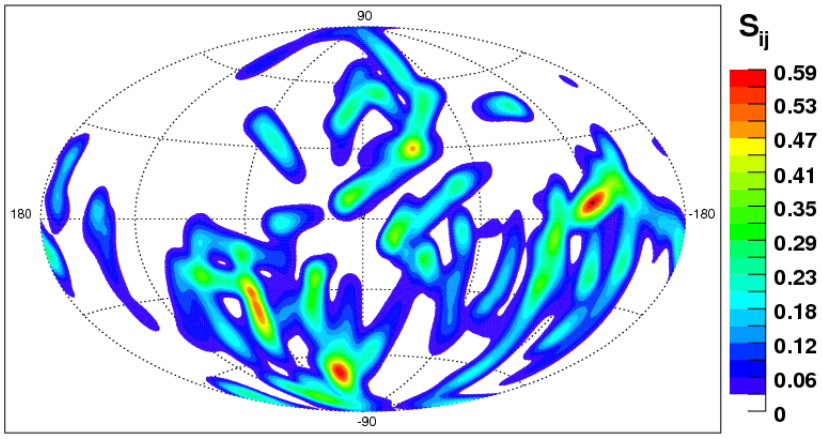
While still recognizing that the results presented in Sections 8.1 and 8.2 are *a posteriori*, and hence no statistical significance is claimed for the peak that is observed upon the analysis of the Pierre Auger Observatory UHECR data, some interpretation of these results in an astrophysical context may be made if it is assumed that they indicate the presence of a source of at least a subset of the events studied here.



(a)



(b)



(c)

Figure 8.8: Skymaps of S_{ij} values following the magnetic spectrometer analysis of (a) the first 27 events, (b) the second set of 27 events and (c) the final 27 events of the complete 69 event data set. In all three cases the analysis parameters are $\Delta\theta_{cut} = 40^\circ$ and $\zeta_{smth} = 5^\circ$.

When analysing the data using parameters of $\Delta\theta_{cut} = 40°$ and $\zeta_{smth} = 5°$, the peak position of the resulting S_{ij} map is located at $(l, b) = (-66.4°, 58.8°)$. It is noted that the Virgo galaxy cluster, which has been considered a possible UHECR source (for example, [180]) owing to its proximity to our galaxy and the presence of the powerful radio galaxy M87, lies approximately 16° from this position. If the Virgo cluster were regarded as the source of these events, the reason for the observed shift in the reconstructed source position would need to be explained. For example, the translation of the arrival directions of all associated events through coherent deflections caused by extragalactic magnetic fields may explain such a shift, although such a deflection would require a large-scale ordered magnetic field to exist between the Milky Way galaxy and the Virgo cluster. A product of the magnetic field strength and distance of the order 100 nG Mpc would be capable of causing such deflections (for example, a 5 nG field strength across the ≈ 20 Mpc distance between our galaxy and the Virgo cluster), although it is currently unclear if such a requirement is plausible.

The results obtained here may be compared with those of [105], in which a different analysis was applied to this data set to search for energy ordering. That analysis found possible evidence of energy ordering for a similar group of events to those considered here - 13 events were considered to be compatible with energy ordering in [105], compared to the 11 events here, with 10 events being common to both studies. The source position was reconstructed in [105] at Galactic co-ordinates $(l, b) = (-106°, 72.5°)$, which is approximately 20° from the peak direction found in this work. This difference may be attributed to two factors: firstly, the angle between the source and each event is considered in [105] to be given by $\theta \propto \frac{1}{E}$, compared to $\sin\theta \propto \frac{1}{E}$ assumed here. At small values of θ these relationships agree well. However, for larger deflections - such as those that are apparent for the group of events considered here - where the value of θ is several tens of degrees for some events, the linear approximation introduces a shift in the position of the source that is reconstructed from the arrival directions of the events. Secondly, the two methods locate a source direction in different ways (the method of [105] directly calculates it from fitting to the arrival directions of the events, while the magnetic spectrometer analysis calculates its position indirectly) and, as a result, slight differences in the two directions are to be expected in any case.

It is of interest to note that the value of $\Delta\theta_{cut} = 40°$ which was found to minimise F_{mock} is double that which was taken to be the optimal value in Chapter 7. This implies that the angular scale over which the arrival directions of the energy ordered events lie is approximately twice that which is expected from the simulations of protons in the GMF models considered in this work. A related indication of this comes from the fitting of Equation 7.5, from which a value of $\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right| \approx \frac{50}{Z}\,\mu$G kpc is derived for the line of sight along the reconstructed source direction. Calculation of the integral of $\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right|$, in the direction of the reconstructed source, for the eight GMF configurations considered in Chapter 7 returned values between approximately $2.3\,\mu$G kpc and $20\,\mu$G kpc. If the cosmic rays are protons, and they undergo deflections predominantly within the Galaxy, then such a large value of $\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right|$ is difficult

to reconcile with the GMF models considered in this work. The high Galactic latitude of the reconstructed source direction implies that the contribution to $\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right|$ from the spiral component of the GMF is small, and that a stronger halo field than was considered here must exist. As the normalisation of the toroidal halo field used in this work ($10\,\mu$G) is already higher than recent observations suggest [167], an underestimate of the strength of the halo field seems unlikely.

Alternatively, if the cosmic rays are heavier nuclei than protons, then a value of $\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right| \approx \frac{50}{Z}\,\mu\mathrm{G\,kpc}$ may be compatible with expectations from the GMF models employed here. For instance, assuming that the events are helium nuclei ($Z = 2$) gives reasonable agreement of the derived value $\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right|$ with the highest values found in the GMF models considered in this work. On the other hand, taking the events to be iron nuclei ($Z = 26$) results in consistency with the lowest value of $\left|\int \boldsymbol{B} \times d\boldsymbol{l}\right|$ given by the GMF models. Thus, the results of Section 8.1 may hint at an intermediate to heavy composition of the events associated with the reconstructed source direction. Such a composition may be suggested by measurements of $\langle X_{max} \rangle$ by the Pierre Auger Observatory [19], although those results do not extend to the energies considered here.

The apparent lack of strong energy ordering in the last 42 events to be detected is curious when compared to the first 27 events. Such an effect is possibly related to the underlying cause of the decrease in the fraction of events that are observed to correlate with local AGN [100]. Alternatively, as discussed in Section 8.1, the removal of the highest energy event results in a peak value that is commonly reproduced by isotropic mock data sets. Thus, despite the apparent consistency in the energy ordering of at least a subset of the associated events (see Figure 8.6), it is possible that the associated events do not all originate at the same source. The presence of the 142 EeV event near the cluster of lower-energy events near the direction of Centaurus A could be a chance occurrence, and the location of any high-energy event near such a cluster may lead to a large S_{ij} peak somewhere on the sky.

It is apparent that, while the results gained here hint at some interesting possibilities, the *a posteriori* nature of the analysis and the small number of events above an energy of 100 EeV mean that firm conclusions currently cannot be drawn. For example, the detection of a second event of a similar energy and arrival direction to the 142 EeV event could lend support to the presence of true energy ordering within the Pierre Auger Observatory data. Only with a larger data set and an *a priori* analysis applied to that data will the presence of any such effect be established with a known statistical significance.

8.4 Conclusions and Future Prospects

While the identification of the sources of ultra high energy cosmic rays remains a challenging task, detectors such as the Pierre Auger Observatory are capable of collecting unprecedented amounts of data which may allow directional cosmic ray astronomy to be performed in the

near future. Data sets at trans-GZK energies are now large enough to permit a variety of analyses to be performed in the search for UHECR sources, one of which has been presented in this work. At very high energies, cosmic rays may suffer deflections in cosmic magnetic fields that are small enough to result in energy ordering of their arrival directions at Earth. The magnetic spectrometer analysis described here is designed to identify such events and reconstruct their source direction using only the arrival directions and energies of a UHECR data set, while requiring only minimal assumptions regarding intervening magnetic fields.

Through applying this analysis to simulated data sets, produced by calculating the arrival directions of UHECRs at Earth after their propagation through representative models of the Galactic magnetic field, it has been shown that the magnetic spectrometer analysis can reconstruct the direction of the source of energy ordered events to within an average of $\approx 6^\circ$, if the assumptions described in Section 6.2 hold.

The magnetic spectrometer analysis, as it has been described here, could be modified for future studies in several ways to improve its performance. Examples of possible changes include:

- Incorporating more stringent criteria in order for events to be considered a pair. This could include setting a stronger constraint on the lower limit to the ratio of event energies (currently $\frac{E_2}{E_1} > 1.02$ is required), as pairs of events with a large energy ratio may be less prone to errors in their reconstructed source position than pairs of events with similar energies. In addition, if an accurate prediction of the deflections suffered by UHECRs prior to reaching Earth can be made (due to better knowledge of the Galactic magnetic field, for example), then a range of $\Delta\theta$ values, dependent on the energy ratio, could be enforced. This is in contrast to the current requirement that $2^\circ \leq \Delta\theta \leq \Delta\theta_{cut}$, where $\Delta\theta_{cut}$ is the same for any value of the energy ratio.

- Currently, the method of finding events that are 'associated' with a peak in an S_{ij} map is rather simplistic. The asymmetry of the reconstructed source direction distribution for a given pair means that its highest value is not in the same position as the nominal source position for that pair. A more robust alternative may be to evaluate the integral of $\frac{dN}{d\Omega}$ for each event pair at the location of a source peak, and apply a threshold cut on this value.

- If a source is found, and associated events with energy ordered arrival directions identified, then a search for more events consistent with that energy ordering may be performed at lower energy. Such searches could be useful in establishing the spectra of events from a single source, if identified, and reveal further information about the intervening magnetic field than could be achieved using only events at trans-GZK energies. However, restricting the initial search to trans-GZK energies is important to limit the source horizon of the events in the data set.

Nevertheless, upon the application of the magnetic spectrometer analysis to UHECR events

detected by the Pierre Auger Observatory, possible evidence of energy ordering of the arrival directions of a subset of those events was found. Through the application of an identical analysis to 500 mock data sets, constructed according to an isotropic arrival direction distribution, it was estimated that the signal found from the analysis of real data has less than a 0.2% chance of arising from the analysis of an isotropic data set, for particular analysis parameters. However, this is a result of *a posteriori* analysis in which the parameters of the magnetic spectrometer analysis were varied to minimise the chance probability and, as such, a statistical significance cannot be claimed.

In order to establish the presence of any such energy ordering, a more thorough study should be performed to find the analysis parameters $\Delta\theta_{cut}$ and ζ_{smth} (which were varied quite coarsely here) and event energy threshold which minimise the value of F_{mock} using a large number of mock data sets (as an inability to distinguish between parameters was limited here by using only 500 such sets). The results of such a study may be used to define an *a priori* analysis to be applied to future UHECR data, thus allowing the presence of any observed energy ordering to be established with a known significance.

The detection of large numbers of UHECR events with current and future detectors, a better understanding of the composition of the CR flux at trans-GZK energies and a more detailed understanding of cosmic magnetic fields will all contribute to the identification of the sources of these particles. The magnetic spectrometer analysis presented here is well-suited to the task of locating UHECR sources, assuming that the magnetic spectrometer assumptions hold. Incorporating more detailed information regarding the expectations of UHECR propagation through cosmic magnetic fields, as it becomes available, may lead to further improvement in the ability of the analysis to identify UHECR source directions.

References

[1] R. W. Clay, B. R. Dawson, *Cosmic Bullets* (Allen & Unwin, St Leonards, NSW, 1997).

[2] A. M. Hillas, *Cosmic Rays* (Pergamon Press, Oxford; New York, 1972).

[3] V. Hess, Physik. Zeitschr. **13** (1912) 1084–1091.

[4] H. Elliot, Y. Sekido, *Early History of Cosmic Ray Studies: Personal Reminiscences with Old Photographs* (D. Reidel Publishing Company, Dordrecht, Holland, 1985).

[5] P. Auger, M. M. Shapiro, *What Are Cosmic Rays?* (University of Chicago Press, Chicago, 1945).

[6] J. Cronin, Proc. of 30th Int. Cosmic Ray Conf. **6** (2009) 3–18.

[7] T. Stanev, *High Energy Cosmic Rays* (Praxis Publishing Ltd., Chichester, UK, 2004).

[8] P. Auger, *et al.*, Rev. Mod. Phys. **11** (1939) 288–291.

[9] T. K. Gaisser, J. Phys. Conf. Series **47** (2006) 15–20.

[10] T. K. Gaisser, *Cosmic rays and Particle Physics* (Cambridge University Press, Cambridge; New York, 1990).

[11] W. Heitler, *The Quantum Theory of Radiation* (Clarendon Press, Oxford, 1954), 3rd ed.

[12] K. Nakamura, *et al.*, J. Phys. G **37** (2010) 075021.

[13] J. Blümer, R. Engel, J. R. Hörandel, Prog. Part. Nucl. Phys. **63** (2009) 293–338.

[14] T. K. Gaisser, A. M. Hillas, Proc. of 15th Int. Cosmic Ray Conf. **8** (1977) 353–357.

[15] M. Unger, *et al.*, Nucl. Instrum. Methods A **588** (2008) 433–441.

[16] R. U. Abbasi, *et al.*, Astrophys. J. **622** (2005) 910–926.

[17] M. Nagano, A. A. Watson, Rev. Mod. Phys. **72** (2000) 689–732.

[18] J. Matthews, Astropart. Phys. **22** (2005) 387–397.

[19] P. Abreu, *et al.* (2011) arXiv:1107.4804.

[20] J. Abraham, *et al.*, Phys. Rev. Lett. **104** (2010) 091101.

[21] L. Anchordoqui, *et al.*, Ann. Phys. **314** (2004) 145–207.

[22] F. Arqueros, J. R. Hörandel, B. Keilhauer, Nucl. Instrum. Methods A **597** (2008) 1–22.

[23] J. Abraham, *et al.*, Nucl. Instrum. Methods A **620** (2010) 227–251.

[24] P. Sokolsky, P. Sommers, B. Dawson, Phys. Rep. **217** (1992) 225–277.

[25] M. Ave, *et al.*, Nucl. Phys. B (Proc. Suppl.) **212** (2011) 356–361.

[26] M. Ave, *et al.*, Astropart. Phys. **28** (2007) 41–57.

[27] J. Abraham, *et al.*, Nucl. Instrum. Methods A **523** (2004) 50–95.

[28] P. Sommers, S. Westerhoff, New J. Phys. **11** (2009) 055004.

[29] D. Newton, J. Knapp, A. A. Watson, Astropart. Phys. **26** (2007) 414–419.

[30] M. Ave, Proc. of 30th Int. Cosmic Ray Conf. **4** (2008) 307–310.

[31] C. Bonifazi, Nucl. Phys. B (Proc. Suppl.) **190** (2009) 20–25.

[32] M. Ave, *et al.*, Nucl. Instrum. Methods A **578** (2007) 180–184.

[33] R. N. Coy, *et al.*, Proc. of 17th Int. Cosmic Ray Conf. **6** (1981) 43–46.

[34] S. Yoshida, *et al.*, J. Phys. G **20** (1994) 651–664.

[35] M. A. Lawrence, R. J. O. Reid, A. A. Watson, J. Phys. G **17** (1991) 733–757.

[36] P. Abreu, *et al.* (2011) arXiv:1107.4809.

[37] A. M. Hillas, *et al.*, Proc. of 12th Int. Cosmic Ray Conf. **3** (1971) 1001–1006.

[38] B. R. Dawson, Proc. of 30th Int. Cosmic Ray Conf. **4** (2008) 425–428.

[39] K. Shinozaki, M. Teshima, Nucl. Phys. B (Proc. Suppl.) **136** (2004) 18–27.

[40] J. Abraham, *et al.* (2009) arXiv:0906.2319.

[41] M. T. Dova, *et al.*, Astropart. Phys. **31** (2009) 312–319.

[42] I. Allekotte, *et al.*, Nucl. Instrum. Methods A **586** (2008) 409–420.

[43] J. Abraham, *et al.*, Astropart. Phys. **33** (2010) 108–129.

[44] J. Pękala, *et al.*, Nucl. Instrum. Methods A **605** (2009) 388–398.

[45] J. Belz, Nucl. Phys. B (Proc. Suppl.) **190** (2009) 5–11.

[46] J. Blümer, New J. Phys. **12** (2010) 035001.

[47] J. Abraham, *et al.*, Phys. Lett. B **685** (2010) 239–246.

[48] P. Abreu, *et al.*, Astropart. Phys. **34** (2011) 368–381.

[49] P. Allison, *et al.*, Proc. of 29th Int. Cosmic Ray Conf. **8** (2005) 307–310.

[50] J. Abraham, *et al.* (2009) arXiv:0906.2189.

[51] P. Sokolsky, G. B. Thomson, J. Phys. G **34** (2007) R401–R429.

[52] J. R. Hörandel, Mod. Phys. Lett. A **22** (2007) 1533–1551.

[53] T. K. Gaisser, T. Stanev, Nucl. Phys. A **777** (2006) 98–110.

[54] R. J. Protheroe, R. W. Clay, Pub. Astron. Soc. Aust. **21** (2004) 1–22.

[55] R. Aloisio, *et al.*, Astropart. Phys. **27** (2007) 76–91.

[56] J. R. Hörandel, Astropart. Phys. **19** (2003) 193–220.

[57] V. Berezinsky, A. Gazizov, S. Grigorieva, Phys. Rev. D**74** (2006) 043005.

[58] R. U. Abbasi, *et al.*, Astropart. Phys. **32** (2009) 53–60.

[59] D. Ivanov, B. Stokes, G. Thompson, *Measurement of the Energy Spectrum by the Telescope Array Surface Detector* (2011), talk presented at 32nd Int. Cosmic Ray Conf., Beijing, China.

[60] J. R. Hörandel, *et al.*, Nucl. Instrum. Methods A **630** (2011) 222–225.

[61] K. Greisen, Phys. Rev. Lett. **16** (1966) 748–750.

[62] G. T. Zatsepin, V. A. Kuz'min, Zh. Eksp. Teor. Fiz. Pis'a Red. **4** (1966) 114.

[63] A. A. Penzias, R. W. Wilson, Astrophys. J.**142** (1965) 419–421.

[64] D. Allard, *et al.*, J. Cosmol. Astropart. Phys. **10** (2008) 033.

[65] M. Takeda, *et al.*, Phys. Rev. Lett. **81** (1998) 1163–1166.

[66] M. Takeda, *et al.*, Astropart. Phys. **19** (2003) 447–462.

[67] K. Shinozaki, *Observation of UHECRs by AGASA* (2006), talk presented at the Quarks-2006 Conf., St. Petersburg, Russia.

[68] R. Abbasi, *et al.*, Phys. Lett. B **619** (2005) 271–280.

[69] R. U. Abbasi, *et al.*, Phys. Rev. Lett. **100** (2008) 101101.

[70] J. Abraham, *et al.*, Phys. Rev. Lett. **101** (2008) 061101.

[71] R. Ulrich, R. Engel, M. Unger, Phys. Rev. D **83** (2011) 054026.

[72] W. D. Apel, *et al.*, Astropart. Phys. **31** (2009) 86–91.

[73] D. J. Bird, *et al.*, Phys. Rev. Lett. **71** (1993) 3401–3404.

[74] K. Shinozaki, Nucl. Phys. B (Proc. Suppl.) **151** (2006) 3–10.

[75] T. Abu-Zayyad, *et al.*, Phys. Rev. Lett. **84** (2000) 4276–4279.

[76] R. U. Abbasi, *et al.*, Phys. Rev. Lett. **104** (2010) 161101.

[77] Y. Tameda, *et al.*, *Measurement of UHECR Mass Composition by TA FD Stereo* (2011), talk presented at 32nd Int. Cosmic Ray Conf., Beijing, China.

[78] J. D. Hague, *et al.*, J. Phys. G **36** (2009) 115203.

[79] R. Clay, B. R. Dawson, G. J. Thornton, C. R. Phys. **5** (2004) 473–482.

[80] J. Abraham, *et al.*, Astropart. Phys. **29** (2008) 243–256.

[81] J. Abraham, *et al.*, Astropart. Phys. **31** (2009) 399–406.

[82] P. Abreu, *et al.* (2011) arXiv:1107.4805.

[83] A. M. Hillas, Ann. Rev. Astron. Astrophys. **22** (1984) 425–444.

[84] D. F. Torres, L. A. Anchordoqui, Rep. Prog. Phys. **67** (2004) 1663–1730.

[85] F. Aharonian, *et al.*, Astron. Astrophys. **425** (2004) L13–L17.

[86] N. Hayashida, *et al.*, Astropart. Phys. **10** (1999) 303–311.

[87] J. A. Bellido, *et al.*, Astropart. Phys. **15** (2001) 167–175.

[88] J. Abraham, *et al.*, Astropart. Phys. **27** (2007) 244–253.

[89] P. Abreu, *et al.*, Astropart. Phys. **34** (2011) 627–639.

[90] A. H. Compton, I. A. Getting, Phys. Rev. **47** (1935) 817–821.

[91] M. Takeda, *et al.*, Astrophys. J. **522** (1999) 225–237.

[92] M. A. Kirn, *et al.*, Astropart. Phys. **22** (2004) 139–149.

[93] R. U. Abbasi, *et al.*, Astrophys. J. **623** (2005) 164–170.

[94] D. S. Gorbunov, *et al.*, Astrophys. J. **577** (2002) L93–L96.

[95] R. U. Abbasi, *et al.*, Astrophys. J. **636** (2006) 680–684.

[96] D. Harari, Proc. of 30th Int. Cosmic Ray Conf. **4** (2008) 283–286.

[97] M.-P. Véron-Cetty, P. Véron, Astron. Astrophys. **455** (2006) 773–777.

[98] J. Abraham, *et al.*, Science **318** (2007) 939–943.

[99] J. Abraham, *et al.*, Astropart. Phys. **29** (2008) 188–204.

[100] P. Abreu, *et al.*, Astropart. Phys. **34** (2010) 314–326.

[101] J. Abraham, *et al.* (2009) arXiv:0906.2347.

[102] R. U. Abbasi, *et al.*, Astropart. Phys. **30** (2008) 175–179.

[103] R. U. Abbasi, *et al.*, Astropart. Phys. **28** (2007) 385–389.

[104] P. Abreu, *et al.*, Astropart. Phys. (2011) in press, arXiv:1111.2472.

[105] G. Giacinti, D. V. Semikoz, Phys. Rev. D **83** (2011) 083002.

[106] G. Giacinti, X. Derkx, D. V. Semikoz, J. Cosmol. Astropart. Phys. **3** (2010) 022.

[107] J. Linsley, L. Scarsi, B. Rossi, Phys. Rev. Lett. **6** (1961) 485–487.

[108] J. Linsley, L. Scarsi, Phys. Rev. **128** (1962) 2384–2392.

[109] J. Linsley, Proc. of 8th Int. Cosmic Ray Conf. **4** (1963) 77–99.

[110] J. Linsley, Phys. Rev. Lett. **10** (1963) 146–148.

[111] R. M. Tennent, Proc. Phys. Soc. **92** (1967) 622–631.

[112] P. R. Blake, H. Ferguson, W. F. Nash, J. Phys. Soc. Jap. **33** (1972) 1197–1206.

[113] D. M. Edge, *et al.*, Proc. of 15th Int. Cosmic Ray Conf. **9** (1977) 137–142.

[114] A. J. Bower, *et al.*, J. Phys. G **9** (1983) 1569–1576.

[115] A. J. Bower, *et al.*, J. Phys. G **9** (1983) L53–L58.

[116] M. Winn, *et al.*, J. Phys. G **12** (1986) 675–686.

[117] C. J. Bell, *et al.*, J. Phys. A **7** (1974) 990–1009.

[118] M. Winn, *et al.*, J. Phys. G **12** (1986) 653–674.

[119] C. J. Bell, J. Phys. G **2** (1976) 857–866.

[120] C. J. Bell, J. Phys. G **2** (1976) 867–880.

[121] L. J. Kewley, R. W. Clay, B. R. Dawson, Astropart. Phys. **5** (1996) 69–74.

[122] A. A. Ivanov, S. P. Knurenko, I. Y. Sleptsov, New J. Phys. **11** (2009) 065008.

[123] A. A. Ivanov, *et al.*, Astrophys. Space Sci. Trans. **6** (2010) 53–57.

[124] A. N. Bunner, K. Greisen, P. B. Landecker, Canad. J. Phys. **46** (1968) S266–S269.

[125] T. Hara, *et al.*, Proc. of 29th Int. Cosmic Ray Conf. **3** (1970) 369–376.

[126] P. Sokolsky, AIP Conf. Proc. **433** (1998) 65–75.

[127] R. M. Baltrusaitis, *et al.*, Nucl. Instrum. Methods A **240** (1985) 410–428.

[128] D. J. Bird, *et al.*, AIP Conf. Proc. **338** (1995) 839–854.

[129] D. J. Bird, *et al.*, Astrophys. J. **441** (1995) 144–150.

[130] N. Chiba, *et al.*, Nucl. Instrum. Methods A **311** (1992) 338–349.

[131] H. Ohoka, *et al.*, Nucl. Instrum. Methods A **385** (1997) 268–276.

[132] N. Hayashida, *et al.*, Phys. Rev. Lett. **73** (1994) 3491–3494.

[133] N. Hayashida, *et al.*, Astropart. Phys. **10** (1999) 303–311.

[134] R. U. Abbasi, *et al.*, Phys. Rev. Lett. **92** (2004) 151101.

[135] J. H. Boyer, *et al.*, Nucl. Instrum. Methods A **482** (2002) 457–474.

[136] T. Nonaka, *et al.*, Nucl. Phys. B (Proc. Suppl.) **190** (2009) 26–31.

[137] H. Tokuno, *et al.*, J. Phys. Conf. Series **293** (2011) 012035.

[138] D. Allard, *et al.*, Proc. of 29th Int. Cosmic Ray Conf. **7** (2005) 71–74.

[139] J. Abraham, *et al.*, Nucl. Instrum. Methods A **613** (2010) 29–39.

[140] J. Abraham, *et al.* (2009) arXiv:0906.2354.

[141] X. Bertou, *et al.*, Nucl. Instrum. Methods A **568** (2006) 839–846.

[142] D. Allard, *et al.*, Proc. of 29th Int. Cosmic Ray Conf. **7** (2005) 287–290.

[143] C. Bonifazi, A. Letessier-Selvon, E. M. Santos, Astropart. Phys. **28** (2008) 523–528.

[144] J. Hersil, *et al.*, Phys. Rev. Lett. **6** (1961) 22–23.

[145] P. Abreu, *et al.* (2011) arXiv:1107.4807.

[146] A. Shalchi, *Nonlinear Cosmic Ray Diffusion Theories* (Springer-Verlag, Berlin, 2009).

[147] M. S. Longair, *High Energy Astrophysics* (Cambridge University Press, Cambridge, 1981).

[148] D. Harari, S. Mollerach, E. Roulet, J. Cosmol. Astropart. Phys. **11** (2010) 033.

[149] V. Ptuskin, J. Phys. Conf. Series **47** (2006) 113–119.

[150] D. Harari, *et al.*, J. High Energy Phys. **3** (2002) 045.

[151] J. Han, Proc. Int. Astron. Union **4** (2008) 455–466.

[152] P. P. Kronberg, Space Sci. Rev. **75** (1996) 387–399.

[153] R. Beck, Space Sci. Rev. **99** (2001) 243–260.

[154] R. Beck, RevMexAA (Serie de Conferencias) **36** (2009) 1–8.

[155] J. C. Brown, Astron. Soc. Pacific Conf. Series **438** (2010) 216–228.

[156] J. H. Simonetti, J. M. Cordes, S. R. Spangler, Astrophys. J. **284** (1984) 126–134.

[157] A. Noutsos, *et al.*, Mon. Not. R. Astron. Soc. **386** (2008) 1881–1896.

[158] T. Stanev, Astrophys. J. **479** (1997) 290–295.

[159] D. Harari, S. Mollerach, E. Roulet, J. High Energy Phys. **1999** (1999) 022.

[160] M. Prouza, R. Šmída, Astron. Astrophys. **410** (2003) 1–10.

[161] J. P. Vallée, Astrophys. J. **681** (2008) 303–310.

[162] X. H. Sun, *et al.*, Astron. Astrophys. **477** (2008) 573–592.

[163] R. Jansson, *et al.*, J. Cosmol. Astropart. Phys. **7** (2009) 021.

[164] A. Waelkens, *et al.*, Astron. Astrophys. **495** (2009) 697–706.

[165] K. Dolag, *et al.*, Astrophys. J. **727** (2011) L4.

[166] J. L. Han, R. N. Manchester, G. J. Qiao, Mon. Not. R. Astron. Soc. **306** (1999) 371–380.

[167] X.-H. Sun, W. Reich, Research in Astron. Astrophys. **10** (2010) 1287–1297.

[168] J. Giacalone, J. Jokipii, Astrophys. J. **520** (1999) 204–214.

[169] A. Chepurnov, A. Lazarian, Astrophys. J. **710** (2010) 853–858.

[170] M. Haverkorn, *et al.*, Astrophys. J. **680** (2008) 362–370.

[171] D. De Marco, P. Blasi, T. Stanev, J. Cosmol. Astropart. Phys. **2007** (2007) 027.

[172] M. Honda, Astrophys. J. **319** (1987) 836–841.

[173] F. Casse, M. Lemoine, G. Pelletier, Phys. Rev. D **65** (2002) 023002.

[174] E. Parizot, Nucl. Phys. B (Proc. Suppl.) **136** (2004) 169–178.

[175] J. W. Armstrong, B. J. Rickett, S. R. Spangler, Astrophys. J. **443** (1995) 209–221.

[176] N. Globus, D. Allard, E. Parizot, Astron. Astrophys. **479** (2008) 97–110.

[177] M. Erdmann, P. Schiffer, Astropart. Phys. **33** (2010) 201–205.

[178] P. J. Rousseeuw, A. M. Leroy, *Robust Regression and Outlier Detection* (John Wiley & Sons, Inc., New York, 1987).

[179] G. L. Cassiday, *et al.*, Nucl. Phys. B (Proc. Suppl.) **14** (1990) 291–298.

[180] A. V. Olinto, Phys. Rep. **333** (2000) 329–348.