8-27-2009

Studies of ultra high energy cosmic rays

John Douglas Hague

John D. Hague
Candidate

Physics & Astronomy
Department

This thesis is approved, and it is acceptable in quality and form for publication:

Approved by the Thesis Committee:

,Chairperson

STUDIES OF ULTRA HIGH ENERGY COSMIC RAYS

by

John Douglas Hague

B.S., Physics, University of Missouri, 2003
B.S., Mathematics, University of Missouri, 2003

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Physics

The University of New Mexico

Albuquerque, New Mexico

August, 2009

Dedication

To my family, with love and grditude.

Acknowledgments

I would like to first thank my academic advisor, Professor John A. J. Matthews without whose support this work would not be possible. Both Micheal Gold and Bernard Becker have contributed substantially to this work.

STUDIES OF ULTRA HIGH ENERGY COSMIC RAYS

by

John Douglas Hague

ABSTRACT OF DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Physics

The University of New Mexico

Albuquerque, New Mexico

August, 2009

STUDIES OF ULTRA HIGH ENERGY COSMIC RAYS

by

John Douglas Hague

B.S., Physics, University of Missouri, 2003
B.S., Mathematics, University of Missouri, 2003

Ph. D., Physics, University of New Mexico, 2009

Abstract

In this thesis, analysis tools for understanding the observed spectrum and arrival directions of ultra high energy cosmic rays are developed and applied to both simulated and observed data.

Two separate statistical tests are applied to the AGASA and preliminary Auger Cosmic Ray Energy spectra in an attempt to find deviation from a pure power-law. The first test is constructed from the probability distribution for the maximum event of a sample drawn from a power-law. The second employs the TP-statistic, a function defined to deviate from zero when the sample deviates from the power-law form, regardless of the value of the power index. Un-binned versions of these and other estimators are developed and applied to simulated cosmic ray spectra. The energy spectrum of Auger events arriving in different regions of the sky is also studied.

The two-point angular correlation function is a traditional method used to search

for deviations from expectations of isotropy. Here, a statistically descriptive three-point method is developed and explored, with the intended application being the search for deviations from isotropy in the highest energy cosmic rays. The sensitivity of this so-called "shape-strength" method is compared to that of a two-point method for a variety of Monte-Carlo simulated anisotropic signals, and studies with anisotropic source signals diluted by an isotropic background are performed. Type I and II errors for rejecting the hypothesis of isotropic cosmic ray arrival directions are evaluated for four different event sample sizes: 27, 40, 60 and 80 events, consistent with near-term data expectations from the Pierre Auger Observatory. In all cases, the ability to reject the isotropic hypothesis improves with event size and with the fraction of anisotropic signal. But while data sets with ~ 40 events should be sufficient for reliable identification of anisotropy in cases of rather extreme (highly anisotropic) data, much larger data sets are suggested for reliable identification of more subtle anisotropies. Overall, the shape-strength method, which can be easily adapted to an arbitrary experimental exposure on the celestial sphere, consistently performs better than the two point method.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 The Origins of CR Observation

To appreciate any serious endeavor, it is invaluable to understand its history and context. There is a large body of historical work in the field of cosmic ray (CR) physics. But, rather than attempt to summarize it here, I will review some selected historical events that I find particularly appealing and relevant to *ultra high* energy cosmic rays.

Simply understanding the origin of the term "cosmic ray" reveals a great deal about the history of the field. Around the turn of the 20^{th} century it was believed that the observed ionization of the atmosphere was caused by radiation from decaying radon in the Earth; "atmospheric electricity" supplies the "ray" in "cosmic ray." Indeed, measurements made in the first decade of the 1900's (at building-level altitudes) could be explained by a model with the Earth as a source. In 1912, however, Victor Hess carried out detailed and higher-altitude balloon-based observations that allowed him to report a four-fold increase over the amount of radiation observed on the ground – clearly inconsistent with an Earth-based source. (His appealingly

adventurous spirit is on display in Figure 1.1.) Later, Robert Millikan proved that these rays are "cosmic" in origin, thus coining the term. [1]

Figure 1.1: Victor Hess in a balloon[2].

As the field of CR physics developed, "atmospheric electricity" came to be understood as extensive air showers, the origins of which could be traced to single particles interacting with the atmosphere; for the highest-energy primaries – in the regime addressed by this thesis – these showers can be tens of square kilometers in size by the time they reach the ground. With the technological advances that led to the assembly of detector-arrays, the reconstruction of the qualities of the primaries from the measured quantities of the shower became possible. The crucial advance for modern ultra high energy detectors was the development of accurate and fast read-out timing capabilities, which allowed the experimenter to confidently identify single, large-area showers and to mathematically reconstruct the shower front.

The University of New Mexico played an important role in the modern development of the field. As the primary member of the Volcano Ranch experiment (1959-1976) set just west of Albuquerque, UNM Researcher John Linsley reported the first cosmic ray primary with energy $\sim 10^{20}$ eV[3]. (Linsley is pictured in Figure 1.3 below checking for snakes in one of the detectors at Volcano Ranch).

Figure 1.2: John Linsley checking for rattle snakes near a Volcano Ranch detector, outside Albuquerque, NM [4].

We now understand that the origin of the low-energy cosmic rays observed by Hess (and scientists at modern and sophisticated facilities) is different than the high energy events observed by Linsley (and at observatories since Volcano Park). But, while observations of low-energy cosmic rays are well-explained by particles accelerated within our solar system, the origin of the cosmic rays observed at the highest energies is still unknown.

This question of ultra high energy cosmic ray-origination has been the subject of decades of theoretical inquiry and observational research by many large, and often international, collaborations. The core of the work presented within these pages is meant to contribute a verse to this dialog by describing empirically-motivated data analysis tools designed to supply direct answers to the questions surrounding ultra high energy cosmic ray origination.

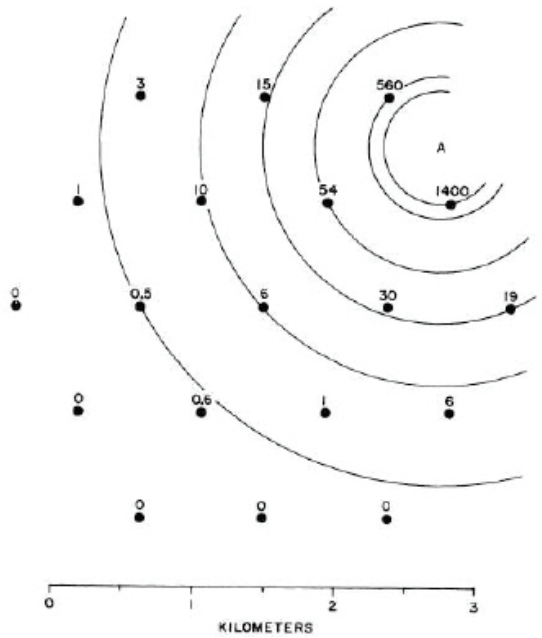


Figure 1.3: A schematic of a shower observed by Volcano Ranch array[5]. The primary that triggered this shower had an energy greater than 10^{20} eV.

1.2 Thesis Summary

At the heart of high energy CR physics lies the question

From where do the the highest energy cosmic rays originate?

It is well beyond the scope of this thesis to supply an answer. But, by allowing the data to speak as loudly as possible, the analysis tools presented here admit an advance toward the resolution of this fundamental issue.

There are three primary observations relevant to the characterization of high energy CR: the energy spectrum, the arrival direction, and the primary composition. Studies of CR anisotropy and the energy spectrum will be the focus of these pages.

As for the composition of CR, topical discussion will be included where relevant, but interested readers should see [6] for a more thorough review.

Starting with Chapter 2, the theoretical context for the observation of the energy spectrum and arrival directions is presented, and the status of current observations – which suggest consistency with a theoretical model wherein charged particle primaries originate from outside our galaxy – is reviewed. As will be clear there, despite the fact that incredibly high energies and (relatively) small magnetic fields are also involved, it is the charge of the primaries that makes CR astronomy so difficult. Given enough statistics (1000's of events) the sources can be identified (as depicted in Figure 2.9, §2.3.2). However, the flux of high energy CR is incredibly low (1 event/km^2/century) and is (presumably) further limited by the so-called GZK effect (see §2.2.2).

Chapter 3 begins with an outline of the particle physics phenomenology of air showers created by CR entering the atmosphere of Earth, and continues with discussions devoted to the reconstruction of the energy and arrival direction of the CR primary from data collected at the Pierre Auger Observatory (Auger). Except for Chapter 6, all remaining chapters are presented as previously published elsewhere. In these Chapters, statistics designed to confirm or deny the presence of flux suppression at the highest energies (i.e. the GZK effect) are developed, explored, and applied (see 4 and 5). Most interestingly, these chapters describe the *Tail Power Statistic*, which has the capacity to provide information about flux suppression *independent* of the underlying power-law index. Chapter 4 (based on a manuscript published in the peer-reviewed journal Astroparticle Physics [7]) explores the binned approach to the so-called TP statistic, and describes the application of this approach to both AGASA and preliminary Auger spectra. The un-binned approach, discussed and applied to simulated CR data sets in Chapter 5, allows more precise parameter estimation, as well as the natural incorporation of systematic and statistical energy

errors into analysis. The results and discussion presented in this Chapter (Chapter 4) appear as published in the peer-reviewed journal Astroparticle Physics [7].

Next, in Chapter 6 an attempt is made to directly compare the spectra observed in different regions of the sky using Auger data. In particular, the characteristic energy at which the spectrum "cuts-off" is studied for each of these spectra.

With Chapter 7, the portion of this thesis dedicated to CR energy spectra is exchanged for the topic of CR anisotropy. In 2008, Auger published a "correlation of [high energy CR] with active galactic nuclei" [8, 9], a claim given careful attention by the entire collaboration. In Chapter 7 a Bayesian scheme to estimate the degree of correlation is introduced, which was later developed and applied as a part of the update of [8, 9]. Appendix A reproduces this work, as intended for publication in the Proceedings of the 31$^{\text{st}}$ ICRC.

In light of the correlation-claim by Auger based on the use of a particular catalog of known objects, in Chapter B four *catalog independent* methods for detecting anisotropy are described and applied to both simulated and actual CR arrival directions. The most interesting and powerful of these methods uses a *shape-strength* parameterization of triplets of arrival directions and can distinguish between possible point-like and string-like anisotropies. This method is developed further in Chapter 8, where it is also closely compared with a similar method that uses doublets of events. Finally, a summary of this thesis is presented in Chapter 9.

Chapter 2

Theoretical Context: Energy and Anisotropy

For the moment, the field of extremely high energy cosmic rays is driven by observational science. While a theoretical understanding of the relevant physics is absolutely essential, there are simply so few events at the relevant energy that many of the most interesting models are difficult to discriminate.

Two paradigms for cosmic ray acceleration, "top-down" and "bottom-up" (described in §2.1), dominate discussion in the field. The latter scenario is strongly favored by contemporary observations that support a model where nucleons are accelerated in extragalactic sources; predictions for the energy spectrum in this case are consistent with the observed spectrum, as discussed in §2.2. Later in §2.3, the theoretical issues surrounding the possible anisotropy of arrival directions of cosmic rays generated by bottom-up mechanisms is described.

2.1 Possible Accelerators

For the highest energy cosmic rays, the center of mass energy is nearly an order of magnitude greater than the energy generated by the Large Hadron Collider. What natural mechanism could possibly accelerate a particle to such an extremely high energy? This is a question at the heart of high energy cosmic ray physics and, until recently, there were two schools of thought used to address it: the "top-down" and "bottom-up" paradigms.

2.1.1 Top-Down Origination

The so-called top-down paradigm attributes the origin of highly energetic cosmic rays to the decay of hypothesized heavy "particles" into standard model particles which are then observed in the atmosphere. As will be clear by the end of this section, these models are not currently favored by the world data set. [1]

Figure 2.1 contains a schematic of some typical top-down generation mechanisms. As depicted starting at the left in Figure 2.1, suppose that at some point very shortly after the formation of the Universe there came to be a distinct region of space-time defined such that, on its boundaries, the Higgs field has null expectation $\langle \phi_H \rangle = 0$. This region – and all of the fields within it – will be causally distinct from the rest of the universe; it is topologically defective. The most viable topologic possibility in this scenario is the zero dimensional case; the monopoles, acting as true magnetic monopoles, could be long-lived and interact with gravity, i.e. behave somewhat like dark matter. Such a topologic defect must then decay into the known paths of the Standard Model, ultimately producing γ-rays, neutrinos, electrons, and nucleons –

[1] This paradigm is described here both because it is theoretically interesting, and because, at the time of the publication of Chapter 4, it had not been as convincingly ruled out.

the particles observed in the atmosphere by CR detectors – with about five percent of the initial energy in nucleons.

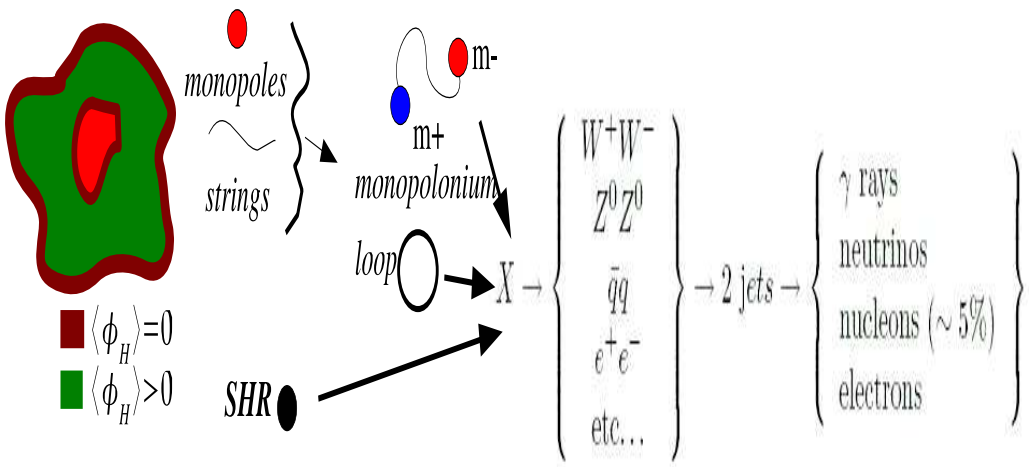


Figure 2.1: Schematic of possible top-down cosmic ray origination. See the text for explanation.

All of the top-down scenarios produce copious amounts of γ-ray photons and, as demonstrated in Figure 2.2, it is the lack of photons in observed data that has helped rule out these models. It is important to emphasize that the upper limits on the fraction of the observed photons shown in Figure 2.2 for various observatories are upper limits on the fraction of photon candidates, since exact event composition cannot be reliably determined for any one event. That said, the limits reported by the Auger Collaboration[10], in particular, tend to disfavor the super heavy dark matter (SHDM) and topological defect (TD) models.

The predicted mass of the top-down "particles" and their tendency to behave like dark matter suggests that they should be relatively near by when they decay, and that they could have energies exceeding 100 EeV. As will be described in §2.2, the consistency of the observed energy spectrum of cosmic rays with particles that originate much further away than our galactic dark matter halo presents another

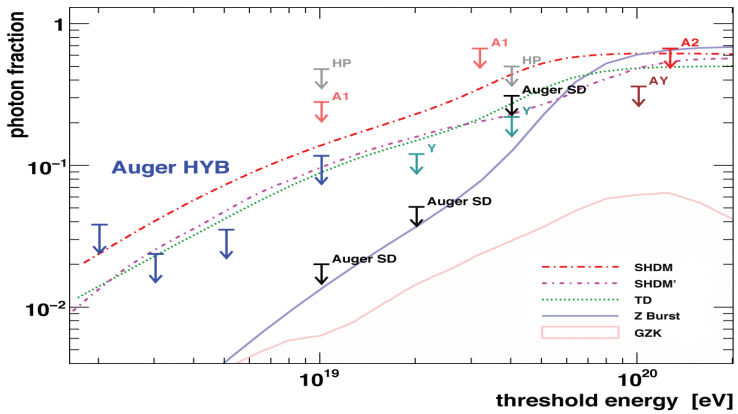


Figure 2.2: The upper limits on the fraction of photon in data collected by different observatories: AGASA(A1, A2), AGASA-Yakutsk(AY), Yakutsk (Y), Haverah Park (HP). In black the limits from the Auger surface detector (AugerSD; see Chapter 3) and in blue the limits above 2, 3, 5, and 10 EeV (AugerHYB). The shaded region shows the expected GZK photon fractionas. Lines indicate predictions from top-down models. Plot and caption adapted from [10].

compelling reason to disfavor this class of models.

2.1.2 Bottom-Up Acceleration

The second, and currently most-favored, explanation for the origin of ultra high energy cosmic rays is that they are charged nucleons accelerated via standard (though extreme) astrophysical mechanisms.

A popular acceleration mechanism has been studied in depth (see [11] and the references therein for some reviews). By far, the simplest and most illustrative example is the so called (non-relativistic) Fermi mechanism. This mechanism is illustrated schematically in Figure 2.3.

Suppose that after each encounter with the shock wave a particle with initial

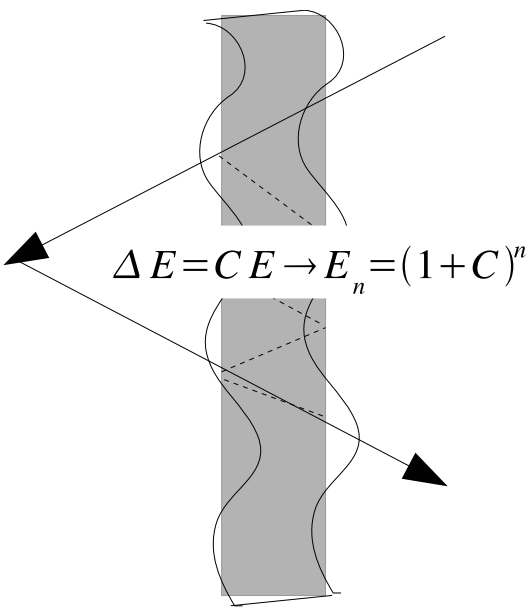


Figure 2.3: A schematic representation of the Fermi mechanism.

energy E_0 gains an energy $\Delta E = CE$ such that, after n encounters,

$$E_n = E_0(1+C)^n. \tag{2.1}$$

The number of encounters to reach a particular energy E is thus found by solving (2.1) for n,

$$n = \frac{\ln(E/E_0)}{\ln(1+C)}.. \tag{2.2}$$

If $(1-P_{esc})^n$ is the probability that particle remains within the acceleration region (for n hits), then the proportion of particles with energy $> E$ is

$$N(\geq E) \propto \sum_{k=n}^{\infty}(1-P_{esc})^k = \frac{(1-P_{esc})^n}{P_{esc}} \Rightarrow \frac{1}{P_{esc}}\left(\frac{E}{E_0}\right)^{-\gamma(P_{esc},C)}, \tag{2.3}$$

where

$$\gamma(P_{esc},C) = \ln\left(\frac{1}{1-P_{esc}}\right)/\ln(1+C) \approx \frac{P_{esc}}{C} \equiv \frac{1}{C}\frac{T_{cycle}}{T_{esc}}. \tag{2.4}$$

By this line of reasoning, the number of events greater than a given energy, i.e. Eq(2.3), is a power-law as a function of energy. The index of the power-law γ in Eq(2.5) is defined in terms of an arbitrary normalization C and the characteristic times of acceleration T_{cycle}, the time spent on each cycle within the acceleration region, and T_{esc}, the average total escape time.

The maximum generation energy for this simple application is given by

$$E \leq E_0(1+C)^{T_{esc}/T_{cycle}} \tag{2.5}$$

So, for a type-II supernovae blowing $10M_\odot$ into the interstellar medium over $T_{FA} \sim 10^3 yrs$, for example, $E_{max} \sim 1 \times 10^{14}$ eV – much less than the 10^{20} eV particle observed by Linsely.

The fact that the observed frequency of cosmic rays as a function of energy very nearly follows a power-law over many orders of magnitude has bolstered support for bottom-up models. But, in general, the simple Fermi mechanism is not sufficient to explain the highest energy cosmic rays. Relativistic generalizations along with detailed accounting of the extreme environments (like the center of active galactic nuclei), however, can produce the desired theoretical effect.

The loose upper limit given in Eq(2.5) shows that the maximum energy is dictated by the time the particle spends in the acceleration region. This time is dominated by the strength of the magnetic field B and the Larmor radius L of a given region of acceleration, allowing us to write the maximum energy as

$$E \approx \beta Z \frac{B}{1\mu\mathrm{G}} \frac{L}{1\mathrm{k}pc} \mathrm{E}eV. \tag{2.6}$$

Figure 2.4 shows a plot of the size of an object versus the strength of the object's magnetic field for various astrophysical acceleration candidates. For protons traveling at nearly the speed of light ($\beta \approx 1$), this plot suggests that active galactic nuclei (AGN) and other extragalactic structures are viable candidates, for instance.

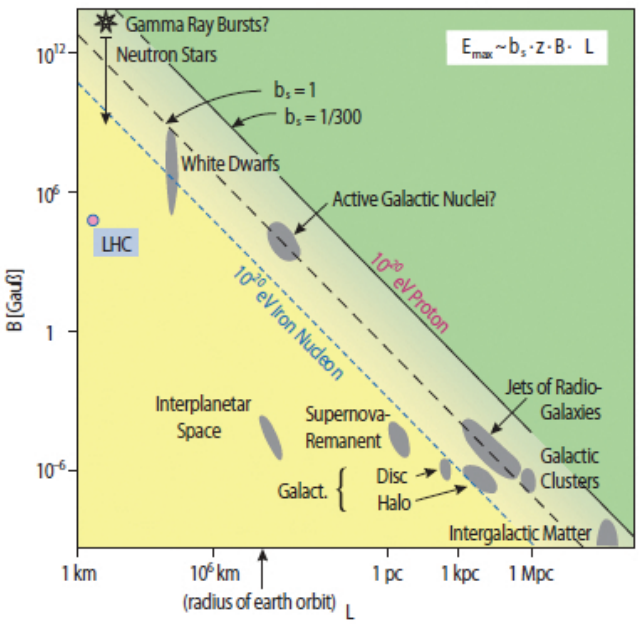


Figure 2.4: Classes of possible sites for acceleration of cosmic rays. This type of plot was developed by A. M. Hillas and this specific plot is from [12].

2.2 The Energy Spectrum

A primary focus of this thesis is the detailed study of the observed energy spectrum at the highest energies. The analysis presented in Chapters 4 and 5, in particular, is intended to confirm (or deny) the presence of a dramatic steepening, or cut-off, in the energy spectrum known as the GZK cut-off, after [13] and [14].

2.2.1 Ultra High Energy

The cosmic ray energy spectrum generally follows the power-law form over many orders of magnitude (see Figure 2.5). The mathematics of the power-law family of distributions has been a major inspiration for this thesis – indeed, it is integral to

the TP-statistic presented in Chapters 4 and 5 below.

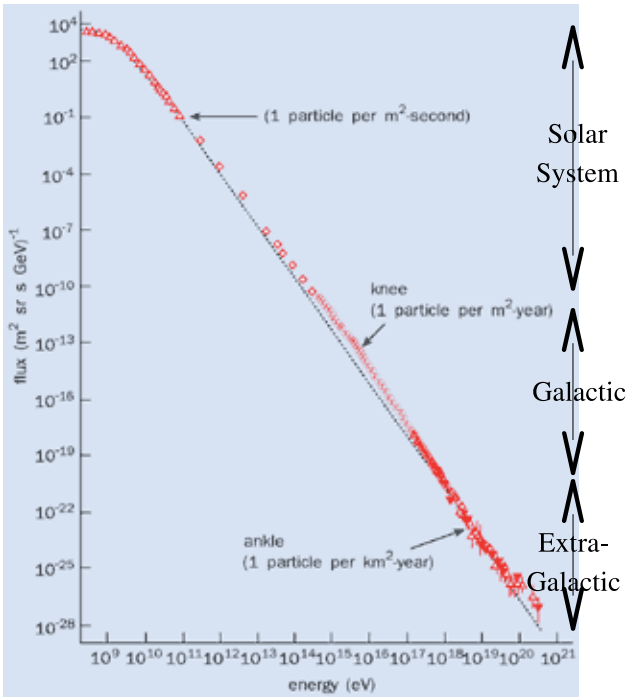


Figure 2.5: The cosmic ray energy spectrum generally follows the power-law form over many orders of magnitude. The scale of the sources of the flux are given very approximately by the black arrows.

The flux at the highest energies is incredibly low (less than 1 particle per square kilometer per year in Figure 2.5). Despite this low flux, many observations at the highest energy have been made, as shown in Figure 2.6 for several experiments. (Note the interesting characteristic energies for the LHC and the Tevatron there, in particular).

Two important facts must be considered while interpreting this figure: first, there are very few events with energy greater than 10 EeV and second, the flux measurements conducted by different collaborations seem to have different normalizations. (Note that the absolute energy scale (horizonatal axis) uncertainties result in scaled-

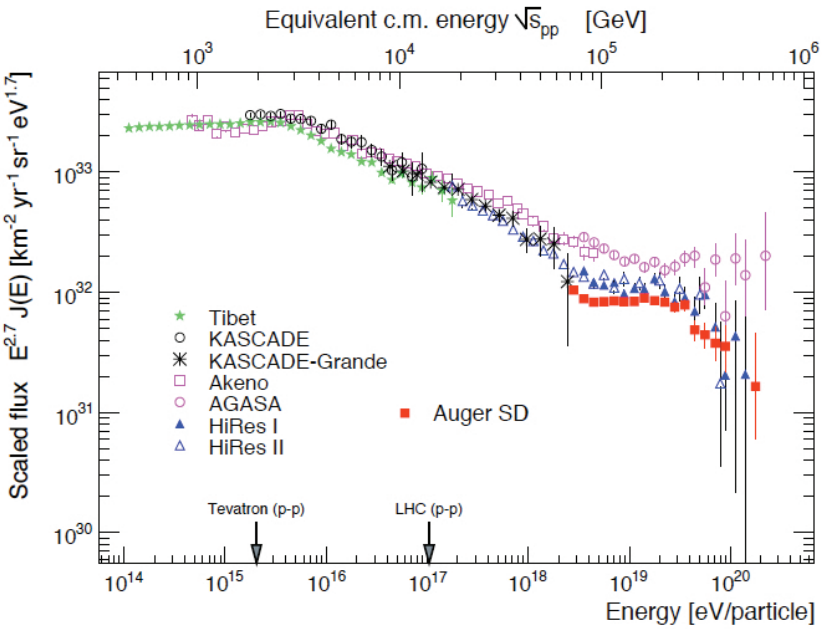


Figure 2.6: The flux of cosmic rays as a function of energy. The top horizontal axis is labeled in terms of the center of mass (c.m.) energy of each event. Notice that the flux steepens near $\sim$ 50 EeV and that there are very few events in excess 100 EeV. This specific plot is from [12].

flux (verticle axis) offsets because of the $E^{2.7}$ factor.) The question of normalization is a purely experimental one; the absolute calibration of modern CR detectors is exceedingly difficult (see Chapter 3), and uncertainties in the absolute values of characteristic energies observed by real detectors can exceed 20%. The deficit of observed events above above 5 EeV, on the other hand, has theoretically-motivated origins, as reviewed in the next section.

2.2.2 The GZK Effect

After the discovery of the cosmic microwave background (CMB) radiation by Penzias and Wilson[15] Greisen, Zatsepin and Kuzmin[13, 14] recognized that very high

energy protons could scatter off of the CMB via the $\Delta(1232)$ resonance.

A schematic representation of the GZK process is shown in Figure 2.7. A proton with energy $\sim 10^{20}$ eV interacts with a 10^{-3} eV photon from the CMB to produce a $\Delta(1232)$ resonance. This particle quickly decays into a proton (or neutron) and a pion. (Note that a neutron will decay into a proton within about 10 Mpc at this energy). This process results in significant loss of energy to the final proton; a proton with initial energy greater than 100 EeV (10^{20} eV) will typically have an energy below 100 EeV within about 50–100 Mpc (see Figure 2.8). Thus, the universe should be opaque to CR with energy greater than the characteristic GZK energy, ~ 50 EeV.

$$p + \gamma_{2.7K} \rightarrow \Delta^{+} \begin{array}{l} \nearrow \; p + \pi^0 \\ \searrow \; n + \pi^+ \end{array}$$

Figure 2.7: A schematic representation of the GZK process. A proton with energy $\sim 10^{20}$ eV interacts with a photon from the CMB to produce a $\Delta(1232)$ resonance. This particle quickly decays into a protons (or neutron) and a pion. This process results in significant loss of energy to the nucleon.

2.3 Anisotropy

The possibility of cosmic ray astronomy – or at least a general detection of anisotropy – has, for many decades, motivated both experimental and theoretical physicists, alike. This is primarily due to the understanding that, if the arrival directions of the CR could be convincingly correlated with individual sources (that had been previously identified via other means), then the task of determining the precise acceleration mechanism and site of origin would be greatly simplified. In this respect, the CR to AGN correlation recently reported by the Auger collaboration (reviewed in

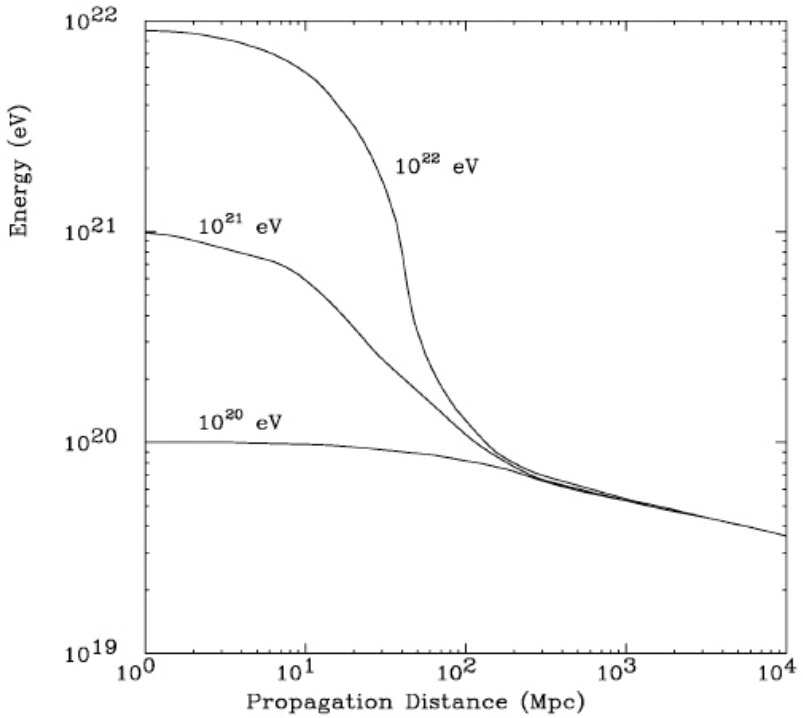


Figure 2.8: The energy of a proton as a function of the distance traveled through the CMB for three different initial energies. No matter what the initial energy, all particles have an energy less than 10^{20} EeV within a distance of about 100 Mpc of their source. This is due to the GZK effect. This plot is from [16].

Chapter A) has generated much excitement. However, the goal of CR astronomy has not yet been reached. The second part of the major analysis in this thesis concerns methods for studying the anisotropy of very high energy CR.

2.3.1 Extra-Galactic Sources

Because very high energy events can overcome the magnitude of typical (though estimated) galactic magnetic fields and escape their large spacial extent, it is expected that these events should arrive from extra galactic sources. This perspective is sup-

ported by the observational reality that CR are highly *isotropic*, except for the small and specific exceptions published by Auger [8, 9] (see below).

2.3.2 Magnetic Fields

An important consideration when attempting to study anisotropy in CR is that of the galactic and intergalactic magnetic fields between the source and Earth, which are able to deflect the arrival direction away from the source location. The amount of deflection can vary a great deal, depending on the (at least) the following quantities: the distance of the source from Earth; the charge of the CR; and the strength, extent and orientation of the galactic and intergalactic magnetic fields. The actual values of all of these quantities are not well known and those for magnetic fields are especially hotly debated. Nevertheless, it is illustrative to assume "typical" values for these parameters, and approximate the spread of the expected deflection as a CR travels from a (nearby) source to Earth as [17]

$$\begin{aligned} \theta_{rms} &= 1.1^\circ Z \left(\frac{E}{10^{20}\text{EeV}}\right)^{-1} \left(\frac{r}{10\text{Mpc}}\right)^{1/2} \left(\frac{l_c}{1\text{Mpc}}\right)^{1/2} \left(\frac{B}{10^{-9}\text{G}}\right) \\ &= 1.1^\circ 1 \left(\frac{.56\times 10^{20}}{10^{20}\text{EeV}}\right)^{-1} \left(\frac{75\text{Mpc}}{10\text{Mpc}}\right)^{1/2} \left(\frac{1\text{Mpc}}{1\text{Mpc}}\right)^{1/2} \left(\frac{10^{-9}\text{G}}{10^{-9}\text{G}}\right) \\ &\approx 6^\circ, \end{aligned} \tag{2.7}$$

where $E = 56$ EeV is the energy of the proton, $r = 75$ Mpc is the distance from Earth to the source, and the the intergalactic magnetic field of strength $B = 1$ nG has coherence length $l_c = 1$ Mpc.

Even with this rough estimate, it is clear that the deflection covers a very large area of the sky; a six degree radius corresponds to many times the size of the full moon. Indeed, one could easily imagine the CR from sources as far apart as 12°

originating from the same source based on this estimate. In this case, if there is a multitude of "6°-sources" in our sky, then CR astronomy may only be possible with many hundreds of events.

Assuming magnetic field strengths are correlated with the presence of matter, then the arrival directs of CRs are also dependent on the intervening distribution of matter. In the local Universe ($r \lesssim 150$ Mpc) the matter is *not* isotropically distributed on the celestial sphere, and nearby galactic clusters like Virgo and Centaurus dominate our neighborhood. Detailed CR propagation simulations have been carried out (see [18]) that take into account this specific local structure (as well as specific magnetic field estimates). A sky map showing the angular deflections predicted by such a model is shown in Figure 2.9. This particular map of deflections is one of the more optimistic in the reviewed literature – showing deflections at about the degree scale – in contrast with other, independent calculations [18].

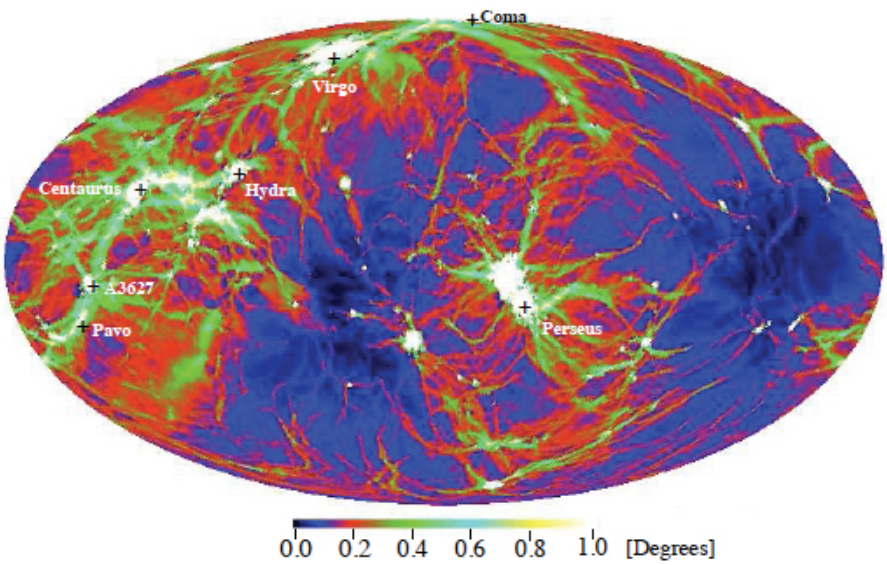


Figure 2.9: Angular deflections of CR as modeled by a detailed simulation of CR propagation in the local Universe. This plot is from [18].

It should be evident from Eq(2.8), Figure 2.9 and Figure 2.10 that the highest energy events offer the best prospect of CR astronomy, no matter what the strength of the intervening magnetic fields; the higher the energy, the smaller the scattering.

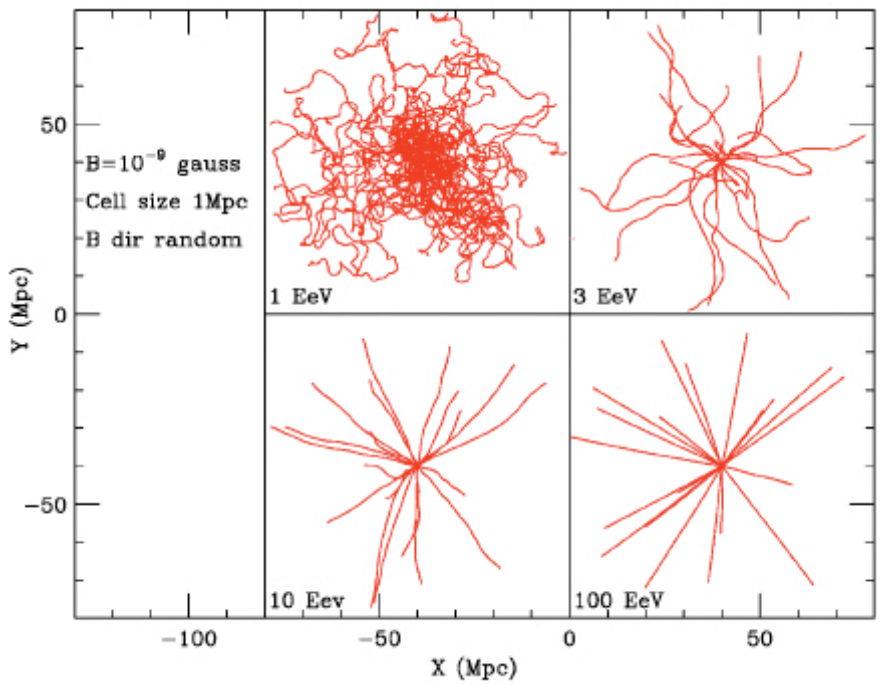


Figure 2.10: A projected view of 20 proton trajectories in the presence of a magnetic field. After 40 Mpc the high energy events have a notably smaller deflection than the lower energies. This plot is from [18].

2.3.3 Tantalizing Suspects

Perhaps the most interesting observation in recent years in the field of CR physics is that of "correlation of [high energy] cosmic rays with nearby objects in the Véron-Cetty and Véron (VCV)] catalog" by Auger[8, 9]. While notable "warm spots" have been observed in the CR sky in the past (see at least [19]), this result stands apart

for at least a few reasons. First, the number of relevant events – 27 with $E \gtrsim 56$ EeV [8, 9] – is nearly double the amount previously available to the scientific community at a comparable energy. Although twenty-seven is a small number from a statistical perspective, the data is in itself quite valuable. Second, unlike with earlier claims, the observers use *two independent* data sets: an initial set of 15 to determine the "interesting" region of parameter space and a final set of 12 to confirm the result. This kind of careful approach is necessitated by the small sample and potentially physically important results.

Third, the particular method of data analysis chosen by Auger is highly suggestive. In the scheme used, CR are correlated with *multiple* active galactic nuclei in the VCV catalog that have been identified by more standard astronomical techniques. Correlations are searched for by scanning over three angular variables – energy, redshift of the VCV object, and angular distance between CR and VCV object – with the most anomalous configuration then tested in the later data. As reported by [8, 9], the parameters which maximize the correlation are 56 EeV, a redshift of 0.018 (~ 75 Mpc) and an angular distance of 3.1°. These parameters are remarkably consistent with the notion that CR originate from active galactic nuclei within ~ 75 Mpc and that the angular deflections are "small." (Of course, the authors of [8, 9] do not make this claim.) However, the scan procedure is known to yield angular and redshift parameters that are significantly smaller than the characteristic values expected from simulations. In addition, the VCV catalog is known to be a statistically incomplete and non-flux-limited sample of AGN. While this fact does not interfere with the confirmation of anisotropy, it does conflict with the notion that "The Sources" of CR have been identified.

One should take care to note that the Auger correlation was observed at only the 1% significance level. Loosely, this means that an observer believes that only one experiment out of 100 experiments performed on a *truly isotropic* sky would give the

same result. [2] In general, within the field of particle physics this level of significance would be classified as "evidence", at best. Furthermore, the meager 27 events on which this evidence is based is barely enough for statistics on the entire sphere. Later in Chapter A the AGN to CR correlation is updated using a total of 58 events.

[2] Would you take a drug that had a 1% chance of resulting in death?

Chapter 3

The Pierre Auger Observatory

The Pierre Auger Observatory is an observatory built by an international collaboration of scientists with the primary goal to observe and understand three main characteristics of ultra high energy cosmic radiation: the energy spectrum, the possibility of anisotropic arrival directions, and the composition of cosmic rays. The Auger detector is specifically designed to make these observations with unprecedented statistics and experimental accuracy; because the CR flux is incredibly low at the highest energies – about 1 particle/km^2/sr/century – the detector is spread over $\sim$3000 km^2 in order to achieve reasonable statistics.

Located near Malargúe Argentina, the observatory consists of an array of (nearly) 1600 water Cherenkov detectors and 4 fluorescence telescope sites. The array of water tanks comprise the Surface Detector (SD), which is used to observe the lateral distribution of particle in a shower at ground level, while four separate batteries of six Fluorescence Detectors (FD) observe lateral shower development via the miniscule fluorescence light emitted as the shower propagates in the atmosphere. Both of these observational techniques (the SD and FD) have advantages and disadvantages. But, it is the skillful combination of the two approaches – into a *hybrid detector* – that

allows Auger to extract both large statistics and small observational uncertainties (at least, as compared with earlier generation observatories). This hybrid concept can achieve a crucial and (nearly) absolute energy calibration along with vastly improved arrival direction determination.

3.1 Extensive Air Showers

3.1.1 The Particle Phenomenology of Air Showers

A cosmic ray primary will interact with a nucleus in the upper atmosphere with a mean free path of a few tens of g/cm^2. This initial interaction (the depth of which is subject to large fluctuations, given the stochastic nature of particle interactions together with the low density of air high in the atmosphere) generates charged pions that produce a cascade of secondary particles known as the hadronic component of the shower. The electromagnetic component of the shower is generated primarily by neutral pions decaying into $e^{\pm}$ pairs which, in turn, radiate photons. A lesser part comes from $\pi^{\pm} \rightarrow \mu^{\pm} \rightarrow e^{\pm}$ decays, see Figure 3.1. As the shower develops, and the number of $e^{\pm}$ pairs increases, the energy of each pair will decrease until the primary source of energy loss is from atmospheric ionization and Compton scattering, while the number of particles in the shower meanwhile decreases.

High energy muons and neutrinos are also products of the shower development. Thus part of the shower, is not detected by the observatory and, as such, is sometimes known as the "missing energy."

A schematic of all three of these component of the air shower is shown in Figure 3.1.

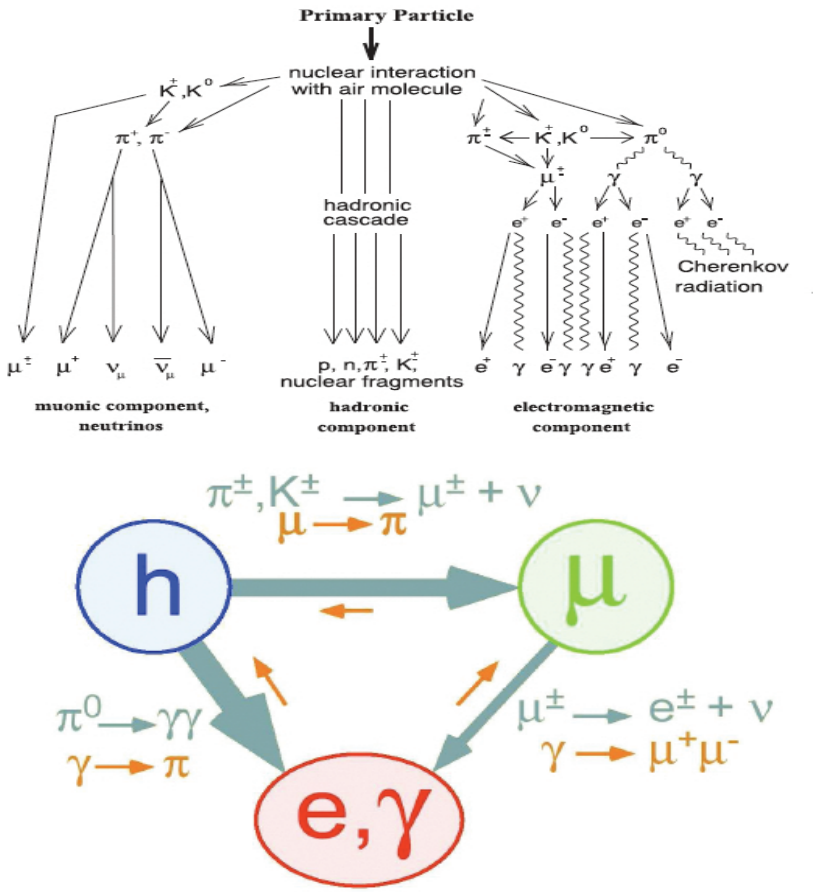


Figure 3.1: *Top:* The particle physics phenomenology of the hadronic, muonic and electromagnetic components of an extensive air shower created by a high energy CR primary (from [12]). *Bottom:* The magnitude and direction of the flow of energy between each of these components is represented by the size and direction of the arrows (from [12]).

3.1.2 Development of the Shower Front

The total distance of air crossed by the shower along its axis of travel is given by the atmospheric slant depth X. This *longitudinal* development of the shower front

is conveniently parameterized by Gaisser and Hillas [20] as

$$N_{GH}(X) = N_{max}\left(\frac{X - X_0}{X_{max} - X_0}\right)^{(X_{max}-X_0)/\lambda} \exp\left(\frac{X_{max} - X}{\lambda}\right), \tag{3.1}$$

where $N(X)$ is the number of charged particles at a particular atmospheric depth X and λ is an interaction length parameter and X_0 is the first interaction parameter. The peak of shower activity at a depth X_{max}. The number of particles at this peak is given as N_{max}.

All of these parameters can be fit for individually observed showers. Since "light" nuclei (p, He) primaries tend to travel deeper into the atmosphere than "heavy" nuclei (Fe), this parameter information – primarily X_{max} – is used for primary composition identification (i.e. distinguishing a proton from He, from C, etc.). For most primaries, the shower maximum is achieved at around 700 g/cm^2, or nearly the depth of the entire atmosphere for Auger, such that many showers can be observed near their maximum development.

But while the majority of the most energetic secondary particles are concentrated near (and along) the axis of shower development, multiple Coulomb scattering and the transverse momentum of these particle will cause some shower components (mostly low energy photons, electrons and muons) to spread away from the shower core. This *lateral development* of the shower can extend a few kilometers from the core and must be incorporated in order to properly reconstruct the shower.

The Nishimura, Kamata and Greisen (NKG) function

$$\rho(r) \propto \left(\frac{r}{r_M}\right)^{s-2}\left(1 + \frac{r}{r_M}\right)^{s-4.5}, \tag{3.2}$$

can be used to approximate the lateral distribution of shower particles as a function of distance to the shower core r, where r_M is the so called Molière radius – about 100 m at ground level for Auger – and s is the so called shower age parameter,

$$s = \frac{3X}{X + 2X_{max}}. \tag{3.3}$$

This particle density can be sampled by the array of surface detectors at "ground depth", the SD, or by fluorescence telescopes that are sensitive to the electromagnetic radiation produced higher in the atmosphere, the FD.

3.2 The Surface Detector

The geographic area of the Auger array of 1600 water Cherenkov (surface) detectors is about 3000 km^2. These tanks are arranged on a triangular grid with a spacing of 1500 m, as shown overlaid on a political map of west central Argentina in Figure 3.2.

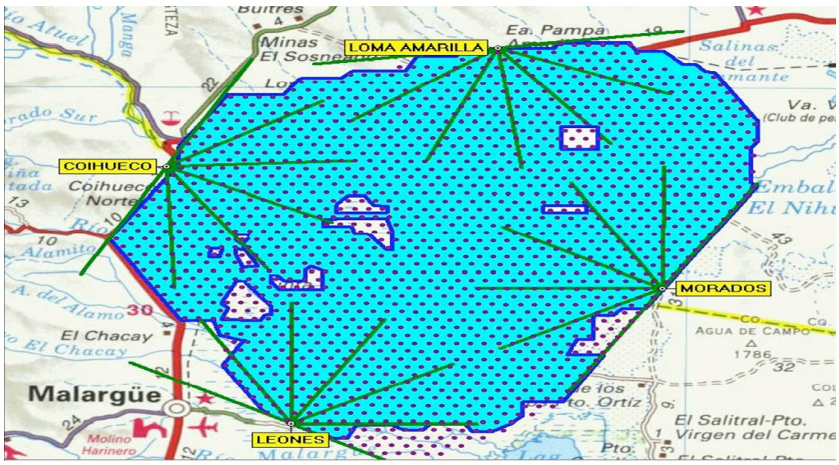


Figure 3.2: The spacial arrangement of the array of surface detectors near Malargüe Argentina. The tanks that have been deployed by the end of 2008 are shaded blue. The 4 fluorescence sights are labeled and the field of view of each telescope is marked in green.

Individual units are designed to measure the light emitted by shower particles traveling through water. Each polyethylene tank, 3.6 m in diameter and 1.55 m in height, is lined with a reflective and protective plastic/Tyvek liner and filled with 12000 liters of highly purified water. The top of the tank is fitted with 3 photo-multiplier tubes mounted so to effectively observe the water (the fiducial volume).

The anode and the diode of each tube are calibrated in order to maximize the dynamic range; the tank is sensitive to showers falling "far" from the tank and will not over-respond when a shower lands directly on top if it. Tank operation relies on autonomous electrical power provided by solar panels connected to two 12 V batteries. Each tank includes a global positioning system unit that provides timing information to within tens of nano seconds for relay to the central data acquisition system (CDAS) in northern Malargúe proper. A schematic of an individual tank can be found in the top panel of Figure 3.3. The bottom panel of the figure shows a tank as deployed on the Argentinean pompas.

When electrons, muons and photons generated by a CR shower enter the water they emit Cherenkov light [1]. Photo-multiplier tubes respond to this light, producing a digital signal in units of charge deposited which is then calibrated using the omni-directional and constant atmospheric muons into *Vertical Equivalent Muons* (VEM). As collected by a photo multiplier tube, a given signal depends (at the very least) on characteristics such as water quality, photo multiplier tube response, and liner reflection. The relation between VEM and muons arriving from all directions will be a stable parameter and is used throughout the detector as a low level threshold.

The selection of signals as candidate surface detector events is based on the implementation of a careful low- and high-level event selection algorithm. The low-level triggers (T1-T3) used early in the data acquisition protocol locate tanks with coincident signals, and search, in particular, for configurations that include nearest-neighbor coincident signals (see Figure 3.4). The T4 trigger is implemented "offline" and is the primary "physical" event selection tool that allows events to be reconstructed. Events with poor quality reconstructions are removed with the T5 trigger.

Following triggers T1-T4 (all triggers are cumulative) the T5 trigger can ensure a

[1] Strictly speaking, the photons themselves do not emit Cherenkov light, they must pair produce, or undergo Compton scattering, to produce fast $e^{\pm}$ pairs that make the Cherenkov signal.

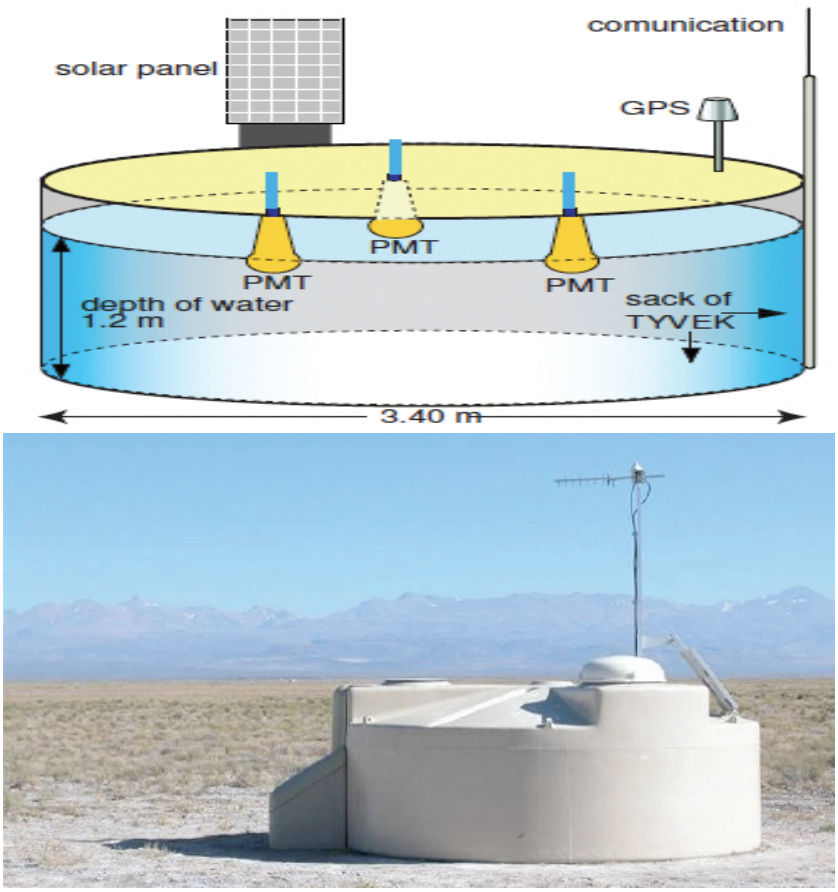


Figure 3.3: *Top:* A schematic of the major components of a water tank detector (from [12]). *Bottom:* A deployed tank on the Argentinean Pampas [21].

greater than 99% probability to reconstruct the event [22] for showers of energy above 3×10^{18} eV and zenith angles below 60°. The SD aperture can also be calculated under this condition using the well defined area of the elementary cells (4.59 km^2), as shown by [22].

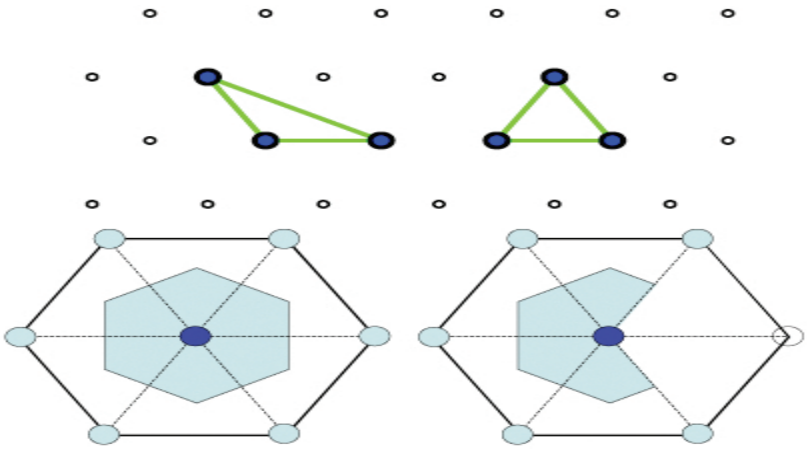


Figure 3.4: *Top:* A low level surface array trigger selects events which have signal in three neighboring tanks. Two examples are shown. *Bottom:* The higher level T5 quality cut accepts events which have signal a hexagonal array of neighboring tanks. This cut allows nearly maximal efficiency in the reconstruction. The shaded areas of active tanks are used to analytically calculate the aperture of the detector [12].

3.2.1 Arrival Direction

Once an event has satisfied at least the T4 condition (or T5 for "good data"), the reconstruction procedure can be implemented to determine the position of the shower core, the arrival direction of the primary CR, and the size of the shower, thereby supplying an estimate the energy of the primary. An in-depth discussion of the reconstruction procedure can be found in [23].

To a first approximation, an extensive air shower may be considered a planar shower-front of particles propagating through the atmosphere along a straight central axis, defined by the arrival direction of the primary CR, perpendicular to the shower front. A schematic of this approximation can be found in Figure 3.5 where the arrival direction axis is represented with the arrow labeled **a**.

The timing information of each tank – obtained via the on-board GPS units –

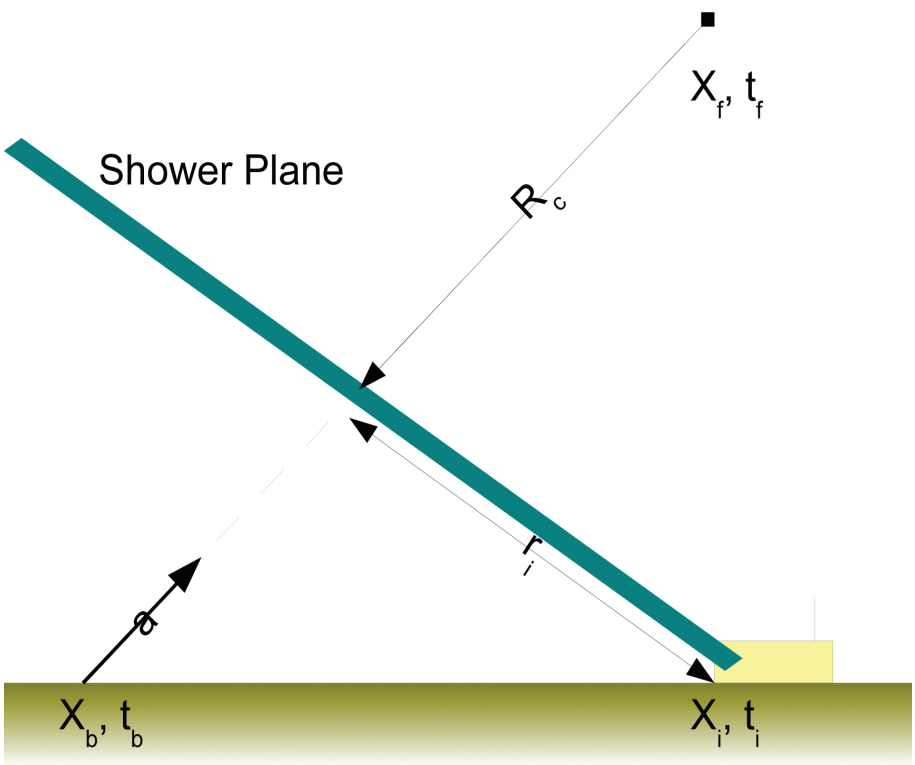


Figure 3.5: A simple schematic of the geometry of an air shower observation by an SD tank (the box).

allows the estimation of the vector **a**. The time $\hat{t}_i$ that the shower plane is expected to pass through a particular SD station located at $\vec{x}_i$ is given by

$$\hat{t}_i = t_b - \frac{\vec{a}}{c} \cdot (\vec{x}_i - \vec{x}_b), \tag{3.4}$$

where the barycenter of the shower – calculated as the weighted average of the triggered tanks – is located at $\vec{x}_b$. This latter quantity is used as the first-order approximation for the shower impact point, such that the time for the shower front to arrive at this location is given by t_b. The expected arrival time $\hat{t}_i$ can then be compared to the observed arrival time t_i via a standard χ^2 function

$$\chi^2 = \sum_{i \in \text{tanks}} \left(\frac{t_i - \hat{t}_i}{\sigma_{t_i}} \right)^2, \tag{3.5}$$

where σ_{t_i} is the station time variance in units of 10^{-9} seconds at time t_i (see [24]), and the sum is carried out over all SD stations (tanks) that have been flagged as involved in the reconstruction. The minimization of the χ^2 functional in Eq(3.4) then occurs with respect to three free parameters: the two components of the unit vector $\vec{a}$ and the time of the shower arrival at $\vec{x}_b$. Note that, with the above-described planar shower front as a first approximation, the estimated arrival directions can be further constrained in the reconstruction chain by introducing more accurate spherical shower front approximations[23].

As an example of SD angular reconstruction, consider the event with "Auger" identification number 200719804025 reconstructed by the surface detector ("SD ID" 3715992) on July 16, 2007 at 05:07:43 in the morning. This event triggered the 15 surface detector units colored in Figure 3.6. The timing information for these tanks is represented by the black squares in Figure 3.12.

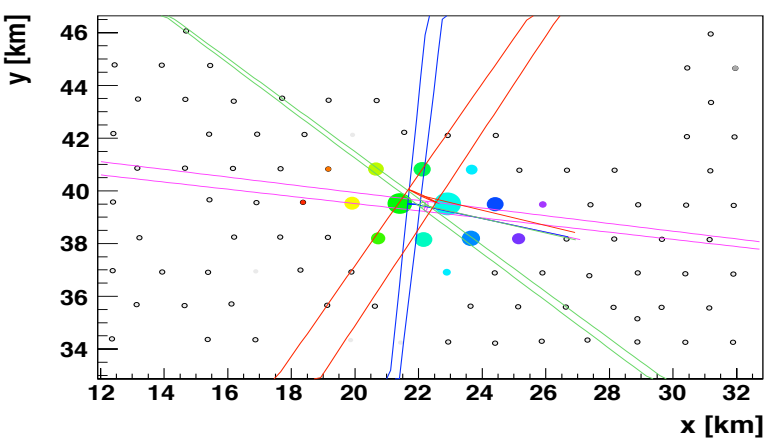


Figure 3.6: A schematic view from above of the surface detector triggering on the example event. The colored and sized circles show the response of each tank. The colored lines correspond to the shower plane for each fluorescence detector (see below).

With the full hybrid reconstruction (see §3.4), the arrival direction of this event

is found to be $(l, b) = (339, -84.6)$ in galactic coordinates. For the SD, alone, the precision achieved in the angular reconstruction is primarily influenced by the precision of the GPS clock on each individual detector unit *and* the shower-to-shower fluctuations in the arrival time of the first particle[25]. This precision can be as good as half of a degree. As can be seen in Figure 3.7, though, the angular precision of the SD is strongly dependent on the number of tanks triggered in the shower and the zenith angle θ of the arrival direction.

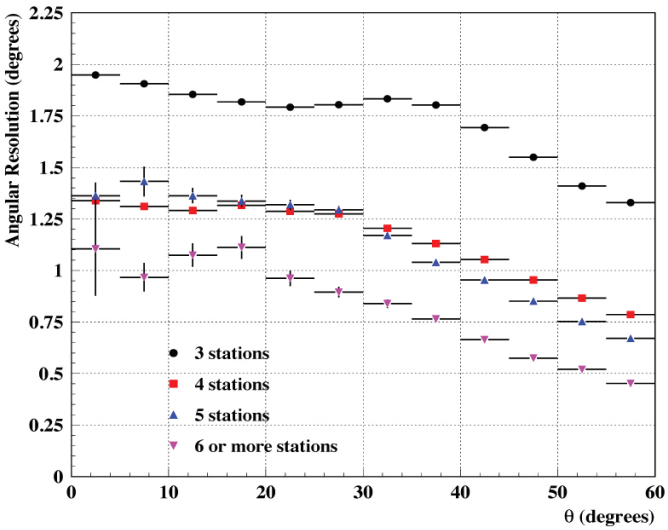


Figure 3.7: The angular resolution of SD events as a function of zenith angle. The angular resolution is plotted for various tank multiplicity (from [25]).

3.2.2 The Energy

In order to estimate the energy of the CR primary one must estimate the size of the shower. The shower size can be extracted from the station signals, via the shower geometry and core position. Specifically, the lateral distribution function introduced

in Eq(3.2) is "convoluted" [2] with the detector response to give the lateral distribution of the shower as a function of distance from the shower core r (in meters):

$$S(r) = S(1000)\left(\frac{r}{1000}\right)^{-\beta}\left(\frac{r+r_0}{1000+r_0}\right)^{-\beta}, \tag{3.6}$$

where $S(r)$ is measured in VEM and $r_0 = 700$ m. The power β is a function of the slope of the lateral distribution, parameterized as a function of the zenith angle of the shower axis (i.e. the arrival direction of the primary)[23]. The value of the signal 1000 m from the core of the shower is the normalization of this function and is the most important SD energy estimator of the shower, as noted below. With it, the standard χ^2 function

$$\chi^2 = \sum_{i\in\text{tanks}}\left(\frac{S_i - S(r_i)}{\sigma_{S_i}}\right)^2, \tag{3.7}$$

supplies the comparison between the signal in the i^{th} station S_i and the expected signal $S(r_i)$ where $\sigma_{S_i} = (0.32+0.42\sec\theta)\sqrt{S_i}$ [12] and $r_i^2 = (\vec{x}_i-\vec{x}_b)^2-(\vec{a}\cdot(\vec{x}_i-\vec{x}_b))^2$ (see Figure 3.5). Note that, implicit in the core position $\vec{x}_b$, the shower size depends on the fit to the parameters of the arrival direction. Further terms and corrections can be added to Eq(3.7) to account for tanks with over saturated photo-multiplier tubes or for tanks which do not trigger, but may be expected to, given the size of the shower[12].

The most important free parameter in Eq(3.7) is the signal 1000 meters from the shower core $S(1000)$. The use of the signal at a given distance from the shower core was first proposed by Hillas [26] and depends primarily on the spacing of the array. For the SD of Auger, the choice of one thousand meters is a good one, since it minimizes the influence of the fit of the lateral distribution function and the shower-to-shower fluctuations [27]. However, the validity of this fit only holds for arrival directions with local zenith less than 60°.

[2] Eq(3.2) shows only the $e^{\pm}$ density. There are also contributions from γ's, $\mu^{\pm}$'s and $\rho(r)$'s and every factor is included in the convolution with the detector response.

The dependence of $S(1000)$ on the zenith angle of the shower can be effectively removed using a so-called *attenuation curve* which relates the size of a shower arriving at a given zenith angle to an equivalent shower arriving at a reference zenith angle (say, $\theta = 38°$). The attenuation curve can be computed empirically via the Constant Intensity Cut (CIC; [28]) method, which ultimately relates $S(1000)$ to the energy of the primary by exploiting two assumptions – both of which are true to a good approximation in the Auger detector for the energies studied here. In the first assumption, CR are assumed isotropically distributed in zenith angle θ, while in the second, the acceptance of the detector is assumed known. Together, these assumptions imply that the the distribution of $\sin^2\theta$ should be uniform such that the differential number of events with signal above a given value per unit of $\sin^2\theta$ should be constant,

$$I(S(1000), \theta) = \frac{dN_{>S(1000)}}{d\sin^2\theta}. \tag{3.8}$$

For a given intensity I (i.e. number of events), the signal at 1000 meters is measured as a function of zenith angle and fitted with a double polynomial function

$$\mathrm{CIC}(\theta) = 1 + a(\cos^2\theta - \cos^2 38°) + b(\cos^2\theta - \cos^2 38°)^2. \tag{3.9}$$

Figure 3.8 shows different attenuation curves for different intensities. For the standard Auger reconstruction the intensity $I = 170$ VEM, $a = 0.91 \pm 0.05$ and $b = 1.22 \pm 0.26$. The quantity labeled $S_{38°}$ there represents an equivalent ($S(1000)$) signal at 38° zenith angle. This value can be calculated for each shower from $S(1000)$ via

$$S_{38°} = \frac{S(1000)}{\mathrm{CIC}(\theta)}, \tag{3.10}$$

allowing the primary CR energy to be calculated using the FD energy (determined via hybrid events, see §3.4).

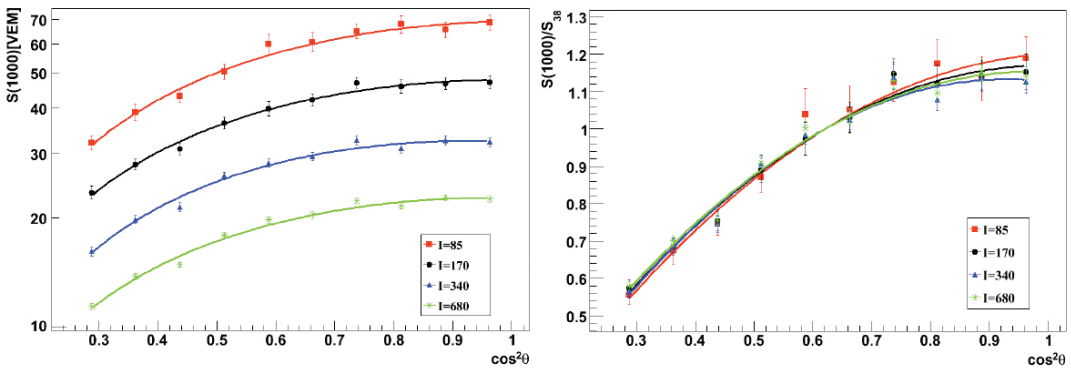


Figure 3.8: *Left:* Attenuation curves (solid lines) for different values of constant intensity I: 85, 170, 340 and 680 VEM. *Right:* Attenuation curves normalized to the signal at $\theta = 38^\circ$, S_{38° (from [21]). Note that the normalized curves are more consistent.

3.3 The Fluorescence Detector

Four separate groups of 6 telescopes overlook the SD and observe lateral shower development via the miniscule fluorescence light emitted during shower propagation in the atmosphere. Together, these constitute the the Fluorescence Detector (FD). The sites hosting each of the four groups are named after the four hills on which they are located (see Figure 3.2 for the geographic location of the FD sites): Los Leones, Los Morados, Loma Amarilla and Coihueco. This arrangement by design grants the fluorescence-detection of any event landing on the SD surface area on a cloudless and moonless night with energy greater than 10^{19} ev.

Each of the 6 telescopes at a given site (all with an individual field of view 28.6° in elevation and 30° in azimuth) are comprised of three main components: an aperture system, a spherical mirror, and a camera of photomultiplier tubes (PMTs) positioned at the focal point of the mirror (see Figure 3.9). The aperture system consists of a circular diaphragm (Schmidt optics [29]) especially designed to minimize comma

aberration; the effect of aberration within the system is less than 0.5°[21]. Near the outside of the site enclosure building, a UV filter both protects the camera system from the elements and transmits the required ultra-violet light (300-400 nm) while blocking the noisy visible light. This UV light is reflected and focused onto the PMT camera by a 3.4 m radius segmented mirror wherein each mirror piece has a reflectivity of $\sim 90\%$ between 300 and 400 nm. Designed for maximum light collection, the PMT camera consists of 440 hexagonal pixels, each instrumented with its own photomultiplier tube[30], and each covering a size of about 1.5°.

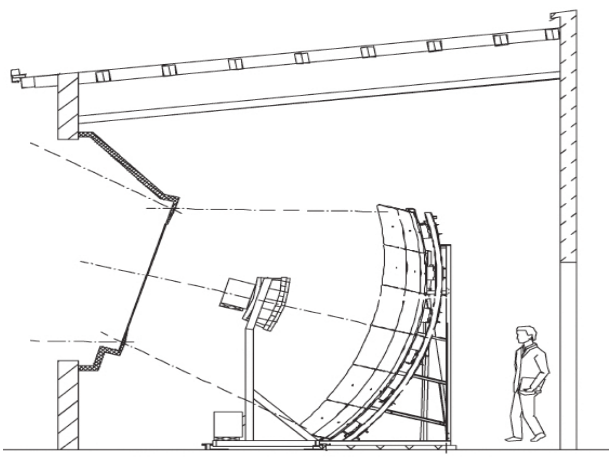


Figure 3.9: A schematic of an FD telescope including (from left to right) the shutters, aperture system (with filter and corrector ring), PMT camera, and spherical mirror (from [21]).

A series of hardware- and software-level triggers reduce the high rate of data collected by, and sent from, the PMT. The first-stage trigger is on the level of the individual pixels; a pixel will trigger if its signal is above a certain threshold. The rate for this trigger is about 100 Hz. The second-level trigger searches for patterns of triggered pixels consistent with a CR shower based on the use of the Hough transform [12] that identifies (spherically) linear alignments of pixels. This trigger rate is typically a few tenths of a Hz per telescope [21]. The third-level trigger is used

for data-readout, making cuts designed to reduce noise and enforce track-length and space-time requirements. The rate for this trigger is about 100 events per hour and each of these events is sent to CDAS for further processing. The last trigger (for a purely FD event) selects those shower events with a properly reconstructed geometry and lateral distribution function; these events are very likely to be true CR showers. This trigger also allows a given event to be considered for hybrid reconstruction; with both SD and FD telescope/site information about this shower, its reconstruction can be greatly improved.

PMT detections correspond to light emitted by N_2 molecules that are excited and ionized by electrons in the shower. Nitrogen molecules give off photons in the spectral range $\sim 300 - 400$ nm when de-excited via two primary modes: the First Negative System of N_2^+ and the Second Positive System of N_2, also called $1N$ and $2P$, respectively[31]. This light is represented schematically in green in the left panel of Figure 3.10.

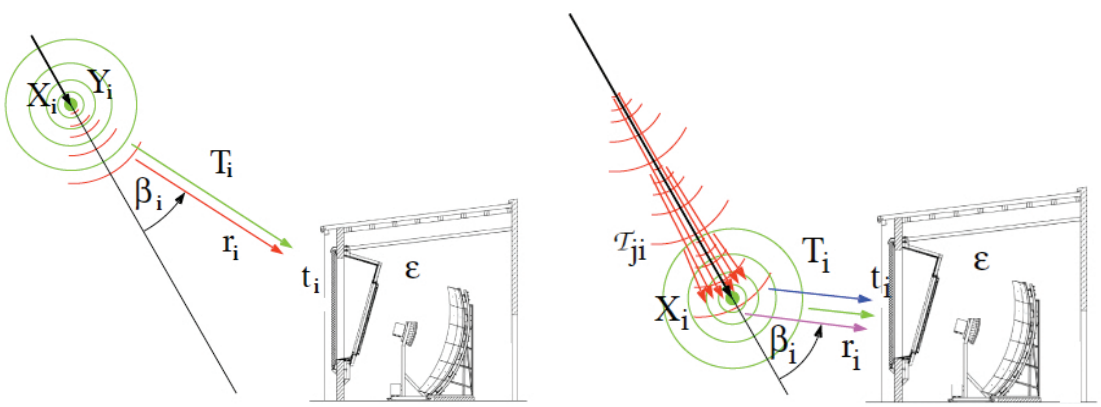


Figure 3.10: Scheme of the different light contributions arriving to an Auger fluorescence telescope: isotropic fluorescence light (green), direct Cherenkov light (red), Rayleigh-scattered Cherenkov light (blue) and Mie-scattered Cherenkov light (magenta). Figure taken from [32].

3.3.1 Arrival Direction

Each pixel in a triggered event refers to a known geometrical pointing direction $\vec{p}_i$ determined through the analysis of star tracks [33]. The 2D surface containing both the shower and the track of triggered pixels defines the shower detector plane (SDP), which can be found empirically by minimizing [12]

$$Q^2 = \sum_{i \in \text{pixels}} \frac{q_i}{\sigma_i^2} \left(\frac{\pi}{2} - \arccos \vec{n}_{SDP} \cdot \vec{p}_i \right), \tag{3.11}$$

where q_i is the signal in the i^{th} pixel, $\sigma_i = 0.35^\circ$ is determined using reconstructed laser shots from the central laser facility, and the free parameter $\vec{n}_{SDP}$ is the vector normal to the SDP. (Note that the geometric pointing direction $\vec{p}_i$ can be converted into a (local) elevation angle $\tilde{\chi}_i$ within the SDP). Timing information provided with each pixel is then used in the standard χ^2 function

$$\chi^2 = \sum_{i \in \text{pixels}} \left(\frac{t_i - \hat{t}_i(\tilde{\chi}_i)}{\sigma_i} \right)^2. \tag{3.12}$$

Here, t_i is the observed time of the measured signal in each pixel, σ_i is the uncertainty in the signal time, and $\hat{t}_i(\tilde{\chi}_i)$ denotes a theoretical expectation calculated from purely geometrical considerations [12] (see Figure 3.11). That is,

$$\hat{t}_i(\tilde{\chi}_i) = T_0 + \frac{R_p}{c} \tan \left(\frac{\chi_0 - \tilde{\chi}_i}{2} \right). \tag{3.13}$$

This expression contains the free parameters to be estimated by the fit, namely, the impact parameter R_p (i.e. the closest approach to the detector), the time T_0 that the "impact" occurs, and the angle between the shower-axis and the ground plane χ_0. Typically, these parameters, which are often highly correlated, have large experimental uncertainties. Including multiple fluorescence sites and/or elements of the SD in the reconstruction can greatly reduce this uncertainty and otherwise minimize correlations.

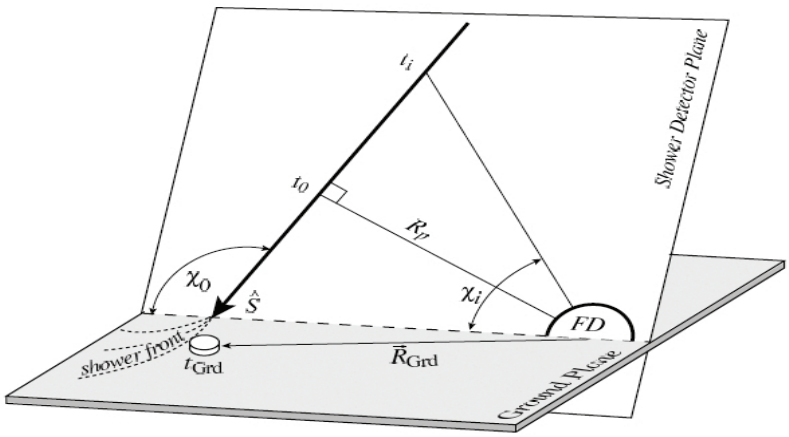


Figure 3.11: The shower detector plane as estimated by the FD. The situation of SD detectors in this geometry is also shown (from [12]).

The left panel of Figure 3.12 shows pixels from one of the telescopes at the Loma Amarillio fluorescence site which had been triggered by the example event previously described in §3.2.1. There, the pixel color indicates the time of triggering, with red for the earliest time and purple for the latest. The timing fit for the event reconstruction is shown in the right panel of Figure 3.12, with each colored point corresponding to the pixel of the same color from the neighboring (left) panel.

3.3.2 The Energy

The energy of any FD-observed shower is proportional to the total flux of light in that shower. Three main components contribute to this flux: the fluorescence light, the direct Cherenkov light, and the scattered Cherenkov light. The fluorescence light is directly proportional to the energy deposit dE/dX of the electromagnetic particles in the air shower. However, uncertainties of as much as 14% can be introduced by temperature, pressure and humidity dependent effects on the absolute atmospheric fluorescence yield.[34, 35] Currently, this particular uncertainty dominates the sys-

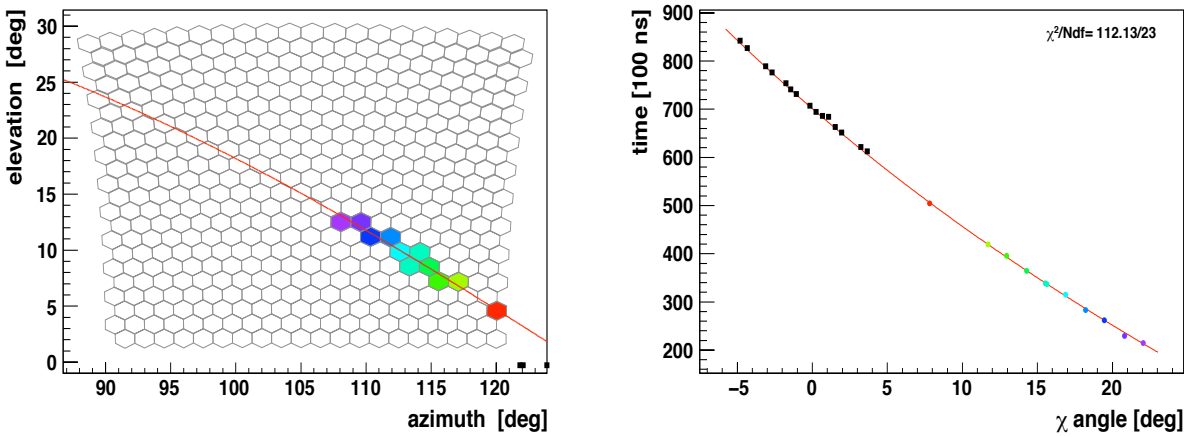


Figure 3.12: The pixels triggered in an eye of the Loma Amarillio fluorescence telescope by the example event described in in §3.2.1.

tematic uncertainty in shower energies observed by Auger.

Further attenuation of the fluorescence light due to transmission through the atmosphere is also of considerable importance to the overall flux (and hence to accurate reconstruction). This attenuation can be classified as the result of either Rayleigh scattering (i.e. a *molecular* component) or Mie attenuation (i.e. scattering off *aerosols* in the atmosphere), both of which are monitored constantly by the Auger Observatory. Finally, direct and scattered Cherenkov light caused by shower electrons radiating in the atmosphere contribute to the total flux, and these components are accounted for in the reconstruction (where appropriate terms in the fitting algorithm are included). All of these light sources are combined to construct the *Cherenkov-fluorescence matrix* which directly relates the signals in a set of pixels to the energy deposit dE/dX of the electromagnetic particles in an air shower [12].

Using a standard χ^2 function, the observed energy deposit dE/dX for each bin of atmospheric depth X can be fit to a which is then integrated over all depths (with

the best-fit parameters) to calculate the FD energy

$$E_{FD} = \int_0^\infty \frac{dE}{dX}(X)dX. \tag{3.14}$$

The function form used for $\frac{dE}{dX}(X)$ is given by the Gaisser-Hillas function (see Eq(3.1)).

For the example event described in §3.2.1, the observed energy deposit and best-fit Gaisser-Hillas function are as shown in Figure 3.13.

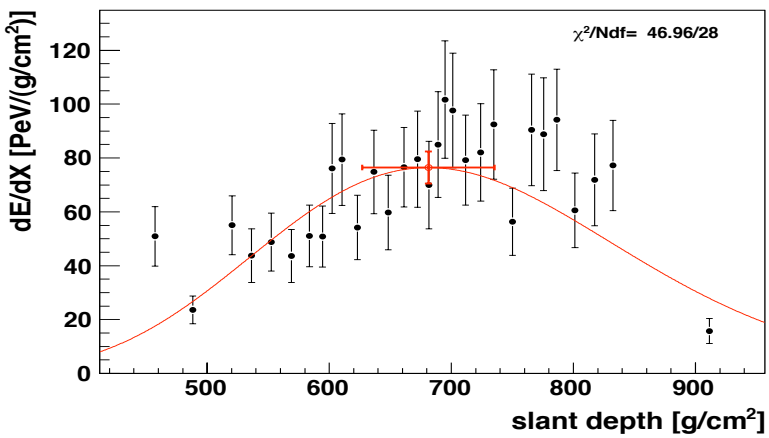


Figure 3.13: The energy deposited in the atmosphere as a function of atmospheric depth for the sample event.

3.4 Hybrid Calibrated Data

An important aspect in the design of the Auger Observatory is the combined operation of the SD and the FD. Figure 3.14 illustrates the amount of information about a single event (the example event described in §3.2.1) that can be accumulated by the combined detector. (There, the reconstructed arrival direction of the event is shown as a red line.) The simultaneous observation of an event by one FD site and a multiplicity of SD tanks is sufficient to qualify it as a *hybrid event*. In Figure 3.14,

15 SD tanks and all four FD sites registered detections for the example event shower, with timing as given in the color gradient used in all previous figures for this event.

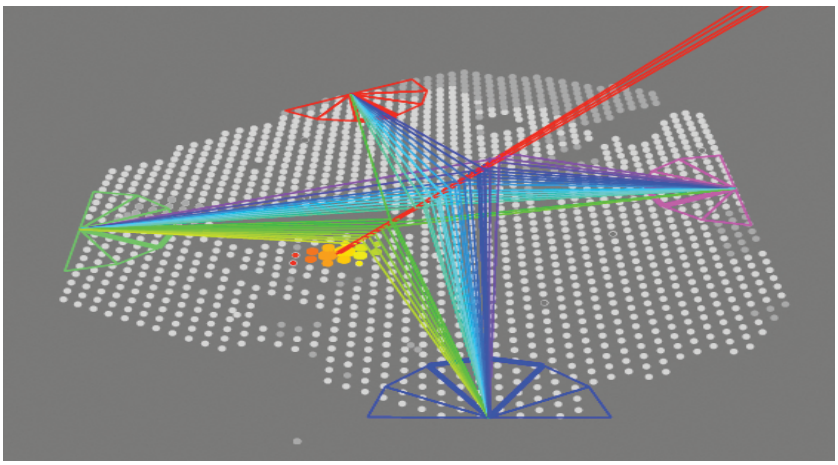


Figure 3.14: The sample event was viewed by the surface detector as well as all four fluorescence detectors.

The SD tanks triggered by a hybrid event act essentially as auxiliary "pixels" in the FD timing fit. Within the overall timing χ^2 function constructed as the sum of Eq(3.5) and Eq(3.12), these extra "pixels" serve to tightly constrain the reconstructed shower geometry. An example of the resulting fit is shown in the right panel of Figure 3.12.

While the precision of the estimated arrival direction of a hybrid event is greatly increased compared to events observed by a single component of the observatory, the energy estimator also benefits substantially from the hybrid scheme. Consider that, where the energy estimate from the SD depends on assumptions about hadronic properties at ultra high energies and the cross section of primaries of unknown type, the FD energy estimator is nearly independent of primary composition and hadronic model (at the 5% level). By integration of the energy deposited in the atmosphere it therefore offers a "direct" estimate of the energy. But, in contrast to the SD which

can collect data 100% of the time, the FD has a small duty cycle ($\sim 10\%$) and can have poor angular reconstruction. Together, though, the SD and FD supply energy estimates based on the calibration of the energies of the more frequent SD events using the FD energy of the events observed in hybrid mode.

The energy calibration is found by calculating the correlation between $\lg S_{38^\circ}$ (see Eq(3.10)) and $\lg E_{FD}$ (see Eq(3.14)) for hybrid events. This correlation is carried out by fitting a graph of the quantities (with 1σ statistical and systematic errors) to the function

$$\lg E_{FD} = \lg a + b \lg S_{38^\circ}. \tag{3.15}$$

For the 661 hybrid events observed between 1 Jan, 2004 and 31 Aug 2007 (graphed in Figure 3.15), the results of this fit (also shown in Figure 3.15) are $[a = 1.49 \pm 0.06(\text{stat}) \pm 0.12(\text{sys})] \times 10^{17}$ eV and $b = 1.08 \pm 0.01(\text{stat}) \pm 0.04(\text{sys})$ with χ^2 per degrees of freedom equal to 1.1 [36]. By replacing E_{FD} with E_{SD} in Eq(3.15) the calibrated energy of every SD event is then be estimated.

Of course, there are a myriad of factors that can contribute relatively large uncertainties in the energy estimator. The *statistical* uncertainty on the observed energy, or energy resolution, arises from the statistical uncertainties in the fitted parameters of the LDF. These errors vary from shower to shower and loosely depend on the energy; at the highest energy ($\gtrsim 10$ EeV) the energy resolution is $\sim 17\%$. The *systematic* uncertainty in the energy estimate originates primarily with uncertainty in the absolute fluorescence yield in the atmosphere. This error corresponds to 14%, in convolution with the other contributing factors outlined in Table 3.1.

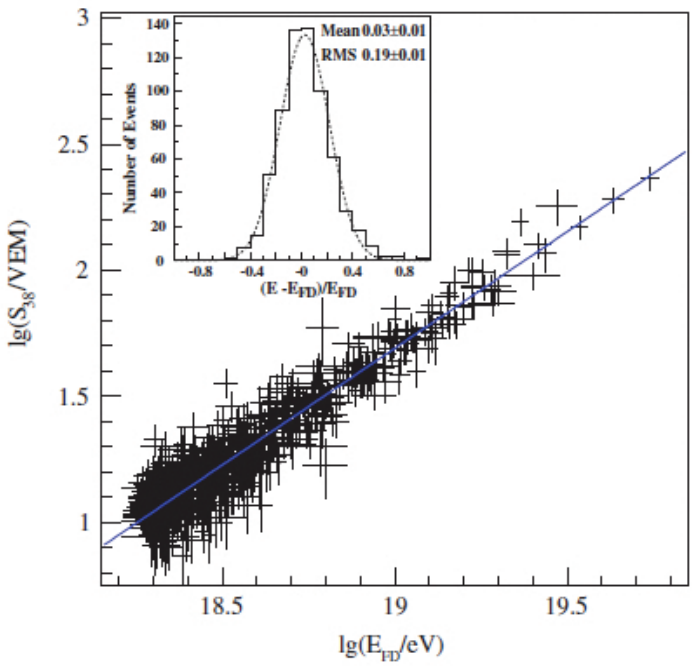


Figure 3.15: The blue line shows the fitted correlation between the SD energy estimator $\lg S_{38}$ and the FD energy estimator $\lg E_{FD}$ for the 661 hybrid events used in [36]. The fractional differences between the two energy estimators are inset. (This plot is from [36].)

Source	$\Delta E_{FD}/E_{FD}$ [%]
Absolute Fluorescence Yield	14
Altitude dependence of Fluo. Yield	10
Pressure dependence of Fluo. Spectrum	1
FD absolute calibration	9.5
FD wavelength dependence response	3
Rayleigh atmosphere	1
Wavelength dependence of aerosol scattering	1
Aerosol phase function	1
spot size and Cherenkov LDF	10
Invisible energy	4
Gaisser-Hillas parameters	3
TOTAL SYST.	23

Table 3.1: The contributions to the systematic energy uncertainties for events observed by Auger. All of these "FD" uncertainties are propagated to the final energy estimate via Eq(3.15) (from [36, 12]).

Chapter 4

Flux Suppression : Binned Analysis

4.1 Chapter Overview

The text in this chapter was published under the title "Power Laws and the Cosmic Ray Energy Spectrum" by the authors J. D. Hague, B. R. Becker, M. S. Gold and J. A. J. Matthews in in the journal *Astroparticle Physics* volume 27, pages 455-464 [7]. This paper introduces the binned statistics used to study the energy spectrum and applies them to the AGASA and *preliminary* Auger spectrum. It is published as *GAP-Note 2006-047* and as arXiv:astro-ph/0610865.

Two separate statistical tests are applied to the AGASA and preliminary Auger Cosmic Ray Energy spectra in an attempt to find deviation from a pure power-law. The first test is constructed from the probability distribution for the maximum event of a sample drawn from a power-law. The second employs the TP-statistic, a function defined to deviate from zero when the sample deviates from the power-law form, regardless of the value of the power index. The AGASA data show no significant

deviation from a power-law when subjected to both tests. Applying these tests to the Auger spectrum suggests deviation from a power-law. However, potentially large systematics on the relative energy scale prevent us from drawing definite conclusions at this time.

4.2 Introduction

Nature offers a wide range of phenomena characterized by power-law distributions: diameter of moon craters, intensity of solar flares, the wealth of the richest people[37] and intensity of terrorist attacks[38], to name a few. These distributions are so-called *heavy-tailed*, where the fractional area under the tail of the distribution is larger than that of a gaussian and there is thus more chance for samples drawn from these distributions to contain large fluctuations from the mean. Anatomical[1] defects aside, the cosmic ray (CR) energy spectrum follows a power-law for over ten orders of magnitude. The predicted abrupt deviation at the very highest energies (the GZK-cutoff[13, 14]) has generated a fury of theoretical and experimental work in the past half century. Recently, Bahcall[39] and Waxman (2003) have asserted that the observed spectra (except AGASA) are consistent with the expected flux suppression above 5×10^{19}eV. However, the incredibly low fluxes combined with as much as $\sim$50% uncertainty in the absolute energy determination means that there has yet to be a complete consensus on the existence of the GZK-cutoff energy.

With this in mind, we consider statistics which suggest an answer to a different question: *Do the observed CR spectra follow a power-law?* Specifically, these studies are designed to inquire whether or not there is a flux deviation relative to the power-law form by seeking to minimize the influence of the underlying parameters.

[1] Well known small deviations from a pure power-law are dubbed "The Knee" and "The Ankle."

The two experimental data sets considered in this study are the AGASA[40] experiment and the preliminary flux result of the Pierre Auger Observatory[41, 42]. The discussion in §A.3 uses these spectra to introduce and comment on the power-law form. The first distinct statistical test is applied to this data in §4.4 where we explore the distribution of the largest value of a sample drawn from a power-law. In §5.5.2 we apply the TP-statistic to the CR flux data. This statistic is asymptotically zero for pure power-law samples *regardless* of the value power index and therefore offers a (nearly) parameter free method of determining deviation from the power-law form. The final section summarizes our results.

4.3 The Data

A random variable X is said to follow a power-law distribution if the probability of observing a value between x and $x + dx$ is $f(x)dx$ where $f(x) = Cx^{-\gamma}$. Normalizing this function such that $\int_{x_{min}}^{\infty} f(x)dx = 1$ gives,

$$f_X(x) = \frac{\gamma - 1}{x_{min}} \left(\frac{x}{x_{min}} \right)^{-\gamma}. \tag{4.1}$$

It is convenient to choose $z = x/x_{min} \Rightarrow dz = dx/x_{min}$, $1 \leq z < \infty$ and doing so yields

$$f_Z(z) = (\gamma - 1)z^{-\gamma}. \tag{4.2}$$

For reference, one minus the cumulative distribution function $F_Z(z)$ is given by,

$$1 - F_Z(z) = \int_z^{\infty} f_Z(y)dy = z^{1-\gamma}. \tag{4.3}$$

Taking the log of both sides of equation (4.1) yields

$$\log f(x) = \log A - \gamma \log x, \tag{4.4}$$

where A is an overall normalization parameter, and suggests a method of estimating γ; the *power index* is the slope of the best fit line to the logarithmically binned data (i.e. bin-centers with equally spaced logarithms). In what follows, we refer to the logarithmically binned estimate [2] of the power index as $\hat{\gamma}$ and assume that the typical χ^2/NDF is indicative of the goodness of fit. The fitting is done with two free parameters, namely A and γ.

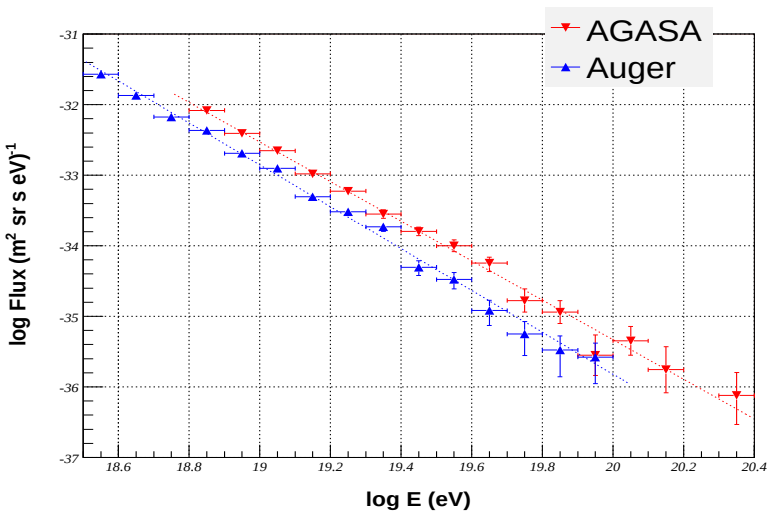


Figure 4.1: This figure displays published AGASA[40] and Auger[41] CR energy spectra. Both axis have logarithmic scales to illustrate the power-law behavior. The vertical axis is the flux J in (m^2 sr sec eV)$^{-1}$ and the horizontal axis is the energy in eV. The best fit lines (see 4.4) have slope $\hat{\gamma}_{AGASA} = 2.80 \pm 0.23$ and $\hat{\gamma}_{Auger} = 2.97 \pm 0.12$ (statistical error only).

The energy flux of two publicly available data sets are shown in Fig. 4.1. The the red point-down triangles represent the $\log_{10}$ of the binned AGASA flux values in units

[2] Unbinned maximum likelihood methods have less error and bias when applied to power-law (or similar) distributions than binned methods[43]. They can also be modified to include energy error and variable acceptance information[44]. Lacking this information, we use the logarithmically binned estimate of γ where necessary. The minimum variance for *any* (unbiased) estimator of γ is given by the Cramer-Rao lower bound; $\sigma_{\hat{\gamma}} \geq \left(\frac{\gamma+1}{N}\right)^{1/2}$.

of $(\mathrm{m}^2\ \mathrm{sr\ sec\ eV})^{-1}$ and the blue point-up triangles correspond to the Auger flux. The vertical error bars on each bin reflect the Poisson error based on the number of events in that bin. The log-binned estimates for each complete CR data set are the slopes of the dashed lines plotted in Fig. 4.1.

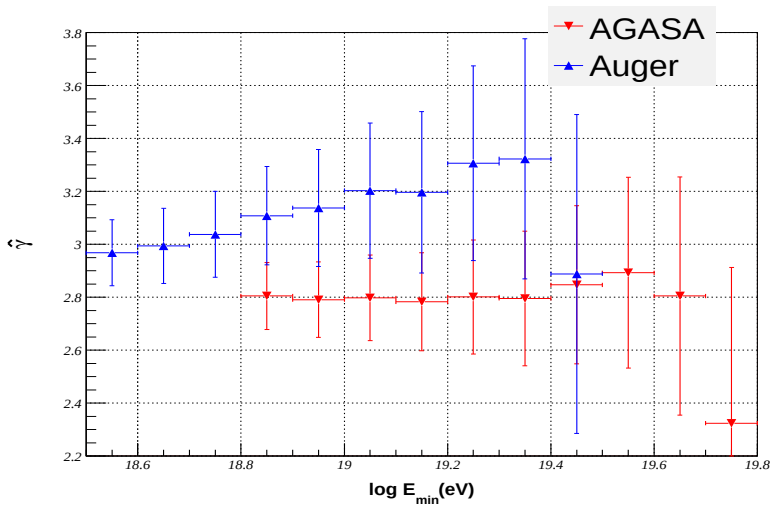


Figure 4.2: To check the stability of $\hat{\gamma}$ we estimate the power index as a function of the minimum energy E_{min} considered for the AGASA and Auger CR data sets; see Fig.4.1. The left most point is the slope of the best fit lines plotted in Fig.4.1. The vertical error bars represent $1\sigma_{\hat{\gamma}}$ deviation.

In order to check the stability of to bound on our estimate, we compute the estimated power index $\hat{\gamma}$ as a function of the minimum energy E_{min} considered for each of the two CR data sets. The left-most blue (red) point in Fig. 4.2 shows $\hat{\gamma}$ for the Auger (AGASA) data taking into account all of the bin values above $\log E_{min} = 18.5$ ($\log E_{min} = 18.8$), the next point to the right represents that for all bins above $\log E_{min} = 18.6$ ($\log E_{min} = 18.9$), and so on. The vertical error bars on these points represent the $1\sigma_{\hat{\gamma}}$ error of the estimate. To ensure an acceptable chi-squared statistic, we demand that at least five bins be considered, thereby truncating E_{min} at $\log E_{min} = 19.4$ for the Auger and $\log E_{min} = 19.7$ for the AGASA data set.

The χ^2/NDF for the left-most points is ~ 0.3 and it increases to ~ 2.5 for the right-most for both experiments. We note that these estimates do not vary widely for the lowest E_{min}'s and that the values of $\hat{\gamma}$ from these experiments are consistent.

The analyses discussed in §4.4 and §5.5.2 will depend on the total number of events in the data set. Since these numbers are not published we use a simple method for estimating them from the CR flux data. If the exposure is a constant function of the energy, then we may take the flux J to be proportional to the number of events in the bin and the exposure η, namely $N = J\eta E_{bin-center} \ln(10)/10$. The Auger exposure is reported to be constant over the energy range reported with $\eta_{Auger} = 5.5 \times 10^{16}$ (m^2 sr sec). The AGASA collaboration report flux data all the way down to $\log E_{min} = 18.5$ but the exposure of the experiment can be considered approximately constant only for energies above $\log E_{min} = 18.8$ (see Fig. 14 of [40]) where $\eta_{AGASA} = 5.1 \times 10^{16}$ (m^2 sr sec). Using this method we get a total of 3567 events with $E \geq 10^{18.5}$ for the Auger flux and 1914 with $E \geq 10^{18.8}$ for the AGASA experiment.

4.4 The Distribution of the Largest Value

As evidence suggestive of a GZK-cutoff, an often cited quantity is the flux suppression, or the ratio of the flux one would expect from a power-law to that actually observed above a given maximum, say, z_{max}. Since $J \propto N$ one may estimate the flux suppression by estimating the number of events N_{sup} out of N_{tot} expected above a given maximum as $N_{sup} = N_{tot}[1 - F_Z(z_{max})] = N_{tot} z_{max}^{1-\gamma}$. Thus, the bin with minimum $= z_{max}$ and maximum $\to \infty$ would have a height $= N_{sup}$ if the data continued to follow a power-law above z_{max}. As a test statistic for this quantity, one may consider the Poissonian probability that the bin height could statistically fluctuate to zero, namely $\mathcal{P}(0, N_{sup}) = \exp[-N_{tot} z_{max}^{1-\gamma}]$.

In this section we derive a similar test statistic based on the distribution of the maximum event from a power-law sample. The statistic discussed here approaches $\mathcal{P}(0, N_{sup})$ for large N_{tot} and allows us to show that the estimation errors associated with $\hat{\gamma}$ are enough to disallow any significant conclusion about the presence of flux suppression for the highest energy CR's.

The form of the power-law distribution allows us to calculate the pdf of the largest value, X_{max}, out of N events. Using the equations (4.1) and (4.3) we can say that the probability that any one value falls between x and $x + dx$ *and* that all of the others are less than it is $f(x)dx \times F(x)^{N-1}$. There are N ways to choose this event and so the probability for the largest value to be between x and $x + dx$ is

$$\pi(x)dx = Nf(x)F(x)^{N-1}dx.$$

In terms of the ratio z, this can be written as

$$\pi(z)dz = N(\gamma - 1)z^{-\gamma}\left(1 - z^{1-\gamma}\right)^{N-1} dz. \tag{4.5}$$

Fig. 4.3 contains a plot of this distribution for $\gamma = 3.0$ with three choices of N. The glaring implication of this plot is that even for "small" N nearly all of the integral of $\pi(z)$ is above $z \sim 10$. This implies that the probability of the maximum energy event falling below 10 times the minimum is very small, for a power-law with these parameters.

Motivated by the location and shape of $\pi(z)$ we consider the probability P that the maximum ratio from a given sample Z_{max} is less than or equal to a particular value [3], say z, in a convenient form as

$$P(Z_{max} \leq z) = \int_1^z \pi(t)dt \; = \; \left[1 - z^{1-\gamma}\right]^N. \tag{4.6}$$

[3]For large N_{tot}, equation (4.6) approaches the Poisson probability mentioned above; $\left[1 - z_{max}^{1-\gamma}\right]^{N_{tot}} \rightarrow \exp[-N_{tot} z_{max}^{1-\gamma}] = \mathcal{P}(0, N_{sup})$.

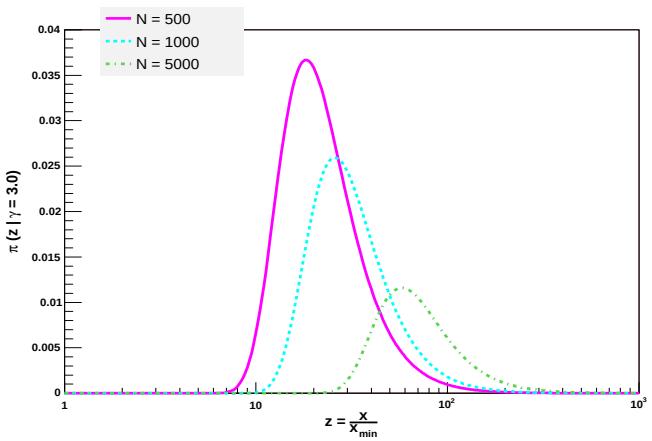


Figure 4.3: A plot of the probability distribution of the maximum of a sample drawn from a power-law with power index $\gamma = 3.0$. This is the distribution $\pi(z)$ defined in equation (4.5) where z is the ratio of the maximum to the minimum. The sample sizes are $N = 500$, 1000 and 5000.

Indeed, with $\gamma = 3.0$ (as in Fig. 4.3), $P(Z_{max} \leq 10) = 6.6 \times 10^{-3}$ for $N = 500$, 4.3×10^{-5} for $N = 1000$ and is 1.4×10^{-23} for $N = 5000$. Another way to say this is that if one were to generate 10^5 *sets* of events, each containing 1000 events drawn from a pure power-law with $\gamma = 3.0$, $\sim 99.99\%$ of these sets would have a maximum element with a value greater than 10 times the minimum. For 500 events/set the fraction decreases to $\sim 99.34\%$. Such simulations were carried out in preparation for this note and the results were consistent with equation (4.6).

To apply this idea to the CR spectrum we consider the following null hypothesis: *The flux of CR's follow a power-law with index $\hat{\gamma}$ for all energies greater than a given minimum.* As a test statistic for this hypothesis we use P, as defined in equation (4.6), with the interpretation that if the null hypothesis is true then P is the probability that the ratio of the maximum energy to the minimum is less than or equal to the observed ratio. Typically, the null hypothesis is rejected at the 5% significance level (S.L.) if $P \leq 0.05$.

To calculate the value of P for the observed data sets we need three pieces of information: the ratio of the maximum observed value to the minimum z_{max}^{obs}, the number of events N_{tot} with values in the interval $[1, z_{max}]$ and a reasonable guess for the power index γ. Since a larger z_{max} will lead to a larger value of P we will conservatively take the highest energy AGASA (resp. Auger) event to fall on the upper edge of the highest energy bin. The method of determining the number of events in each bin is described in §A.3 and here the parameter N_{tot} represents the total number above a given minimum. We will use the logarithmically binned estimates and errors of $\hat{\gamma}$ discussed in §A.3.

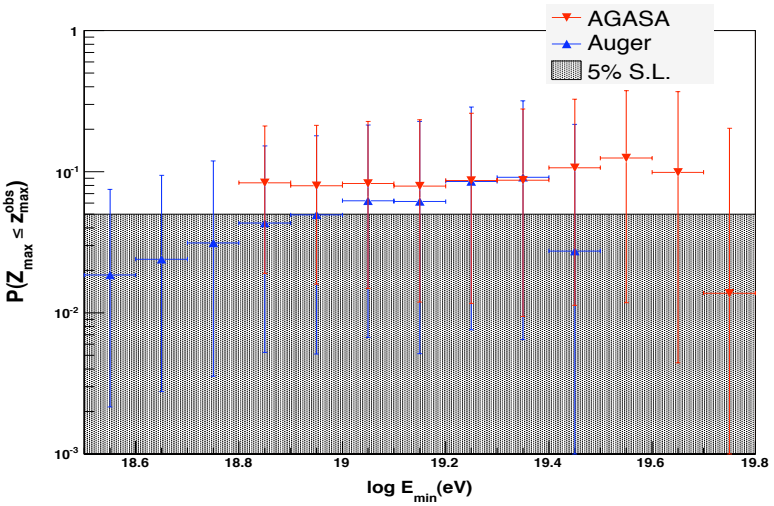


Figure 4.4: A plot of the probability that the maximum of a sample drawn from a power-law will be less than or equal to the maximum observed by the Auger ($E_{max} = 10^{20}$, blue point-up) and AGASA ($E_{max} = 10^{20.4}$, red point-down) experiments as a function of the minimum energy considered. The vertical error bars represent the effect of a $1\sigma_{\hat{\gamma}}$ deviation and the hatched area shows the 5% significance level.

The plot in Fig. 4.4 shows $P(Z_{max} \leq z_{max}^{obs})$ given N_{tot} and $\hat{\gamma}$ as a function of minimum energy considered for each of the CR data sets in Fig. 4.2. In particular, for each E_{min} the values of N_{tot}, z_{max}^{obs} and $\hat{\gamma} \pm \sigma_{\hat{\gamma}}$ are estimated from the CR flux and the

resulting P are plotted for the Auger (blue) and AGASA (red) data. For example, the left-most Auger point represents the probability that if $N_{tot} = 3567$ events are drawn from a power-law with $\hat{\gamma} = 2.97^{+0.12}_{-0.12}$ then there is a $1.9^{+5.7}_{-1.6}\%$ chance that the maximum log-ratio $\log z_{max}$ would be less than or equal to that reported by the Auger experiment, namely $\log z^{obs}_{max} = \log 10^{20}/10^{18.5}(\text{eV})$. Taken at face value, one may reject the null hypothesis at the 5% S.L. for this data set. The left-most AGASA point represents the same probability for the complete set of AGASA data, namely $P(\log Z_{max} \leq \log 1.6) = 8.4^{+13}_{-6.5}\%$ for $N_{tot} = 1914$ events drawn from a power-law with $\hat{\gamma} = 2.80^{+0.13}_{-0.13}$. Thus we cannot reject the null hypothesis for the AGASA data.

The upper (lower) vertical error bars depicted in Fig. 4.4 represent the value of P if we have under (over) estimated the power index by $\sigma_{\hat{\gamma}}$, that is if $\gamma = \hat{\gamma} \pm \sigma_{\hat{\gamma}}$, keeping the log-ratio and the total number of events constant. (The possible errors in the total number of events are on the order of a few percent and are negligible.) Since the fitting scheme considers successively lower energy bins, the points (and errors) for each experiment plotted in Fig. 4.4 are highly correlated. The upper error bars fall above the 5% S.L. for all minimums considered and therefore the statistical error associated with $\hat{\gamma}$ is enough to disallow rejection of the power-law hypothesis.

The biggest systematic measurement uncertainty in the CR data is the calibration of the energy. This uncertainty leads to an error in the reported absolute energy values of $\sim$ 30% for the AGASA [40] data and as much as $\sim$50% for the highest energy events in the Auger data set. Since the probability considered here depends only on the ratio of the observed energies, it is independent of any constant systematic uncertainty in the energy determination. However, this probability is sensitive to energy errors which vary over the range considered and will thus cause uncertainty in z^{obs}_{max}.

For example, if we take the maximum to be 50% higher (but hold $\hat{\gamma} = 2.97$ and $N_{tot} = 3567$ constant) the value of P represented by the left most Auger point in

Fig. 4.4 changes from 1.9% to 17%. Thus the large uncertainty in z_{max}^{obs} combined with the errors associated with $\hat{\gamma}$ implies that the preliminary Auger data set does not suggest sufficient evidence to reject the pure power-law hypothesis for all events above $E_{min} = 10^{18.5}$(eV).

4.5 The TP-Statistic

Considering the error and extra degree of freedom associated with γ, an analysis of a distribution's adherence to the power-law form without reference to, or regard for, this parameter is could lead to enhanced statistical power. First proposed by V. Pisarenko and D. Sornette [4], the so-called *TP-statistic*[45, 46] is a function of random variables that (in the limit of large N) tends to zero for samples drawn from a power-law, regardless of the value of γ. (TP stands for *tail power*, as oppossed to TE, also introduced in [45, 46], which stands for *tail exponential*.) This section will describe the TP-statistic and apply it to the CR data.

The raw moments of the pdf equation (4.1) are [37]

$$\langle z^m \rangle_Z = \int_1^\infty z^m f_Z(z) dz \rightarrow \begin{cases} \infty & m \geq \gamma \\ \frac{\gamma-1}{\gamma-1-m} & m < \gamma. \end{cases} \tag{4.7}$$

Thus power-laws with $\gamma \leq 3$ have a finite mean but an infinite variance (in the limit of large N) and sample statistics created from these moments are not particularly helpful. However, taking the natural logarithm of z allows the integrals to converge and one may write (for all $\gamma \geq 2$ and $m = 0, 1, 2, \ldots$),

$$\nu_m = \langle \ln^m z \rangle_Z = \frac{m!}{(\gamma - 1)^m}. \tag{4.8}$$

[4] They studied earthquake and financial return data.

The TP-statistic is calculated by noting that $\nu_1^2 - \nu_2/2 = 0$. Therefore, if we use the sample analog of these quantities, namely

$$\hat{\nu}_m = \frac{1}{N}\sum_{i=1}^{N} \ln^m \frac{x_i}{x_{min}} \tag{4.9}$$

then we can define (for all $x_i \geq u$),

$$TP(u) = \left(\frac{1}{N}\sum_{i=1}^{N} \ln \frac{x_i}{u}\right)^2 - \frac{1}{2N}\sum_{i=1}^{N} \ln^2 \frac{x_i}{u}. \tag{4.10}$$

By the law of large numbers this sample statistic tends to zero as $n \to \infty$, independent of the value of γ. The TP-statistic allows us to test for a power-law like distribution without comment about the value of the power index. Furthermore, for any one sample we can vary u from the sample minimum X_{min} to the sample maximum X_{max} and calculate the TP-statistic over the range of x in the sample.

Given complete event lists one may use equation (4.10) to calculate the TP-statistic for the unbinned data. Since only the binned CR flux is publicly available we adapt the statistic to a binned analysis and apply it first to an example distribution with a cutoff and then to the CR data sets.

4.5.1 An Example

In order to build intuition about the TP-statistic and its variance before studying the CR data, we first apply this statistic to simulated event sets drawn from both a pure power-law distribution and a similar distribution with a cut-off. The cut-off pdf is chosen so that it mimics a power-law for the lowest values but has an abrupt (and smooth) cut-off at a particular value, say x_{cut}. The functional form we will use here is

$$g(x) = B(\gamma, x_{min}, x_{cut}) \frac{x^{-\gamma}}{e^{x - x_{cut}} + 1}. \tag{4.11}$$

The normalization of this pdf is $B(\gamma, x_{min}, x_{cut})$, the value of which must be computed numerically. Fig. 4.5 contains a logarithmically binned histogram of 3000 events drawn from a pure power-law (black circles) with $x_{min} = 1.0$ and $\gamma = 3.0$, and two pdf's in the from of equation (4.11); the magenta squares have $\log x_{cut} = 1.0$ and the green triangles have $\log x_{cut} = 1.5$. While arbitrary, the values of these parameters are chosen to be similar to the AGASA and Auger data (see Fig.4.1).

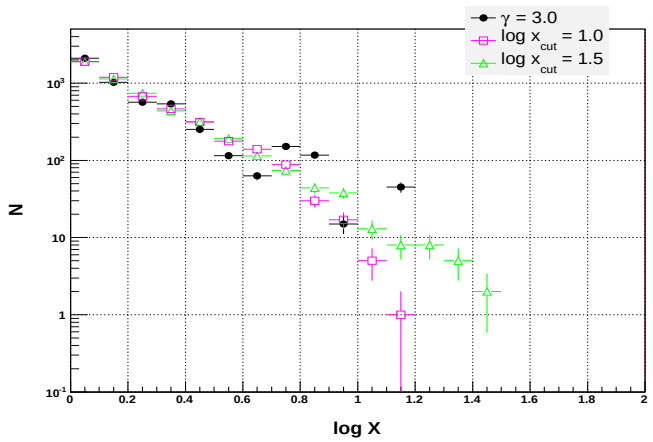


Figure 4.5: Logarithmically binned histogram of 3000 events drawn from a pure power-law with $\gamma = 3.0$ and two power-laws with a cut, see equation (4.11). The magenta squares are drawn from the distribution with $\log x_{cut} = 1.0$ and the green triangles have $\log x_{cut} = 1.5$. As noted in the text, while arbitrary, the values of these parameters are chosen to be similar to the AGASA and Auger data (see Fig.4.1).

If we write the sorted (from least to greatest) values from a sample as $\{X_{(1)}, X_{(2)}, \ldots, X_{(N)}\}$, the solid black line in Fig. 4.6 is created by calculating $TP(u = X_{(j)})$ for each value of the 3000 events drawn from the pure power-law histogram in Fig. 4.5. The circles represent the mean of the the statistic within the i^{th} bin, say $\overline{TP}_i$, and the vertical error bars show the root-mean-squared deviation of the statistic within the bin. Note that the total number of events considered by the statistic decreases quickly from left to right which leads to a bias in and an increasing variance of the statistic.

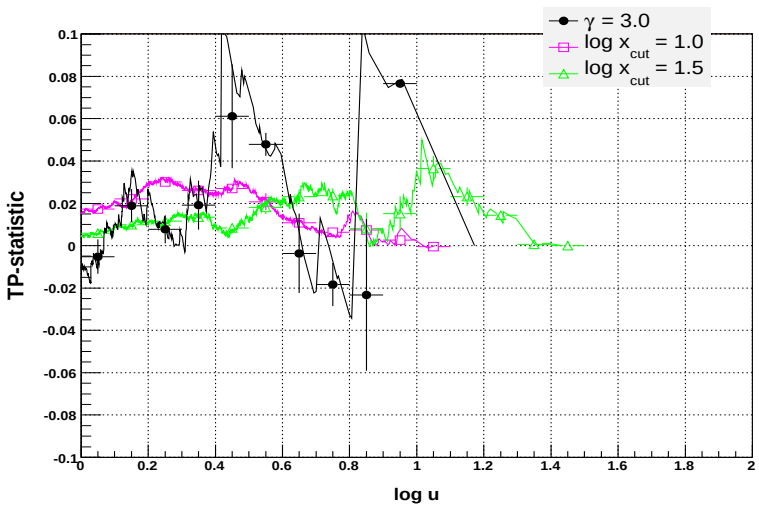


Figure 4.6: The TP-statistics, defined in equation equation (4.10), as a function of minimum value "u" for the 3 sets of 3000 events plotted in Fig. 4.5. Also plotted is the mean of the TP-statistic within each of the logarithmically spaced bins which is referred to in the text as $\overline{TP}$. The vertical error bars represent the RMS deviation of the statistic within each bin. Parenthetically, with increased statistics, say 10,000 events, the distinct characteristics of the TP-statistic for a pure power-law, a power-law with a cutoff $\log x_{cut} = 1.0$ or a power-law with a cutoff $\log x_{cut} = 1.5$ become more clearly different.

The jagged magenta line Fig. 4.6 shows the most obvious deviation from the power-law form; it is systematically offset from zero for nearly all minima of the data set. Of course, with 3000 events the histograms (see Fig. 4.5) are enough to distinguish between these two distributions. But the TP-statistic allows us to see this deviation by considering the entire data set (the left most magenta point in Fig. 4.6), not just by analyzing the events in the upper most bins. The green line in the figure shows $TP(u)$ for events drawn from equation (4.11) with $\log x_{cut} = 1.5$. The histogram for this set is not as clearly different from the power-law as the magenta points and neither is the TP-statistic; the left-most green point shows no more deviation from zero than the power-law. However, as the minimum increases

(and nears x_{cut}) the statistic moves away from zero (more noise not withstanding) and suggests that the data above the minimum deviate from the power-law.

It is important to note that the TP-statistic is positive for both of the cutoff distributions. Recall that for a pure power-law, $\nu_1^2 - \nu_2/2 = 0$. The cutoff distribution, however, lacks an extended tail and will therefore have a smaller second log-moment ν_2 as compared with (the square of) the first log-moment ν_1 and will thus result in a positive TP-statistic. A distribution with an enhancement, rather than a cutoff, in the tail would result in a negative TP-statistic, since it would have a larger second log-moment (i.e. a larger "variance"). See the Appendix (§4.7) for a detailed discussion of the TP-statistic applied to the double power-law.

To quantify the significance of the TP-statistics' deviation from zero, 10^4 sets of 3000 events were generated for each of the three distributions discussed in this section. For each set we calculate the *mean* TP-statistic $\overline{TP}$ within each of the logarithmically spaced bins. The resulting distribution of $\overline{TP}$'s within each bin is then fitted to a gaussian.

The black circles in Fig. 4.7 represent the mean of the gaussian fit to the distribution of $\overline{TP}$'s within each bin for a power-law and the error bars on the points represent the fitted 1σ deviation of the $\overline{TP}$'s. We interpret the left-most of these points in the following way: for 3000 events drawn from a power-law the "expected value" of $\overline{TP}$ in the first bin is effectively indistinguishable from zero, as expected.

Though the statistic itself does not depend on γ, the variance on this value does. The reason for this is that the variance of the $\overline{TP}$'s depends on the average total number of events greater than a given minimum, which is influenced by γ. In this case the total number of events per set for minima in the first bin is at least a few thousand and the variance of the $\overline{TP}$'s is $\sigma_{\overline{TP}} \sim 0.005$. These errors increase from left to right since each successively higher bin will contain $\overline{TP}$'s based on fewer and

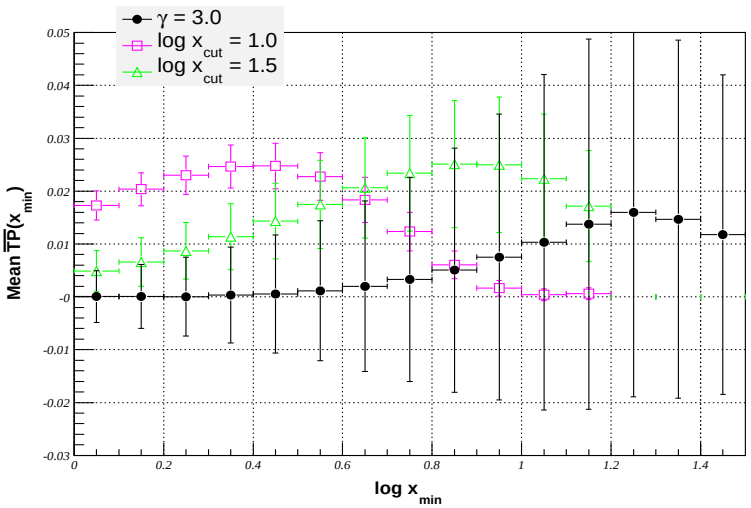


Figure 4.7: The fitted mean and 1σ deviation of the $\overline{TP}$'s (see definition in text) within each bin for the three distributions described in the text. This plot is the result of 10^4 simulated sets of events, where Fig.4.6 is one example, and where each set contains 3000 events.

fewer events.

The magenta squares represent the fitted mean $\overline{TP}$ as a function of x_{min} for sets drawn from a power-law with a cut-off at $\log x_{cut} = 1.0$. They deviate from zero for all but the largest x_{min}. Furthermore, this offset is statistically significant for the lowest few bins of x_{min}, where the statistic reflects the deviation from power-law considering most of the events in the set. The green triangles show the fitted means for the $\log x_{cut} = 1.5$ distribution. They also display some deviation from zero, but they are not as significant since they fall near the 1σ errors for the pure power-law distribution.

Indeed, one may inquire as to which of the bins deviate the most from the simulated power-law. This is equivalent to asking, "above what minimum do the data generated from this cut-off distribution maximally deviate from a pure power-law?"

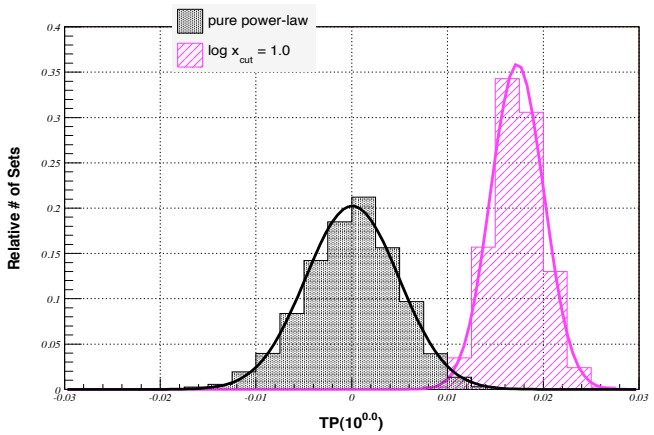


Figure 4.8: The distribution of the $\overline{TP}$'s in the first bin of Fig.4.7 (the bin with minimum $\log x_{min} = 0.0$) for the simulated pure power-law (black, shaded) and a power-law with a cut-off at $\log x_{cut} = 1.0$ (magenta, hatched). For these distributions $P_{TP} = 0.00218$ (see equation (4.12)).

To quantify this deviation, we use a P-value given by

$$P_{TP} = 1 - \frac{1}{\sqrt{2\pi}} \int_{-\beta}^{\beta} e^{-t^2/2} dt, \tag{4.12}$$

where

$$\beta = \frac{|\mu_1 - \mu_2|}{\sqrt{\sigma_1^2 + \sigma_2^2}}, \tag{4.13}$$

μ_i is the mean of the fitted gaussian and σ_i is the standard deviation. We reject the pure power-law hypothesis (at the 5% S. L.) if $P_{TP} \leq 0.05$. The mean of the gaussian fit to the distribution of $\overline{TP}$'s for the power-law in the bin with minimum $\log x_{min} = 0.0$ is $\mu_1 = (0.0056 \pm 5.1) \times 10^{-3}$ with a standard deviation $\sigma_1 = 4.9 \times 10^{-3}$. The mean of the fitted gaussian for the $\log x_{cut} = 1.0$ distribution in this bin is $\mu_2 = (1.7 \pm 0.28) \times 10^{-2}$ with a standard deviation $\sigma_2 = 2.8 \times 10^{-3}$. Therefore, the significance level of the deviation is $P_{TP} = 2.18 \times 10^{-3}$ and we can reject the pure power-law hypothesis for this distribution. The distribution of the $\overline{TP}$'s for this bin is plotted in Fig. 4.8 for the pure power-law (black, shaded) and the $\log x_{cut} = 1.0$

(magenta, hatched) pdf. The maximum deviation for the $\log x_{cut} = 1.5$ pdf occurs in the bin with minimum $\log x_{min} = 0.4$ and the corresponding distributions of $\overline{TP}$ are plotted in Fig. 4.9. The significance of this deviation is lower; $P_{TP} = 0.298$.

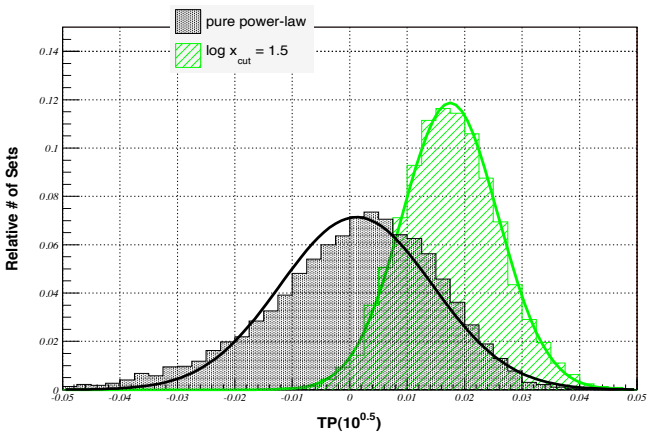


Figure 4.9: The distribution of the $\overline{TP}$'s in the fifth bin of Fig.4.7 (bin with minimum $\log x_{min} = 0.4$) for the simulated pure power-law (black, shaded) and a power-law with a cut-off at $\log x_{cut} = 1.5$ (green, hatched). For these distributions $P_{TP} = 0.298$ (see equation (4.12)).

4.5.2 The Cosmic Ray Data

In order to apply the TP-statistic to the CR data, Monte-Carlo simulations were conducted and analyzed in a manner similar to that discussed in §4.5.1; we generate 10^4 sets of events from the reported flux and the resulting distribution of $\overline{TP}$ (within each bin) is fitted to a gaussian. Since the significance of deviation from zero depends on both the power index and the number of events, we will compare each of the Auger and AGASA data sets with a unique power-law. We will take the AGASA experiment to have 1916 events above $\log E_{min} = 18.8$ and we will compare the resulting TP-statistics with those of a power-law with the same minimum and $\gamma = 2.80$. The Auger spectrum has a power-index estimate of 2.97 considering all of the

data above $\log E_{min} = 18.5$ and a total of 3570 events, so we will therefore compare the TP-statistics arising from the Auger flux to those of a pure power-law with these parameters.

The application of this scheme to the AGASA spectrum is plotted in Fig. 4.10 in red triangles. The black circles represent average TP-statistic value for data drawn from a pure power-law with $\hat{\gamma}_{AGASA}$. Both plots have $N = 886$ events per sky. The error bars on each point represent the 1-sigma deviation of the gaussian fit to the distribution of the mean TP-statistic. Since the AGASA values do not significantly deviate from zero (or the power-law values) this plot suggests that the AGASA distribution does not significantly deviate from a pure power-law. The most significant deviation occurs in the bin with minimum $10^{19.2}$(eV) and gives $P_{TP} = 0.161$, which is consistent with the P-value for this bin discussed in §4.4. These distributions are plotted in Fig 4.11.

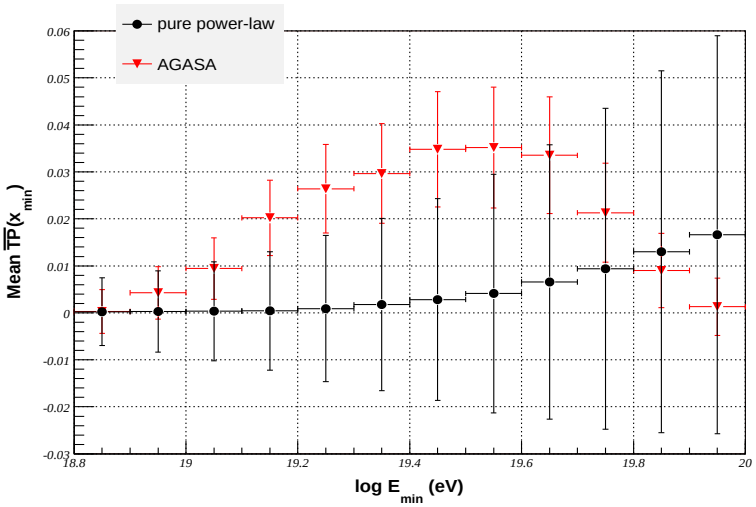


Figure 4.10: The fitted mean and 1σ deviation of the $\overline{TP}$'s (see definition on the text) within each bin for the AGASA spectrum (red triangles) and a pure power-law distribution (black circles). This plot is the result of 10^4 simulated sets of events where each set contains 1916 events and the power-law has index $\gamma = \hat{\gamma}_{AGASA}$.

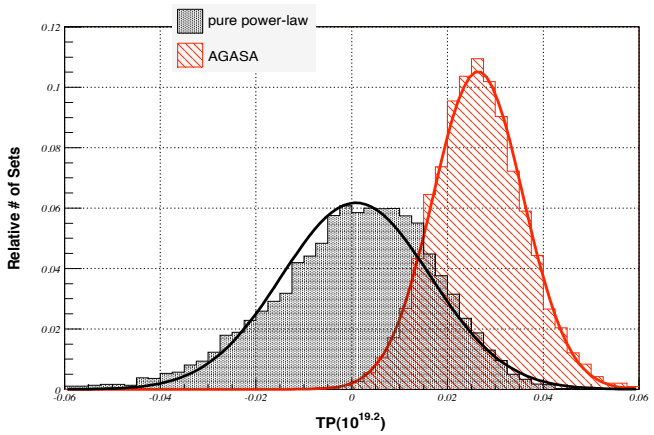


Figure 4.11: The distribution of $\overline{TP}$'s in the fifth bin of Fig.4.10 (the bin with minimum $E_{min} = 10^{19.2}$(eV)) for the pure power-law (black, shaded) and the AGASA spectrum (red, hatched). For these distribution $P_{TP} = 0.161$ (see equation (4.12)).

The simulation results from the Auger spectrum are plotted in Fig. 4.12. This plot shows deviation from a power-law for the lowest minimums considered. For the bin with minimum $\log E_{min} = 18.6$ we find $P_{TP} = 1.54 \times 10^{-4}$. Thus we can say that the Auger spectrum with energies greater than $10^{18.6}$(eV) deviate from a power-law by $\sim 3.78\sigma$, where $\sigma^2 = \sigma_1^2 + \sigma_2^2$. The distribution of $\overline{TP}$'s for this minimum energy is plotted in Fig. 4.13.

Since the TP-statistic nearly eliminates the need to estimate γ, the biggest systematic uncertainty in analyzing the CR data with the TP-statistic is likely to be errors in the event energies. Similar to the P-value discussed in §4.4, it is only the relative energy errors which can effect the result, since the TP-statistic depends only on the ratio. However, any elongation of the observed spectrum brought about by this relative uncertainty effect the TP-statistic. Without further study of the CR energy systematics, we cannot draw a conclusion from the $\sim 3.78\sigma$ deviation in Fig. 4.13.

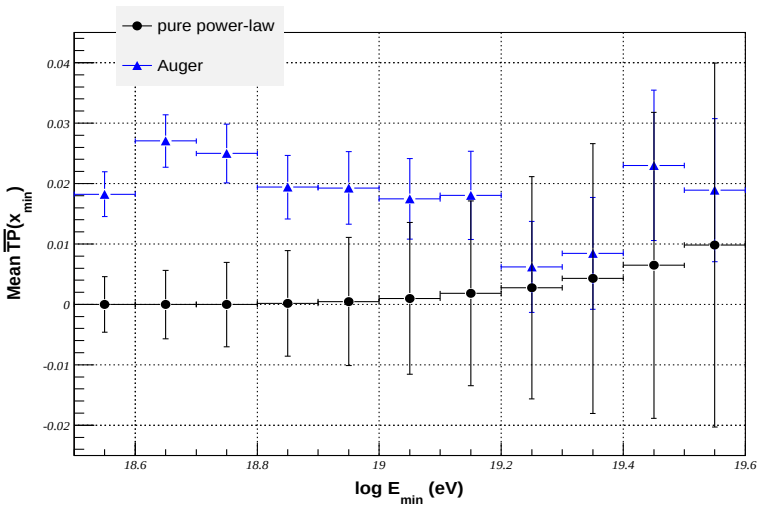


Figure 4.12: The fitted mean and 1σ deviation of the $\overline{TP}$'s (see definition on the text) within each bin for the Auger spectrum (blue triangles) and a pure power-law distribution (black circles). This plot is the result of 10^4 simulated sets of events where each set contains 3570 events and the power-law has index $\gamma = \hat{\gamma}_{Auger}$.

4.6 Summary

In §A.3 we use the reported (AGASA and Auger) CR fluxes to discuss the power-law form and illustrate the logarithmically binned estimates of the power index γ. The probability P that the maximum value of a sample drawn from a power-law is less than or equal to a particular value is defined in equation (4.6). Using reasonable estimates for γ, N_{tot} and z_{max}^{obs} from the CR data sets we calculate P in §4.4. The value of P is used to test the null hypothesis that these data sets follow a power-law . The AGASA data give no reason to reject the hypothesis; $P_{AGASA} \sim 8.4\%$ for the data with $\log E(eV) \geq 18.8$. The Auger data give more reason to reject the null hypothesis, $P_{Auger} \sim 1.9\%$ for the data with $\log E(eV) \geq 18.5$. However, consideration of the errors on $\hat{\gamma}$ prevent any solid conclusion.

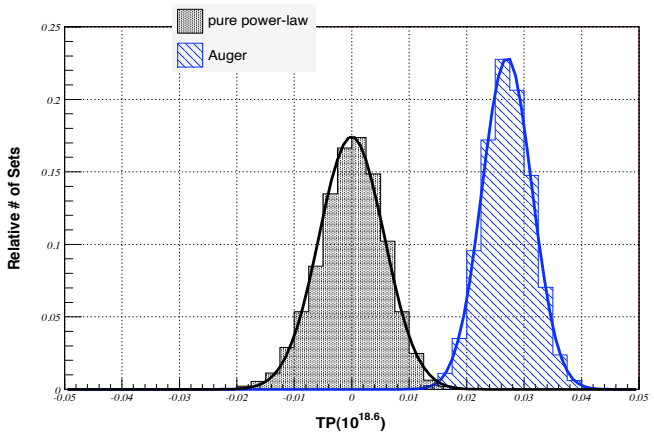


Figure 4.13: The distribution of $\overline{TP}$'s in the second bin of Fig.4.10 (the bin with minimum $E_{min} = 10^{18.6}$(eV)) for the pure power-law (black, shaded) and the Auger spectrum (blue, hatched). For these distribution $P_{TP} = 1.54 \times 10^{-4}$ (see equation (4.12)).

For the purpose of statistical analysis it would be useful to eliminate, or at least minimize, the importance of γ. The TP-statistic tends (asymptotically) to zero regardless of the value of γ and is the subject of §5.5.2. We apply the TP-statistic to the CR data sets using a Monte-Carlo method described in §4.5.2. The AGASA data give a value of $P_{TP} = 0.161$ for energies greater than $10^{19.2}$(eV). a value consistent with the P-value discussed in §4.4 (Fig.4.4). The preliminary Auger flux results in a TP-statistic with more significant deviation from the power-law form: $P_{TP} = 1.54 \times 10^{-4}$ for $E_{min} = 10^{18.6}$(eV). Comparing this value with the P-value for this bin derived in §4.4, namely $P \sim 2 \times 10^{-2}$, illustrates the power of the method based on the TP-statistic which is essentially independent of gamma. Better understanding of the relative errors on the CR energies should lead to a definitive conclusion on the question of a cut-off in the CR spectrum.

4.7 Appendix

In §4.5.1 we state that the TP-statistic will be distinctly positive for distributions which contain a tail-suppression and negative for distributions which contain a tail-enhancement (relative to the pure power-law form). In this section we numerically compute the TP-statistic for a "double power-law" distribution and describe the parameter space associated with this statistic.

Consider the following probability distribution function:

$$f(x) = \begin{cases} A(x_{min}, x_{bend}, \gamma, \delta)x^{-\gamma} & x_{min} \leq x < x_{bend} \\ B(x_{min}, x_{bend}, \gamma, \delta)x^{-\delta} & x_{bend} \leq x < \infty, \end{cases} \tag{4.14}$$

where $A(x_{min}, x_{bend}, \gamma, \delta)$ and $B(x_{min}, x_{bend}, \gamma, \delta)$ are chosen such that

$$\lim_{x \to x_{bend}^{+}} f(x) = \lim_{x \to x_{bend}^{-}} f(x)$$

and

$$\int_{x_{min}}^{\infty} f(x)dx = 1.$$

This distribution follows a power-law with index γ for $x_{min} \leq x < x_{bend}$, and δ for $x \geq x_{bend}$.

Given the parameter set $\{x_{min}, x_{bend}, \gamma, \delta\}$, we define the TP-statistic for this distribution as

$$TP(u) = \left[\int_u^{\infty} \ln\left(\frac{x}{u}\right) f(x)dx\right]^2 - \frac{1}{2}\int_u^{\infty} \ln^2\left(\frac{x}{u}\right) f(x)dx. \tag{4.15}$$

For $u \geq x_{bend}$ and/or $\gamma = \delta$ equation (4.15) is identically zero since it is equal to $\nu_1^2 - 1/2\nu_2$ (see equation (4.8)). However, equation (4.15) is non-trivial when $x_{min} \leq$

$u < x_{bend}$ and $\gamma \neq \delta$. In what follows, we calculate $TP(u)$ for $x_{min} \leq u < x_{bend}$ and various values of x_{bend} and δ with $x_{min} = 1$ and $\gamma = 3$ fixed.

Fig.4.14 contains a plot of $\log f(x)$ versus $\log x$ with $\delta = \gamma \pm 1$ for several choices of $\log x_{bend}$ (namely, for $\log x_{bend}$ varying from 1 to 2 in steps of 0.2). The red curves correspond to $\gamma < \delta = 4$ (tail-suppression) and the blue curves have $\gamma > \delta = 2$ (tail-enhancement). The TP-statistic for each of these distributions is shown in Fig.4.15 as a function of u. Examination of Fig.4.15 suggests the following conclusions for a given γ and δ:

- $TP(u)$ is positive for all values of u and x_{bend} if and only if $\gamma < \delta$, and it is negative if and only if $\gamma > \delta$.
- For x_{bend} much greater than x_{min}, $TP(u = x_{min})$ is approximately zero. Specifically, as $x_{bend}/x_{min} \to \infty$, $TP(x_{min}) \to 0$.
- The location of the maximum deviation, say u_0 where

$$\left(\frac{\partial}{\partial u} TP(u)\right)_{u=u_0} = 0, \tag{4.16}$$

  is highly correlated with the location of the bend x_{bend}. Indeed, we have found that there is a linear relationship between $\log u_0$ and $\log x_{bend}$ and that this relationship is independent of whether γ is less than or greater than δ.
- The maximum deviation of the TP-statistic, i.e. $TP(u_0)$, is independent of $\log x_{bend}$.

To isolate the effects of power index choice, consider the family of distributions where $\log x_{bend} = 1.0$ is fixed but δ is allowed to vary. Since the integrals in equation (4.15) only converge if $\delta \geq 2$, the minimum δ we can choose is $\delta = 2$. There is no upper bound on δ so we vary this parameter over the interval $2 \leq \delta < 3$ in steps of 0.2 and over the interval $3 < \delta < 10$ in steps of 0.5. Fig.4.16 contains a plot of

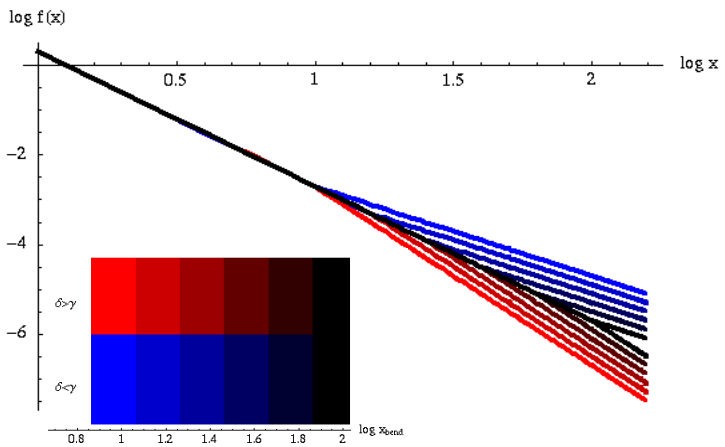


Figure 4.14: A plot of $\log f(x)$ (see equation (4.14)) versus $\log x$ with $\delta = \gamma \pm 1$ for several choices of $\log x_{bend}$ (namely, for $\log x_{bend}$ varying from 1 to 2 in steps of 0.2). The red curves correspond to $\gamma < \delta = 4$ (tail-suppression) and the blue curves have $\gamma > \delta = 2$ (tail-enhancement). The more black the color of the curve, the larger $\log x_{bend}$.

$\log f(x)$ versus $\log x$ with $\log x_{bend} = 1.0$ and $\gamma = 3$. The blue curves have $2 \leq \delta < 3$ (i.e. $\delta - \gamma < 0$) and the red curves have $3 < \delta < 10$ (i.e. $\delta - \gamma > 0$). The more black the color of these curves, the closer δ is to γ.

Fig.4.17 contains a plot of $TP(u)$ for the distributions plotted in Fig.4.16. As noted earlier, $TP(u) > 0$ if and only if $\delta - \gamma > 0$ and $TP(u) < 0$ if and only if $\delta - \gamma < 0$. The colored points on these curves show where each curve maximally deviates from zero; the coordinates of these points are $\{u_0, TP(u_0)\}$ for each curve (see equation (4.16)). These points show a weak dependence of $\log u_0$ on δ, for a given $\log x_{bend}$.

The value of the maximum deviation $TP(u_0)$ also shows dependence on δ. In Fig.4.18 we plot $TP(u_0)$ versus $\delta - \gamma$ for each of the points in Fig.4.17. These plots suggest the following:

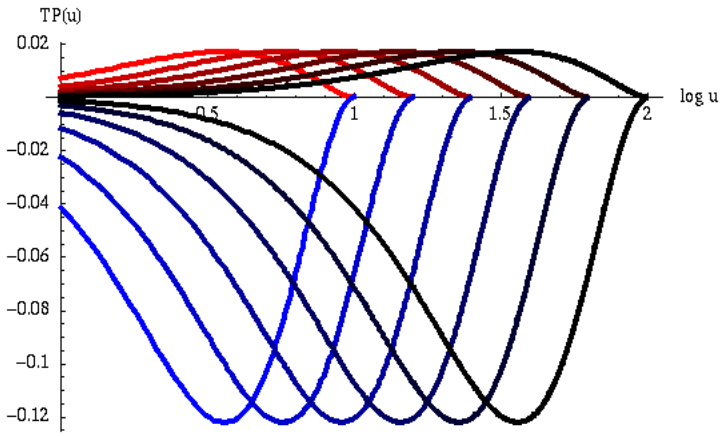


Figure 4.15: A plot of $TP(u)$ (see equation (4.15)) for each of the distributions plotted in Fig.4.14. Those distributions with tail-suppression (red) have $TP(u) > 0$ and those with tail-enhancement (blue) have $TP(u) < 0$.

- For $-1 \leq \delta - \gamma \lesssim 1$ (blue and black), a small change in δ will lead to a large change in $TP(u_0)$.
- If $\delta - \gamma \gg 1$ (bright red), however, a large change in δ will result in a small change in $TP(u_0)$. This case is of particular interest since a large δ will mimic the cutoff distribution defined in equation (4.11).
- By inspection of Fig.4.18 we note that $TP(u_0) \approx 0.025$ for $\delta - \gamma \gg 1$.
- Comparison with Fig.4.15 suggests that the limiting value of $TP(u_0)$ is roughly independent of x_{bend}.

The studies described in this section show that the TP-statistic can distinguish tail-suppressed ($\delta - \gamma > 0$) from tail-enhanced ($\delta - \gamma < 0$) distributions, i.e. $TP(u) > 0$ if and only if $\delta - \gamma > 0$ and $TP(u) < 0$ if and only if $\delta - \gamma < 0$. Furthermore, they show that in the limiting case of $\delta - \gamma \gg 1$ the most important parameter in determining u_0 is x_{bend} but that the limiting value of $TP(u_0)$ is roughly independent

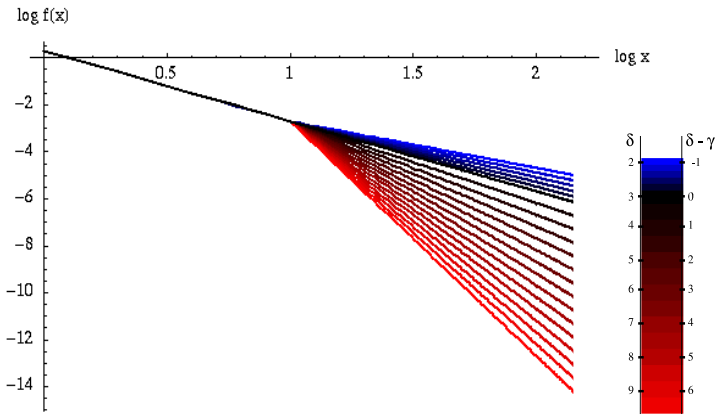


Figure 4.16: A plot of $\log f(x)$ (see equation (4.14)) versus $\log x$ with $\log x_{bend} = 1.0$ and $\gamma = 3$. The blue curves have $2 \leq \delta < 3$ (i.e. $\delta - \gamma < 0$) and the red curves have $3 < \delta < 10$ (i.e. $\delta - \gamma > 0$). The more black the color of these curves, the closer δ is to γ.

of x_{bend} and $\delta - \gamma$.

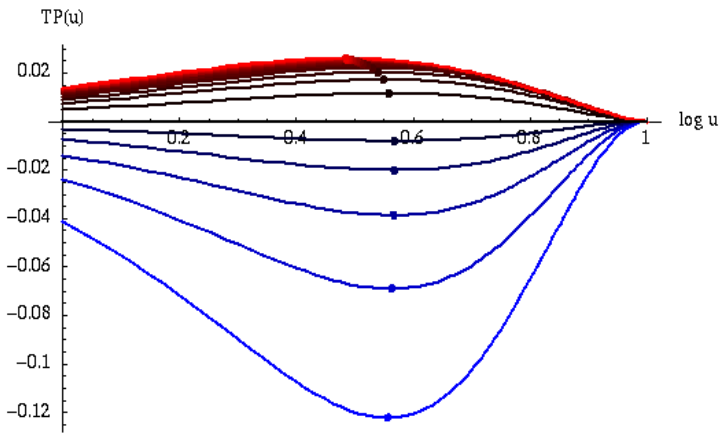


Figure 4.17: A plot of $TP(u)$ (see equation (4.15)) for the distributions plotted in Fig.4.16. The colored points on these curves show where each curve maximally deviates from zero; the coordinates of these points are $\{u_0, TP(u_0)\}$ for each curve (see equation (4.16)).

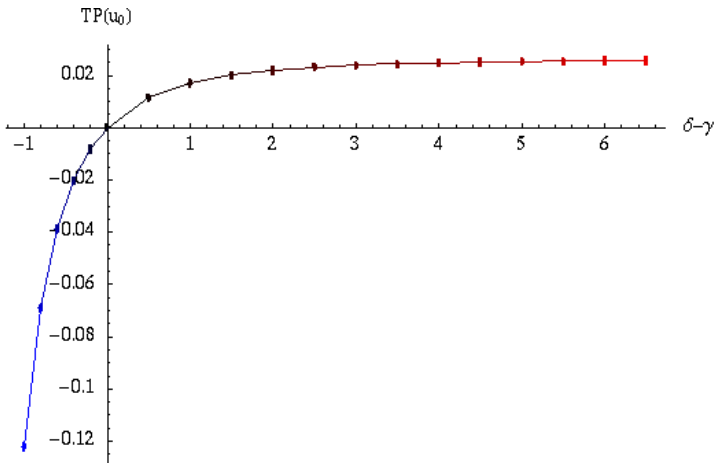


Figure 4.18: A plot of $TP(u_0)$ (see equations (4.14) and (4.16)) versus $\delta - \gamma$ for each of the points in Fig.4.17. Note that $TP(u_0) \approx 0.025$ for $\delta - \gamma \gg 1$.

Chapter 5

Flux Suppression : Unbinned Analysis

5.1 Chapter Overview

The text in this chapter was submitted under the title "Statistical Tools for Analyzing the Cosmic Ray Energy Spectrum" by the authors J. D. Hague, B. R. Becker, M. S. Gold and J. A. J. Matthews in in the journal *Astroparticle Physics*. This paper introduces the un-binned statistics used to study the energy spectrum and applies them to a *simulated* Auger spectrum. It is published as *GAP-Note 2007-070* and as arXiv:0808.2685.

In this paper un-binned statistical tools for analyzing the cosmic ray energy spectrum are developed and illustrated with a simulated data set. The methods are designed to extract accurate and precise model parameter estimators in the presence of statistical and systematic energy errors. Two robust methods are used to test for the presence of flux suppression at the highest energies: the Tail-Power statistic and a likelihood ratio test. Both tests give evidence of flux suppression in the simulated

data. The tools presented can be generalized for use on any astrophysical data set where the power-law assumption is relevant and can be used to aid observational design.

5.2 Introduction

The observation of suppression in the flux of the highest energy cosmic rays (CRs) has been of central interest to astro-particle physics since the prediction of the GZK-effect[13, 14] in 1966. Most recently both the Auger[47] and the HiRes[48] detectors have released results favoring the observation of flux suppression at a 6σ and 5σ level of confidence, respectively.

With this in mind, we describe a set of statistical tools designed to extract the most accurate and precise information concerning the flux of the highest energy cosmic rays. By binning the data we can only lose information[43] (see §5.7) and therefore our statistical tools use an un-binned maximum likelihood approach[49, 44, 37, 50] to answer two related statistical questions: *Is there flux suppression at the highest energies?* and, if yes, *What are the characteristic cut-off energy and shape parameters?*

In detail we first generate a toy data set using the CRPropa package[51], as in §5.3.2. We then fit this simulated data to the three models described in §6.4.1. The un-binned maximum likelihood fit is outlined in §5.4.1 and methods for incorporating systematic and statistical energy errors are described in §5.4.2 and §5.4.3 respectively. In §5.5 we describe several statistical tools for hypothesis testing: the Kolmogorov-Smirnov test, the tail power statistic[45, 7, 47], and a likelihood ratio test[52].

Though we cast our discussion in terms of cosmic ray energies, it is worth noting that these tools can be applied to any astrophysical data set where deviations from

the power-law hypothesis are relevant, e.g. the galaxy correlation function[53] or gamma ray astronomy[54].

5.3 CRPropa Data Set and Models

5.3.1 Input from the HiRes and Auger Observatories

Both the HiRes[48] and Auger[47] observatories have reported spectra and fit parameters for various power-law models. The collaborations use binned fitting methods. They fit the spectrum over many orders of magnitude in energy but we summarize here the model parameters[1] relevant only to the highest energies. The best fit double power-law parameters reported by HiRes[48] are $\gamma = 2.81 \pm 0.03(\text{stat}) \pm 0.02(\text{sys})$, $E_{\text{b}} = 10^{1.75 \pm 0.04}(\text{stat})$ and $\delta = 5.1 \pm 0.7(\text{stat})$. For the same model Auger[47] reports $\gamma = 2.62 \pm 0.03(\text{stat}) \pm 0.02(\text{sys})$, $E_{\text{b}} = 10^{1.6}(\text{fixed})$ and $\delta = 4.14 \pm 0.42(\text{stat})$. Fitting to the Fermi power-law Auger[47] finds $\gamma = 2.56 \pm 0.06(\text{stat})$, $E_{\frac{1}{2}} = 10^{1.74 \pm 0.06}(\text{stat})$ and $w_c = 0.16 \pm 0.04(\text{stat})$.

5.3.2 A Toy CR Data Set

To illustrate the methods in this note we use un-binned proton primary cosmic ray, CR, arrival energies (in EeV$\equiv 10^{18}$eV) as simulated by the package CRPropa[51] with input spectral index $\gamma_{\text{IN}} = 2.6$, $E_{\text{min}} = 10$ EeV and $E_{\text{max}} = 2000$ EeV. We draw 5×10^3 events to act as a *toy* data set from a modern CR detector.

The CRPropa toy data set is similar size and shape to the flux reported by these observatories but the results of this study do not, otherwise, reflect any information about any physical data set. The probability distribution function (p.d.f.) of the

[1] See §6.4.1 and Table 6.1 for the definition of these parameters.

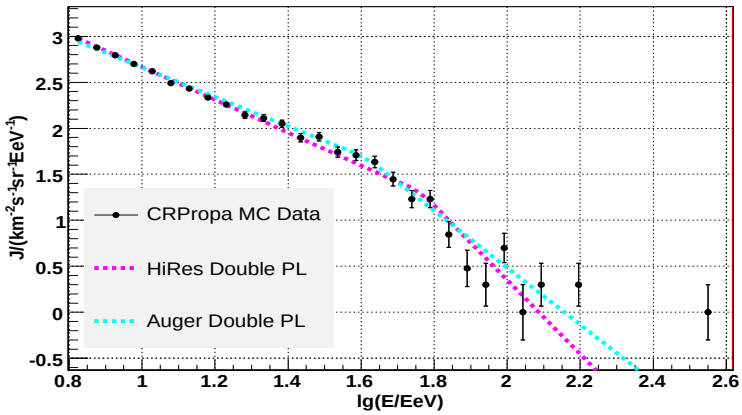


Figure 5.1: The differential flux as simulated by 5×10^3 events from the CRPropa toy set with parameters $\gamma_{\text{IN}} = 2.6$ and $E_{\text{max}} = 2000$ EeV (see §5.3.2). The p.d.f. of the best fit double power-laws reported by HiRes[48] and Auger[47] are the dashed lines.

best fit double power-laws reported by HiRes[48] and Auger[47] are shown in Figure 5.1 along with the CRPropa toy data.

The CRPropa propagation simulation is implemented by first generating proton CR primaries with initial energies according to a power-law "at the source," propagating them through a simulated Universe and then observing the final energy. The spacial extent of the sources is simulated as a uniform distribution of discrete sources on a grid with 10 Mpc steps extending to a distance of 4.07 Gpc, (from redshift $z = 0.0$ to $z = 2.73$). Nuclei traveling over many megaparsecs from these sources will suffer significant energy loss in an expanding Universe filled with the cosmic microwave background, CMB, radiation. As a result, the highest energy flux is much less than one would expect from a power-law alone. This suppression is known as the GZK-effect[13, 14].

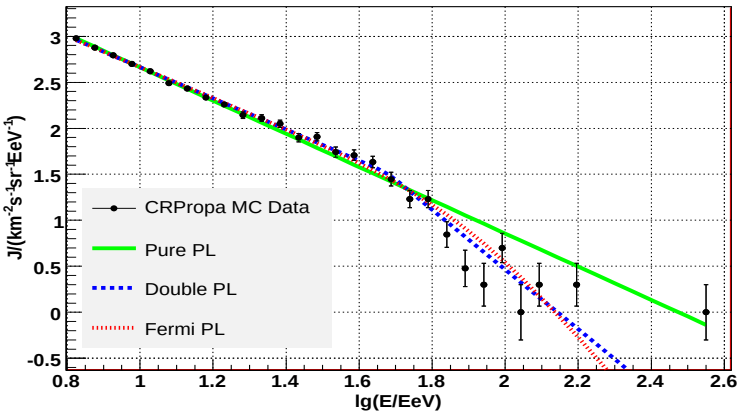


Figure 5.2: The differential flux as simulated by 5×10^3 events from the CRPropa toy set with parameters $\gamma_{\mathrm{IN}} = 2.6$ and $E_{\mathrm{max}} = 2000$ EeV. The best fit models are described in §6.4.1.

5.3.3 Power-Law Models

The fundamental probability distribution function governing the pure power-law assumption, denoted f_{P}, is shown in Table 6.1: $f_{\mathrm{P}} = (\gamma - 1){E_{\mathrm{min}}}^{\gamma-1}E^{-\gamma}$. The parameter γ is referred to as the *spectral index*. Here the sub-scripted-P stands for Pure-power-law.

For the highest energy CRs, the interesting observation would be to confirm or deny deviation from the power-law form at the highest magnitudes, i.e. the GZK-cutoff. We therefore study two toy models that mimic a pure power-law for lower energies but exhibit flux suppression above a given energy. The first is a double power-law (DP) with two spectral indexes, γ below E_{b} ("b" for bend or break) and $\delta > \gamma$ above. The point at which this p.d.f. reaches half the value it would have if the pure power-law continued above E_{b} is given by ${E_{\frac{1}{2}}}^{\mathrm{dp}} = 2^{\frac{1}{\delta-\gamma}}E_{\mathrm{b}}$, see [55] for a discussion of this quantity. Both HiRes[48] and Auger[47] have analyzed their data using this model.

We also study a toy p.d.f. where the cut-off is a "Fermi-like" Power-law (FP)[47, 7]. The advantage of fitting with this toy model is that the location parameter $E_{\frac{1}{2}}$ is a parameter in the fit.

All three p.d.f.'s are normalized on the interval $[E_{\rm min}, \infty)$, i.e. $\langle\rangle_{\rm M} \equiv \int_{E_{\rm min}}^{\infty} f_M(t)dt = 1$ for each of the models $\mathrm{M} \in \{\mathrm{P}, \mathrm{DP}, \mathrm{FP}\}$. The first element of the parameter vector $\theta_1 \equiv E_{\rm min}$ is fixed for the fit (see §5.4) and then varied to estimate the stability (see §5.5.1). Thus the power-law has one free parameter and the other models have three; low energy spectral index, location of cut-off and "steepness" of cut-off.

Model	$N_{\rm dof}$	Normalization	Function
P	1	$(\gamma-1){E_{\rm min}}^{\gamma-1}$	$E^{-\gamma}$
DP	3	$\frac{\gamma-1}{E_{\rm b}}\left\{\left(\frac{E_{\rm b}}{E_{\rm min}}\right)^{\gamma-1}+\frac{\gamma-1}{\delta-1}-1\right\}^{-1}$	$\left(\frac{E}{E_{\rm b}}\right)^{-\gamma}$ $E_{\rm min} \le E < E_{\rm b}$ $\left(\frac{E}{E_{\rm b}}\right)^{-\delta}$ $E_{\rm b} \le E$
FP	3	$\langle\rangle_{\rm FP}^{-1}$, numerically	$E^{-\gamma}\left[1+\left(\frac{E}{E_{\frac{1}{2}}}\right)^{1/(w_c \ln 10)}\right]^{-1}$

Table 5.1: The model designation (Model = Pure power-law, Double Power-law or Fermi Power-law), number of free parameters, normalization, and form of the function used to fit the simulated fluxes used in this study.

5.4 Fitting the Data

We take an un-binned maximum log-likelihood approach to estimating the best-fit parameters of each model. The method constructed here is designed to extract the maximum possible statistical information about these parameters. For the ideal detector we assume that the observed energies are known with infinite precision.

5.4.1 Ideal Detector

We find estimates of the parameters in each model by maximizing,

$$\mathcal{L}_{\mathrm{M}}(\vec{\theta}) = \sum_{i=1}^{N} \ln \left\{ f_{\mathrm{M}}(E_i; \vec{\theta}) \right\}, \tag{5.1}$$

where the sum is carried out over the event energies and $\theta_1 \equiv E_{\min}$ is fixed. The global maximum of this function $\mathcal{L}_{\mathrm{M}}(\hat{\theta})$ determines the best parameter estimates, $\hat{\theta}$. The the function is maximized using Minuit[56] with the MIGrad option.

To determine the one degree of freedom error estimate[49] for a parameter we vary the parameter (with the others fixed at $\hat{\theta}$) until $-2\Delta\mathcal{L}_{\mathrm{M}} = 1$. The two degrees of freedom error estimates[49] are determined by varying two parameters with the other fixed and choosing the contour such that $-2\Delta\mathcal{L}_{\mathrm{M}} \geq 2.3$. For the toy data set, we plot these contours and the asymmetric one degree of freedom error estimates in §5.9: Figure 5.11 and 5.12.

5.4.2 Systematic Energy Error

The errors on the observed energy E_{obs} of an event from a real CR detector are considerable and must be included in any realistic analysis of a spectrum. For our purposes, these errors take the two canonical forms; *statistical* and *systematic*, i.e. $E_{obs} \pm \sigma_{stat} \pm \sigma_{sys}$.

The systematic errors energy errors of a CR detector reflect the uncertainties in the absolute calibration of the detector. At the highest energies the systematics are the dominant contribution to the overall uncertainty of an event's energy. For example, the two fluorescence detectors Auger[47] and Hires[48] report uncertainties of 22% and 17% respectively.[2] The shift in energy due to the systematic error can

[2]With its hybrid detector the Auger reduces the systematic error to between 7% and 15%[47].

be asymmetric, i.e. $\sigma^{+}_{sys} \neq \sigma^{-}_{sys}$, and energy dependent, see Eq(5.2), but it effects every event at a given energy the same way; a shift up or down. For the Monte-Carlo (MC) data sets we model the systematic detector energy errors using:

$$\frac{\sigma(E;\vec{p})}{E} = p_1 + p_2 \lg(E). \tag{5.2}$$

Here we choose symmetric systematically-shifted energies such that the energy of the k^{th} event is $E^{\pm}_k = E_k \pm \sigma(E_k;\vec{p}_{sys})$. For the systematic errors we choose $p_1 = 0.05$ and $p_2 = 0.10$.

To account for this in the parameter estimation procedure, we shift each energy up or down and carry out the methods in §5.4.1. The difference between the parameter estimates of a shifted set and those of the centered set gives "systematic" errors of the parameter estimates.

5.4.3 Statistical Energy Error

To model the statistical energy errors of the detector we assume that the *true* energy of the cosmic ray has a 68% chance of being within the interval $(E_{obs} - \sigma_{stat}, E_{obs} + \sigma_{stat})$. The observed energy has been "smeared" from the true value; $E_{obs} = E_{true} + Y$ where Y is drawn from a normal distribution with mean 0 and variance σ_{stat}. Note that while the true energies can only be found on $[E_{\min}, \infty)$, there is a nonzero probability for the (after smearing) observed energy to be less than $E_{\min}$; E_{obs} lives on the interval $(-\infty, \infty)$. This edge effect near $E_{\min}$ can be accounted for by assuming that the true distribution of energies follows a power-law well below $E_{\min}$ and then re-normalizing the convolution technique used in Howell[44]. See §5.8 for further discussion. For the integrand, three factors are necessary:

1. The model to be fitted, $f_{\mathrm{M}}(t;\vec{\theta})$ (see §6.4.1). By letting $\theta_0 = 0.1E_{\min}$ we are assuming that the power-law extends below the observed $E_{\min}$.

2. A normal distribution $G(t; E_{\rm obs}, \sigma_{stat}(t;\vec{p}))$ with mean $E_{\rm obs}$ and variance $\sigma_{stat}(t;\vec{p})$ to reflect the statistical energy errors.

3. The acceptance of the CR detector as a function of the true energies $\Omega(t)$. Since we are using MC data we choose $\Omega(t) = 1$ for simplicity.

The convolution is calculated by integrating over all possible *true* energies (t):

$$g_{\rm M}(E_{\rm obs};\vec{\theta},\vec{p}) = \int_{0.1E_{\rm min}}^{\infty} f_{\rm M}(t;\vec{\theta})\, G(t; E_{\rm obs}, \sigma_{stat}(t;\vec{p}))\, \Omega(t)\, dt. \tag{5.3}$$

Re-normalizing so that the observed energies define a p.d.f., we numerically calculate the p.d.f. to be:

$$\tilde{f}_{\rm M}(E_{\rm obs};\vec{\theta},\vec{p}) = \frac{g_{\rm M}(E_{\rm obs};\vec{\theta},\vec{p})}{\int_{E_{\rm min}}^{\infty} g_{\rm M}(y;\vec{\theta},\vec{p})\, dy}, \tag{5.4}$$

and we must modify the likelihood found in Eq(5.1) accordingly:

$$\tilde{\mathcal{L}}_{\rm M}(\vec{\theta}) = \sum_{i=1}^{N} \ln\left\{\tilde{f}_{\rm M}(E_i;\vec{\theta})\right\}. \tag{5.5}$$

By finding the parameters $\hat{\theta}$ which maximize Eq(5.5) we can be confident that we are accounting for the statistical uncertainty inherent in data collected by a realistic detector. To model statistical errors in our toy data set, we parameterize σ_{stat} as in Eq(5.2) with $p_1 = 0.15$ and $p_2 = 0$.

5.5 Evaluating the Fit

In this section we outline ways to evaluate the fit of a candidate model to the data set. The Kolmogorov-Smirnov statistic can be used to extract a best fit minimum energy $\hat{E_{\rm min}}$ and, with its corresponding p-value, evaluate the "absolute goodness of fit" of a candidate model (see §5.5.1). The relevant question for CR physics is not

whether a particular model is a good fit to the data but rather whether the flux exhibits suppression (relative to the single power-law form) at the highest energies. To address this question directly we use two statistics with well defined p-values: the Tail-Power statistic (see §5.5.2), which can give information about tail suppression in standard deviations, and a likelihood ratio that allows rejection of the single power-law hypothesis in favor of a suppressed candidate model (see §5.5.3).

5.5.1 Kolmogorov Statistic

While the minimum value of the likelihood function will indeed give the best value of the fit parameters, this fit may nonetheless be poor. The typical[43, 50] method for evaluating goodness of fit is the Kolmogorov-Smirnov test[49]. The relevant statistic for this test is the KS distance:

$$D_{\rm KS}(E_{\rm min}) = \max_{E \geq E_{\rm min}} |F_{\rm fit}(E) - F_{\rm data}(E)|\,, \tag{5.6}$$

where, $F_{\rm fit}$ and $F_{\rm data}$ are the cumulative distribution functions (c.d.f.) of the best fit model and the data respectively. The maximum distance between the c.d.f.'s is taken over all energies in the fitted data set, $E \geq E_{\rm min}$. By stepping over $E_{\rm min}$ and re-minimizing Eq(5.1) at each step to determine the best fit parameters, we can calculate $D_{\rm KS}$ as a function of $E_{\rm min}$. The value of $\hat{\theta}_0 \equiv \hat{E_{\rm min}}$ that minimizes $D_{\rm KS}$ can be taken as the best estimate of the minimum energy above which the model holds[50].

To test how well a particular model fits the data we must simulate many MC data sets drawn from the best fit model p.d.f. with the same number of events as the original data. The fraction of sets $p_{\rm KS}$ with $D_{\rm KS}$ greater than that of the data gives the suitable p-value; if $p_{\rm KS} \ll 1$ then it is unlikely that the data are drawn from the model under consideration, and in this way the KS test statistic $p_{\rm KS}$ can rule out the different candidate models[50].

5.5.2 Tail Power Statistic

The Tail-Power (TP) statistic is similar to the KS statistic discussed above, however it has, at least, three advantages over $p_{\rm KS}$ when testing the power-law assumption;

1. The TP statistic and it's corresponding p-value $p_{\rm TP}$ are nearly independent of the value of the spectral index γ,
2. The asymptotic behavior of the TP statistic is known, and therefore no simulations are required to calculate the corresponding p-value $p_{\rm TP}$,
3. If TP > 0 the deviation suggests flux *suppression in the tail* and if TP < 0 the deviation suggests flux *enhancement in the tail*[7] and
4. $p_{\rm TP}$ offers an unambiguous p-value in standard deviations.

This "measure of power-law-ness" has been developed and studied elsewhere (see [45, 47, 7]) and here we expand its use to the un-binned case.

The sample TP statistic is defined as [45]:

$$\hat{\tau}(E_{\rm min}) = \hat{\nu}_1^2(E_{\rm min}) - \frac{1}{2}\hat{\nu}_2(E_{\rm min}), \tag{5.7}$$

where:

$$\hat{\nu}_n(E_{\rm min}) = \frac{1}{N_>} \sum_{E_i > E_{\rm min}} \ln^n \frac{E_i}{E_{\rm min}} \tag{5.8}$$

and the sum is carried out over all $N_>$ events with energy greater than a given minimum. If the data are drawn from a pure power-law then $\hat{\tau}(E_{\rm min})$ will tend to zero as $N \to \infty$, regardless of the value of γ[43].

We may approximate the asymptotic joint distribution of $\hat{\nu}_1$ and $\hat{\nu}_2$ as a bivariate Gaussian $f_{\nu_1\nu_2}(\nu_1, \nu_2)$. The asymptotic mean and variance of ν_1 are $\frac{1}{\gamma-1}$ and $\frac{1}{N(\gamma-1)^2}$,

and of ν_2 are $\frac{2}{(\gamma-1)^2}$ and $\frac{20}{N(\gamma-1)^4}$. The random variables ν_1 and ν_2 are highly correlated; the correlation coefficient is $\rho = \frac{2}{\sqrt{5}}$, *independent of* γ. Thus, for a given N and γ, we calculate the p.d.f. of τ to be,

$$f_{TP}(\tau; N, \gamma) = \int_{-\infty}^{\infty} f_{\nu_1\nu_2}(t, 2(t^2 - \tau))dt. \tag{5.9}$$

The analytic "location" $\langle\tau\rangle_{TP} \sim 0$ and "shape" $\langle\sigma_\tau\rangle_{TP} = \sqrt{\langle\tau^2\rangle_{TP} - \langle\tau\rangle^2_{TP}} \sim N^{-1/2}(\gamma - 1)^{-2}$ parameters of this distribution are consistent with simulation generated values. We measure the p-value $p_{\rm TP}$ for the TP statistic in units of standardized deviation,

$$p_{\rm TP}(E_{\rm min}) = \frac{\hat{\tau}(E_{\rm min}) - \langle\tau\rangle_{TP}}{\langle\sigma_\tau\rangle_{TP}}. \tag{5.10}$$

A spectrum with flux suppression in the tail (like that in the Fermi-like model) will result in a *positive* significance[7].

The application of Eq(5.10) to the toy CR data set (see §5.3.2) is plotted in Figure 5.3. The top panel shows the (pure power-law) spectral index as a function of $E_{\rm min}$. A spectral index which increases as $E_{\rm min}$ increases is indicative of flux suppression. The red, left leaning hatching shows the variation of $\hat{\gamma}$ due to a $\pm 1\sigma$ *systematic* shift in the energies (see §5.4.2) while the opposite, blue hatching shows the *statistical* error of the estimator $\hat{\gamma}$, see §5.4.1. The bottom panel shows the resulting TP statistic significance $p_{\rm TP}(E_{\rm min})$ in standard deviations. Notice that while the systematic errors can be significant for the measured spectral index, they do not effect the TP statistic. Since we must estimate the spectral index to compute $p_{\rm TP}$, we also propagate the statistical errors on $\hat{\gamma}$ to the tail power statistic.

To test the effectiveness of this statistic, we apply it to a series of simulated data sets drawn from both the Fermi and double power-law models. For all the models we set[3] $E_{\rm min} = 1.0$EeV, $\gamma = 2.75$ and either $\delta = 4.75$ or $w_c = 0.10$. We vary each characteristic cut-off energy, either $E_{\rm b}$ or $E_{\frac{1}{2}}$, in three steps $\lg(E_{\rm cut}/E_{\rm min}) =$

[3]These values are similar to the Auger[47] and HiRes[48] best fit values.

0.5, 1.0, and 1.5. The total number of events in the data set is varied in four steps $\lg(N) \sim 2.5, 3.0, 3.5, 4.0$. For each of these twelve sets of parameter choices we make 10^3 Monte-Carlo realizations and plot the mean and RMS of $p_{\mathrm{TP}}(E_{\mathrm{min}} = 1.0)$ in Figure 5.4.

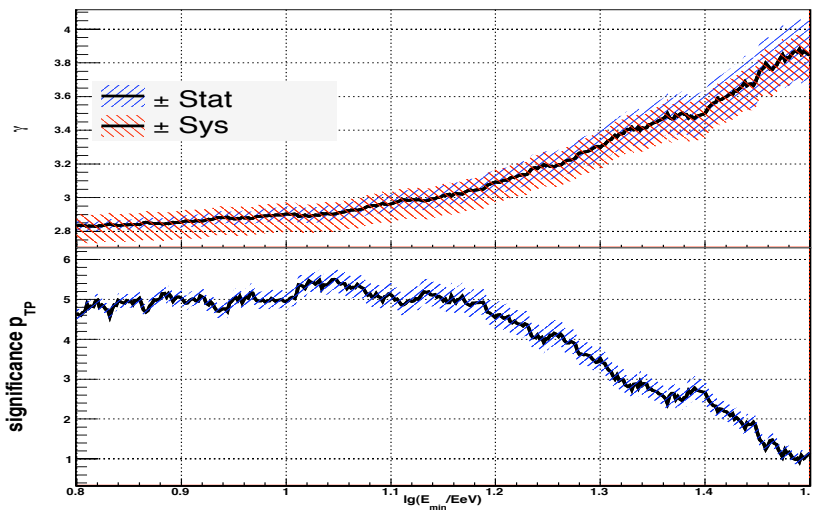


Figure 5.3: *Top* The best fit (see §5.4.1, Eq(5.1)) spectral index $\hat{\gamma}$ as a function of $\lg E_{\mathrm{min}}$ for the the toy CR data set (see §5.3.2) fit to the pure power-law model (P). *Bottom* The resulting TP statistic significance $p_{\mathrm{TP}}(E_{\mathrm{min}})$ in standard deviations as a function of the minimum energy E_{min}, see Eq(5.10). Both plots give strong evidence of flux suppression of the highest energy MC events.

Based on Figure 5.4 we can see that the best way to evaluate a data set with a potential for tail suppression is to collect as much data with E_{min} as close to the expected cut-off as possible. The experimenter may use Figure 5.4, or one like it, to help tune observation parameters, i.e. collecting time on a gamma ray source or size of a CR detector, in advance of the observation and in anticipation of flux suppression of a certain type. Note, however, that one should choose an E_{min} *prior* to analyzing a data set to avoid a penalty for scanning in this parameter.

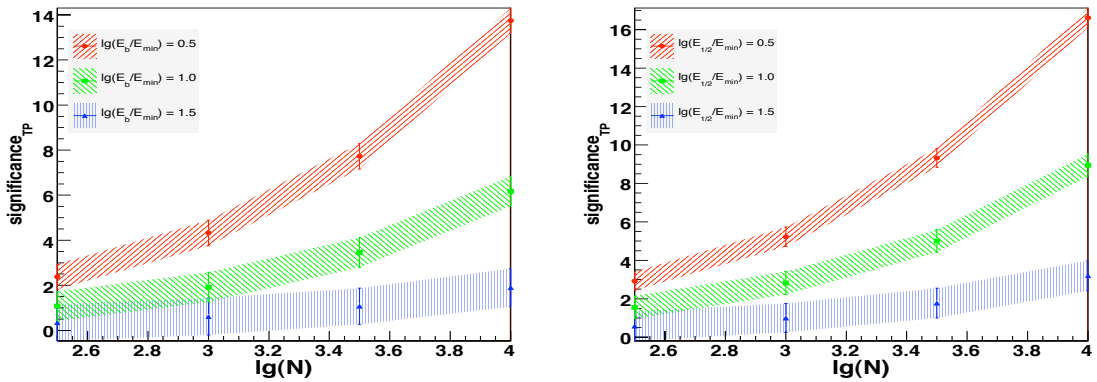


Figure 5.4: The tail power significance, $p_{\rm TP}(E_{\rm min} = 1.0)$ as a function of the ($\log_{10}$ of the) number of events in each Monte-Carlo realization. Each plot style represents a different choice of $\lg(E_{\rm cut}/E_{\rm min}) = 0.5$, 1.0, or 1.5. *Left*, the double power-law, $E_{\rm cut} \equiv E_{\rm b}$. *Right*, the Fermi power-law, $E_{\rm cut} \equiv E_{\frac{1}{2}}$.

5.5.3 Model Discrimination

Here we introduce a likelihood ratio test designed to discriminate candidate suppressed models (DP and FP) from the pure power-law. We define two log-likelihood ratios; for each model M:

$$\mathcal{R}_{\rm M} = \sum_{i=1}^{N} \{\ell_{\rm M}(E_i) - \ell_{\rm P}(E_i)\} = \mathcal{L}_{\rm M} - \mathcal{L}_{\rm P}, \tag{5.11}$$

where $\ell_{\rm M}(E_i) = \ln f_{\rm M}(E_i;\hat{\theta})$ with M either DP (double power-law) or FP (Fermi-like), and $\ell_{\rm P}(E_i) = \ln f_{\rm P}(E_i;\hat{\theta})$ for the pure power-law likelihood per event (see Table 6.1 and Eq(5.1)). Note that each suppressed model is fit *independently* of the pure power-law best fit. The asymptotic variance of $\mathcal{R}$ can be estimated by the sample value:

$$\sigma^2_{\mathcal{R}} = \frac{1}{N}\sum_{i=1}^{N}\left\{[\ell_{\rm M}(E_i) - \ell_{\rm P}(E_i)] - \left[\frac{\mathcal{L}_{\rm M} - \mathcal{L}_{\rm P}}{N}\right]\right\}^2, \tag{5.12}$$

The hypothesis of the pure power-law is *nested* within the hypothesis of a sup-

pressed power-law. As a consequence, $|\mathcal{R}|/\sigma_{\mathcal{R}} \to 0/0$ as $N \to \infty$ and the distribution of $\mathcal{R}/\sigma_{\mathcal{R}}$ is not Gaussian[50]. The correct p-value is calculated as the integral of a χ^2 function[57, 50]:

$$p_{\mathcal{R}}(z^2) = \frac{1}{\sqrt{2\pi}} \int_{z^2}^{\infty} t^{-1/2} e^{-t/2} dt, \tag{5.13}$$

where $z^2 = \mathcal{R}_{\mathrm{M}}^2 / \left(2N\sigma_{\mathcal{R}}^2\right)$.

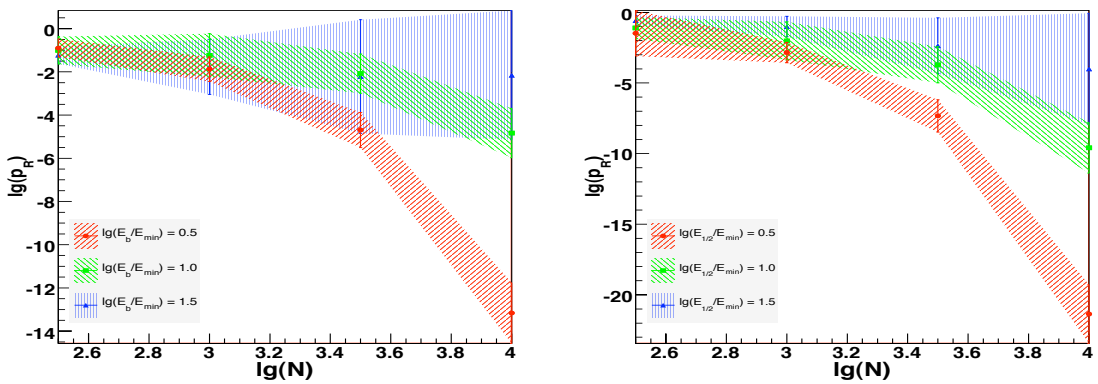


Figure 5.5: The log of the likelihood ratio significance, $p_{\mathcal{R}}$ as a function of the ($\log_{10}$ of the) number of events in each Monte-Carlo realization. Each plot style represents a different choice of $\lg(E_{\mathrm{cut}}/E_{\mathrm{min}}) = 0.5$, 1.0, or 1.5. *Left*, the double power-law, $E_{\mathrm{cut}} \equiv E_{\mathrm{b}}$. *Right*, the Fermi power-law, $E_{\mathrm{cut}} \equiv E_{\frac{1}{2}}$.

We interpret this p-value in the following way: if $p_{\mathcal{R}}$ is "small" then the best fit model M may be preferred over the best fit pure power-law. By small we mean that, *a priori* and rather arbitrarily, we may choose to reject the single power-law in favor of the model if $p_{\mathcal{R}} \leq 10^{-3}$. This quantity tells us only whether a given suppressed model is better than the pure power-law. It says nothing about how well any of the fits actually represent the data.

For each of the twelve sets of parameter choices used in Figure 5.4, we plot the mean and RMS of $p_{\mathcal{R}}$ in Figure 5.5. As before, we see that the best way to reject the power-law in favor of the suppressed model is to collect as much data with E_{min}

as close to the expected cut-off as possible. Note that for $\lg(E_{\rm cut}/E_{\rm min}) = 1.5$ the distribution of likelihood ratios is strictly positive and highly peaked near zero; the mean and RMS are not good reflections of this distribution.

5.6 Summary and Conclusion

In this paper we describe a set of statistical tools designed to extract the most accurate and precise information about the flux of the highest energy cosmic rays. We show how to use the un-binned likelihood method described in §5.4.1 to fit a data set to the three model distributions described in §6.4.1. Techniques for incorporating the systematic and statistical errors associated with a real CR detector into the likelihood method are described in §5.4.2 and §5.4.3 respectively. In §5.5 we describe p-values useful for extracting information about flux suppression. We show in §5.5.2 and §5.5.3 how an experimenter might use an *a priori* estimate of the cut-off energy to maximize an observational setup for detecting flux suppression.

The collection of these statistical tools are the primary result of this paper. To answer the questions posed in the introduction for a given data set we suggest the following steps:

1. Estimate the best fit parameters $\hat{\theta}$ of the model;
   (a) The estimates $\hat{\gamma}$, $\hat{E_{\rm b}}$ or $\hat{E_{\frac{1}{2}}}$ and $\hat{\delta}$ or $\hat{w_c}$ are determined via the likelihood Eq(5.1),
   (b) The estimate of the minimum energy $\hat{E_{\rm min}}$ is that which minimizes the Kolmogorov distance $D_{\rm KS}$ (see §5.5.1).
2. Shift the energies up and down according to the systematic uncertainty described in §5.4.2 and repeat step (1). The resulting shift in parameter estimates

gives the systematic uncertainty of those estimates.

3. Obtain the model parameter estimates using the methods in §5.4.3 to incorporate the statistical error of each event energy.

4. Test the model hypothesis;

   (a) The absolute goodness of fit for any of the models can be evaluated using $p_{\rm KS}$ in §5.5.1,

   (b) The Tail-Power statistic $p_{\rm TP}$ can be used to reject the single power-law hypothesis (nearly independently of the spectral index estimate, see §5.5.2)

   (c) The single power-law may be rejected in favor of a specific alternative model using $p_{\mathcal{R}}$, here we study the double and Fermi power-law distributions (see §5.5.3).

The best estimates for the *characteristic cut-off energy and shape parameters*, determined via steps (1), (2) and (3), are $\hat{E}_{\rm b}$ or $\hat{E}_{\frac{1}{2}}$ and $\hat{\delta}$ or $\hat{w}_c$ respectively. The presence of *flux suppression at the highest energies* can be evaluated using step (4).

By applying these methods to the toy Monte-Carlo set of CRPropa events we illustrate in §5.9 how the procedure may be implemented on an actual CR detector, i.e. a detector with systematic and statistical event energies. Suppression in the tail is clear in Figure 5.9 and Figure 5.10; the tail power statistic is 4.6σ and the p-value for the double (Fermi) power-law is $\lg p_{\rm DP} = -2.7$ ($\lg p_{\rm FP} = -1.9$).

The methods are sufficient and robust. Indeed, many of them have been applied by the Auger collaboration which reports suppression with 6σ confidence[47]. These tools serve as a basis for further investigation of the CR spectrum such as evidence for more detailed spectral information. They can be applied to any data set, astrophysical or otherwise, to provide information both about data already collected and help to optimize future observations for detecting tail suppression.

5.7 Binned vs. Un-Binned

The statistical superiority of an un-binned maximum likelihood estimate of the pure power-law spectral index to the logarithmically binned least-χ^2 method often used has been established in [43] and expanded upon more recently in [44, 37, 50, 7, 52]. In this section we compare the binned to the un-binned fitting method for the two suppressed models, i.e. the double and Fermi power-laws (see §6.4.1).

To calculate the binned estimators we minimize a $\chi^2(\vec{\theta})$ function that relates the logarithmically binned (width w) histogram of the data to that expected by a model. The function is[4],

$$\chi^2(\vec{\theta}) = \sum_{i=1}^{N_b} \left(\frac{\lg Y_i^{\rm data} - \lg Y_i^{\rm fit}(\vec{\theta})}{\sigma_i^{\rm data}} \right)^2 , \tag{5.14}$$

where N_b is the number of bins, $Y_i^{\rm data}$ is the number of events in the $i^{\rm th}$ bin b_i and σ_i is determined by Gaussian errors when $Y_i^{\rm data} > 10$ and Poissonian errors when $Y_i^{\rm data} \leq 10$. We minimize with respect to the parameters $\vec{\theta}$ (with $\theta_0 \equiv E_{\rm min}$ fixed) using the number of events in a bin expected by the model M,

$$Y_i^{\rm fit}(\vec{\theta}) = N \int_{10^{b_i - w/2}}^{10^{b_i + w/2}} f_{\rm M}(t;\vec{\theta}) dt.$$

To study the asymptotic bias and error produced by the two estimation techniques we draw 10^5 sets of 5×10^3 events from a pure power-law and separately from a double distribution. For each Monte-Carlo set we estimate the best fit model parameters $\hat{\theta}$ using both the likelihood Eq(5.1) and the χ^2 Eq(5.14) methods. The un-binned estimator of the pure power-law spectral index (see §5.4.1) has been shown[43, 44]

[4]For the case of the single power-law $\lg f_{\rm P} = \lg C - \gamma \lg E$ where C is the normalization. Thus the binned fitting method reduces to fitting the $\log_{10}$ of the (error weighted) bin heights to a straight line with slope γ. This technique is often used to mitigate the effects of the heaviness of the power-law tail but un-binned methods are more accurate and precise.

to have an error estimate within $\sim 1\%$ of the Cramer-Rao lower bound for a sample with as few as ~ 100 events.

In Figure 5.6 and Figure 5.7 we plot the results of the simulations. We can conclude that the un-binned fitting method is most important when fitting a power-law in the tail of a distribution; the binned estimator performs nearly as well as the un-binned for the double power-law parameters γ and $E_{\rm b}$. The (binned) methods used to report parameters like the "ankle" and the "knee" in [47] and [48] are sufficient but limited by the bin width.

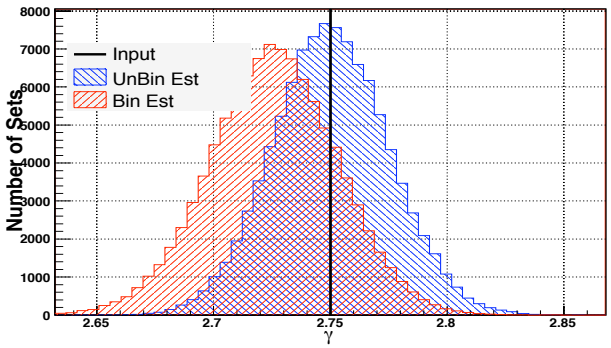


Figure 5.6: For each of 10^5 sets of 5×10^3 events drawn from a pure power-law with index $E_{\min} = 1.0$ and $\gamma = 2.75$ we estimate the spectral index using the binned Eq(5.14) and un-binned Eq(5.1) methods. The bias and error of the un-binned estimator is 0.0002 and 0.0247 and that of the binned is -0.024 and 0.0272.

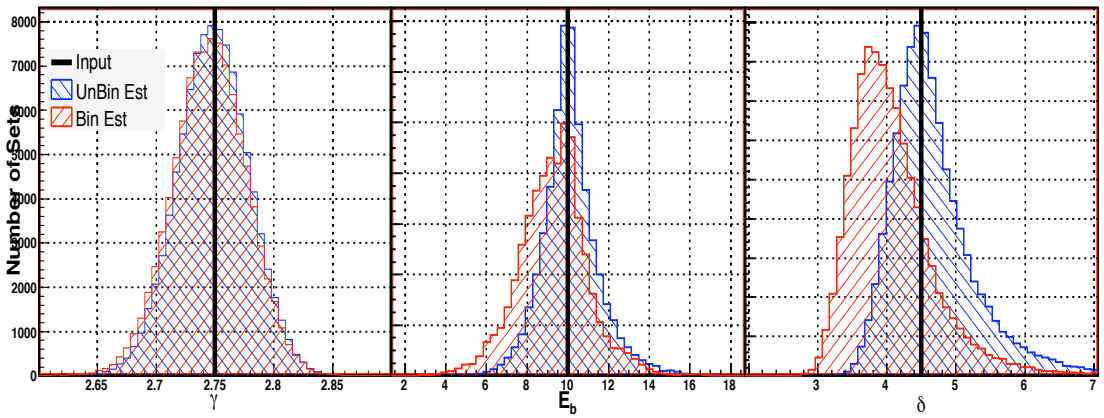


Figure 5.7: For each of 10^5 sets of 5×10^3 events drawn from a double power-law with parameters $\{\gamma, E_{\rm b}, \delta\} = \{2.75, 10.0, 4.5\}$ we estimate the spectral index using the binned Eq(5.14) and un-binned Eq(5.1) methods. The bias and error of the un-binned estimators are $\{-0.002, 0.13, 0.16\}$ and $\{0.03, 1.4, 0.60\}$ and those of the binned are $\{-0.005, -0.71, -0.43\}$ and $\{0.03, 1.7, 0.60\}$.

5.8 Statistical Error: Monte-Carlo Example

To illustrate the effect the statistical energy smearing has on a pure power-law we generate 9000 MC events from a power-law distribution with $E_{\rm min} = 1.0$ and $\gamma = 2.75$. A histogram of these events is represented by the black filled circles plotted in Figure 5.8. By minimizing Eq(5.1), we calculate the estimated spectral index for this data to be $\hat{\gamma} = 2.742 \pm 0.019$ (with $\hat{E_{\rm min}} = 1.0$, see §5.5.1). A power-law with these parameters is plotted as the dashed line in Figure 5.8.

To each MC event E_i we then add a random number Y_i drawn from a normal distribution with mean zero and variance $0.2E_i$. The new events are histogram-ed with blue open circles in Figure 5.8. We fit these events by maximizing a likelihood with

$$\int_{E_{\rm min}}^{\infty} f_{\rm M}(t;\vec{\theta})\, G(t; E_{\rm obs}, \sigma_{stat}(t;\vec{p}))\, dt. \tag{5.15}$$

(compare with Eq(5.3)) as the p.d.f. and we find that $\hat{\gamma} = 2.749 \pm 0.020$. The smearing does not effect the estimated spectral index, though it does increase the error of the estimate. The dashed curve in Figure 5.8 shows Eq(5.15) evaluated at the best fit values. Notice that the histogram of the smeared energies deviates from the un-smeared case near $\lg E \sim 0$. In §5.4.3 we account for this edge effect at the low energy end by assuming that the true energies follow the power-law well below the observed minimum energy; in constructing the likelihood we choose $0.1E_{\rm min}$ for the lower rage of integration (compare Eq(5.15) with Eq(5.3)) and we re-normalize to ensure a true p.d.f. (see Eq(5.4)).

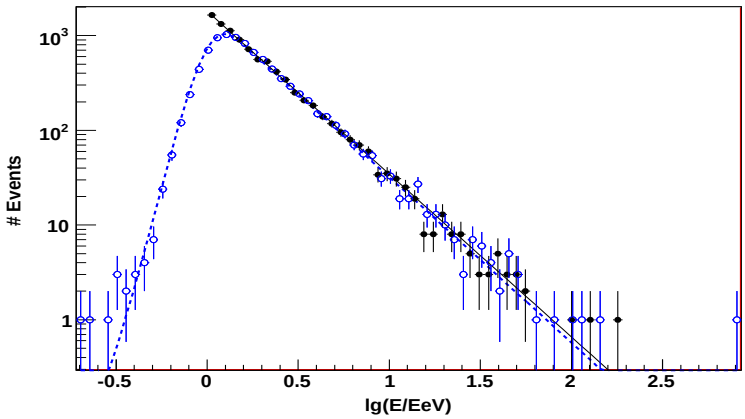


Figure 5.8: An example of a pure power-law before and after smearing. A histogram of 9000 events drawn from a single power-law with $E_{\min} = 1.0$EeV and $\gamma = 2.75$ is plotted in black filled circles. The best fit (using Eq(5.1)) power-law for these events is plotted in solid black. The blue open circles are a histogram of these events after being smeared by a Gaussian with variance $0.2E$ (see §5.8). The blue dashed curve shows the best fit using Eq(5.15). To account for the edge effect near $\lg E \sim 0$ we use the methods in §5.4.3, namely Eq(5.3).

5.9 Results of CRPropa Toy Set

By applying the statistical tools presented in this paper (summarized by steps (1)-(4) [5] in §5.6) to the toy set of 5×10^3 CRPropa events (see §5.3.2) we illustrate how the tools might be implemented on an actual CR detector. By construction, this toy set has parameter estimates and, more importantly, errors estimates and hypothesis test p-values that are numerically comparable with those reported by Auger[47] and HiRes[48].

In preparation for this paper we generated 14 CRPropa simulations of $\sim 2 \times 10^5$ events with different injection spectral indexes, $\gamma_{\text{IN}} = (2.0, 2.1, \ldots, 2.6)$, and with

[5] Note that since we are not interested in the absolute goodness of fit for any of these *toy* models to this *toy* data set, we do not perform step (4a) of §5.6.

different values of maximum generation energy, $E_{\rm max}/{\rm EeV} = (400, 2000)$. The (after propagation) estimated characteristic break point energy, i.e. $\hat{E}_{\frac{1}{2}}$ or $\hat{E}_{\rm b}$, is found to be independent of the spectral index at the site of generation, $\gamma_{\rm IN}$. The estimated spectral index $\gamma_{\rm OUT}$ is found to be linearly related to the input spectral index $\gamma_{\rm IN}$ with linear slope ~ 1. The high energy estimated shape parameters, δ and w_c, are more sensitive to the maximum generation energy (at the sources) than they are to $\gamma_{\rm IN}$.

In Figs. 5.9 and 5.10 we plot the toy data set and the best fit models in two (non-binned) ways not commonly seen in the CR literature. The first is a *rank-frequency* plot. For each event (black filled circle) we plot $\lg E$ along the horizontal axis and the log of the number of events with energy greater than E along the vertical. For each of the models (see §6.4.1), the vertical axis is $\lg(N_{tot}(1 - F(E)))$ where $F(E)$ is the model cumulative distribution function. From the rank-frequency plot we derive an instructive visualization tool in Figure 5.10; we plot the difference between the number of events above a given energy for the toy set $N_{>}^{\rm obs}$ and that expected by the best fit models $N_{>}^{\rm exp}$.

The best fit pure power-law parameters for the toy set described in §5.3.2 are $E_{\rm min} = 6.31 \pm 0 \pm^{0.82}_{0.82}$ and $\gamma = 2.83 \pm^{0.03}_{0.03} \pm^{-0.07}_{+0.10}$ where the first error is statistical and the second systematic. The tail power significance $p_{\rm TP}$ is 4.6σ. The best fit double power-law parameters for the toy set are $E_{\rm min} = 6.31 \pm 0 \pm 0.82$, $\gamma = 2.71 \pm 0.03 \pm^{-0.06}_{+0.10}$, $E_{\rm b} = 45.7 \pm^{2.3}_{4.1} \pm 9.9$ and $\delta = 4.30 \pm 0.26 \pm^{-0.11}_{+0.20}$. The correlation coefficients are $\rho_{\gamma E_{\rm b}} = 0.18$, $\rho_{\gamma\delta} = -0.15$ and $\rho_{E_{\rm b}\delta} = 0.32$, see Figure 5.11. The likelihood ratio significance is $\lg p_{\mathcal{R}} = -2.7$. The best fit Fermi power-law parameters for the toy set are $E_{\rm min} = 6.31 \pm 0 \pm 0.82$, $\gamma = 2.69 \pm 0.03 \pm^{-0.06}_{+0.09}$, $E_{\frac{1}{2}} = 78.6 \pm^{6.8}_{7.6} \pm^{18.6}_{19.1}$ and $w_c = 0.139 \pm^{0.024}_{0.029} \pm^{0.005}_{0.008}$. The correlation coefficients are $\rho_{\gamma E_{\frac{1}{2}}} = 0.61$, $\rho_{\gamma w_c}$ and $\rho_{E_{\frac{1}{2}} w_c} = -0.07$, see Figure 5.12. The likelihood ratio significance is $\lg p_{\mathcal{R}} = -1.9$.

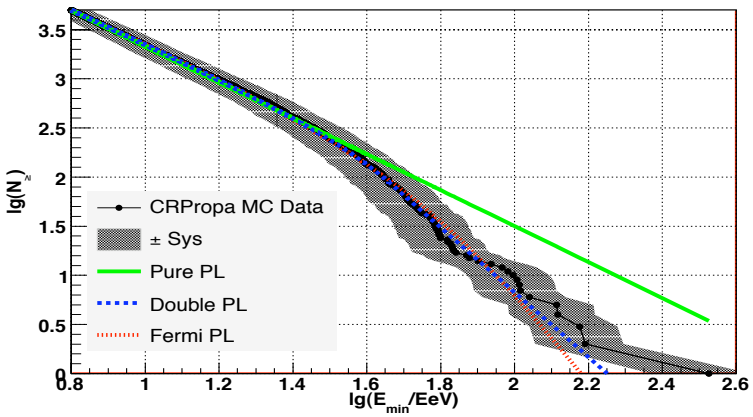


Figure 5.9: A rank-frequency plot as simulated by 5×10^3 events from the CRPropa set with parameters $\gamma_{\rm IN} = 2.6$ and $E_{\rm max} = 2000$ EeV. For each event (black filled circle) we plot $\lg E$ along the horizontal axis and log-number of events with energy greater than E along the vertical. The models are described in §6.4.1.

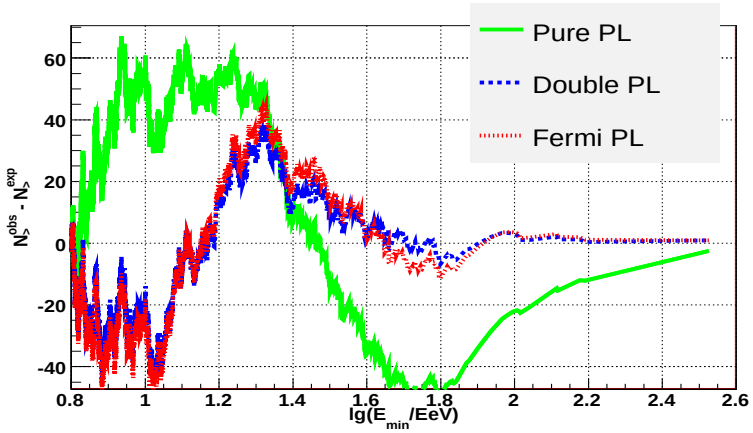


Figure 5.10: Using the rank-frequency plot (see Figure 5.9) we plot the difference between the number of events above a given energy for the toy set $N_>^{\rm obs}$ and that expected by the best fit models $N_>^{\rm exp}$. Note that at $\lg E_{\rm min}/{\rm EeV} \sim 1.7$, there are at least forty fewer events observed than expected by the pure power-law fit, i.e. flux suppression.

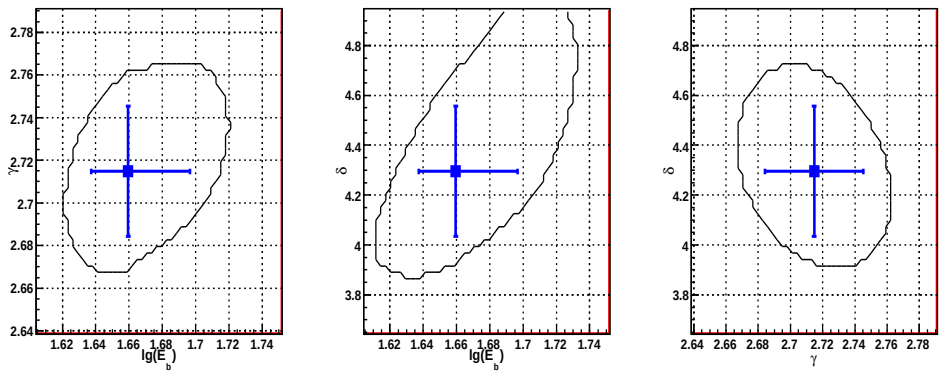


Figure 5.11: The change in log-likelihood $-2\Delta\mathcal{L}_{\mathrm{DP}}$ (see §5.4.1) as a function of the parameters γ, E_{b} and δ of the double power-law. The data set is the toy set described in §5.3.2. The best estimate for each parameter is plotted as a blue box, the asymmetric one degree of freedom error estimates ($-2\Delta\mathcal{L}_{\mathrm{DP}} = 1$) are plotted as solid blue lines and the black contour defines the two degree of freedom error estimate ($-2\Delta\mathcal{L}_{\mathrm{DP}} \geq 2.30$).

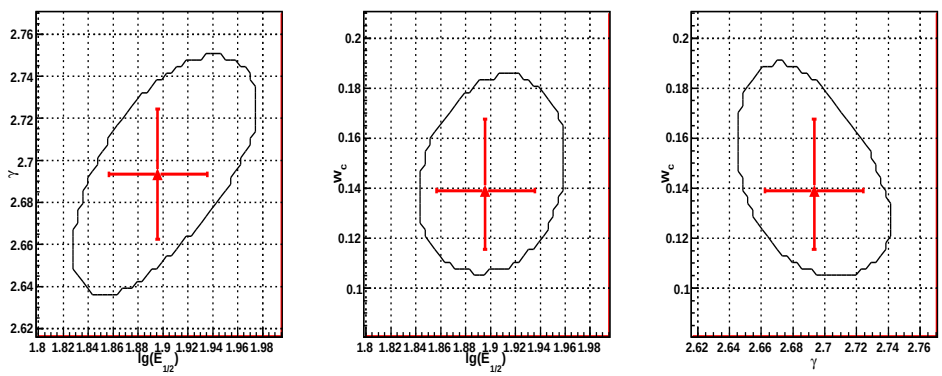


Figure 5.12: The change in log-likelihood $-2\Delta\mathcal{L}_{\mathrm{FP}}$ (see §5.4.1) as a function of the parameters γ, $E_{\frac{1}{2}}$ and w_c of the Fermi power-law. The data set is the toy set described in §5.3.2. The best estimate for each parameter is plotted as a blue box, the asymmetric one degree of freedom error estimates ($-2\Delta\mathcal{L}_{\mathrm{FP}} = 1$) are plotted as solid blue lines and the black contour defines the two degree of freedom error estimate ($-2\Delta\mathcal{L}_{\mathrm{FP}} \geq 2.30$).

Chapter 6

Connecting Flux Suppression with Arrival Directions

6.1 Chapter Overview

The text in this chapter was intended for submission under the title "The Flux at different 'Angles"' by the authors J. D. Hague, B. R. Becker, M. S. Gold and J. A. J. Matthews in on the internal Auger technical notes server *GAP-Notes*. This paper is intended as a first attempt at extracting and comparing spectral information for events with different arrival directions in the sky (and the temporal coordinate). This work has not been previously published.

The spectrum in six different "angular" coordinates is studied; zenith, azimuth, time, declination, angular distance from the galactic plane, the super galactic planes and Centaurus A. We bin the events by angle into equal event number bins and fit the fluxes to the single and double power-laws. We concentrate on characterizing the consistency of the highest energy flux between the angular bins. We find anomalous high energy flux for the azimuth between -3 and 89 degrees and within 44 degrees of

Centaurus A.

6.2 Introduction

Of primary interest to the Auger Observatory (PAO) is the study of the spectrum the highest energy cosmic rays. The PAO has reported strong evidence in favor of both the GZK-cutoff [36] and anisotropy [8, 9] at the highest energies. With this in mind, we endeavor here to qualatatively describe the flux of cosmic rays (CRs) in different regions of the sky.

We begin with a brief description of how the energy is obtained from the observed quantities with an emphasis on the zenith angle dependence of the reconstructed energy. The methods we apply to this data set are described in §6.4; we fit spectra from four angular bins in six different "angles"[1] to two different flux models. We summarize our observations in §B.2.4 and a graphical and numerical summary of the fit results can also be found there.

6.3 Data Set

6.3.1 Observer

For this study we use the $\overline{\text{Off}}\underline{\text{line}}$ Observer surface detector data[58] that satisfy the following standard cuts: Bad Period "\$5 < 1", T5 "\$10 == 1 || \$10 == 3", Zenith angle θ "\$17 > 0 && \$17 < 60" and Energy $E/\text{EeV} \geq 10^{0.5}$.

At the time of writing, the energies of each event in the Observer file[58] have been reconstructed using parameters derived for ICRC '05[59]. The reconstruction

[1] Time is not an angle.

methods and parameters have since been refined and updated to include more contemporary data. We use the parameters and methods found in the Physics Review Letters article[36].

6.3.2 Constant Intensity Cut

To calculate the energy using the Constant Intensity Cut (CIC) method we need two measured quantities; the zenith angle θ in degrees and the signal one thousand meters from the shower core S_{1000} in vertical equivalent muons VEM. A histogram of these two measured quantities for this data set can be found in Figure 6.1.

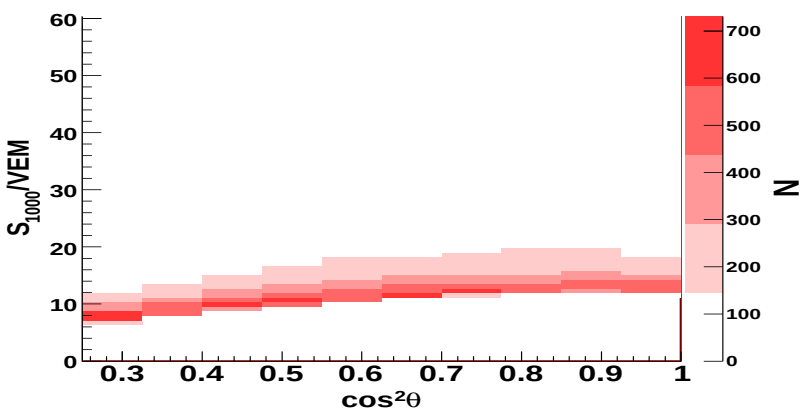


Figure 6.1: A two dimensional histogram of the data as a function of the square of the cosine of the zenith angle $\cos^2(\theta)$ and signal S_{1000}. These two quantities are used to reconstruct the energy.

Using the attenuation curve f_{Att} we can "reconstruct" the energy as

$$E(S_{1000}, \theta; a, b, A, B) = A \left(\frac{S_{1000}}{f_{Att}(\theta; a, b)} \right)^B , \tag{6.1}$$

where $f_{Att}(\theta; a, b) = 1 + ax + bx^2$ with $x = \cos^2(\theta) - \cos^2(38^o)$, $a = 0.92$, $b = -1.13$, $A = 0.149$EeV and $B = 1.080$[36]. The CIC method makes two assumptions about

the flux of CRs; first the isotropy hypothesis, that "the number of events of a given energy arriving from a given solid angle around a direction in the sky is the same for all directions"[36], and second the hypothesis of the energy *independence* of Eq(6.1). The parameters a and b are are determined by fitting the attenuation curve to the data for a given intensity, $I_0 = 0.24/\,\mathrm{km}^2\,\mathrm{sr}\,\mathrm{yr} \approx 9\mathrm{EeV} \approx 47\mathrm{VEM}$[60]. Note that we use the exposure $\Omega = 7000\,\mathrm{km}^2\,\mathrm{sr}\,\mathrm{yr}$ quoted in [36]. The parameters A and B are derived using hybrid events.

In Figure 6.2 we compare the energy of each event reported in the Observer with the energy obtained using Eq(6.1). We note the the energies are highly corelated and that the relative difference in the energies can be substantial; $\sim 20\%$ of the events have thier energy shifted by more than $\sim 10\%$.

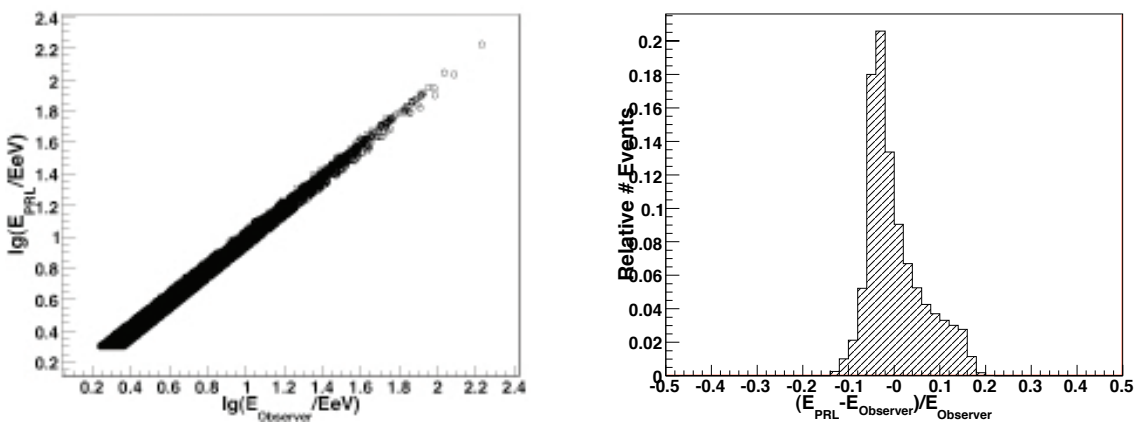


Figure 6.2: *Left:* For each event we plot the log of the energy reported in the Observer versus the log of the energy obtained using Eq(6.1). *Right:* The distribution of the relative difference between the energy reported in the Observer[58] and the energy obtained using Eq(6.1).

The final and most substantial cut that we apply to the data set is that the energy obtained using Eq(6.1) be greater than $10^{0.5}$EeV. There are a total of 17,361 events that satisfy all of our cut criteria.

6.4 Method

6.4.1 Models

In this study we fit various subsets of the energy spectrum to two models; the single and double power-laws, see Table 6.1. The single power-law has one free parameter; the spectral index γ. The double power-law has three; a low energy index γ, a characteristic break energy $E_{\rm b}$/EeV and a high energy index δ. We use the unbinned maximum likelihood methods outlined in [7, 52] and Chapter 5 to estimate the best fit parameters of each model. The statistical error of each parameter is given by parameter values such that the change in the likelihood function is one-half.

Model	$N_{\rm dof}$	Normalization	Function
P	1	$(\gamma-1){E_{\rm min}}^{\gamma-1}$	$E^{-\gamma}$
DP	3	$\frac{\gamma-1}{E_{\rm b}}\left\{\left(\frac{E_{\rm b}}{E_{\rm min}}\right)^{\gamma-1}+\frac{\gamma-1}{\delta-1}-1\right\}^{-1}$	$\left(\frac{E}{E_{\rm b}}\right)^{-\gamma}$ $E_{\rm min}\le E<E_{\rm b}$ $\left(\frac{E}{E_{\rm b}}\right)^{-\delta}$ $E_{\rm b}\le E$

Table 6.1: The model designation (Model = Pure power-law or Double Power-law), number of free parameters, normalization, and form of the function used to fit the fluxes used in this study.

6.4.2 "Angles"

Associated with each energy is, of course, the arrival direction of the CR. We study the energy spectrum in five different angular coordinates. For each angle we bin the events into (approximately) equal event number bins and fit the energy flux of the events in each bin to both the pure and double power-laws. We use four bins with ~ 4350 events per bin for each of these angles. The angles we study can be divided into two classes, they are:

- **The Local Frame** This coordinate system is the natural one for the observatory.
  - *Zenith* angle $0^{\rm o} \leq \theta \leq 60^{\rm o}$. The determination of the energy, see §6.3, is directly dependent on θ.
  - *Azimuthal* angle $-180^{\rm o} < \phi \leq 180^{\rm o}$, measured from due north.
  - *Time* of arrival 26/1/04≤D/M/Y≤1/6/08.

  In principle, we expect these fluxes to be independent of the angular bin and they therefore offer a good cross check for the data.

- **The Global Frames** These coordinate systems are based on astronomically convenient parametrizations of the sphere.
  - Equatorial coordinates, *declination* $-90^{\rm o} < {\rm Dec} \leq 25^{\rm o}$.
  - Galactic plane,*magnitude of the galactic latitude* $0^{\rm o} < |b_{\rm G}| \leq 90^{\rm o}$.
  - Super-Galactic plane, *magnitude of the super-galactic latitude* $0^{\rm o} < |b_{\rm SG}| \leq 90^{\rm o}$.
  - Centaurus A, *angular distance from Centaurus A* $0^{\rm o} < |{\rm Cent\ A}| \leq 180^{\rm o}$.

6.4.3 Metrics

We know that the single power-law is a bad fit to the data, see [36], and we are therefore most interested in the characteristic location and shape of the flux of the highest energy events.

The three parameters in the double power-law fit can describe a flux with a sharp cut-off but they all conspire to effect the high energy flux. Therefore, we use as our first metric, a combination of the double power-law parameters. The energy in EeV

at which the double power-law model reaches half the value it would have if the low energy index continued above $E_{\rm b}$ is given by

$$E_{\frac{1}{2}} = 2^{\frac{1}{\delta-\gamma}} E_{\rm b}. \tag{6.2}$$

Since it depends on $E_{\rm b}$ and $\delta - \gamma$, this parameter includes information about both the location and the shape of a particular double power-law parameter set. Writing the error matrix for the parameters $E_{\rm b}$, δ and γ as

$$\begin{pmatrix} \sigma_\gamma^2 & \sigma_\gamma\sigma_{E_{\rm b}} & \sigma_\gamma\sigma_\delta \\ \sigma_\gamma\sigma_{E_{\rm b}} & \sigma_{E_{\rm b}}^2 & \sigma_{E_{\rm b}}\sigma_\delta \\ \sigma_\gamma\sigma_\delta & \sigma_{E_{\rm b}}\sigma_\delta & \sigma_\delta^2 \end{pmatrix}, \tag{6.3}$$

we can write the statistical error on $E_{\frac{1}{2}}$ as

$$\left(\frac{\sigma_{E_{\frac{1}{2}}}}{E_{\frac{1}{2}}}\right)^2 = \left(\frac{\sigma_{E_{\rm b}}}{E_{\rm b}}\right)^2 - \frac{2(\sigma_{E_{\rm b}}\sigma_\delta - \sigma_\gamma\sigma_{E_{\rm b}})}{E_{\rm b}} \left(\frac{\lg 2}{(\delta-\gamma)^2}\right) + (\sigma_\gamma^2 - 2\sigma_\gamma\sigma_\delta + \sigma_\delta^2) \left(\frac{\lg 2}{(\delta-\gamma)^2}\right)^2.$$

By calculating the error weighted average of $E_{\frac{1}{2}}$ over the bins of each angle we obtain a χ^2/DoF for each angle. If this value is near unity we can be reasonably sure that the spectra are consistent within each angle.

The second metric we use is motivated solely by "expected" physics; we tabulate the number of events with energy greater than 56EeV $N_{E>56\rm EeV}$ for each bin of each angle.

6.5 Conclusion

In Figs. 6.4 and 6.6 we plot a representation of the flux for each bin of the angles studied. For each event in a set we plot the $\log_{10}$ of the energy along the horizontal axis. Along the vertical axis we plot the $\log_{10}$ of the "cumulative flux" Φ times a

scale factor $E^2_{(k)}$. Let $E_{(1)} < E_{(2)} < \ldots < E_{(k)} < \ldots < E_{(N)}$ be the ordered set of energies under consideration. For the $k^{\rm th}$ event energy $E_{(k)}$ we plot

$$\lg(E^2_{(k)}\Phi) = \lg\left(E^2_{(k)}\frac{N-k+1}{\Omega}\right), \tag{6.4}$$

where $\Omega = 7000\,{\rm km}^2\,{\rm sr}\,{\rm yr}$ is the exposure quoted in [36], N is the total number of events in the set being plotted and k is the event's *rank*. These types of *rank-frequency* plots allow one a detailed view of the flux of the highest energy events. In each of the plots in Figs. 6.4 and 6.6 we show the flux of all of the 17,361 events that satisfy our cut criteria in solid black. The flux for each bin of the angle under consideration is plotted in broken and colored lines.

The numerical results of fitting the models to the fluxes in each angular bin are summarized in Tables 6.3 and 6.4. The bin-widths, chosen such that the number of events in each bin is ~ 4350, are listed in the first column of the tables. In the second column we tabulate the best fit spectral index γ of the single power-law fit to each bin. The next three columns list the best fit double power-law parameters γ, $E_{\rm b}$ and δ. The χ^2/DoF for all of these fits are summarized in Table 6.2.

The zenith angle is probably the most important of the angles to check for consistency since the energy is so intimately tied to it, see Eq(6.1). In our analysis we demand only that there be four bins with approximately equal numbers of events. The fact that the 2nd and 3rd zenith bins meet at $\sim 38^{\rm o}$ means that there are approximately the same number of events with $\theta < 38^{\rm o}$ as there are with $\theta > 38^{\rm o}$. This is consistent with the method of reconstruction, i.e. the "isotropy" assumption in §6.3, for events with $E \geq 10^{0.5}$EeV (see also Chapter 3).

All of the other coordinates show little difference between the spectra observed in each bin, with two exceptions. The first is the azimuthal coordinate (i.e. the cardinal of the arrival direction). The events in the $-3^\circ < \phi < 89^\circ$ bin have a markedly higher characteristic cutoff energy $E_{\frac{1}{2}}$. This means that there are somewhat more ultra high

Angle	$\bar{E_{\frac{1}{2}}}$	χ^2/DoF
θ	55 ± 2	$3.26/3 = 1.09$
ϕ	54 ± 2	$30.25/3 = 10.08$
Time	55 ± 2	$8.03/3 = 2.68$
UTC	54 ± 2	$2.86/3 = 0.95$
NSec	53 ± 2	$22.55/3 = 7.52$
NSecMod	54 ± 2	$20.33/3 = 6.78$
Sec	55 ± 2	$8.04/3 = 2.68$
SecMod	53 ± 3	$2.96/3 = 0.99$
Dec	52 ± 1	$9.79/3 = 3.26$
$\lvert b_{\rm gal}\rvert$	57 ± 2	$2.66/3 = 0.89$
$\lvert b_{\rm sgal}\rvert$	55 ± 2	$5.33/3 = 1.78$
Cent A	53 ± 1	$11.46/3 = 3.82$

Table 6.2: The error weighted average $\bar{E_{\frac{1}{2}}}$ for each "angle" studied. The errors are statistical. See §6.4 for the description of the models and methods.

energy events arriving from the north-west quadrant of the local coordinate system[2]. We know of no physical explanation for this deviation – CR "should be" isotropic in this coordinate at this energy. This excess flux could be due to a statistical fluctuation – $N_{E>56\rm EeV}$ is a meager 14 for this quadrant – or, possibly, an anomaly in the reconstruction algorithm.

The second anomalous $E_{\frac{1}{2}}$ measurement occurs within 44° of Cen A. This excess flux would be generally consistent with an excess of high energy events arriving from a source near Cen A. Indeed, much work is underway within Auger to observe such an excess via traditional angular methods. The excess reported here is, however, called into question by the anomalous flux in the north-west cardinal coordinate and observers should take care to check for anisotropies in all coordinate systems.

[2]This is the coordinate system one uses with a compass!

"Angle"	Interval	γ	γ	$E_{\rm b}$/EeV	δ	$E_{\frac{1}{2}}$/EeV	$N_{E>56\rm EeV}$
Zenith	$0^{\rm o} \leq \theta < 27^{\rm o}$	2.78 ± 0.03	2.72 ± 0.03	30 ± 4	3.69 ± 0.25	61 ± 6	11
	$27^{\rm o} \leq \theta < 38^{\rm o}$	2.75 ± 0.03	2.69 ± 0.03	40 ± 4	4.52 ± 0.47	59 ± 4	11
	$38^{\rm o} \leq \theta < 49^{\rm o}$	2.83 ± 0.03	2.77 ± 0.03	38 ± 3	4.86 ± 0.55	52 ± 3	7
	$49^{\rm o} \leq \theta < 60^{\rm o}$	2.83 ± 0.03	2.78 ± 0.03	37 ± 5	4.25 ± 0.42	59 ± 7	7
Azimuth	$-180^{\rm o} \leq \phi < -92^{\rm o}$	2.75 ± 0.03	2.67 ± 0.03	29 ± 3	4.13 ± 0.30	47 ± 3	8
	$-92^{\rm o} \leq \phi < -3^{\rm o}$	2.82 ± 0.03	2.76 ± 0.03	38 ± 3	4.72 ± 0.53	55 ± 2	7
	$-3^{\rm o} \leq \phi < 89^{\rm o}$	2.80 ± 0.03	2.77 ± 0.03	58 ± 11	4.60 ± 0.72	84 ± 6	14
	$89^{\rm o} \leq \phi < 180^{\rm o}$	2.82 ± 0.03	2.76 ± 0.03	31 ± 3	3.95 ± 0.30	55 ± 4	7
Time	28/1/2004 $\leq$ D/M/Y $<$ 7/3/2006	2.77 ± 0.03	2.71 ± 0.03	35 ± 3	4.23 ± 0.38	56 ± 3	10
	7/3/2006 $\leq$ D/M/Y $<$ 16/2/2007	2.80 ± 0.03	2.73 ± 0.03	31 ± 3	4.15 ± 0.33	51 ± 3	6
	16/2/2007 $\leq$ D/M/Y $<$ 18/10/2007	2.82 ± 0.03	2.77 ± 0.03	37 ± 3	4.43 ± 0.45	56 ± 3	10
	18/10/2007 $\leq$ D/M/Y $<$ 1/6/2008	2.79 ± 0.03	2.75 ± 0.03	40 ± 6	4.03 ± 0.40	69 ± 6	10
Dec	$-90^{\rm o} \leq$ Dec $< -50^{\rm o}$	2.76 ± 0.03	2.69 ± 0.03	38 ± 3	5.15 ± 0.57	50 ± 2	6
	$-50^{\rm o} \leq$ Dec $< -30^{\rm o}$	2.80 ± 0.03	2.75 ± 0.03	38 ± 3	4.35 ± 0.44	59 ± 4	9
	$-30^{\rm o} \leq$ Dec $< -10^{\rm o}$	2.82 ± 0.03	2.76 ± 0.03	25 ± 3	3.47 ± 0.19	66 ± 9	11
	$-10^{\rm o} \leq$ Dec $< 25^{\rm o}$	2.81 ± 0.03	2.75 ± 0.03	35 ± 4	4.06 ± 0.36	59 ± 4	10

Table 6.3: The best fit model parameters for each bin of the Theta coordinate. The errors are statistical. See §6.4 for the description of the models and methods.

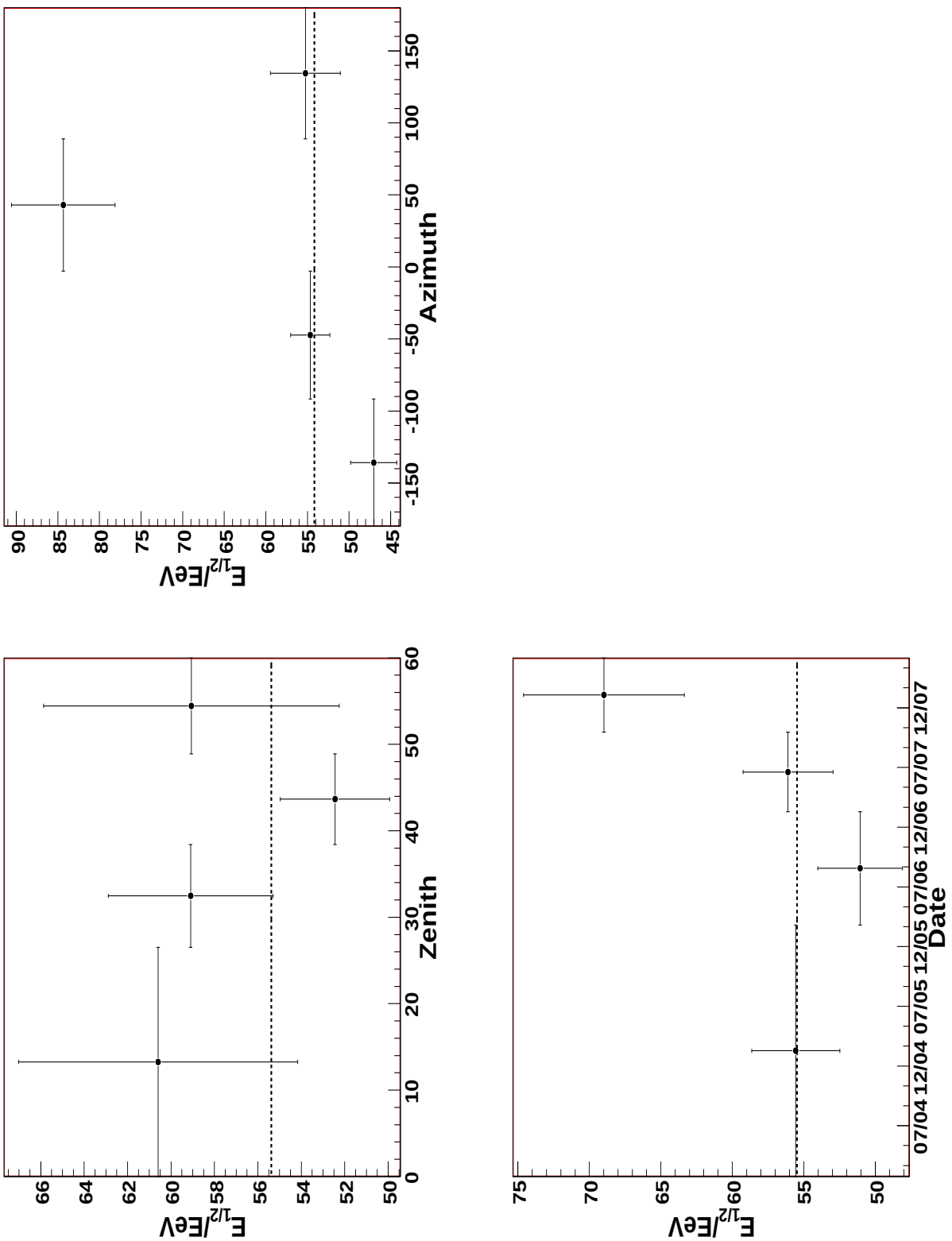
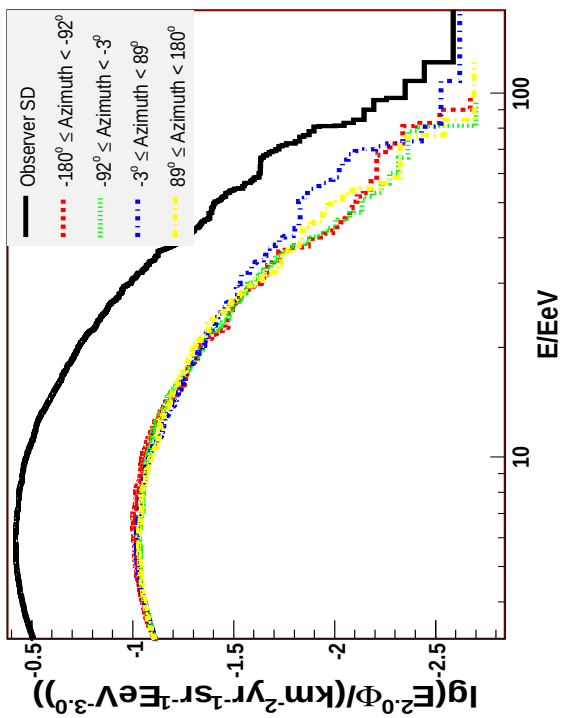
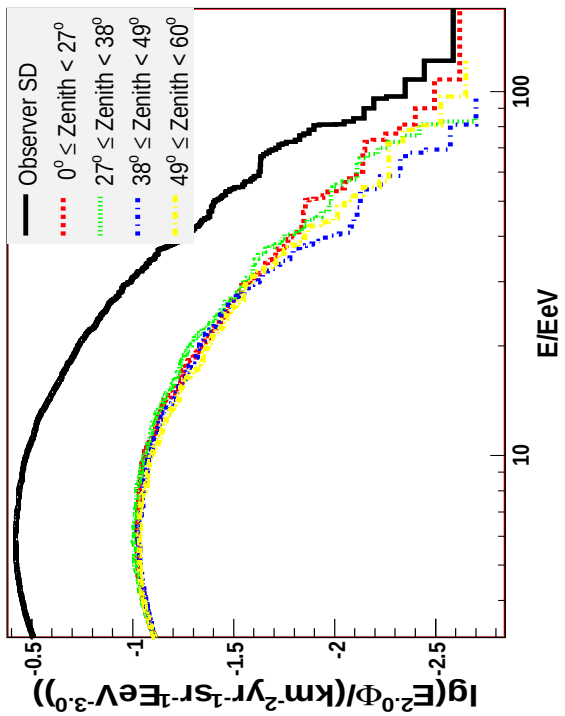


Figure 6.3: The value and statistical error of $E_{\frac{1}{2}}$, see Eq(6.2) and Table 6.2, for the local coordinates.





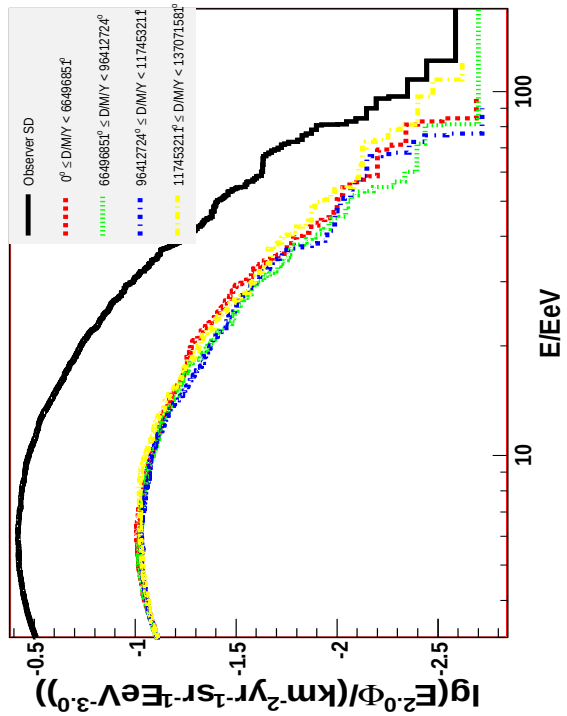


Figure 6.4: The cumulative flux, see Eq(6.4), for the local coordinates.

"Angle"	Interval	γ	γ	$E_{\rm b}$/EeV	δ	$E_{\frac{1}{2}}$/EeV	$N_{E>56{\rm EeV}}$
Dec	$-90^{\rm o} \leq$ Dec $< -50^{\rm o}$	2.76 ± 0.03	2.69 ± 0.03	38 ± 3	5.15 ± 0.57	50 ± 2	6
	$-50^{\rm o} \leq$ Dec $< -30^{\rm o}$	2.80 ± 0.03	2.75 ± 0.03	38 ± 3	4.35 ± 0.44	59 ± 4	9
	$-30^{\rm o} \leq$ Dec $< -10^{\rm o}$	2.82 ± 0.03	2.76 ± 0.03	25 ± 3	3.47 ± 0.19	66 ± 9	11
	$-10^{\rm o} \leq$ Dec $< 25^{\rm o}$	2.81 ± 0.03	2.75 ± 0.03	35 ± 4	4.06 ± 0.36	59 ± 4	10
Gal. Lat.	$0^{\rm o} \leq \lvert b_{\rm gal}\rvert < 14^{\rm o}$	2.80 ± 0.03	2.75 ± 0.03	43 ± 4	4.81 ± 0.60	60 ± 3	8
	$14^{\rm o} \leq \lvert b_{\rm gal}\rvert < 30^{\rm o}$	2.78 ± 0.03	2.72 ± 0.03	39 ± 3	4.52 ± 0.47	58 ± 3	11
	$30^{\rm o} \leq \lvert b_{\rm gal}\rvert < 48^{\rm o}$	2.80 ± 0.03	2.74 ± 0.03	31 ± 3	3.83 ± 0.29	59 ± 5	10
	$48^{\rm o} \leq \lvert b_{\rm gal}\rvert < 90^{\rm o}$	2.80 ± 0.03	2.74 ± 0.03	35 ± 3	4.31 ± 0.39	54 ± 3	7
S. Gal. Lat.	$0^{\rm o} \leq \lvert b_{\rm sgal}\rvert < 14^{\rm o}$	2.81 ± 0.03	2.76 ± 0.03	28 ± 4	3.59 ± 0.23	64 ± 7	12
	$14^{\rm o} \leq \lvert b_{\rm sgal}\rvert < 30^{\rm o}$	2.78 ± 0.03	2.72 ± 0.03	37 ± 5	4.13 ± 0.38	60 ± 4	9
	$30^{\rm o} \leq \lvert b_{\rm sgal}\rvert < 49^{\rm o}$	2.79 ± 0.03	2.72 ± 0.03	38 ± 2	4.92 ± 0.54	52 ± 2	7
	$49^{\rm o} \leq \lvert b_{\rm sgal}\rvert < 90^{\rm o}$	2.81 ± 0.03	2.75 ± 0.03	34 ± 4	4.21 ± 0.38	55 ± 3	8
Cent A	$0^{\rm o} \leq$ Cent A $< 44^{\rm o}$	2.80 ± 0.03	2.76 ± 0.03	38 ± 5	3.88 ± 0.35	70 ± 6	14
	$44^{\rm o} \leq$ Cent A $< 67^{\rm o}$	2.78 ± 0.03	2.71 ± 0.03	37 ± 3	5.22 ± 0.59	49 ± 2	3
	$67^{\rm o} \leq$ Cent A $< 93^{\rm o}$	2.83 ± 0.03	2.77 ± 0.03	38 ± 4	4.79 ± 0.55	54 ± 2	6
	$93^{\rm o} \leq$ Cent A $< 180^{\rm o}$	2.77 ± 0.03	2.70 ± 0.03	26 ± 2	3.54 ± 0.20	58 ± 6	13

Table 6.4: The best fit model parameters for each bin of the Theta coordinate. The errors are statistical. See §6.4 for the description of the models and methods.

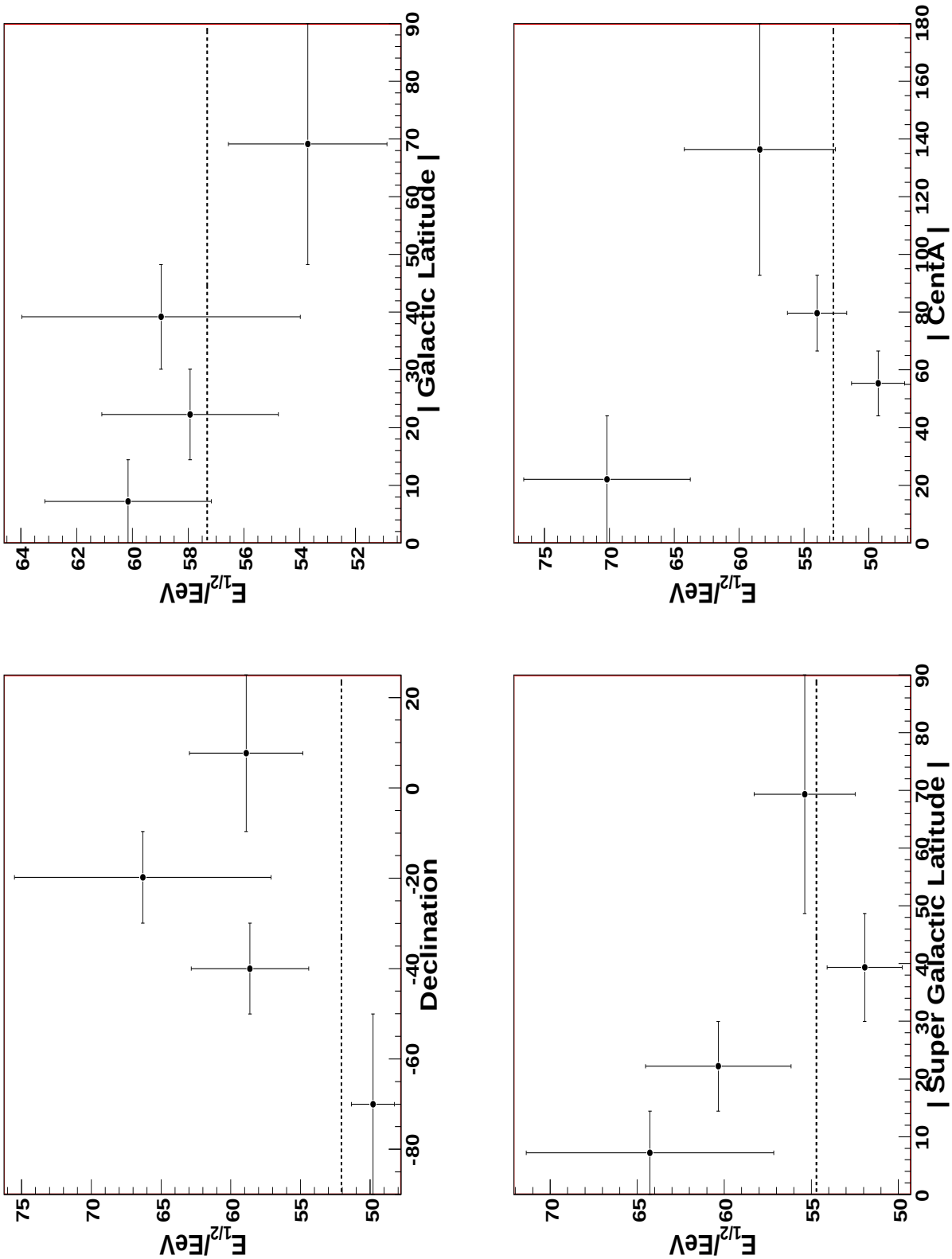


Figure 6.5: The value and statistical error of $E_{\frac{1}{2}}$, see Eq(6.2) and Table 6.2.

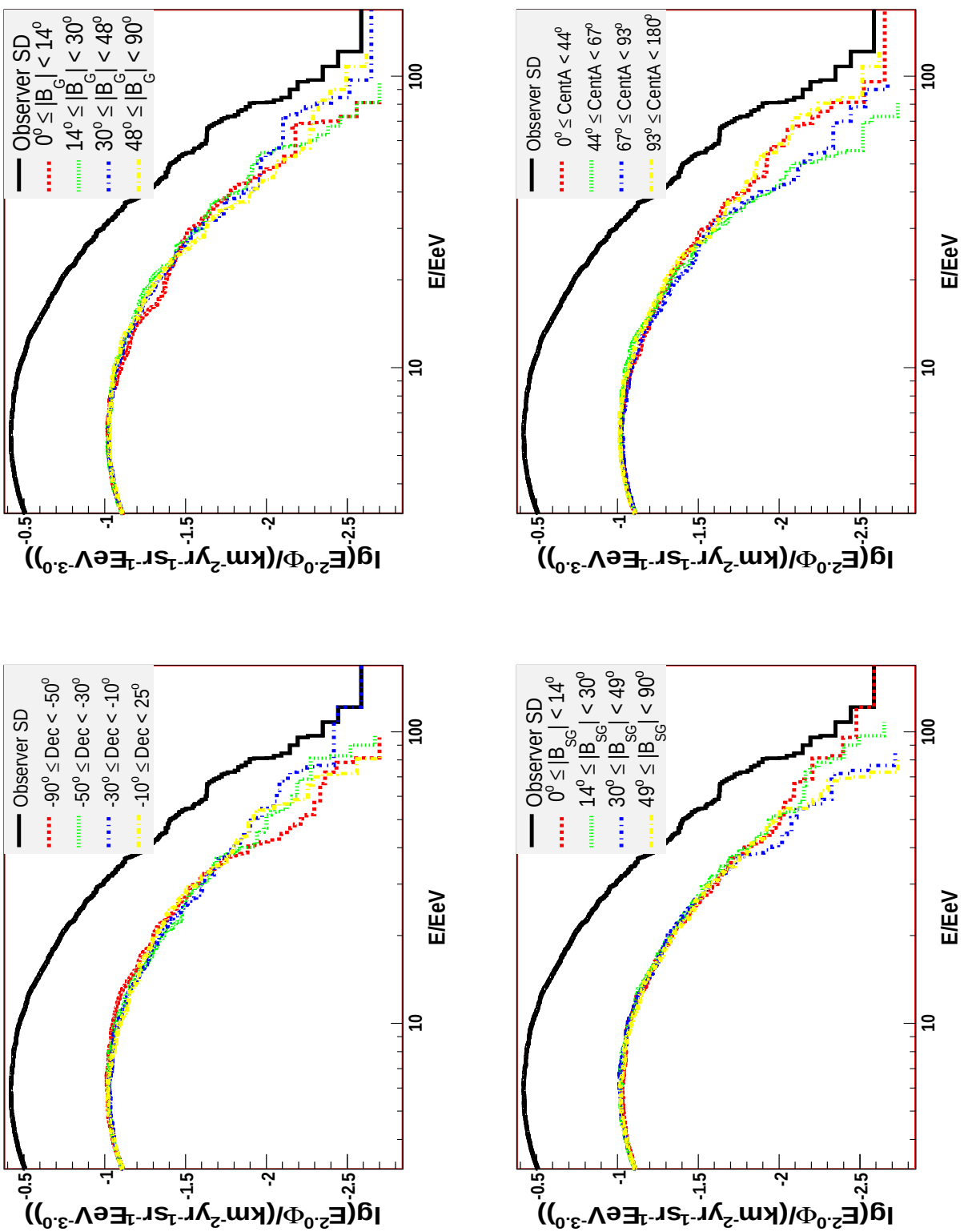


Figure 6.6: The cumulative flux, see Eq(6.4), for the global coordinates.

Chapter 7

Anisotropy : Estimating Little-p

7.1 Chapter Overview

The text in this chapter was published under the title "The Flux at different 'Angles"' by the authors B. R. Becker, S. BenZvi, B. Connolly, M. S. Gold, J. D. Hague, J. A. J. Matthews and S. Westerhoff in on the internal Auger technical notes server as *GAP-Note-2007-097*. This paper was written just after the so called "running prescription" passed[61], which allowed "the correlation of ultra high energy CR with Active Galactic Nuclei"[8]. The primary goal of this work is to estimate the observed value of the correlation *little-p* in post-scan data. This work was not widely used at the time it was published, but has since become the analysis basis for some of the key results of later Auger work (see §7.7).

7.2 Introduction

Given that the *running-prescription*[62] has passed[61], *viz.* that 6 of the first 8 events correlated, it is timely to ask what values of little-p are consistent with our observations. For this note the most relevant parts of the *prescription* are:

- n, the total number of events with energy greater than 56 EeV,
- k, the number of events (out of n) that correlate with the angular region of the sky defined in the *prescription*, *i.e.* $\gamma_{max} = 3.1^o$ and $z_{max} = 0.018$,
- **little-p**, the fractional probability of any specific outcome *e.g.* the per cosmic ray probability that it passes the *prescription* and
- **little-p_{bg}** $= 0.21$, the probability for an accidental (or chance) cosmic ray passing the *prescription*. If the cosmic rays follow the acceptance of the Auger experiment, then little-p=0.21.

In this note we use the *alternative prescription* conventions outlined in Connolly *et.al.*[62], which does not "pass" until $k/n = 8/11$. However, the discussion of our estimate of little-p is relevant to both schemes since it does not depend on how one chooses to reject the isotropic (null) hypothesis.

The likelihood of getting k correlations out of n events is the binomial term[62],

$$L(k|n,p) = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}. \tag{7.1}$$

In this note will choose a flat prior, $\pi(p) = 1$ for $0 < p < 1$. Intuitively, this choice reflects a neutral stance as to how many events k we *expect* out of the total n. We leave the discussion of different choice of prior to future communique.

Following Bayes' Theorem [49] we can write our belief about the parameter little-p *given* that k out of n events correlate as,

$$Post(p|n,k) = \frac{(n+1)!}{k!(n-k)!}p^k(1-p)^{n-k}. \tag{7.2}$$

This *posterior* distribution tells the whole story about ones belief about little-p after the data is collected. Indeed, no matter the evidence reported by the experimenter, "... the consumer of that result will almost certainly use it to derive some impression about the value of the parameter. This will inevitably be done, either explicitly or intuitively, with Bayes' theorem ..."[1]

7.3 Estimating little-p

In order to gain intuition about $Post(p|n,k)$ we can quantify its characteristic values; "location" $\hat{p}$ and "shape" $\delta p_{\pm}$. The maximum of eq(7.2) occurs at $\hat{p} = k/n$. We may consider $\hat{p}$ the estimated value of little-p and associate it with the location of $Post(p|n,k)$. We determine the error of this estimate (*i.e.* the shape of $Post(p|n,k)$) in two different ways.

The first is based on the expected error in the observed number of events k. In the limit of large n, the variance of k for $L(k|n,p)$ (see eq(7.1)) is $(\delta k)^2 = np(1-p)$. From the fact that $\hat{p} = k/n$ we can write $\delta p \sim (\delta k)/n$, which simplifies to

$$\delta p = \sqrt{\frac{k(n-k)}{n^3}}. \tag{7.3}$$

This gives a approximate and symmetric error for our estimate of little-p.

The second method directly uses the posterior and gives an asymmetric interval about $\hat{p}$. We define the 1-standard-deviation (*i.e.* 1σ) confidence interval for $\hat{p}$ such

[1] Quoted from the PDG[49] §*32 Statistics*.

that $Post(p|n,k)$ (eq(7.2)) has the same value when it is evaluated at lower-limit and at the upper-limit **and** that the integral of $Post(p|n,k)$ from the lower- to the upper-limit is equal to the (symmetric) area under the standard normal curve. Numerically, we find the interval $\hat{p} - \delta p_- \leq p \leq \hat{p} + \delta p_+$, such that δp_- and δp_+ satisfy:

$$Post(\hat{p} - \delta p_- | n,k) = Post(\hat{p} + \delta p_+ | n,k) \tag{7.4a}$$

$$\int_{\hat{p}-\delta p_-}^{\hat{p}+\delta p_+} Post(p|n,k)dp = \int_{-1}^{1} e^{-t^2/2}/\sqrt{2\pi}dt \sim \%68.3. \tag{7.4b}$$

The interpretation of this interval is that *given* that k out of n events correlate, there is a %68.3 chance that the true little-p falls between $\hat{p} - \delta p_-$ and $\hat{p} + \delta p_+$.

7.4 Consistency with Background

Under the scheme discussed in this note, the proper method of accepting or rejecting the background (isotropic) hypothesis is outlined in [63]. To this end, equation five on page seven of [63] defines a test statistic R',

$$R' = \frac{\frac{k!(n-k)!}{(n+1)!}}{0.21^k(1-0.21)^{n-k}} = \frac{1}{Post(p_{bg} = 0.21|n,k)}. \tag{7.5}$$

This illustrates the connection between isotropic hypothesis rejection (R') and our posterior belief about little-p given k and n ($Post(p|n,k)$).

For completeness, we review the three possible outcomes[63] for this test statistic; if (1) $R' \geq 95$ then we can reject the isotropic hypothesis with a probability $\alpha = \%1$, (2) $0.051 < R' < 95$ then we can neither accept nor reject the hypothesis of isotropy, (3) then $R' < 0.051$ then we can accept the hypothesis with a probability $\beta = \%5$.

7.5 Results

s Here we present a graphical and numerical summary of the aforementioned methods applied to the post-*Prescription*-definition[62] P.A.O. data. For example, *given* that 6 out of 8 events correlate, the first row of Table 7.1 tells us the following;

- little-$\hat{p}$ (3^{rd} column): the "best estimate" of little-p is $\hat{p} = k/n = 0.750$,
- $^{+\delta p_+}_{-\delta p_-}$ (4^{th} column): the probability that the true value of little-p is in the interval $0.750^{+0.122}_{-0.157}$ is approximately %68.3, see eqs(7.4a, 7.4b),
- δp (5^{th} column): the symmetric error of $\hat{p}$, see eq(7.3) and
- R' (6^{th} column): $R' = 74.14$ and thus we can neither accept nor reject the isotropic hypothesis (see [63] for details).

k	n	little-$\hat{p}$	$^{+\delta p_+}_{-\delta p_-}$	δp	R'
6	8	0.750	$^{+0.122}_{-0.157}$	0.153	74.14
7	10	0.700	$^{+0.122}_{-0.145}$	0.145	85.31
8	11	0.727	$^{+0.113}_{-0.137}$	0.134	$270.83^{\dagger}$
8	13	0.615	$^{+0.121}_{-0.131}$	0.135	47.69

Table 7.1: For k out of n events correlating we calculate little-$\hat{p}$ is the estimated value, $^{+\delta p_+}_{-\delta p_-}$ is the (asymmetric) %68.3 confidence interval, δp is a symmetric confidence interval based on δk and R' is the test statistic for rejection of the isotropic hypothesis. **Note** that one may only read one row of this table to make inferences about the data. †This means that $R' \geq 95$, at this point we can reject the isotropic hypothesis[63], with $\alpha = \%1$ probability.

7.6 Conclusion

The posterior distribution $Post(p|n,k)$ defined in eq(7.2) and plotted for the Auger data in Fig.7.1 tell the whole story of ones belief(s) about little-p given that k out of n events correlate. One may use only one row of Table 7.1 to gain a numerical understanding of the location and shape of this posterior distribution. Despite this, it is interesting to note that δp slowly decreases as n increases. Also of note is that the *alternative prescription*[63] "passes" at 8/11 but at 8/13 the evidence against the null hypothesis has decreased.

7.7 Chapter Post-Script

As mentioned in the overview, this work on estimating little-p has been an important building block for the analysis used to monitor the signal with data that arrived after the start of the prescription. In particular, one may construct an estimate of little-p and its uncertainty for each new event that arrives. Each new estimate is correlated with the prior estimate in such a way that ones current understanding of the observed correlation is given by the current estimate of little-p. The update of the correlation with AGN using data collected after [8] is included as an appendix to this thesis (see Chapter A) and a figure showing the evolution of the observed value of little-p can be found in right panel of Figure A.1.

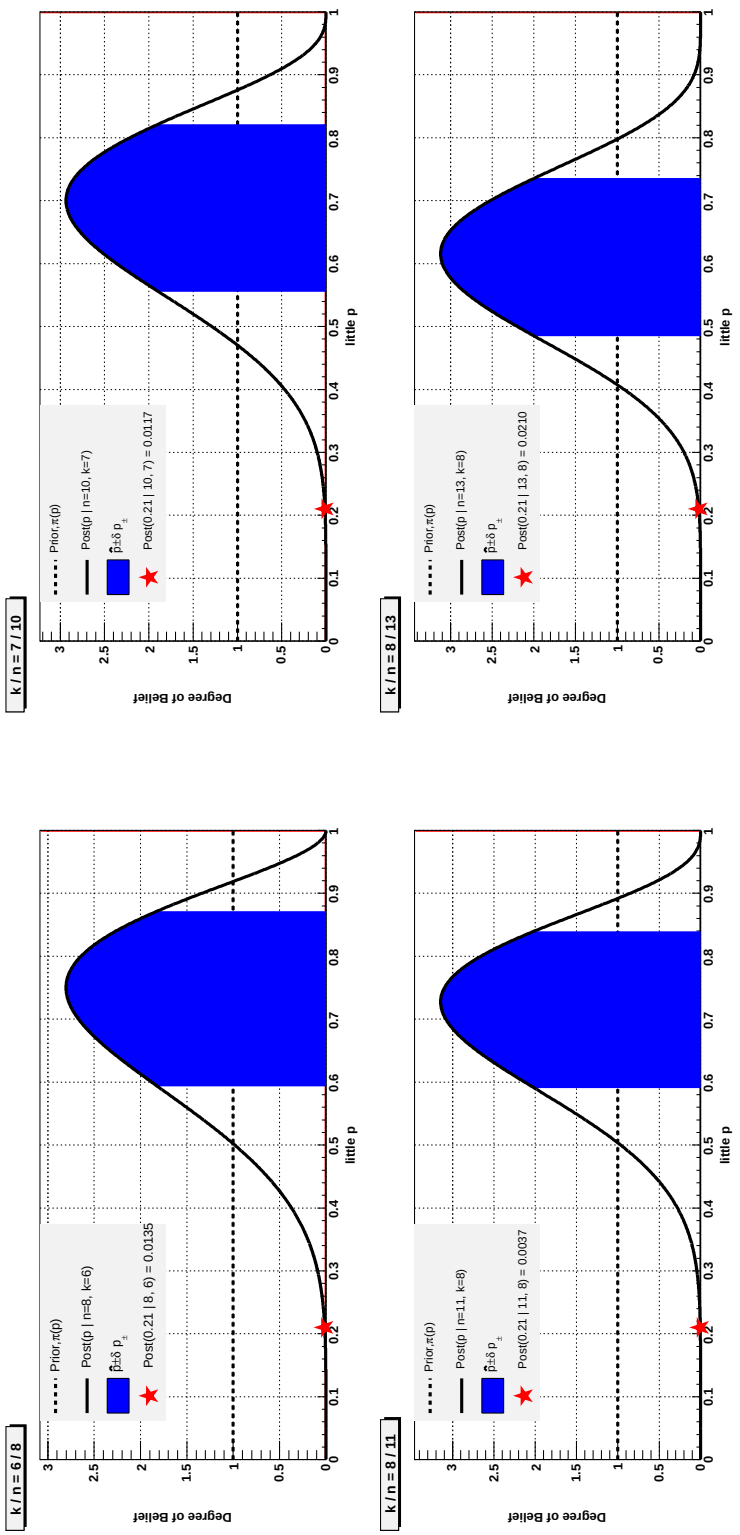


Figure 7.1: For each n and k, *i.e.* each plot; the prior distribution $\pi(p)$ is the dashed black line (see §A.2), the posterior distribution for little-p $Post(p|n,k)$ is the solid black line (see eq(7.2)), the %68.3 confidence interval for little-p is plotted in solid blue (see eqs(7.4a, 7.4b)) and $1/R'$ is plotted as a red star (see eq(7.5)).

Chapter 8

Anisotropy : A Three-Point Cosmic Ray Anisotropy Method

8.1 Chapter Overview

The text in this chapter was submitted under the title "A Three-Point Cosmic Ray Anisotropy Method" by the authors J. D. Hague, B. R. Becker, M. S. Gold and J. A. J. Matthews in in the journal *Journal of Physics G: Nuclear methods*. This paper describes and tests the shape-strength method for measuring anisotropy. It is published as arXiv:0905.4488[64].

The two-point angular correlation function is a traditional method used to search for deviations from expectations of isotropy. In this paper we develop and explore a statistically descriptive three-point method with the intended application being the search for deviations from isotropy in the highest energy cosmic rays. We compare the sensitivity of a two-point method and a "shape-strength" method for a variety of

Monte-Carlo simulated anisotropic signals. Studies are done with anisotropic source signals diluted by an isotropic background. Type I and II errors for rejecting the hypothesis of isotropic cosmic ray arrival directions are evaluated for four different event sample sizes: 27, 40, 60 and 80 events, consistent with near term data expectations from the Pierre Auger Observatory. In all cases the ability to reject the isotropic hypothesis improves with event size and with the fraction of anisotropic signal. While ~ 40 event data sets should be sufficient for reliable identification of anisotropy in cases of rather extreme (highly anisotropic) data, much larger data sets are suggested for reliable identification of more subtle anisotropies. The shape-strength method consistently performs better than the two point method and can be easily adapted to an arbitrary experimental exposure on the celestial sphere.

8.2 Introduction

Cosmic rays with energies above 10 EeV (10^{19} eV) have been observed[5, 40, 48, 36]. However, the sources of these cosmic rays (CR) are unknown and the physics responsible for accelerating CR to these energies is at best conjecture. Evidence supporting an extra-galactic origin of these CR is the observation of energy flux suppression consistent with the GZK-effect[13, 14] by the High Resolution Fly's Eye Experiment[48] and the Pierre Auger Observatory (Auger)[36]. The primary evidence supporting the astrophysical origin of these CR (as opposed to, say, heavy relic decay) is the lack of an observable flux of photons by Auger[65, 66, 10] and the lack of neutrinos observed by ANITA[67, 68].

If the sources are astrophysical, expectations for asymmetries in the arrival directions increase at the very highest CR energies because the local ($\lesssim 100$ Mpc) universe is very anisotropic[69, 70] and the GZK-effect[48, 36] at these energies implies that the sources are local. Observation of an anisotropy in the arrival directions

of CR would be an important step towards identifying the sources of these ultra high energy particles.

Evidence for structure (anisotropy) in the arrival directions has been reported [71, 72, 73, 74, 8, 9, 75, 76]. The most compelling observational evidence consistent with astrophysical expectations of anisotropy is arguably the 27 events with energy greater than 57 EeV recently reported by Auger in [8, 9]. Using the Véron-Cetty – Véron (VCV) catalog[69], the active galactic nuclei (AGN) maximum redshift and correlation angle chosen by Auger defined a limited area (effectively 21%) of the sky[9]. Reported at the 1% significance level, the Auger AGN to CR correlation signal is evidence for a flux of CR enhanced near known low-redshift extra-galactic objects[9].

As even the largest experiments accumulate the very highest energy CR only slowly, [1] the development of new, more sensitive, techniques to search for deviations from isotropy is of particular interest. In contrast to the catalog dependent method used by [8, 9, 77], in this paper we study the effectiveness of two catalog independent methods. Catalog independent techniques avoid the penalty factors for scans over many different catalogs and/or the need to restrict the CR data based on limited sky coverage of a catalog. The first catalog independent technique is a binned two-point (*2-Pt*) angular correlation method (§8.3.2). We also introduce a new three-point method which uses a shape and a strength parameter (*S-S*, or Shape-Strength) for each triplet of events (§8.3.3). Both methods are compared throughout via the binned-likelihood analysis described in §8.3.1.

Arguably, the primary impediment to definitive CR source identification is the small number of ultra-high energy events (those near and above the GZK cut-off, which are most likely to be anisotropic). While lower energy events are more abun-

[1] For example, the Auger event rate for CR above the GZK knee, $\sim$ 56 EeV, is on the order of two per month[36].

dant, their sources are likely to be further away, where the universe is isotropic. Furthermore, galactic/intergalactic magnetic fields are likely to wash out any correlation with the sources of lower energy events (neglecting the possible effects of magnetic field caustics[78, 79]). Thus, as one decreases the minimum observed energy one expects to include events which dilute any high energy anisotropy signal. Furthermore, there is typically significant error in the value of an observed energy (as much as 25%[36]). We therefore pay careful attention to the performance of the methods under variation in the total number of events and dilution factor (signal to isotropic background) for different types of signals in §8.4.1.

8.3 Methods

The 2-Pt (§8.3.2) and S-S (§8.3.3) methods are compared using the analysis paradigm described in §8.3.1 When needed for a concrete example, we use the largest currently operating observatory (Auger) for representative data set sizes and sky exposure[8, 9]. The methods presented here, however, can be applied to a spherical data set of any size and with an arbitrary experimental exposure.

8.3.1 Analysis Paradigm

We use a similar analysis paradigm for both the 2-Pt and S-S methods to calculate a p-value for rejecting the isotropic (null) hypothesis, H_{iso}. Each method uses binned *parameters* to compute a pseudo-log-likelihood test statistic Σ_P, "pseudo-" because the bins are correlated. The correlation does not effect the final answer because the p-value is derived by comparing the distribution of the Σ_P in a test sky to that of identically analyzed isotropic skies. The flatness of the distribution of p-values for isotropic test skies has been verified. The parameter space for the 2-Pt method is

the angular distance between two events. For the S-S method the parameter space is two dimensional. In neither case is the parameter space scanned to determine an optimal value. Instead, we compare the entire observed distribution to that expected by isotropy.

For a given set of cosmic ray events (referred to here as a *sky*) we compute Σ_P by comparing the binned distribution of the test sky's parameter(s) to the parameter(s) distribution expected from an isotropic sky. The probability for observing n_{obs} doublets (2-Pt) or triplets (S-S) from the test sky in the i^{th} parameter bin, given that you expect[80] to see n_{exp} from an isotropic sky, is approximated[81] by a Poisson distribution $P_i(n_{obs}|n_{exp}) = {n_{exp}}^{n_{obs}} e^{-n_{exp}}/n_{exp}!$. The pseudo-log-likelihood is $\Sigma_P = \sum_i \ln P_i(n_{obs}|n_{exp})$, where the sum is carried out over the bins of the parameter space. The ratio of the number of isotropic skies with Σ_P less than that of the test sky to the total number of simulated isotropic skies gives the p-value for the test sky.

In the following discussion $\vec{r}_k$ is defined as the arrival direction of the k^{th} event in a sky. This (unit) vector has Cartesian coordinates $\{r_x, r_y, r_z\}$ when projected from the galactic sphere.

8.3.2 Two-Point Correlation

The 2-Pt correlation distribution is calculated by computing the number of event pairs in a test sky as a function of the angular distance between any two events, $\theta = \cos^{-1}(\vec{r}_j \cdot \vec{r}_k)$ (see [75, 76, 82] for similar methods). We use 5° bins for $\theta \in [0^\circ, 180^\circ)$, so that the pseudo-log-likelihood is the sum over all angular scales, $\Sigma_P = \sum_\theta \ln P_\theta(n_{obs}|n_{exp})$.

8.3.3 Shape-Strength

This method involves an eigenvector decomposition, or principle component analysis, of the arrival directions of all sets of triplets found in the data set. It is inspired primarily by Fisher [83] (see also [84, 85]) but differs in that we decompose all subsets of triplets in a sky to obtain a test statistic.

For each triplet we calculate the components of the symmetric (3×3) orientation matrix $\mathbf{T}$[83]. In Cartesian coordinates, $T_{ij} = \frac{1}{3} \sum_{k \in triplet} (r_i r_j)_k$ for $i, j \in \{x, y, z\}$. The largest eigenvalue of $\mathbf{T}$, τ_1, results from a rotation of the triplet about the *principle* axis $\vec{u}_1$. The middle and smallest eigenvalues correspond to the *major* $\vec{u}_2$ and *minor* $\vec{u}_3$ axis respectively. The left panel of Figure 8.1 shows a graphical illustration of these eigenvectors. The eigenvalues satisfy $\tau_1 + \tau_2 + \tau_3 = 1$ and $\tau_1 \geq \tau_2 \geq \tau_3 \geq 0$, and thus there are only two independently varying parameters for any triplet.

It is convenient and statistically descriptive to work with a *shape*, γ, and a *strength*, ζ, parameter[83];

$$\gamma = \lg \left\{ \frac{\lg(\tau_1/\tau_2)}{\lg(\tau_2/\tau_3)} \right\} \tag{8.1}$$

$$\zeta = \lg(\tau_1/\tau_3) \tag{8.2}$$

As ζ increases from 0 to ∞ the events in the triplet become more concentrated. Generally, as γ increases from $-\infty$ to $+\infty$ the shape of the triplet transforms from elliptical, i.e. strings, to symmetric about $\vec{u}_1$, i.e. point source. See the right panel of Figure 8.1 for a schematic representation. In Figure 8.2 we show the how the variation of the ellipticity of a source on the galactic sphere effects the shape-strength parameter space.

To compute the test statistic Σ_P using this method we sum over sixty bins for $\gamma \in [-3.0, 3.0)$ and seventy-five bins for $\zeta \in [0.0, 15.0)$, i.e. $\Sigma_P = \sum_{\gamma\zeta} \ln P_{\gamma\zeta}(n_{obs}|n_{exp})$. We have checked that this parameter range is sufficient to cover event sets like those expected by Auger and that little is gained by enlarging the range.

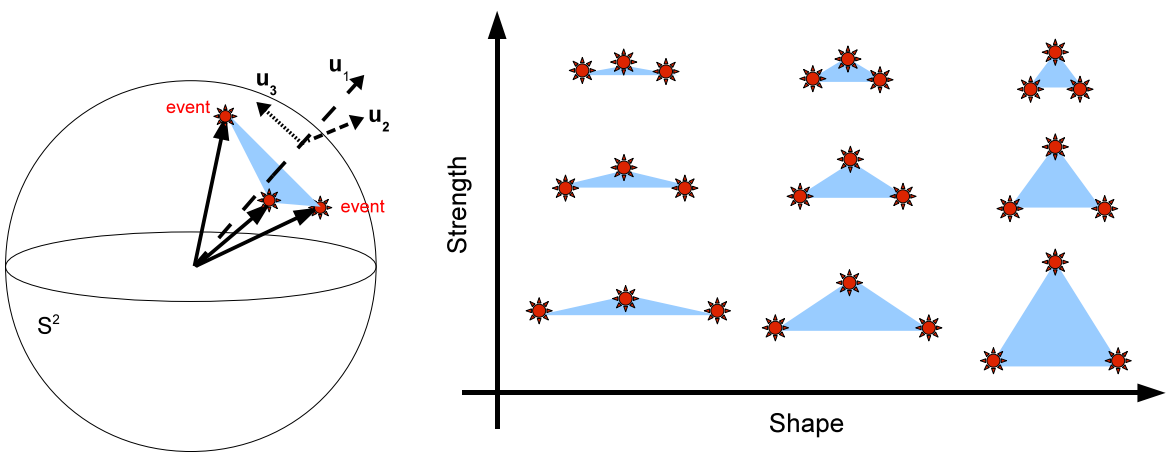


Figure 8.1: *Left:* The eigenvectors of a triplet of events on the sphere (S^2) are the principle axis $\vec{u}_1$, the major axis $\vec{u}_2$ (pointing into the page) and the minor axis $\vec{u}_3$. The eigenvalues of these vectors are used to compute this triplet's shape and strength. *Right:* An intuitive interpretation of the shape and strength parameters. As the strength parameter ζ increases from 0 to ∞, the events become more concentrated. As the shape parameter γ increases form $-\infty$ to $+\infty$, the events become more rotationally symmetric or less elongated.

8.4 Results

In order to gain confidence in, and intuition about, the S-S method we apply it (and the 2-Pt correlation) to an astro-physically motivated simulated (mock) data set in §8.4.1. The results of applying the S-S and 2-Pt methods to the 27 most energetic events from Auger[8, 9] are reported in §8.4.2.

8.4.1 Mock Signals

In weighing the effectiveness of a method for rejecting the isotropy hypothesis H_{iso} for a given CR sky we are interested in the probabilities for two types of testing errors[49]. A type I error is the probability α (significance or p-value) of rejecting H_{iso} given that H_{iso} is *true*; in practice it should be chosen a priori. In this analysis we choose the 1% significance level. One percent is arbitrary and is chosen to be the same as the value used in [8, 9]. One could choose, for example, 0.1% but this would require more data and/or a higher fraction of anisotropic signal to be detected. For each method this choice corresponds to a unique ${\Sigma_P}^{\alpha}$; we determine ${\Sigma_P}^{\alpha}$ such that the ratio of the number of isotropic skies with Σ_P less than ${\Sigma_P}^{\alpha}$ is $\alpha = 1\%$. We use the 10^4 isotropic skies to determine the upper bound of the signal region of likelihood space.

A type II error is the probability, β, of accepting H_{iso} (i.e. of rejecting the mock, or toy, signal hypothesis H_{sig}) given that H_{iso} is *false*. This value is dependent on the choice of H_{sig}. As we are interested in the effectiveness of accepting the signal hypothesis, we use the quantity $1 - \beta$, called the *power* of the method[49]. By applying the ensemble of each mock signal to each method we estimate the power as the ratio of mock signal skies with $\Sigma_P < {\Sigma_P}^{\alpha}$ to the total number of mock skies. As a heuristic measure we will describe a method's power as "good" if it is at least 90%, i.e. a high probability ($1 - \beta > 0.90$) to observe an anisotropy when there is an anisotropy in the data, and questionable if it is less than 90%.

Of significant physical interest is the ability of these catalog *independent* methods to detect signals generated from a catalog *dependent* map. To this end, we have studied simulated data sets generated from subsets of the VCV[69] galaxy catalog. We consider only galaxies with reshift $z \leq 0.02$ and we weight each galaxy either by a $1/z^2$ acceptance factor or not at all. We simulate events arriving from these galaxies

with a random component given by a two dimensional Gaussian centered on the galaxy and with deviation $\sigma = 3^\circ$. These choices are arbitrary in the sense that they describe some subset of nearby AGN with events smeared by a few times the angular resolution of Auger[9]. It should also be noted that the redshift weighted map (see Figure 8.3) is highly anisotropic, consisting of a number of small to medium scale clumps on the celestial sphere, and is likely to yield multiple events per sky within these groups. In contrast, the unweighted VCV map (see Figure 8.4) is notably more dispersed on the sphere.

The true CR data is likely to contain a *mixture* of background events and signal events. To explicitly study this dilution effect has on the power we separately construct mock ensembles in which each sky has a certain ratio, r, of signal events to the total number of events, with $r = 0.2, 0.4, \ldots, 1.0$. Notice that, because our methods use all the triplets or doublets in a given sky, the mixture Σ_P distributions are not a simple sum of the signal and isotropic Σ_P distributions.

Detection power is also strongly effected by the number of (high energy) events in a sky. The effect can be similar to those of signal dilution in that the power is decreased. We generate ensembles of 10^4 skies with 27, 40, 60, and 80 events per sky from the VCV catalog. Results for all combinations of source *purity* and number of (mock) data events are reported in Figures 8.3 and 8.4. The dark blue regions in lower plots of Figure 8.3 show that at least $40 - 80$ events with $(60 - 40)\%$ signal is required to achieve a high detection power, $1 - \beta \sim 90\%$, for the redshift weighted VCV maps. The un-weighted VCV maps in Figure 8.4 require a nearly pure signal and 60 or more events to have high detection power.

In general, where one method is good (power, $1 - \beta \gtrsim 90\%$) so is another; the methods are correlated. However, the S-S method consistently performs better than the two point correlation for the types of signals discussed here.

8.4.2 Auger Data

It is of interest to apply these techniques to experimental data. The largest public ultra high energy data set is the 27 events that form the basis of the Auger result reporting evidence for anisotropy (at the 1% significance level) in the highest energy CRs[8, 9]. The p-values obtained are: $p \sim 3\%$ for the 2-point method[2] and $p \sim 0.2\%$ for the S-S method. Thus, of these two methods only the S-S method would pass a requirement of $p < 1\%$ as evidence of anisotropy. Note that these events are known to be anisotropic – by the methods described in [8, 9, 76] – and therefore the p-values reported here reflect only on the *statistical methods* described in this paper.

8.5 Conclusion

In this paper we have introduce a shape-strength method for testing isotropy on the unit sphere. We have shown that this method uses pattern-descriptive parameters and can consistently out-perform the two-point correlation method. By simulating artificial and astrophysically motivated signals of various sizes and purity we can gauge how this method might perform on real data. The S-S method out performs the two-point method for all of our parameter choices.

The S-S method was found to detect the redshift weighted VCV toy signal (having significant small scale anisotropies) with at least $\sim 50\%$ signal purity and about 60 events in $> 90\%$ of the Monte Carlo simulations. We also wish to emphasize from the analysis of the diluted mock signals that when the signal to all ratio $r \gtrsim 50\%$ we can expect that a redshift weighted "VCV-like" CR signal should be identified with power $\gtrsim 50\%$ by both methods for data sets with $\gtrsim 40$ events. The unweighted

[2]We note that the 2-point method used here differs from the auto-correlation analysis performed on the 27 events in [76].

VCV toy signal (which is more diffuse on the sphere) is only reliably detected with greater than 80 events and 80% signal purity.

In agreement with qualitative expectations, this analysis demonstrates quantitatively how *both* signal purity and the total number of events dramatically effect the signal detection power. Furthermore, while sources types with significant small scale anisotropy can be detected with modest signal purity and total number of events, analysis of more subtle anisotropies suggest that either high purity signals or, more likely, much more data are needed for reliable identification.

8.6 Acknowledgements

We wish to thank members of the Auger collaboration for generous feedback and review of this paper. We also wish to thank Tim Thomas and the University of New Mexico Center for High Performance Computing[86] for their generous donations of CPU processing power. This work is supported by DOE grant DE-FR02-04ER41300.

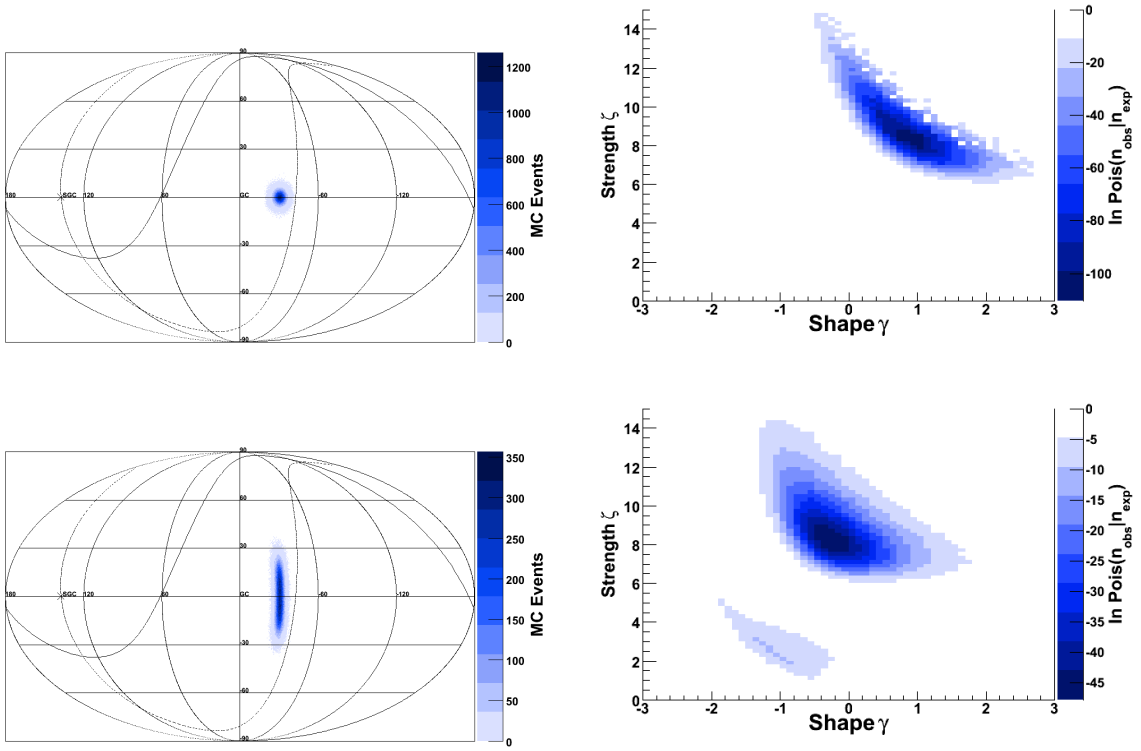


Figure 8.2: *Left column:* Histogram of 10^4 skies of 27 Monte-Carlo cosmic rays simulated from a single source centered on $\{l, b\} = \{-30.0, 0.0\}$ in galactic coordinates. We use the Fisher-Bingham distribution[87] on the sphere with $\kappa = 400.0$ to generate these events. For the spherically symmetric point-like (top) distribution we use $\beta = 0.0$. For the elliptically shaped (bottom) distribution we use $\beta = 200.0$ with the major axis pointing perpendicular to the galactic plane. See [87] for a detailed description of the parameters κ and β. *Right column:* The ensemble average (over all 10^4 sets of 27 event skies) of the $\ln P(n_{obs}|n_{exp})$ parameter space of the S-S method for the point-like (top) and elliptically shaped (bottom) toy anisotropies. In the bottom (right) panel one can see the relatively small deficit of triplets generated from the source with $\gamma \sim 1$ and $\zeta \sim 2$ in addition to the large excess of triplets with $\gamma \sim 0$ and $\zeta \sim 8$. The deficit arises from the non-uniform isotropic exposure of Auger[8, 9] and the excess from the simulated source. Both features contribute to the pseudo-likelihood where no distinction is made between excess and deficit of triplets. These two signals can be consistently detected with both the 2-Pt and S-S methods.

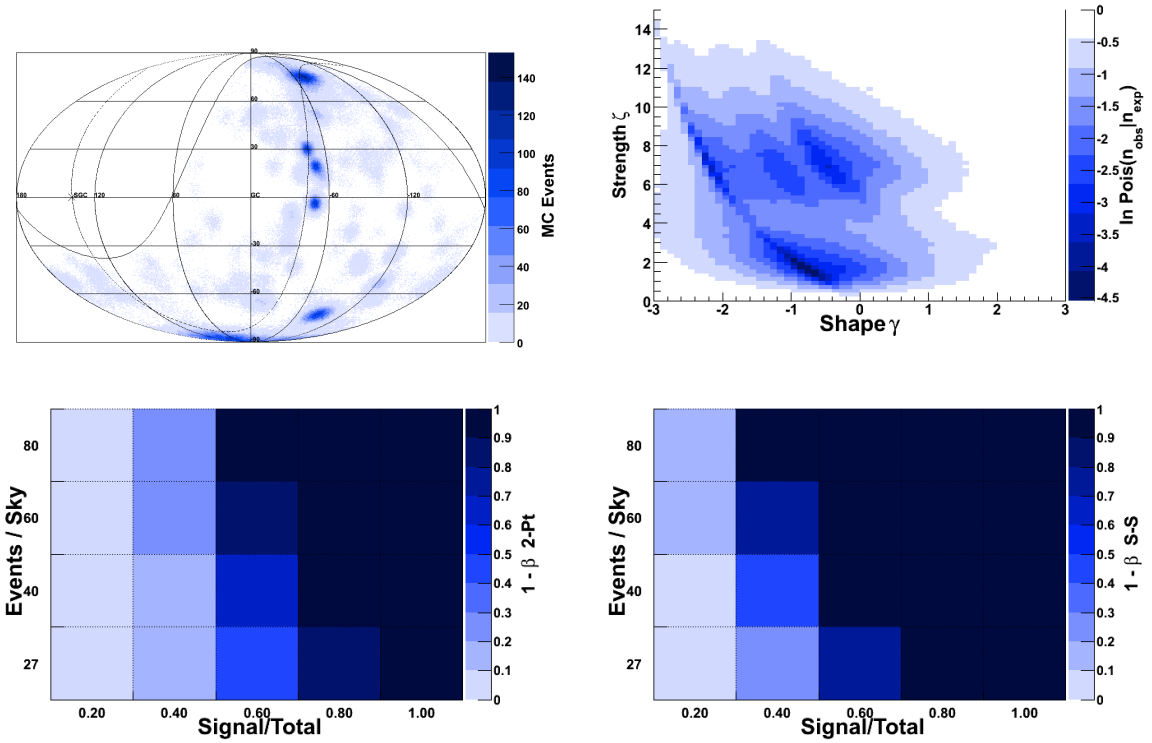


Figure 8.3: *Top left:* Histogram of 10^4 skies of 27 Monte-Carlo cosmic rays simulated from the VCV[69] catalog. We select objects with redshift $z_{max} \leq 0.020$ and they are weighted by $1/z^2$. Each simulated CR is drawn from a collection of 2D-Gaussian probability distributions centered on the the catalog sources, with deviation $\sigma = 3°$. (See §8.4.1.) *Top right:* Using the VCV catalog we plot the ensemble average of the $\ln P(n_{obs}|n_{exp})$ parameter space of the S-S method. *Bottom row:* Using the VCV ensemble files we can study the detection power $1-\beta$ as a function of the number of events per sky and fraction of each sky containing signal events using both the 2-Pt (*left*) and the S-S (*right*) methods.

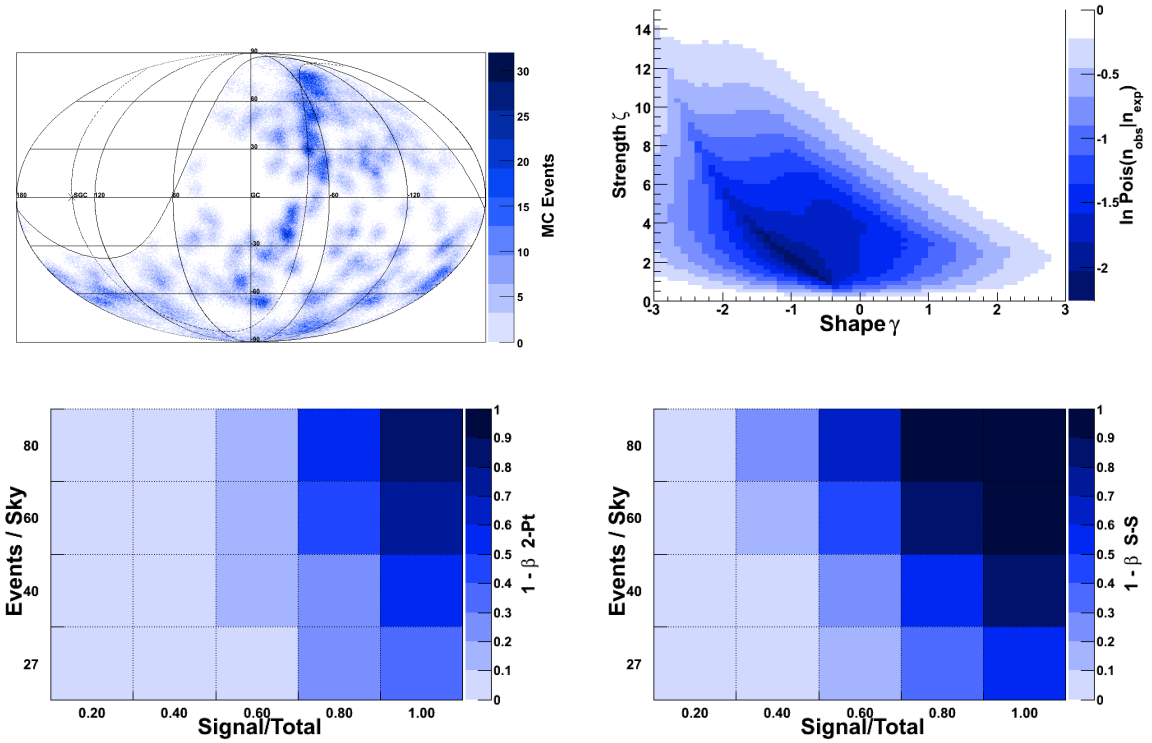


Figure 8.4: *Top left:* Histogram of 10^4 skies of 27 Monte-Carlo cosmic rays simulated from the VCV[69] catalog. We select objects with redshift $z_{max} \leq 0.020$ and they are not weighted. Each simulated CR is drawn from a collection of 2D-Gaussian probability distributions centered on the the catalog sources, with deviation $\sigma = 3°$. (See §8.4.1.) *Top right:* Using the VCV catalog we plot the ensemble average of the $\ln P(n_{obs}|n_{exp})$ parameter space of the S-S method. *Bottom row:* Using the VCV ensemble files we can study the detection power $1 - \beta$ as a function of the number of events per sky and fraction of each sky containing signal events using both the 2-Pt (*left*) and the S-S (*right*) methods.

Chapter 9

Summary

This thesis presents an exploration of two issues with fundamental importance to the study of high energy CR: the energy spectrum, and anisotropy in the CR sky. As an introduction to these topics, the theoretical context for the observation of the energy spectrum and arrival directions is discussed in Chapter 2. An outline of the particle physics phenomenology of air showers created by CR entering the atmosphere of Earth is given in §3.1, while the remaining sections of Chapter 3 are devoted to the reconstruction of the energy and arrival direction of the CR primary from data collected at the Pierre Auger Observatory.

In Chapter 4 two separate statistical tests are applied to both AGASA and preliminary Auger Cosmic Ray Energy spectra in an attempt to find deviation from a pure power-law. The first test is constructed from the probability distribution for the maximum event of a sample drawn from a power-law. The second employs the TP-statistic, a function defined to deviate from zero when the sample deviates from the power-law form, regardless of the value of the power index. The AGASA data show no significant deviation from a power-law when subjected to both tests. Applying these tests to the Auger spectrum suggests deviation from a power-law.

However, potentially large systematics on the relative energy scale prevent us from drawing definite conclusions at this time.

Un-binned statistical tools for analyzing the cosmic ray energy spectrum are developed and illustrated with a simulated data set in Chapter 5. These tools are designed to extract accurate and precise model parameter estimators in the presence of statistical and systematic energy errors. Two robust methods are used to test for the presence of flux suppression at the highest energies: the Tail-Power statistic and a likelihood ratio test. Both tests supply evidence of flux suppression in the simulated data. The tools presented can be generalized for use on any astrophysical data set where the power-law assumption is relevant and can be used to aid observational design.

In Chapter 5 the spectrum in six different "angular" coordinates is studied; zenith, azimuth, time, declination, angular distance from the galactic plane, the super galactic planes and Centaurus A. These events are binned by angle into equal event number bins and the fluxes fit to single and double power-laws. With a concentration on the characterization of the consistency of the highest energy flux between the angular bins, anomalous high energy flux for the azimuth between -3 and 89 degrees and within 44 degrees of Centaurus A is found.

Chapter 7 relates a manuscript written just after the so-called "running prescription" was passed[61], the results of which the Auger collaboration relied upon in the publication, "the correlation of ultra high energy CR with Active Galactic Nuclei"[8]. The primary goal of this work is the estimation of the observed value of the correlation *little-p* in post-scan data. Although not widely used at the time of its initial release, this work has since become the analysis-basis for the results of some of the later Auger anisotropy work (see Chapter A).

The two-point angular correlation function is a traditional method to search for

deviations from expectations of isotropy. In Chapter 8, a statistically descriptive three-point method is developed and explored, with the intended application being the search for deviations from isotropy in the highest energy cosmic rays. (The background and documentation for this three-point method can be found in Chapter B) As part of this work, the sensitivity of the so-called "shape-strength" method is compared to that of a two-point method for a variety of Monte-Carlo simulated anisotropic signals. In addition, studies are performed with anisotropic source signals diluted by an isotropic background. Type I and II errors for rejecting the hypothesis of isotropic cosmic ray arrival directions are evaluated for four different event sample sizes: 27, 40, 60 and 80 events, consistent with near-term data expectations from the Pierre Auger Observatory. In all cases, the ability to reject the isotropic hypothesis improves with event size and with the fraction of anisotropic signal. But while data sets with ~ 40 events should be sufficient for reliable identification of anisotropy, in cases of rather extreme (highly anisotropic) data, much larger data sets are suggested for reliable identification of more subtle anisotropies. The shape-strength method consistently performs better than the two point method and can be easily adapted to an arbitrary experimental exposure on the celestial sphere.

Together, the analysis tools developed in this thesis afford a much more precise view of high energy CR, their characteristic energies, and their origins. Armed with these tools, the Auger collaboration–and the scientific community, in general–is now in a position to more fully understand the nature of these energetic particles from outer space.

Appendix A

Anisotropy : Correlation with AGN

A.1 Chapter Overview

The text in this chapter has been submitted to the proceeding of the 31st annual International Cosmic Ray Conference (ICRC) in Łódź, Poland under the title "Correlation of the Highest Energy Cosmic Rays with Nearby Extragalactic Objects in Pierre Auger Observatory Data"[88]. I am the "primary author" of this text (including figures and numerical values), but it the final language and content are the result of a massive, active and quite essential collaboration. Indeed, much of the actual language is a balance of many – often divergent – opinions. In producing this document (chapter) I learned as much about the collaborative endeavor as the scientific.

The complete text is included here as an augmentation to Chapter 7; to show how the estimation of little-p (simply p below) is incorporated into a the complete update if the AGN signal[8, 9]. In particular the work of Chapter 7 is used to create

the right panel of Figure A.1. The final text is reviewed in §A.7.

Abstract: We update the analysis of correlation between the arrival directions of the highest energy cosmic rays observed by the Pierre Auger Observatory and the positions of nearby active galaxies.

A.2 Introduction

Using data collected between 1 January, 2004 and 31 August, 2007, the Pierre Auger Observatory has reported [8, 9] evidence of anisotropy in the arrival directions of cosmic rays (CR) with energies exceeding $\sim$ 60 EeV (1 EeV is 10^{18} eV). The arrival directions were correlated with the positions of nearby objects from the 12th edition of the catalog of quasars and active galactic nuclei (AGN) by Véron-Cetty and Véron [69] (VCV catalog). This catalog is not an unbiased statistical sample, since it is neither homogeneous nor statistically complete. This is not an obstacle to demonstrating the existence of anisotropy if CR arrive preferentially close to the positions of nearby objects in this sample. The nature of the catalog, however, limits the ability of the correlation method to identify the actual sources of cosmic rays. The observed correlation identifies neither individual sources nor a specific class of astrophysical sites of origin. It provides clues to the extragalactic origin of the CR with the highest energies and suggests that the suppression of the flux (see [36] and [48]) is due to interaction with the cosmic background radiation.

In this article we update the analysis of correlation with AGN in the VCV catalog by including data collected through 31 March, 2009. We also analyse the distribution of arrival directions with respect to the location of the Centaurus cluster and the radio source Cen A. Alternative tests that may discriminate among different populations of source candidates are presented in a separate paper at this conference [89].

A.3 Data

The data set analyzed here consists of events observed by the Pierre Auger Observatory prior to 31 March, 2009. We consider events with zenith angles smaller than 60°. The event selection implemented in the present analysis requires that at least five active nearest-neighbors surround the station with the highest signal when the event was recorded, and that the reconstructed shower core be inside an active equilateral triangle of detectors. The integrated exposure for this event selection amounts to 17040 km^2 sr yr ($\pm 3\%$), nearly twice the exposure used in [8, 9].

In [8, 9] we published the list of 27 events with $E > 57$ EeV. Since then, the reconstruction algorithms and calibration procedures of the Pierre Auger Observatory have been updated. The lowest energy among these same 27 events is 55 EeV according to the latest reconstruction. Reconstructed values for the arrival directions of these events differ by less than 0.1° from their previous determination. There are now 31 additional events above the energy threshold of 55 EeV. The systematic uncertainty of the observed energy for events used here is $\sim 22\%$ and the energy resolution is $\sim 17\%$ [90, 91]. The angular resolution of the arrival directions for events with energy above this threshold is better than 0.9° [92].

A.4 Update of the correlation with AGN

To avoid the negative impact of trial factors in *a posteriori* analyses, the statistical significance of the anisotropy reported in [8, 9] was established through a test with independent data. The parameters of the test were chosen by an exploratory scan using events observed prior to 27 May, 2006. The scan searched for a correlation of CR with objects in the VCV catalog with redshift less than $z_{\max}$ at an angular scale $\psi_{\max}$ and energy threshold E_{th}. The scan was implemented to find a minimum of the

probability P that k or more out of a total of N events from an isotropic flux are correlated by chance with the selected objects at the chosen angular scale, given by

$$P = \sum_{j=k}^{N} \begin{pmatrix} N \\ j \end{pmatrix} {p_{\rm iso}}^{j} (1 - p_{\rm iso})^{N-j} \ . \tag{A.1}$$

We take $p_{\rm iso}$ to be the exposure-weighted fraction of the sky accessible to the Pierre Auger Observatory that is within $\psi_{\rm max}$ degrees of the selected potential sources. The minimum value of P was found for the parameters $\psi_{\rm max} = 3.1^\circ$, $z_{\rm max} = 0.018$ and $E_{\rm th} = 55$ EeV (in the present energy calibration). The probability that an individual event from an isotropic flux arrives within the fraction of the sky prescribed by these parameters by chance is $p_{\rm iso} = 0.21$.

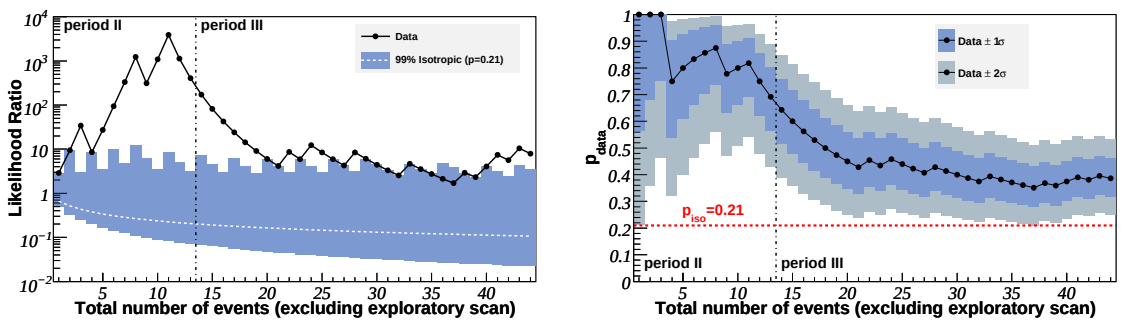


Figure A.1: Monitoring the correlation signal. *Left:* The sequential analysis of cosmic rays with energy greater than 55 EeV arriving after 27 May, 2006. The likelihood ratio $\log_{10} R$ (see Eq(A.2)) for the data is plotted in black circles. Events that arrive within $\psi_{\rm max} = 3.1^\circ$ of an AGN with maximum redshift $z_{\rm max} = 0.018$ result in an up-tick of this line. Values above the area shaded in blue have less than 1% chance probability to arise from an isotropic distribution ($p_{\rm iso} = 0.21$). *Right:* The most likely value of the binomial parameter $p_{\rm data} = k/N$ is plotted with black circles as a function of time. The 1σ and 2σ uncertainties in the observed value are shaded. The horizontal dashed line shows the isotropic value $p_{\rm iso} = 0.21$. The current estimate of the signal is 0.38 ± 0.07. In both plots events to the left of the dashed vertical line correspond to period II of Table A.1 and those to the right, collected after [8, 9], correspond to period III.

Table A.1: A numerical summary of results for events with $E \geq 55$ EeV. See the text for a description of the entries.

Period	Exposure	GP	N	k	$k_{\rm iso}$	P
I	4390	unmasked	14	9	2.9	
		masked	10	8	2.5	
II	4500	unmasked	13	9	2.7	2×10^{-4}
		masked	11	9	2.8	1×10^{-4}
III	8150	unmasked	31	8	6.5	0.33
		masked	24	8	6.0	0.22
II+III	12650	**unmasked**	**44**	**17**	**9.2**	$\mathbf{6 \times 10^{-3}}$
		masked	35	17	8.8	2×10^{-3}
I+II	8890	unmasked	27	18	5.7	
		masked	21	17	5.3	
I+II+III	17040	unmasked	58	26	12.2	
		masked	45	25	11.3	

Of the 27 events observed prior to 31 August, 2007, 13 were observed after the exploratory phase. Nine of these arrival directions were within the prescribed area of the sky, where 2.7 are expected on average if the flux were isotropic. This degree of correlation provided a 99% significance level for rejecting the hypothesis that the distribution of arrival directions is isotropic.

The left panel of Figure A.1 displays the likelihood ratio of correlation as a function of the total number of time-ordered events observed since 27 May, 2006, i.e. excluding the data used in the exploratory scan that lead to the choice of parameters. The likelihood ratio R is defined as (see [93] and [94])

$$R = \frac{\int_{p_{\rm iso}}^{1} p^k (1-p)^{N-k}\, dp}{{p_{\rm iso}}^k (1-p_{\rm iso})^{N-k+1}} \ . \tag{A.2}$$

This quantity is the ratio between the binomial probability of correlation – marginalized over its range of possible values and assuming a flat prior – and the binomial probability in the isotropic case ($p_{\rm iso} = 0.21$). A sequential test rejects the isotropic hypothesis at the 99% significance level (and with less than 5% chance of incorrectly accepting the null hypothesis) if $R > 95$. The likelihood ratio test indicated a 99%

significance level for the anisotropy of the arrival directions using the independent data reported in [8, 9]. Subsequent data neither strengthen the case for anisotropy, nor do they contradict the earlier result. The departure from isotropy remains at the 1% level as measured by the cumulative binomial probability ($P = 0.006$), with 17 out of 44 events in correlation.

In the right panel of Figure A.1 we plot the degree of correlation ($p_{\rm data}$) with objects in the VCV catalog as a function of the total number of time-ordered events observed since 27 May, 2006. For each new event the best estimate of $p_{\rm data}$ is k/N. The 1σ and 2σ uncertainties in this value are determined such that the area under the posterior distribution function is equal to 68% and 95%, respectively. The current estimate, with 17 out of 44 events that correlate in the independent data, is $p_{\rm data} = 0.38$, or more than two standard deviations from the value expected from a purely isotropic distribution of events. More data are needed to accurately constrain this parameter.

The correlations between events with $E \geq 55$ EeV and AGN in the VCV catalog during the pre- and post- exploratory periods of data collection are summarized in Table A.1. The left most column shows the period in which the data was collected. Period I is the exploratory period from 1 January, 2004 through 26 May, 2006. The data collected during this period was scanned to establish the parameters which maximize the correlation. Period II is from 27 May, 2006 through 31 August, 2007 and period III includes data collected after [8, 9], from 1 September, 2007 through 31 March, 2009. The numbers in bold correspond to period II+III and give the results for the post-exploratory data (see Figure A.1). The exposure for each period is listed in units of km^2 sr yr and has an uncertainty of 3%. If the region of the sky within 12° of the galactic plane (GP) is included in the analysis then the third column is marked "unmasked" (and $p_{\rm iso} = 0.21$), if not then it is marked "masked" (and $p_{\rm iso} = 0.25$). The average number of events from an isotropic flux expected to correlate is listed as

$k_{\rm iso} = Np_{\rm iso}$, where N is the total number of events observed during each period. k is the number of events that arrive within 3.1° of an AGN with a redshift of 0.018. The cumulative binomial probability (see Eq(A.1)) is shown in the right most column. We do not include this value for any row containing period I because this period was used to determine the correlation parameters for the rest of the table and cannot, therefore, be interpreted as a statistical significance.

Note that during period I+II (reported in [8, 9]), 18 out of 27 events arrive within 3.1° of an AGN in the VCV catalog with redshift less than 0.018. [1] There are 31 additional events (during period III) above the specified energy threshold, 8 of which have arrival directions within the prescribed area of the sky, not significantly more than the 6.5 events that are expected to arrive on average if the flux were isotropic.

While the degree of correlation with objects in the VCV catalog has decreased with the accumulation of new data, a re-scan of the complete data set shows that the values of $\psi_{\rm max}$, $z_{\rm max}$ and $E_{\rm th}$ that characterise the correlation have not changed appreciably from the values reported in [8, 9].

A.5 A posteriori analyses

In this section we further analyze the complete set of 58 events with energy larger than 55 EeV collected before 31 March, 2009.

To complement the information given in Table A.1 over different angular scales, we plot in Figure A.2 the distribution of angular separations between the arrival directions of the 58 events with $E > 55$ EeV and the position of the closest object in the VCV catalog within redshift $z_{\rm max} \leq 0.018$. The cumulative distribution is plotted

[1] Two additional events correlate within a slightly larger angular distance, as reported in [8, 9]. Here we restrict the analysis to the parameters chosen to monitor the correlation signal.

in the left panel and the differential distribution is plotted in the right. The average distribution expected for 58 events drawn from an isotropic flux is also shown. In the right panel the 13 events with galactic latitudes $|b| < 12^\circ$ have been shaded. Note that only 1 of these 13 events is within 3° of a selected AGN. Incompleteness of the VCV catalog due to obscuration by the Milky Way or larger magnetic bending of CR trajectories along the galactic disk are potential causes for smaller correlation of arrival directions at small galactic latitudes.

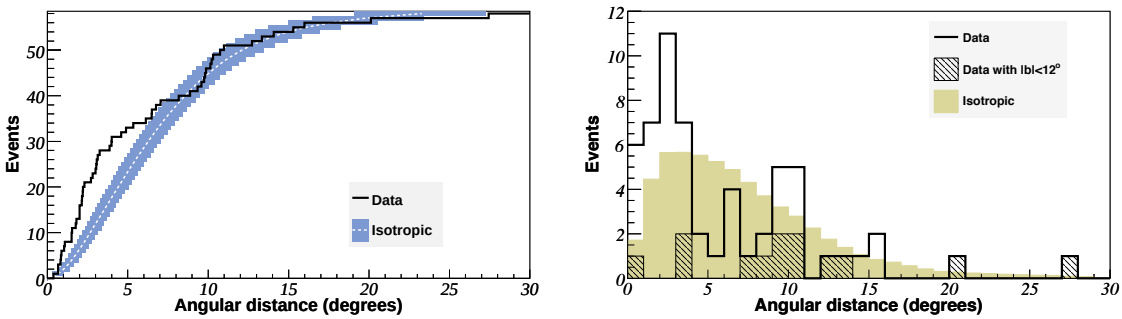


Figure A.2: The distribution of angular separations between the 58 events with $E > 55$ EeV and the closest AGN in the VCV catalog within 75 Mpc. *Left:* The cumulative number of events as a function of angular distance. The 68% the confidence intervals for the isotropic expectation is shaded blue. *Right:* The histogram of events as a function of angular distance. The 13 events with galactic latitudes $|b| < 12^\circ$ are shown with hatching. The average isotropic expectation is shaded brown.

An excess of events as compared to isotropic expectations is observed from a region of the sky close to the location of the radio source Cen A ($(l, b) = (-50.519.4$ [95]). In Figure A.3 we plot the distribution of events as a function of angular distance from Cen A. In a Kolmogorov-Smirnov [96] test 2% of isotropic realizations have maximum departure from the isotropic expectation greater than or equal to the maximum departure for the observed events. The excess of events in circular windows around Cen A with the smallest isotropic chance probability corresponds to a radius of 18°, which contains 12 events where 2.7 are expected on average if the

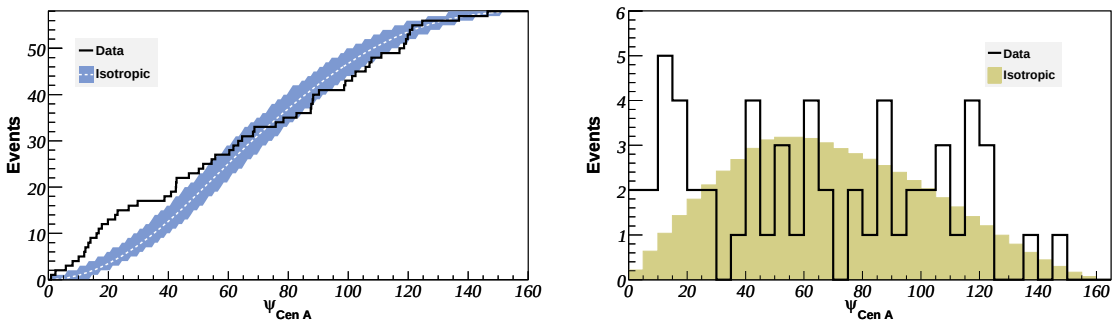


Figure A.3: *Left:* The cumulative number of events with $E \geq 55$ EeV as a function of angular distance from Cen A. The average isotropic expectation with approximate 68% confidence intervals is shaded blue. *Right:* The histogram of events as a function of angular distance from Cen A. The average isotropic expectation is shaded brown.

flux were isotropic. The (differential) histogram of angular distances from Cen A is in the right panel of Figure A.3.

By contrast, the region around the Virgo cluster is densely populated with galaxies but does not have an excess of events above isotropic expectations. In particular, a circle of radius 20° centred at the location of M87 ($(l, b) = (76.274.5$ [95]) does not contain any of the 58 events with energy $E > 55$ EeV. This is a region of relatively low exposure for the Pierre Auger Observatory and only 1.2 event is expected on average with the current statistics if the flux were isotropic.

A.6 Discussion

With data collected by the Pierre Auger Observatory between 1 January, 2004 and 31 March, 2009, we have updated the analysis reported in [8, 9] of correlation between the arrival directions of the highest energy cosmic rays and the positions of nearby objects from the 12th edition of the VCV catalog of quasars and active galactic nuclei. The total number of events above 55 EeV is 58. A subset of 44 events are

independent of those used to determine the parameters ($\psi_{\rm max} = 3.1^\circ$, $z_{\rm max} = 0.018$ and $E_{\rm th} = 55$ EeV) with which we monitor the correlation signal (see Table A.1 for more details). 17 of these 44 events correlate under these parameters. This correlation has a less than 1% probability to occur by chance if the arrival directions are isotropically distributed. The evidence for anisotropy has not strengthened since the analysis reported in [8, 9]. The degree of correlation with objects in the VCV catalog appears to be weaker than suggested by the earliest data.

We note that there is an excess of events in the present data set close to the direction of the radio source Cen A, a region dense in potential sources. This excess is based on *a posteriori* data but suggests that the region of the sky near Cen A warrants further study.

Additional data are needed to make further progress in the quest to identify the sites of ultra high energy CR origin. Alternative tests that may discriminate among different populations of source candidates are presented in a separate paper at this conference [89].

A.7 Chapter Post-Script

The biggest (psychological) obstacle for this paper is coping with the fact the the signal strength has decreased (drastically) with the accumulation of new data. Of course, the new data present a bit of a set back for the most popular current interpretation of the sources of CR – that they originate from *VCV* AGN – and this accounts, I think, for the "defensive tone" of the language.

While the picture of a disappearing signal is clear (see Figure A.1), it is worth noting that the presentation strives to emphasize the positive [2] aspects. For exam-

[2]Data is neutral; we are *always* biased.

ple, Figure A.3 and Figure A.2 use **all 58** events; the events used to establish the correlation parameters are **included**! By including the original 27 (pre-scan) events Figure A.2 and any numbers derived from it – derived either numerically or simply by looking at and interpreting the plot – are *biased*. This is because nearly half of the CR in the 58 were scanned against the AGN for the most anomalous correlation.

By definition, Cen A (as in Figure A.3, or Virgo) is a particular source. From an observational perspective one might as well choose *any* point in the sky. In this case, using all 58 events only includes a bias on the energy threshold $E \sim 55$ EeV which can be loosely justified by other means, i.e. the GZK-effect. Indeed, in preparation for this article many other "points" on the sky were studied by the author, including; the supergalactic plane, the galactic plane, the equitorial plane, Fornax A, the galactic center etc. They were all rejected for publication because they produce null results.

Appendix B

Anisotropy : Some Catalog Independent Methods

B.1 Chapter Overview

The material presented in this chapter was original published in a series of *GAP-Notes* in 2008 [97, 98, 99, 100, 101]. It is included here as documentation and background for the paper presented in Chapter 8, since *GAP-Notes* are typically not public.

B.2 Study of Anisotropy Metrics

B.2.1 Section Overview

This section was originally published as [97].

In this note we test four metrics of anisotropy in the arrival direction of a set of cosmic ray (CR) events. These metrics are compared and tested for sensitivity to selected anisotropic distributions as defined by the anisotropy working group. [102].

The metrics studied here are:

1. "2pt"– standard two-point correlation function;
2. "Rayleigh" – modified two-point Rayleigh metric as introduced by reference [103] but analyzed differently as described below;
3. "R^2"– simple three-point directional vector sum (a measure of triangular area); and
4. "SS" – a three-point metric derived from a principle component analysis, described below.

These are described in more detail in section B.2.3. We note that the two-point metrics have been studied previously.[104]

B.2.2 Methods

For each metric we determine the significance (p-value) for rejecting the hypothesis of isotropy in test samples generated by the anisotropy working group [102]. Each of these test sets consists of 1000 sub-sets of sixty events each ("skies"). These sets are compared to a reference set of "isotropic" (convoluted with the PAO coverage map) skies (10,000 skies also of sixty events each).

1. ISOTROPY
2. DIPOLEz

3. IRAS0020-ISO

4. VCV0020

5. VCVflat0020

For each metric we follow the same procedure.

1. Compare the (binned) distribution of the metric between the test and isotropic reference skies.

2. For the i^{th} bin, calculate the probability of observing n_{obs} event sets from a Poisson distribution with mean given by the number of event sets expected from an isotropic sky, n_{exp};

$$P_i(n_{obs}|n_{exp}) = \frac{n_{exp}^{n_{obs}} e^{-n_{exp}}}{n_{exp}!}. \tag{B.1}$$

3. The sum over all bins is taken as a pseudo-log-likelihood[1] for both the data set and all of the isotropic skies;

$$\Sigma_P = \sum_{i=1}^{N_{\text{bins}}} \ln P_i(n_{obs}|n_{exp}). \tag{B.2}$$

4. The ratio of the number of isotropic skies with Σ_P less than that of the data set to the total number of simulated isotropic data sets gives the p-value.

[1] *Pseudo* since the bins are correlated. The correlation does not effect the final answer because the p-valueis derived by comparing the distribution of the Σ_P in a test sky to that of identically analyzed isotropic skies. The flatness of the distribution of p-values for isotropic test skies has been verified.

B.2.3 Metrics

2-pt

The 2-pt correlation distribution is calculated by computing the number of event pairs in a data set as a function of the angular distance between the two events. We use 5^{o} angular bins.

Rayleigh

This method is based on the procedure introduced in [103]. Here we have used the same angular-scale (angular binning) as for the 2-pt correlation. We have then calculated the significance according to the same procedure as the other statistics (as described above.)

The following two methods involve *triplets* of events, i.e. sets of three, as their basic building blocks. A primary resource for the motivation behind the construction of these statistics can be found in [83]. In the following discussion let (x_i, y_i, z_i), with $i = 1, 2, 3$, be the Cartesian coordinates of a set of three arrival directions

R^2

For each set of triplets in the data set we calculate the square of the vector sum of the three arrival directions $\vec{r_i}$ using their Cartesian coordinates,

$$R^2 = \sum_{i=1}^{3} \left(x_i^2 + y_i^2 + z_i^2 \right). \tag{B.3}$$

The metric R^2 can take any value on the interval $[0, 9]$ and we choose to bin its distribution in steps of 0.2.

SS

The metric for this method is derived from the principle component analysis of directional vectors taken in triplets. We calculate the orientation matrix to be[83]

$$\mathbf{T} = \frac{1}{3}\begin{pmatrix} \sum x_i^2 & \sum x_i y_i & \sum x_i z_i \\ \sum x_i y_i & \sum y_i^2 & \sum y_i z_i \\ \sum x_i z_i & \sum y_i z_i & \sum z_i^2 \end{pmatrix}. \tag{B.4}$$

We obtain the eigenvalues and eigenvectors of this matrix. [83]. The largest eigenvalue τ_1 results from a rotation of the triplet about the *principle* axis $\vec{u}_1$ (note the similarity to R^2 in §B.2.3). The middle and smallest eigenvalues correspond to the *major* $\vec{u}_2$ and *minor* $\vec{u}_3$ axis respectively.

The eigenvalues satisfy[2]

$$\tau_1 + \tau_2 + \tau_3 = 1,\ \tau_1 \geq \tau_2 \geq \tau_3 \geq 0. \tag{B.5}$$

Therefore there can be only two independently varying eigenvalues for any triplet. It is convenient to work with a "shape"

$$\gamma = \lg\left\{\frac{\lg(\tau_1/\tau_2)}{\lg(\tau_2/\tau_3)}\right\} \tag{B.6}$$

and a "strength"

$$\zeta = \lg(\tau_1/\tau_3) \tag{B.7}$$

parameter. As the strength parameter ζ increases from zero to infinity the the events under consideration become more concentrated. If the shape parameter $\gamma \ll 0$, then the events make a "banana" shape on the sky. As γ approaches zero the events become more aligned ("strings"). For $0 < \gamma < \infty$ the events become more circular ("point source" or "dipole"). To gain intuition about the scope of this method we summarize some of the properties of the eigenvalues in Table B.1 [83]. We make a

[2]Fisher[83] uses slightly different notation, i.e. $\tau_3 \equiv$ principle axis.

Relative magnitudes	Type	Features
$\tau_1 = \tau_2 = \tau_3 = 1$	uniform	
τ_1 large; τ_2, τ_3 small		
(i) $\tau_2 \neq \tau_3$	uni-modal if $\|R\| \sim 3$	concentration at one end of $\vec{u}_1$
	bimodal otherwise	concentration at both ends of $\vec{u}_1$
(ii) $\tau_2 \approx \tau_3$	uni-modal if $\|R\| \sim 3$	concentration at one end of $\vec{u}_1$
	rotational symmetry	about $\vec{u}_1$
τ_3 large; τ_2, τ_1 large		
(i) $\tau_2 \neq \tau_1$	girdle	girdle plan spanned by $\vec{u}_1$, $\vec{u}_2$
(ii) $\tau_2 \approx \tau_1$	symmetric girdle	rotational symmetry about $\vec{u}_1$

Table B.1: Some descriptions for the relative magnitudes of the eigenvalues of Eq(B.4). The values in this table are strictly true only for data sets without a coverage map. Reproduced from [83].

metric from the scatter plot of the γ versus ζ distribution. This is the shape-strength ("SS") metric.

B.2.4 Results and Conclusions

In Figs. B.21, B.22, B.23, B.24 we show the correlations between the methodsfor the various test sets. We note that we do not have sufficient statistics in the test sets to measure the probabilities for the VCV0020 set and therefore these are not shown.

The sensitivity of the various metrics is shown in table B.2. We note that we do not have sufficient statistics in the test sets to measure the probabilities for the VCV0020 set. This study suggests that the SS metric is the most sensative to anisotropy. We also can see that the IRAS0020-ISO is difficult to distinguish from isotropic with any of the four methods considered here.

B.2.5 Figures

test	2-pt	Rayleigh	R^2	SS (γ, ζ)
ISOTROPY	0.1010 ± 0.0095	0.1090 ± 0.0099	0.0920 ± 0.0091	0.0900 ± 0.0090
DIPOLEz	1.0000 ± 0.0000	0.9950 ± 0.0022	1.0000 ± 0.0000	0.9950 ± 0.0022
IRAS0020-ISO	0.1320 ± 0.0107	0.2730 ± 0.0141	0.1070 ± 0.0098	0.2470 ± 0.0136
VCV0020	1.0000 ± 0.0000	1.0000 ± 0.0000	0.9970 ± 0.0017	1.0000 ± 0.0000
VCVflat0020	0.9720 ± 0.0052	1.0000 ± 0.0000	0.7970 ± 0.0127	0.9990 ± 0.0010

Table B.2: The fraction of skies that are inconsistent with isotropy at the 10% level.

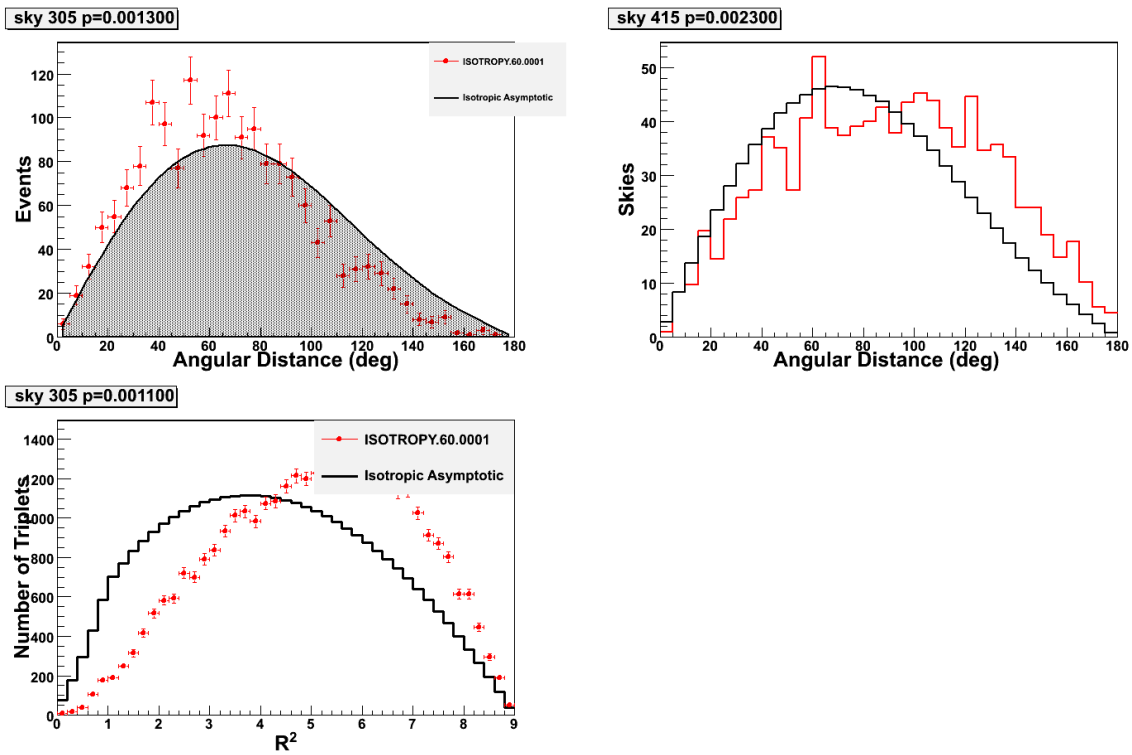


Figure B.1: Least isotropic sky compared to average isotropic metric.

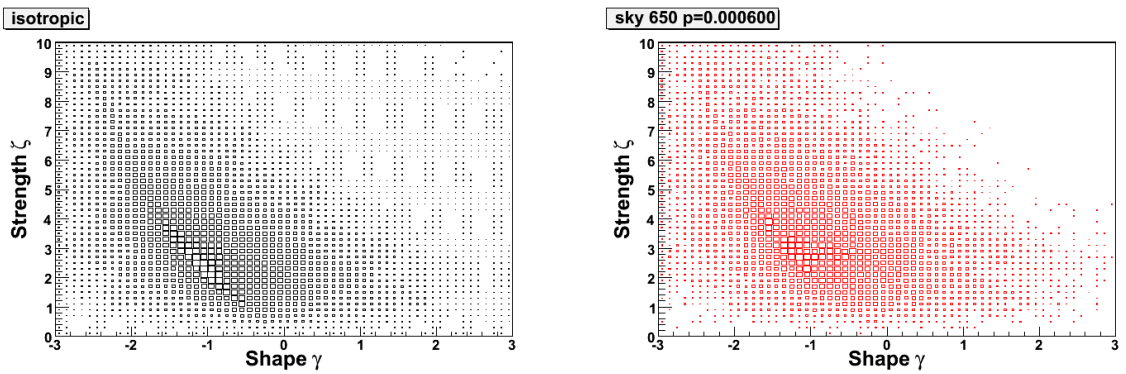


Figure B.2: Least isotropic sky (top) compared to average isotropic sky (bottom) for the SS metric.

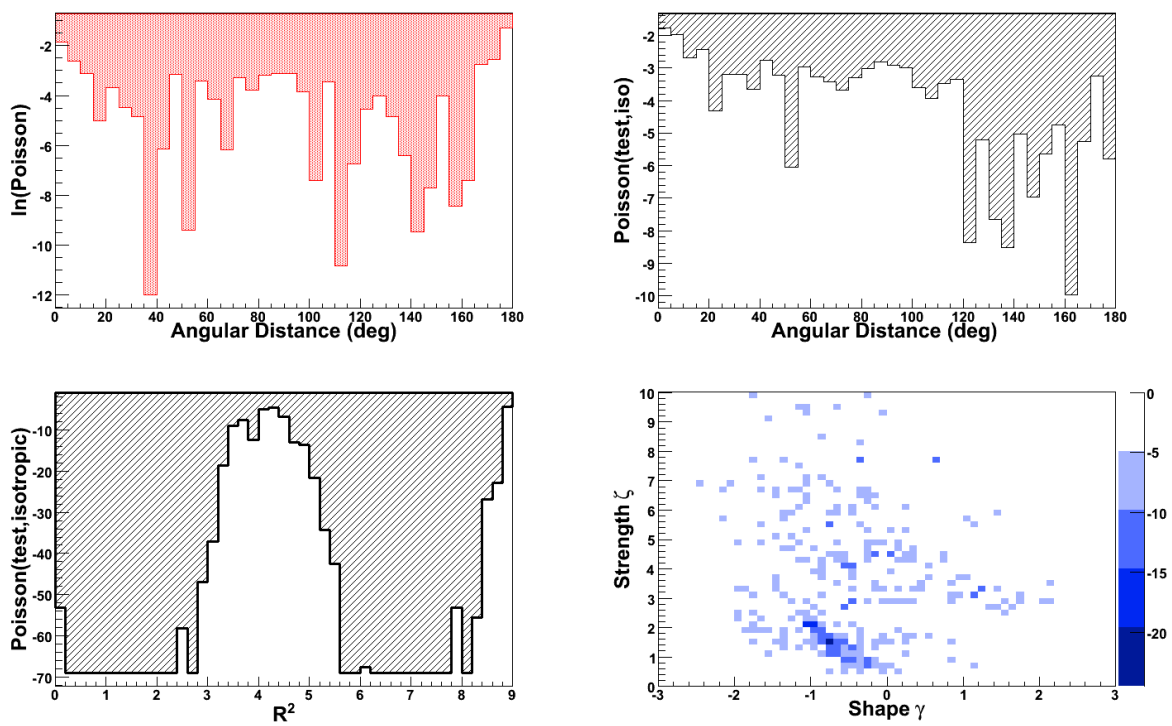


Figure B.3: Least isotropic sky Poissson values by bin.

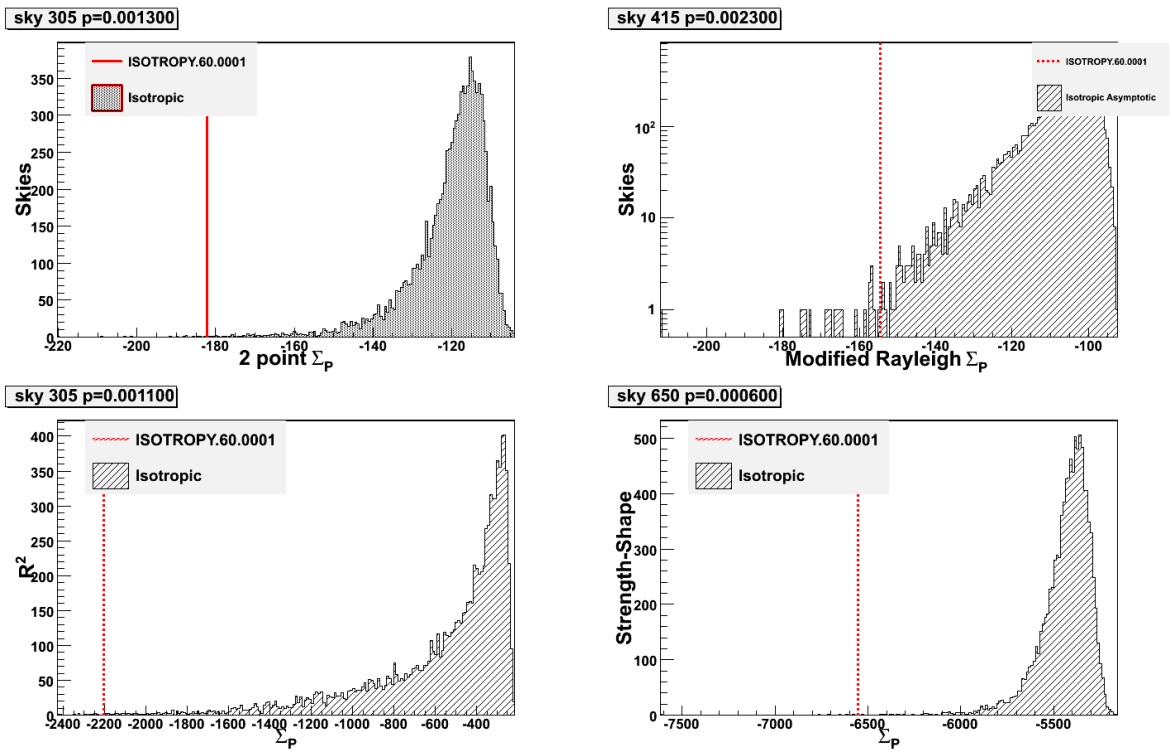


Figure B.4: Distribution of Σ_p for isotropic skies and for the least isotropic test sky (indicated by the red line). Each plot is the Σ_p frequency distribution (number of skies versus Σ_p. Please note that the bottom two plots are mis-labeled on the vertical axis. The plots (clockwise, starting from top right) are for the metrics: 2pt,Rayleigh, SS, R^2.

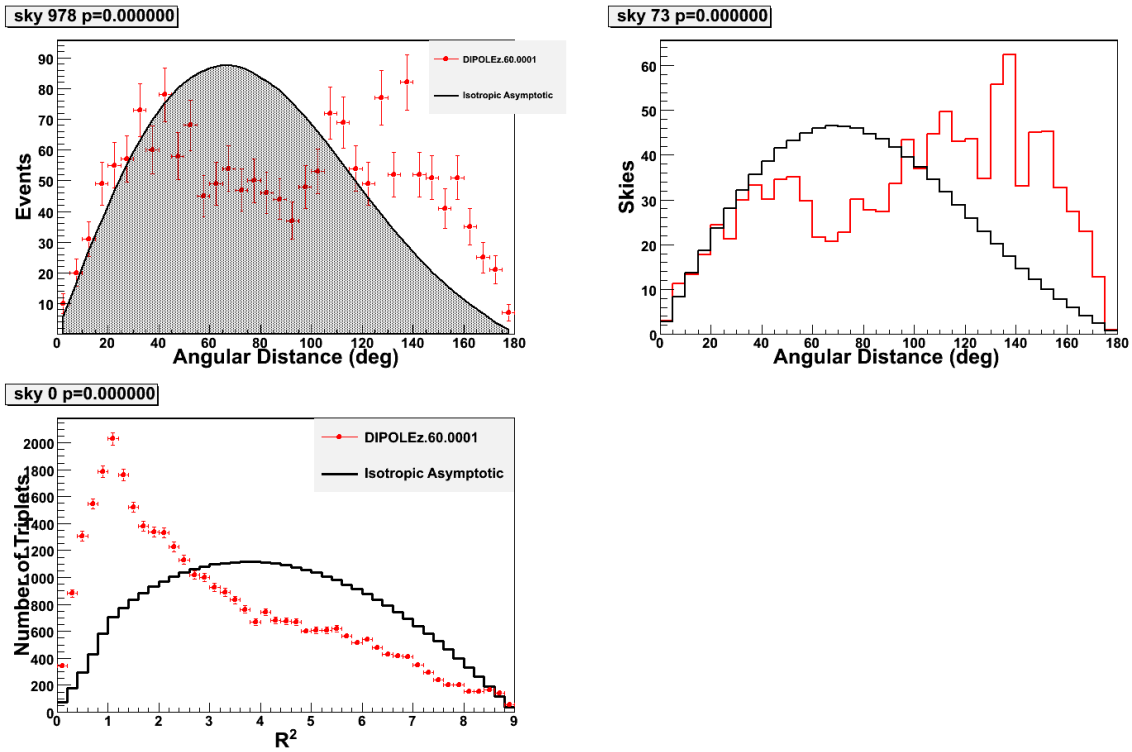


Figure B.5: Least isotropic sky compared to average isotropic metric.

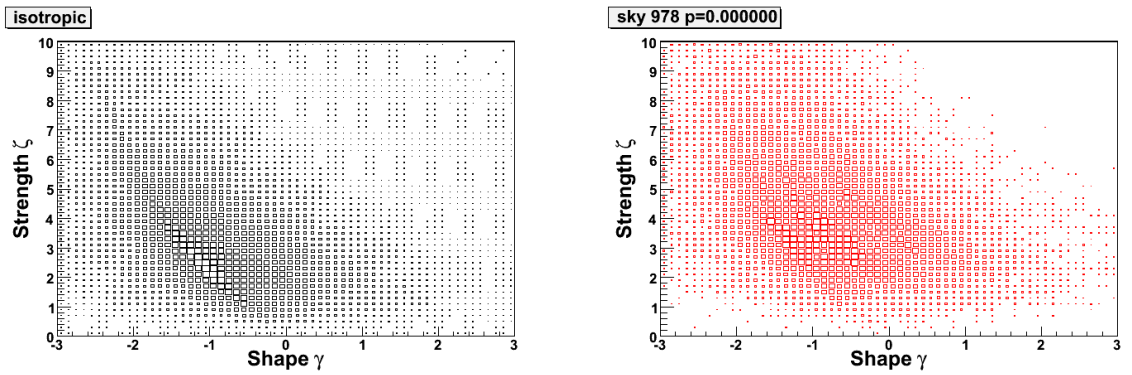


Figure B.6: Least isotropic sky (top) compared to average isotropic sky (bottom) for the SS metric.

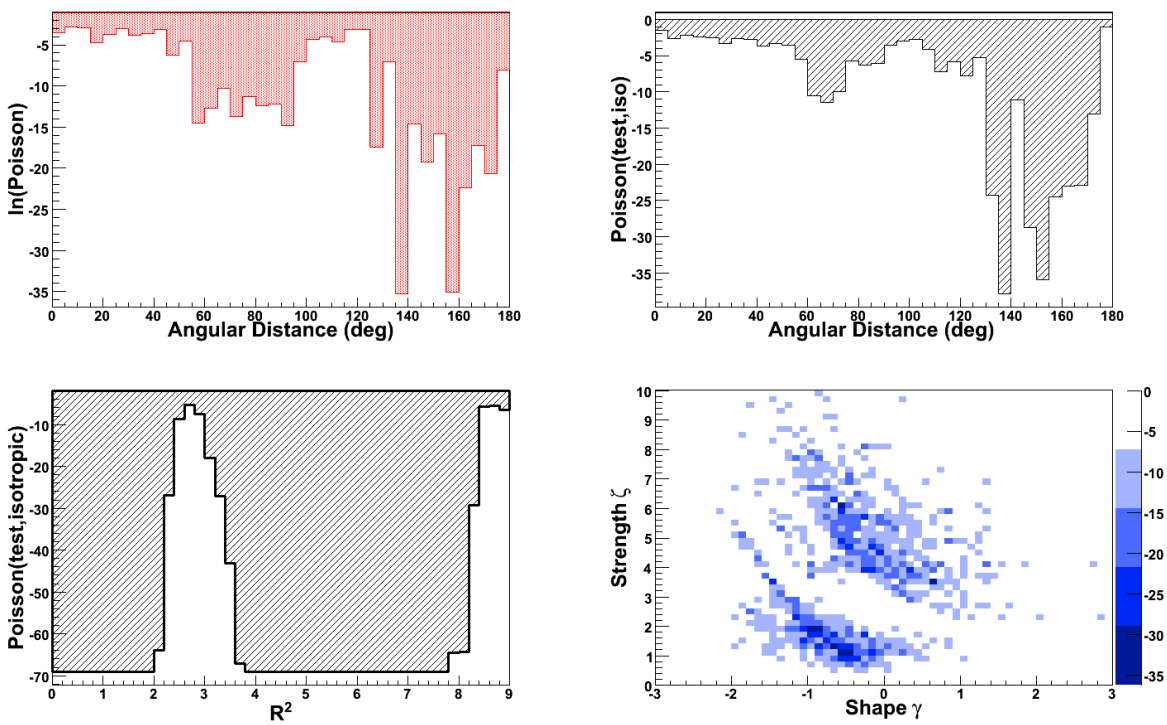


Figure B.7: Least isotropic sky Poissson values by bin.

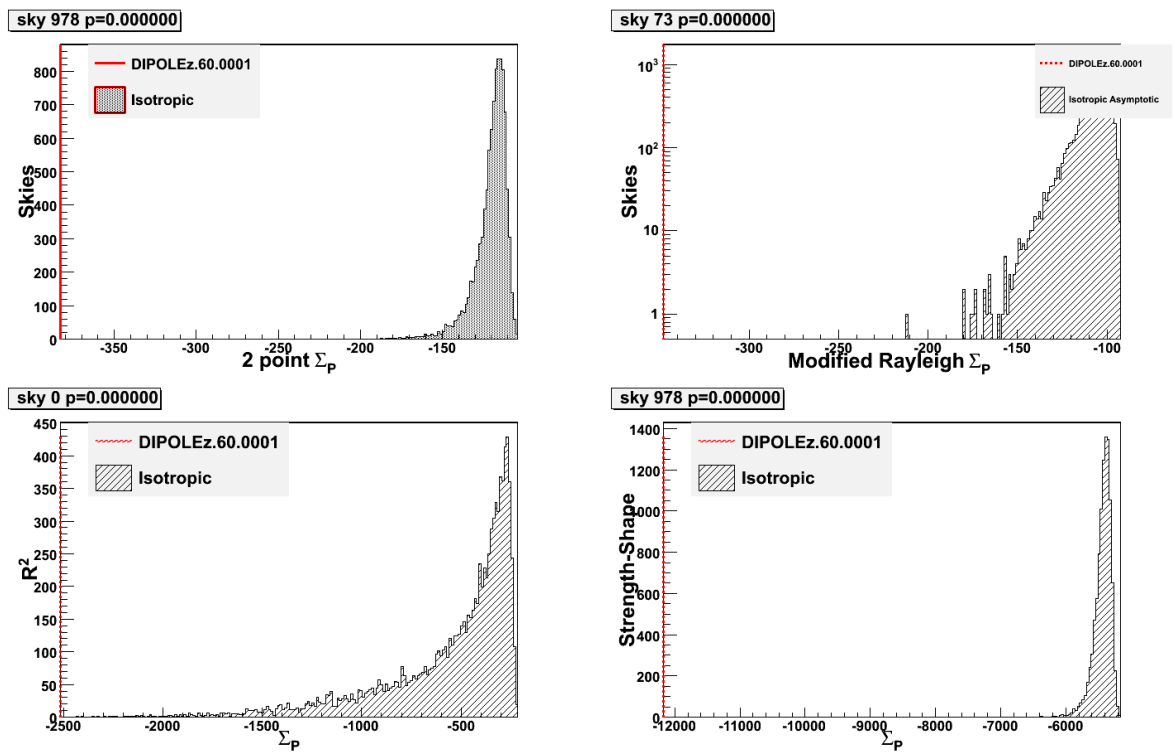


Figure B.8: Distribution of Σ_p for isotropic skies and for the least isotropic test sky (indicated by the red line). Each plot is the Σ_p frequency distribution (number of skies versus Σ_p. Please note that the bottom two plots are mis-labeled on the vertical axis. The plots (clockwise, starting from top right) are for the metrics: 2pt,Rayleigh, SS, R^2.

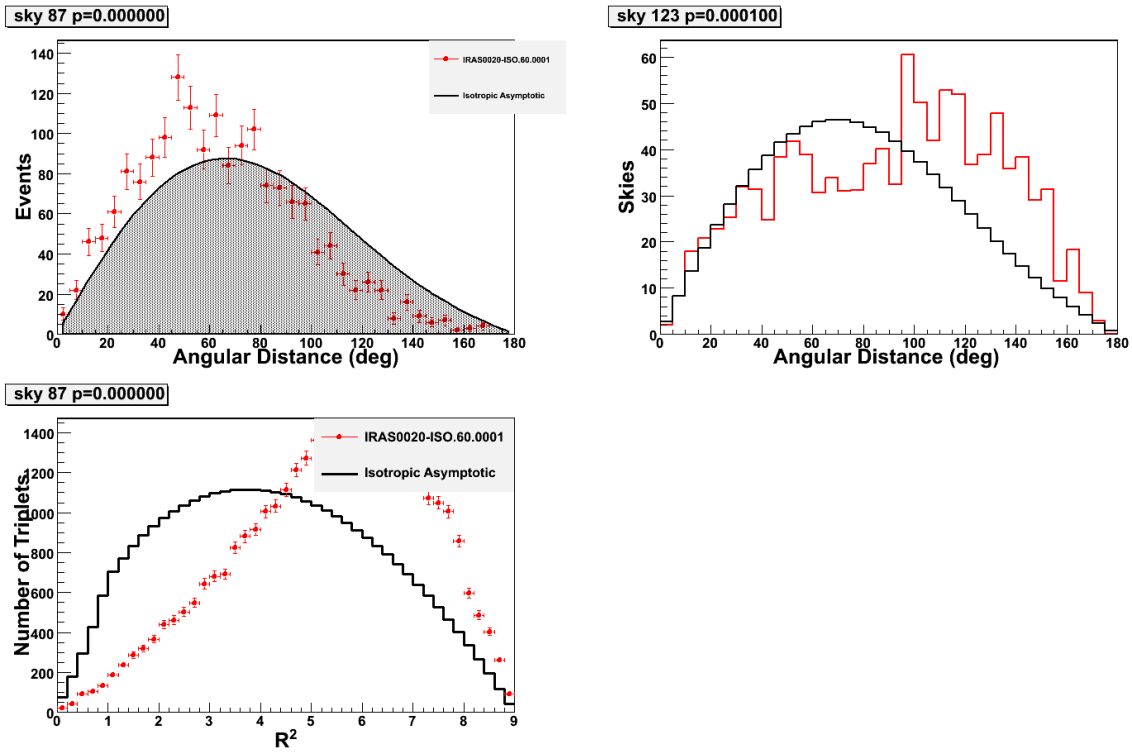


Figure B.9: Least isotropic sky compared to average isotropic metric.

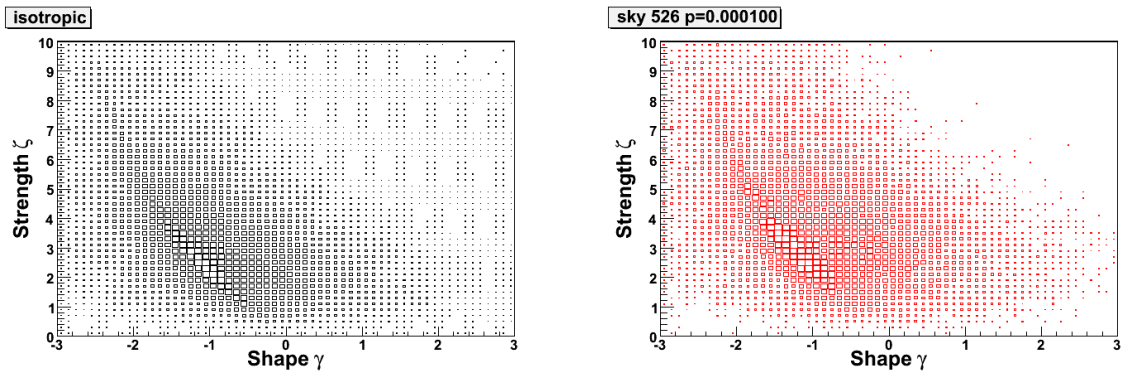


Figure B.10: Least isotropic sky (top) compared to average isotropic sky (bottom) for the SS metric.

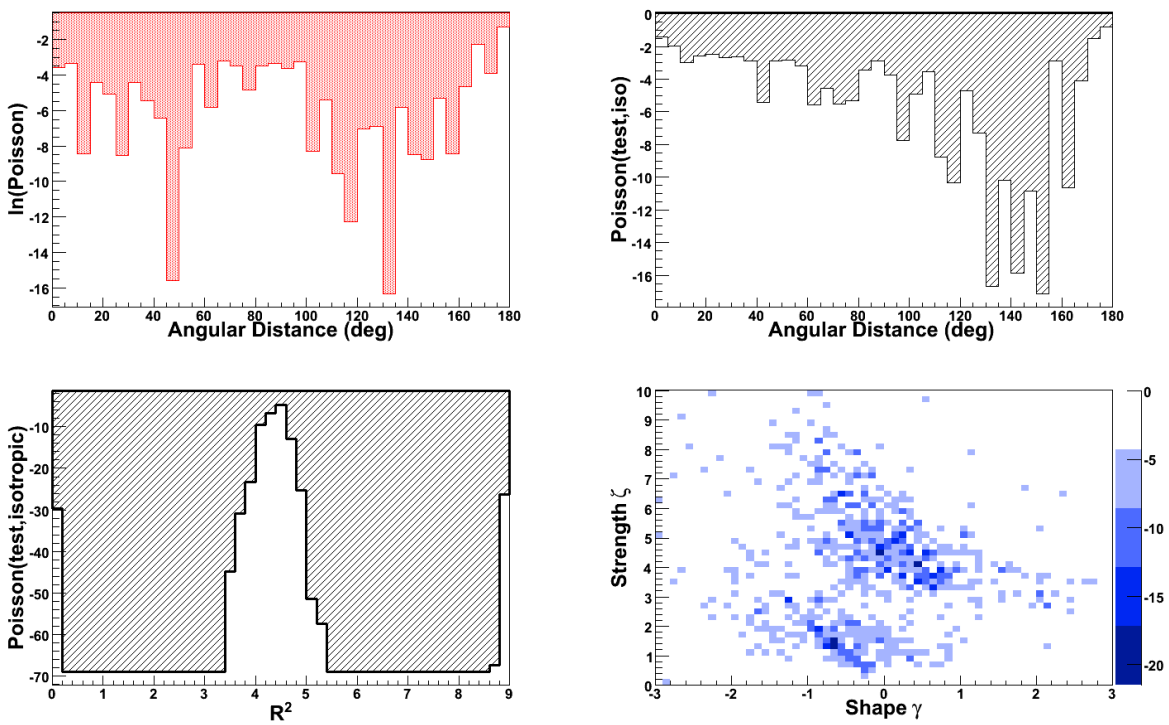


Figure B.11: Least isotropic sky Poissson values by bin.

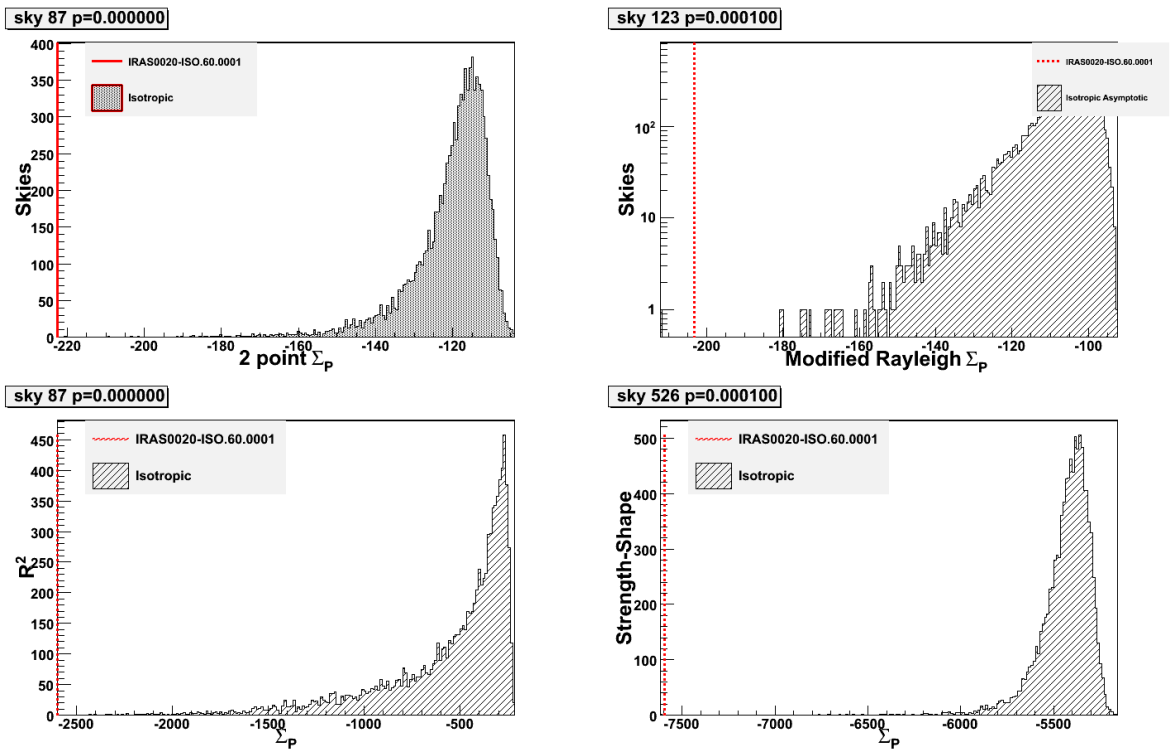


Figure B.12: Distribution of Σ_p for isotropic skies and for the least isotropic test sky (indicated by the red line). Each plot is the Σ_p frequency distribution (number of skies versus Σ_p. Please note that the bottom two plots are mis-labeled on the vertical axis. The plots (clockwise, starting from top right) are for the metrics: 2pt,Rayleigh, SS, R^2.

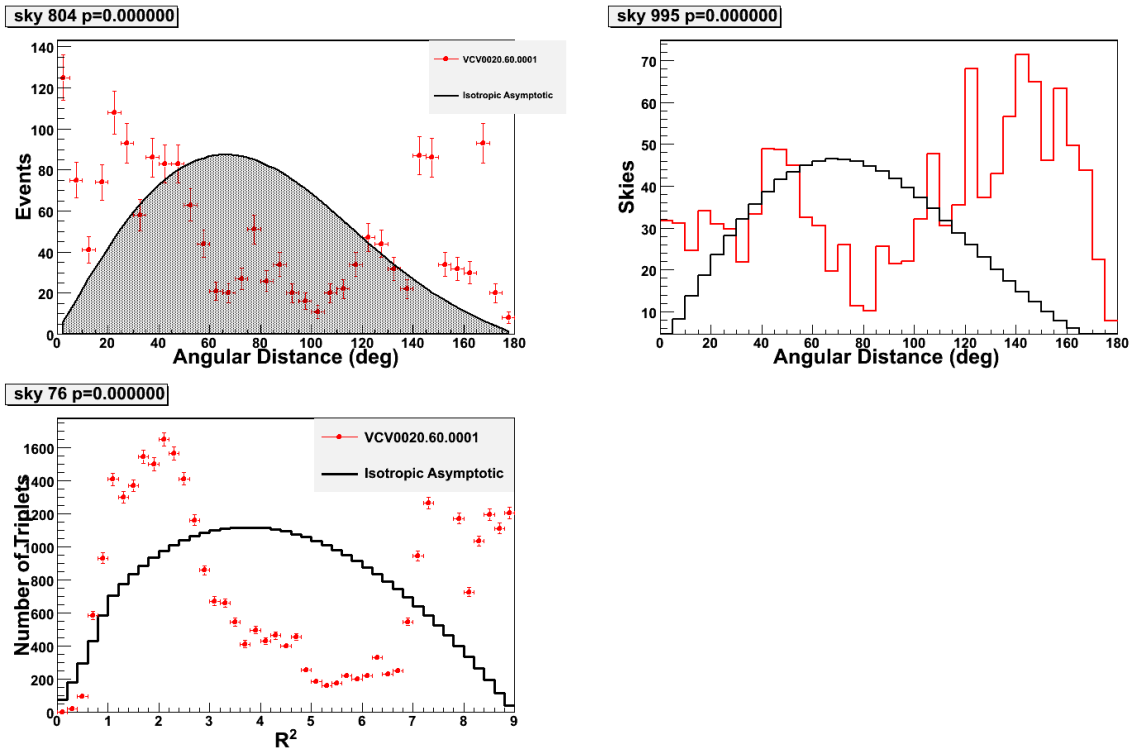


Figure B.13: Least isotropic sky compared to average isotropic metric.

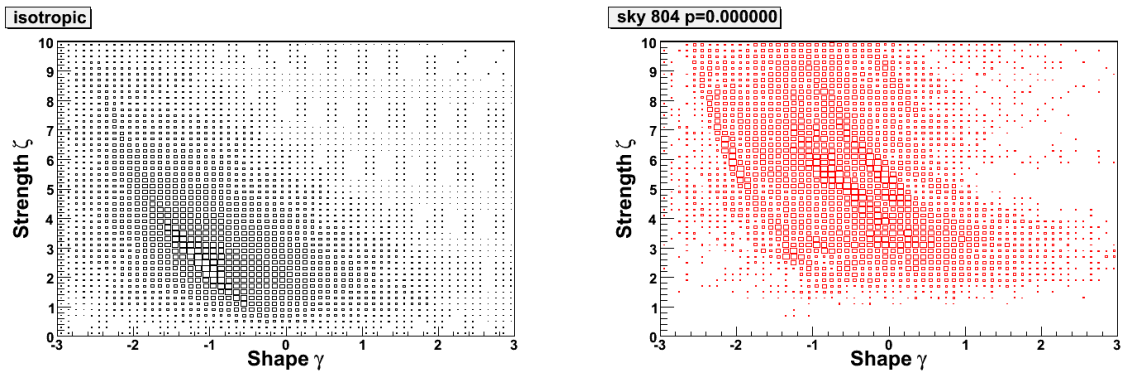


Figure B.14: Least isotropic sky (top) compared to average isotropic sky (bottom) for the SS metric.

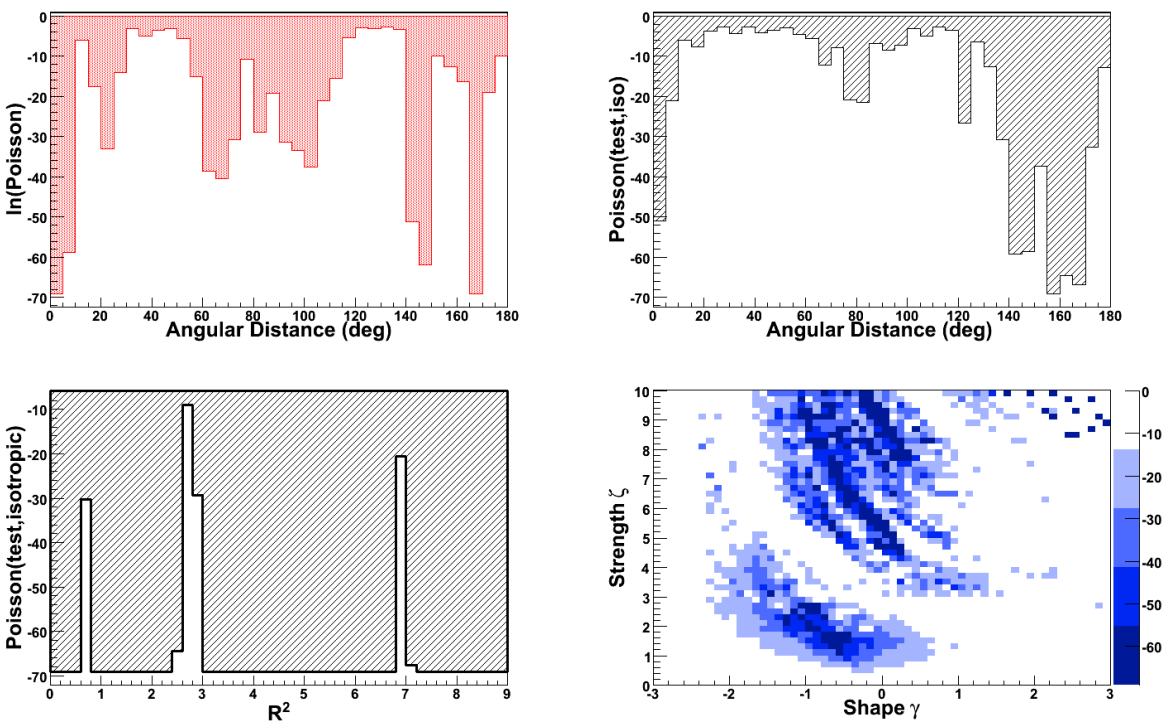


Figure B.15: Least isotropic sky Poissson values by bin.

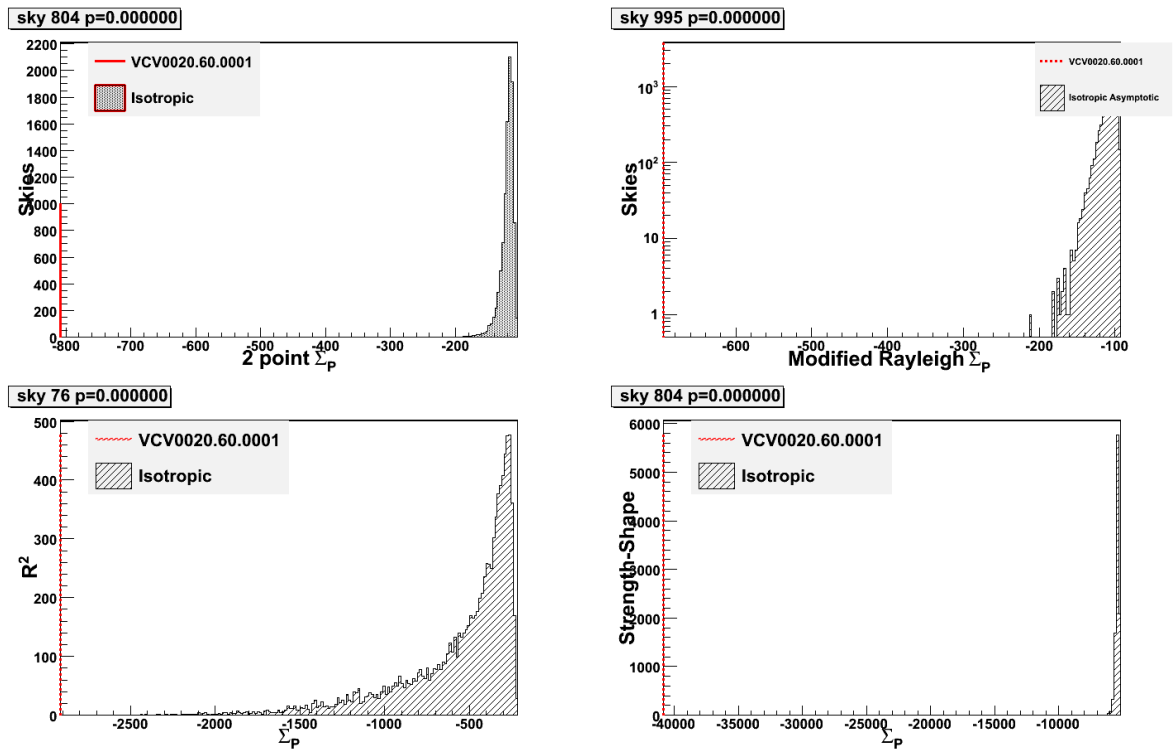


Figure B.16: Distribution of Σ_p for isotropic skies and for the least isotropic test sky (indicated by the red line). Each plot is the Σ_p frequency distribution (number of skies versus Σ_p. Please note that the bottom two plots are mis-labeled on the vertical axis. The plots (clockwise, starting from top right) are for the metrics: 2pt,Rayleigh, SS, R^2.

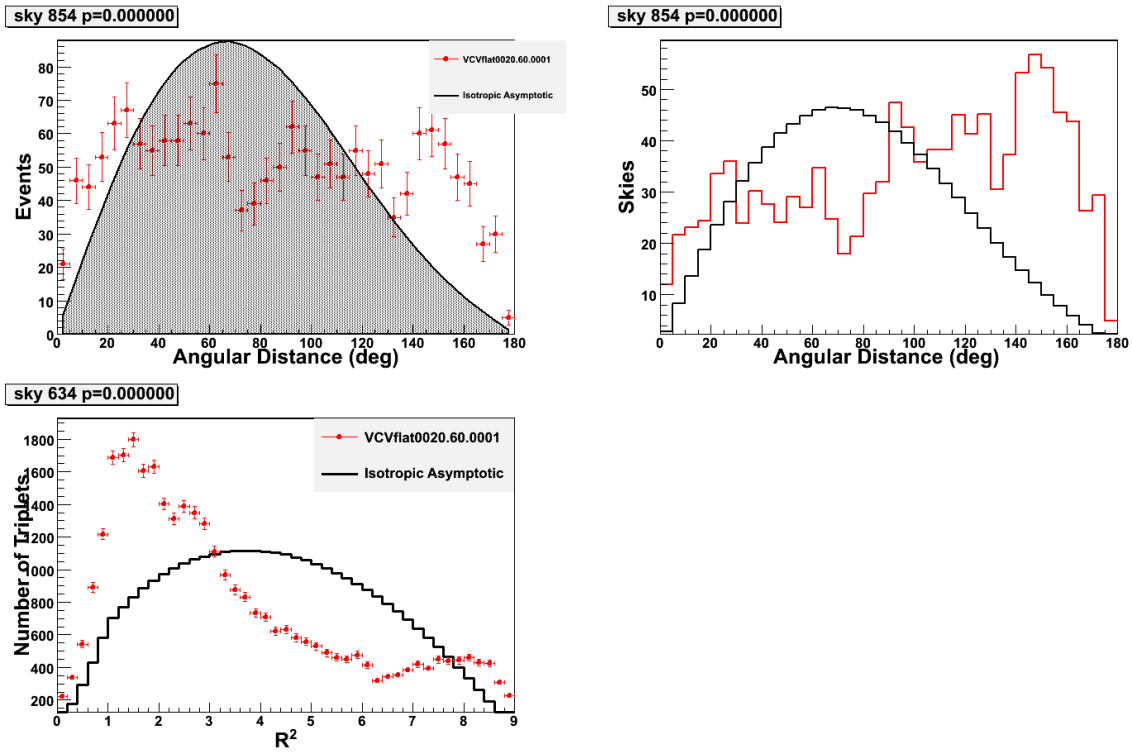


Figure B.17: Least isotropic sky compared to average isotropic metric.

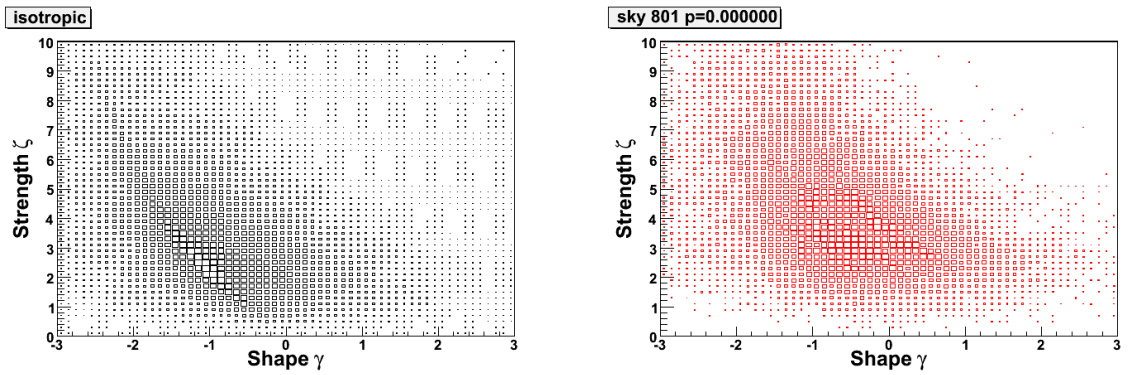


Figure B.18: Least isotropic sky (top) compared to average isotropic sky (bottom) for the SS metric.

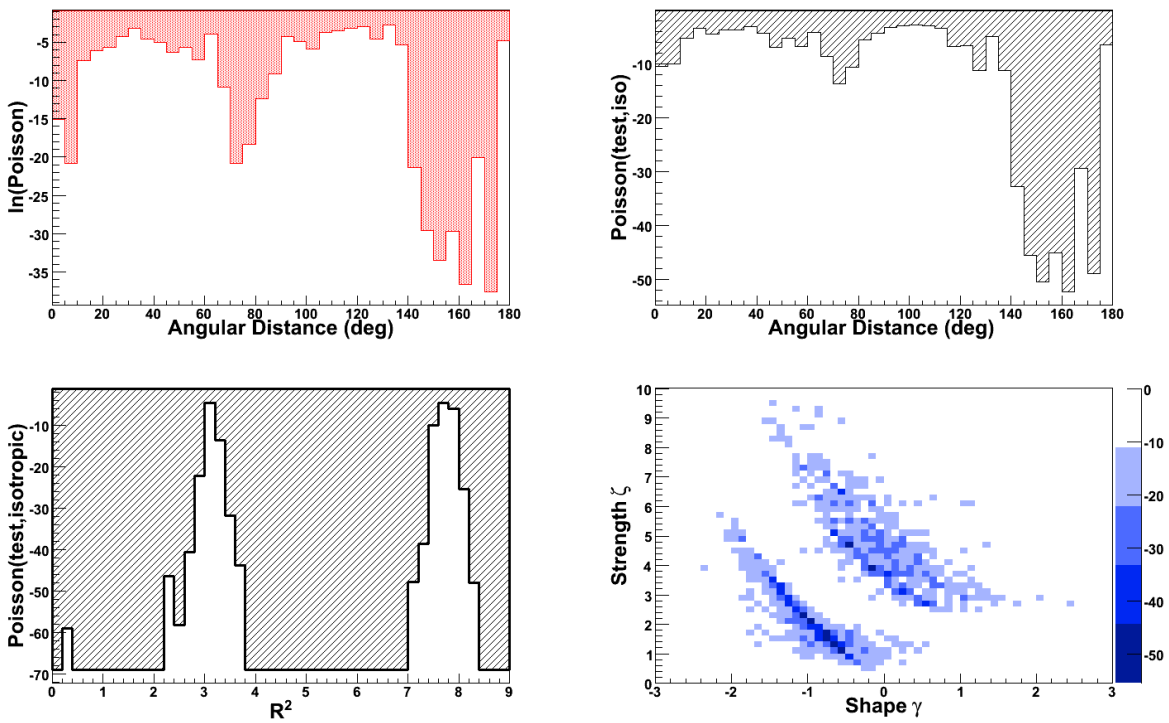


Figure B.19: Least isotropic sky Poissson values by bin.

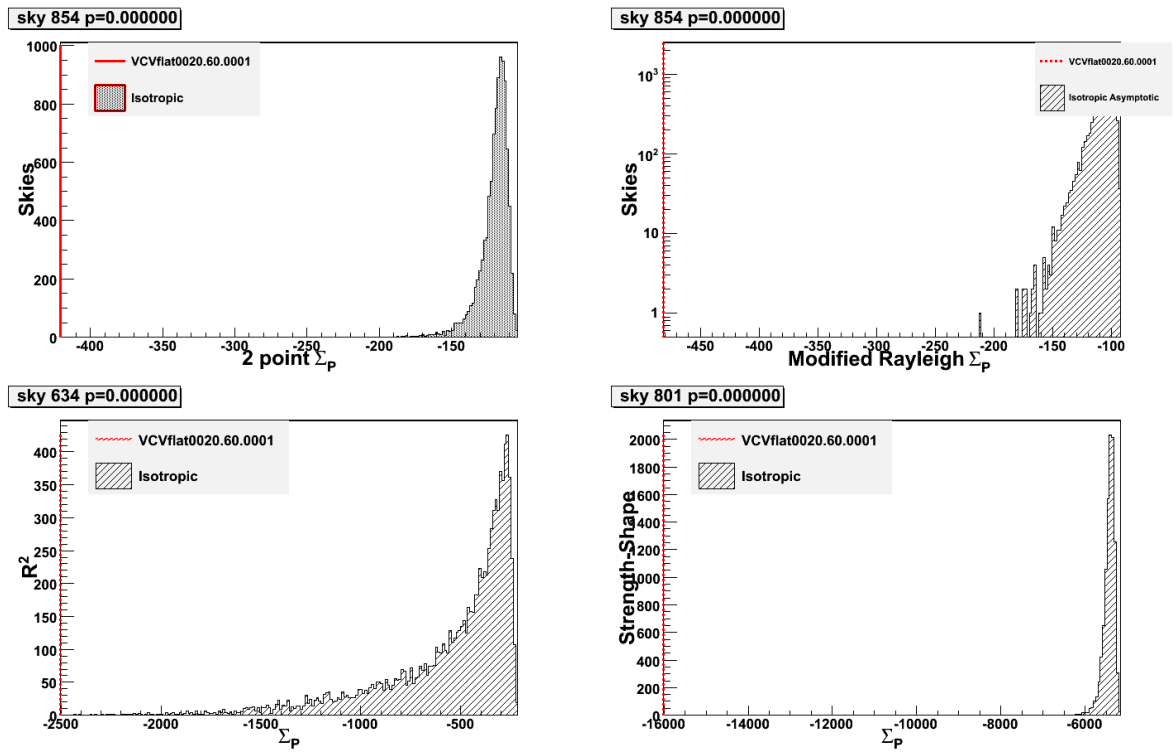


Figure B.20: Distribution of Σ_p for isotropic skies and for the least isotropic test sky (indicated by the red line). Each plot is the Σ_p frequency distribution (number of skies versus Σ_p. Please note that the bottom two plots are mis-labeled on the vertical axis. The plots (clockwise, starting from top right) are for the metrics: 2pt,Rayleigh, SS, R^2.

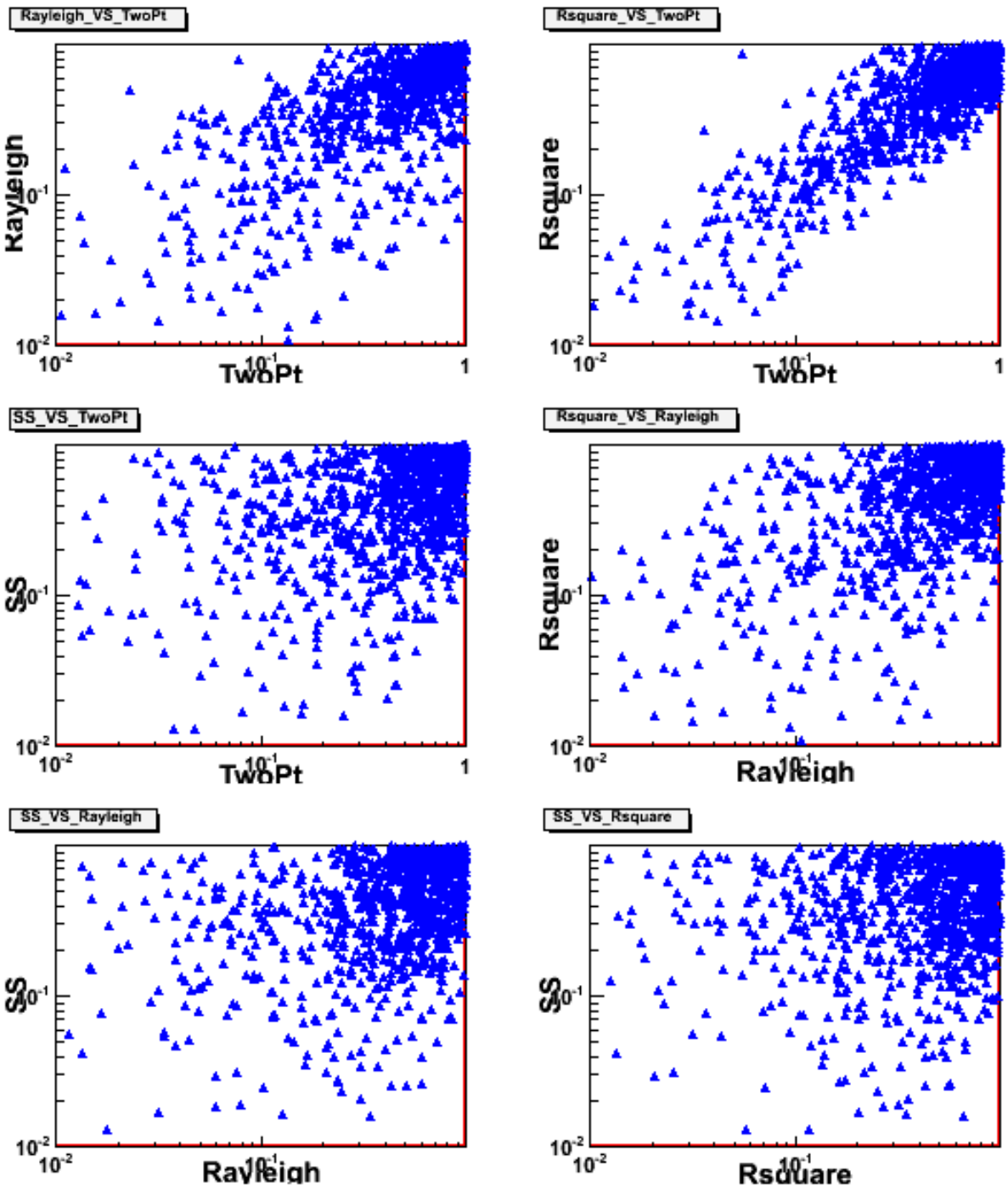


Figure B.21: Correlations for set ISOTROPY.60.0001

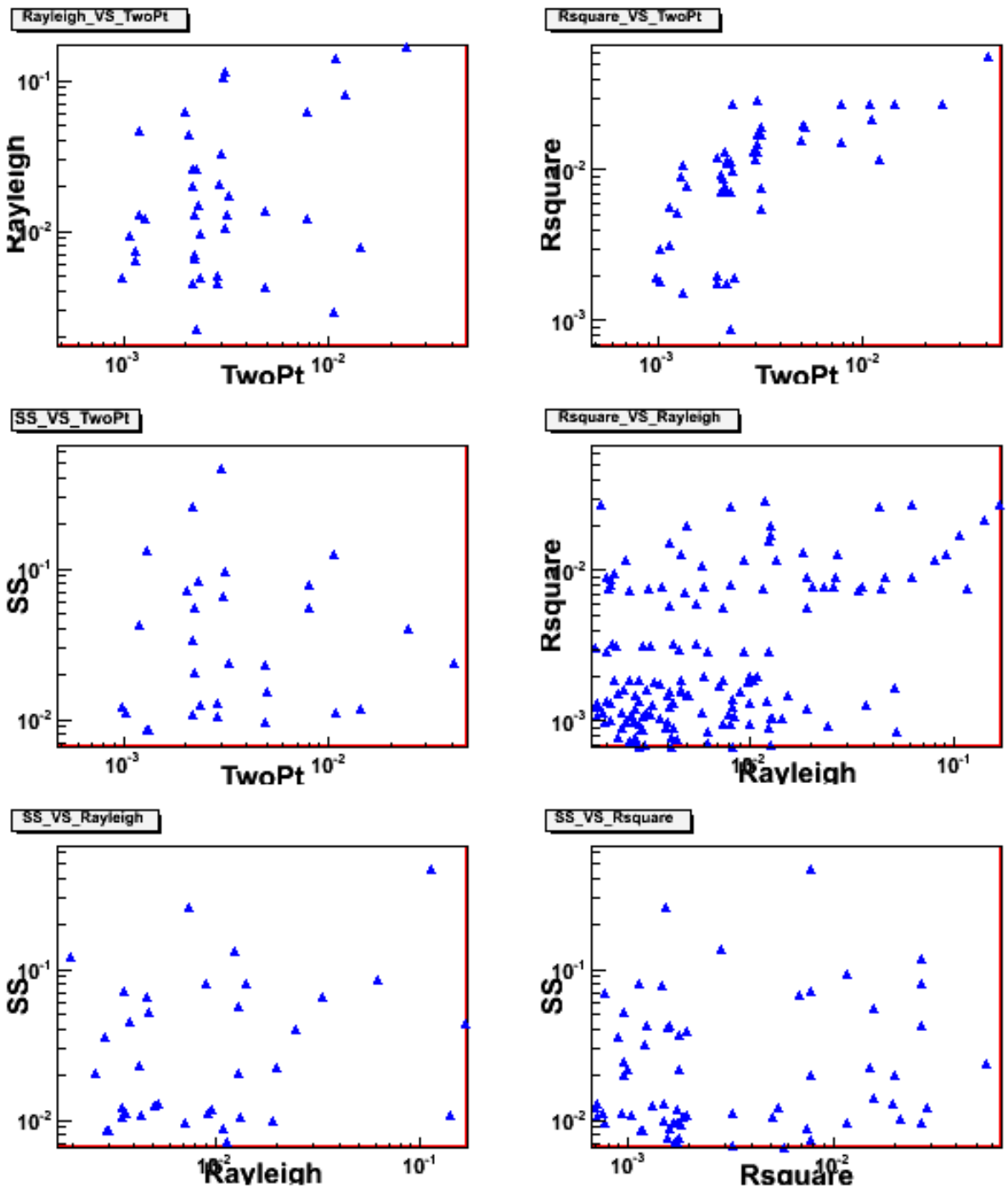


Figure B.22: Correlations for set DIPOLEz.60.0001

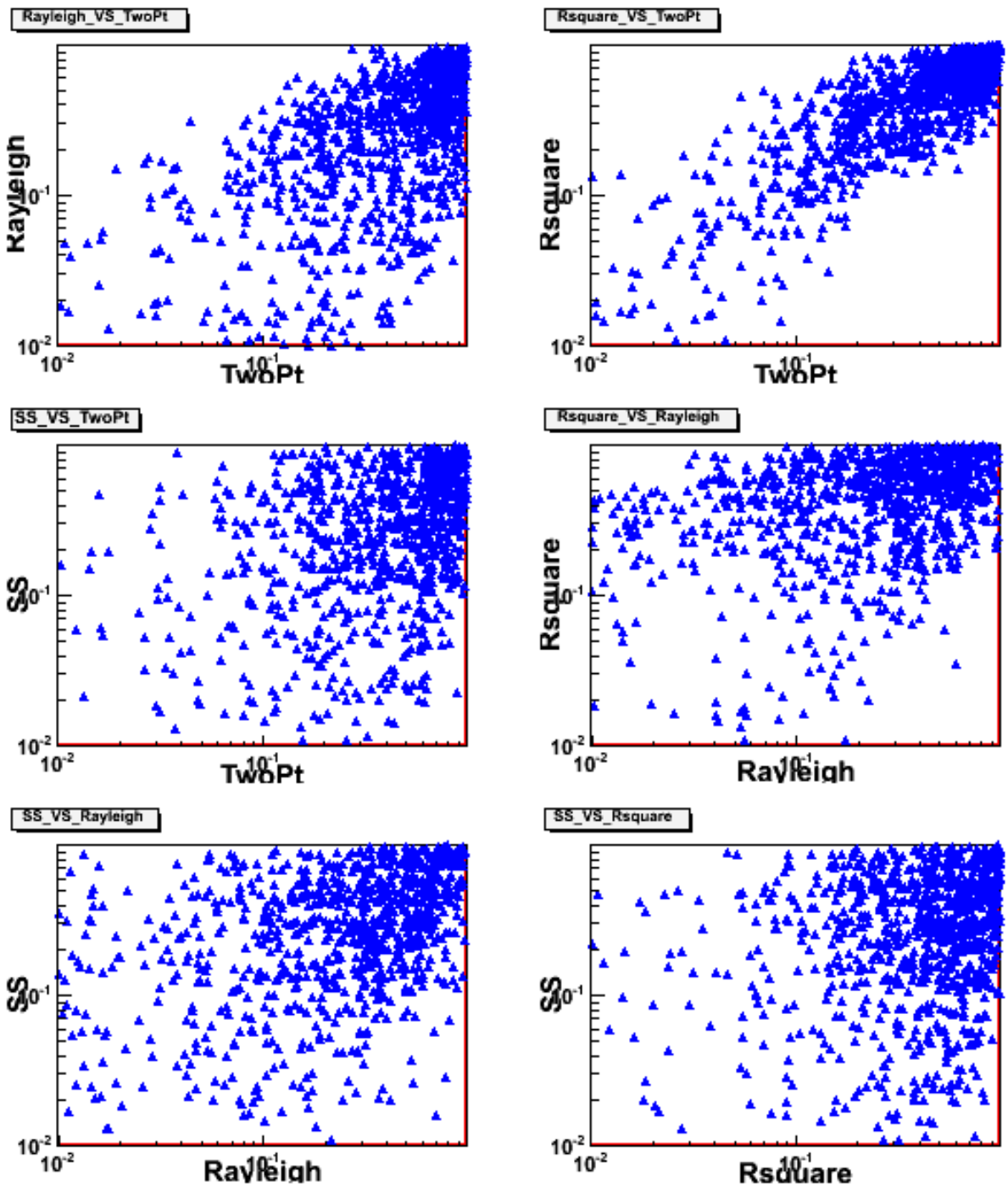


Figure B.23: Correlations for set IRAS0020.60.0001

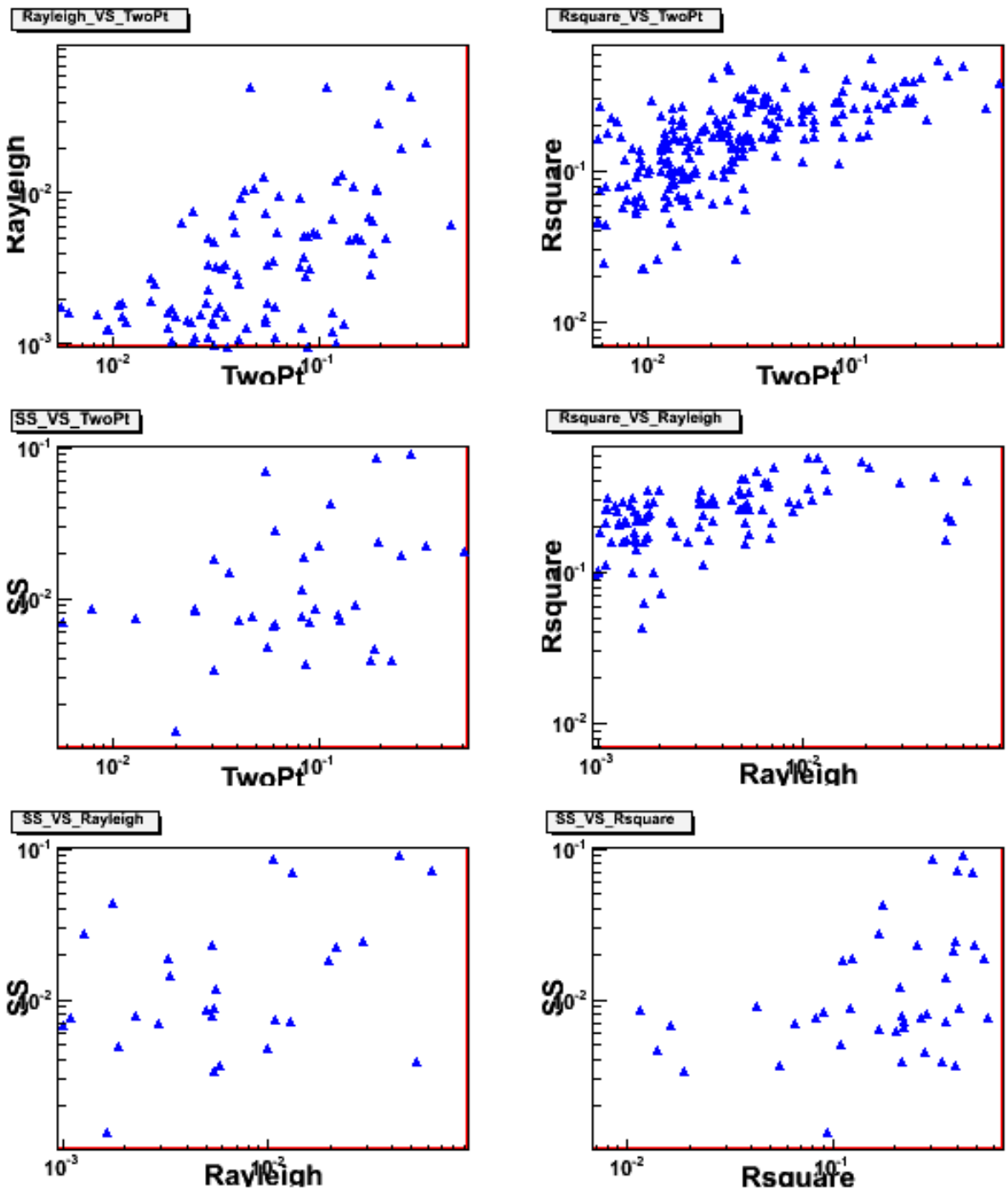


Figure B.24: Correlations for set VCVflat.60.0001

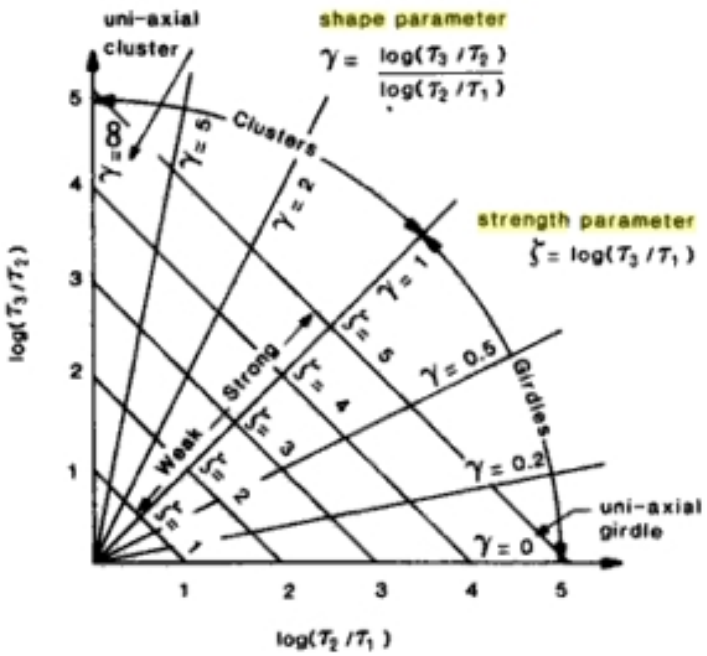


Figure B.25: Interpretation of shape and strength metrics. Reproduced from [83].

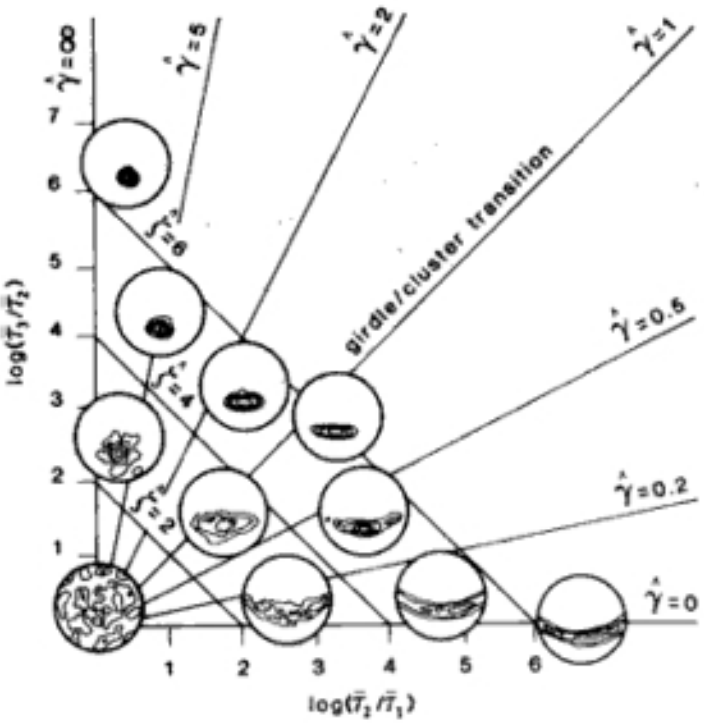


Figure B.26: Interpretation of shape and strength metrics. Reproduced from [83].

B.3 Mock Signal Detection Efficiency

B.3.1 Section Overview

This section was originally published as [98].

The Type I and II errors for rejecting isotropic cosmic ray arrival directions are discussed and calculated using the four metrics described in GAP-2008-081 applied to 16 mock signal ensembles. We conclude that the resultant (R^2) metric is a poor identifier of anisotropy. The other metrics have high efficiency for some of the mock signals and are correlated with each other. We explicitly study mock data sets of three sizes (20, 40 and 60 events) that have been diluted with background events such that the fraction of signal is 20%, 40%, 60%, 80% and 100%. We conclude that both dilution and the total number of events dramatically effect the signal detection efficiency.

B.3.2 Analysis

In this note we study 4 "metrics"[105] for measuring cosmic ray anisotropy: **2pt**; the standard two-point correlation function (i.e. see [106, 107]); **Rayleigh**; the modified two-point Rayleigh metric (see [106, 107] and [3]); **Resultant**; or R^2 is a simple three-point directional vector sum (a measure of spherical triangular area); and **SS**; or Shape-Strength, a three-point metric derived from a principle component analysis (i.e. see [109, 83, 84, 85]).

For a given set of CR arrival directions (a "sky") and for each metric independently we compute a pseudo-log-likelihood, Σ_P, by comparing the test sky metric's

[3]We note that this metric may be coordinate system dependent, see [108]. We use the "$+z$ choice" in galactic coordinates for the vector distance between events, where as [106, 107] uses equatorial coordinates.

to an ensemble of 2×10^4 isotropic skies with the same number of events and weighted by the Auger exposure, see [105] for a detailed discussion. Each of the four pseudo-log-likelihoods serves as the test statistic for testing the (null) isotropy hypothesis, H_{iso}.

To weigh the effectiveness of a metric for rejecting isotropy in a given CR sky, we are interested in the probabilities for two types of hypothesis testing errors[49]:

Type I The probability, α, of rejecting H_{iso} given that H_{iso} is *true*. This value is called the *significance* and in practice it should be chosen *a priori*. In this analysis we study two choices, $\alpha_{01} = 0.01$ and $\alpha_{001} = 0.001$, corresponding to the 1% and 0.1% significance levels respectively[4]. For each metric the choice of α corresponds to a unique Σ_P; we find the ${\Sigma_P}^{\alpha_{01}}({\Sigma_P}^{\alpha_{001}})$ such that the ratio of the number of isotropic skies with Σ_P less than ${\Sigma_P}^{\alpha_{01}}({\Sigma_P}^{\alpha_{001}})$ is $\alpha_{01}(\alpha_{001})$. We use the 10^4 skies in the ISOTROPY ensemble described in Table B.4 to determine the upper bound of the signal regions ${\Sigma_P}^{\alpha_{01}}$ and ${\Sigma_P}^{\alpha_{001}}$.

Type II The probability, β, of rejecting H_{iso} (i.e. of accepting the mock signal hypothesis H_{sig}) given that H_{iso} is *false*. This value is dependent on the choice of H_{sig} and might be called the "efficiency of accepting" H_{sig} [5]. By applying the ensemble of each mock signal to each metric we estimate $\beta_{\alpha_{01}}(\beta_{\alpha_{001}})$ as the ratio of mock signal skies with $\Sigma_P < {\Sigma_P}^{\alpha_{01}}(\Sigma_P < {\Sigma_P}^{\alpha_{001}})$ to the total number of mock skies. As a heuristic measure we will describe a metric's efficiency as "good" if it is at least 90%, i.e. $\beta \gtrsim 0.90$, and questionable if it is less than about 90%.

In this study we compute the Type II error probability for each of the four metrics

[4]These values for α are chosen to be the same as the value used in [8] ($\alpha_{01} = 0.01$) and ten times smaller ($\alpha_{001} = 0.001$).

[5]The quantity $1-\beta$ is called the *power* of the metric for rejecting the signal hypothesis[49].

applied to all sixteen of the mock signal ensemble files described in Table B.4. See FIGs. B.31, B.32, B.33 and B.34 for examples of the distributions of Σ_P (for 60 event data sets) used to compute ${\Sigma_P}^{\alpha_{01}}$ and ${\Sigma_P}^{\alpha_{001}}$. The resulting $\beta_{\alpha_{01}}$'s and $\beta_{\alpha_{001}}$'s are plotted for each metric and each signal type in Figure B.27. The files with an "ISO" suffix contain a mixture of mock signal and background events in each sky.

To explicitly study the effect that such a mixture has on β we separately construct mock ensemble files in which each sky has a certain ratio, r, of signal events to the total number of events,

$$\{\text{Mock Mixture Data}\} = r\times\{\text{Mock Signal Data}\}+(1-r)\times\{\text{Mock Isotropic Data}\}, \tag{B.8}$$

with $r = 0.2, 0.4, 0.6$ and 0.8. We use four mock signal types, DIPOLEz, QUADRAPOLEx, VCV0020 and VCVflat0020 and each file contains 10^4 skies with sixty events per sky. Notice that because our metrics use *all* triplets or doublets in a given sky the mixture Σ_P distributions are *not* a simple sum of the signal and isotropic Σ_P distributions. See Figure B.28 for the efficiencies for detecting anisotropy in these diluted mock signal skies.

Detection efficiency is also strongly effected by the most important number in all of CR astronomy; the number of (high energy) events in a sky [6]. The effect can be similar to those of signal dilution in that the detection efficiency is decreased. With this in mind we perform a study similar to that described in previous paragraph except that we (separately) use twenty and forty events per sky. We use four mock signal types, DIPOLEz, QUADRAPOLEx, VCV0020 and VCVflat0020 and each file contains 10^4 skies. See Figure B.29 and Figure B.30 for the efficiencies for detecting anisotropy in the twenty and forty events-per-sky diluted mock signal skies respectively.

[6] Under the ridiculous assumption that Auger was built *solely* for ultra-high energy CR astronomy, each of the twenty-seven events used in [8] cost $\sim \$3.5 \times 10^6$.

ISOTROPY	Isotropy weighted by exposure (like the rest of the cases!).
DI/QUADRA/OCTU-POLE	Either a dipole, quadrapole or octu-pole pointing in either the $+z$ or $+x$ direction in equatorial coordinates. For DIPOLEzn see [102].
FiveEll10-10-25ISO	See [102].
FiveEll15-5-25ISO	See [102].
IRAS0020-10ISO	See [102].
IRAS0020-ISO	30 events coming from ISOTROPY, 30 events from IRAS redshift $z_{max} \leq 0.020$. Weighted by $1/z^2$ and selection function. NO IRAS MASK. Angular randomization: 2D-Gaussian, $\sigma = 3^o$.
VCV0020	60 events from VCV $z_{max} \leq 0.020$. Weighted by $1/z^2$. Angular randomization: 2D-Gaussian, $\sigma = 3^o$.
VCV0020-ISO	30 events coming from ISOTROPY, 30 events from VCV $z_{max} \leq 0.020$. Weighted by $1/z^2$. Angular randomization: 2D-Gaussian, $\sigma = 3^o$.
VCVflat0020	60 events from VCV $z_{max} \leq 0.020$. NO z weighted (flat). Angular randomization: each CR is drawn from a collection of 2D-Gaussian probability distributions centered at the catalogue sources, with $\sigma = 3^o$.
VCVflat0020-ISO	30 events coming from ISOTROPY, 30 events from VCV $z_{max} \leq 0.020$. NO z weighted (flat). Angular randomization: each CR is drawn from a collection of 2D-Gaussian probability distributions centered at the catalogue sources, with $\sigma = 3^o$.

Table B.3: A reproduction of the list and description of the 16 mock files provided by [102] and used in this study. Each of these mock files consists of 10^4 skies of 60 events each. The files with an "ISO" suffix contain a mixture of mock signal and background events in each sky. Sky plots of these ensembles can be found at [110].

B.3.3 Observations

The efficiencies for each of the four metrics and all 16 mock signal ensembles (see Table B.4) are the first primary result of this note and are plotted in Figure B.27. We note as a code/consistency cross-check that the ISOTROPY ensemble is consistent with H_{iso}, i.e. $\beta \approx 0\%$. For 60 event data sets, all of the metrics are poor ($\beta \lesssim 50\%$) at detecting DIPOLEx, DIPOLEzn, OCTUPOLE, OCTUPOLEx, IRAS0020-10ISO, IRAS0020-ISO and VCVflat0020-ISO. Both DIPOLEz and DIPOLEx favor the **2pt** metric, though DIPOLEx is the more difficult of the two to detect. The VCV0020 and VCVflat0020 mock signal ensembles favor the **Rayleigh** and **SS** metrics. Generally speaking, where one method is good ($\beta \gtrsim 90\%$) so is another; the metrics are correlated. The exception to this generality is the VCV0020-ISO mock signal ensemble where the shape strength metric is good, while the other metrics are not as good.

We explicitly study the effect that signal dilution of a 60 event sky has on β in Figure B.28. We notice, as expected, that as the amount of signal is increased, the efficiency also increases monotonically. Diluting the signal dramatically decreases the efficiency of detection, i.e. $\beta \lesssim 60\%$ for nearly all the studied ensembles. The **2pt**, **Rayleigh** and **SS** metrics can, however, detect the mock signal VCV0020 with $\sim 90\%$ efficiency with only 60% of the events in each (60 event) sky being mock signal events. We notice that the VCV0020-ISO and VCVflat0020-ISO mock signal ensembles contain an even mixture ($r = 0.5$) of signal and background events and that the efficiencies for these signal types plotted in Figure B.27 are consistent with those plotted in Figure B.28.

By comparing and contrasting FIGs. B.28, B.29 and B.30 we can get qualitative information about how the total number of events in a sky (equal to 60, 40 and 20 events respectively) can effect the signal detection efficiency. We notice that *only*

the VCV0020 is detectable with twenty events in a sky. The mock signal types VCVflat0020 and DIPOLEz need both large numbers of events (40 to 60) and high signal purity ($r \gtrsim 80\%$). The QUADRUPOLEx signal type can possibly be detected *either* with large statistics (60 events) but $r < 80\%$ signal purity *or* with modest statistics (40 events) and fairly high ($r \gtrsim 80\%$) purity.

We notice that there is a trade-off between α and β; a smaller *a priori* choice of α (that is, a higher level of confidence for rejecting H_{iso}) will result in a smaller efficiency β of accepting a given mock signal type. An example of this can be seen in the efficiencies of the **SS** metric for detecting the VCV0020-ISO signal type, plotted in Figure B.27, where $\beta_{\alpha_{01}} \gtrsim 95\%$'s but $\beta_{\alpha_{001}} \lesssim 80\%$.

B.3.4 Conclusion

Both the Type I and Type II errors probabilities are relevant for weighing one metric against another. The first is chosen a priori and defines the signal region for Σ_P. One desires the metric that gives the most narrow isotropic distribution so that the signal region is as large as possible. The second gives the efficiency of detecting a specific type of anisotropy. It is *highly* dependent on the type of anisotropy being tested (H_{sig}) and it therefor has the potentiality, when used with multiple metrics, for discriminating possible signal scenarios in the actual (and sole) observed arrival directions data set. Thus, it is desirable to choose a metric such that its distribution of Σ_P has little overlap with the isotropic Σ_P distribution and, thus, a large efficiency β. See FIGs. B.31, B.32, B.33 and B.34 for examples of the distributions of Σ_P and determination of the signal regions. The median p-valueused by the anisotropy working group[102] reflects some combination of Type I and II errors.

An important conclusion to draw from this analysis is that *both* signal purity and the total number of events can dramatically effect the signal detection efficiency.

We note that the **Resultant** metric is poor for detecting all the mock signals, i.e. $\beta \lesssim 80\%$, and it will be dropped from further consideration. We also wish to emphasize from the analysis of the diluted mock signals that when the signal to noise ratio $r \gtrsim 50\%$ we can expect that "VCV-like" CRs should be identified with $\beta \gtrsim 50\%$ by the **2pt**, **Rayleigh** and **SS** metrics for sixty events data sets. Furthermore this study points to the need for Northern Auger (with $\sim 8\times$ the aperture of Auger South) to have high detection efficiency for a broad range of possible anisotropic signals.

B.3.5 Acknowledgements

We wish to thank Tim Thomas and the University of New Mexico Center for High Performance Computing[86] for their generous feedback and copious floating point operation donations. A special thanks to Lorenzo Cazon and the other organizers of the anisotropy working group for providing the mock signal files and a standardized analysis paradigm, both of which we have found highly educational.

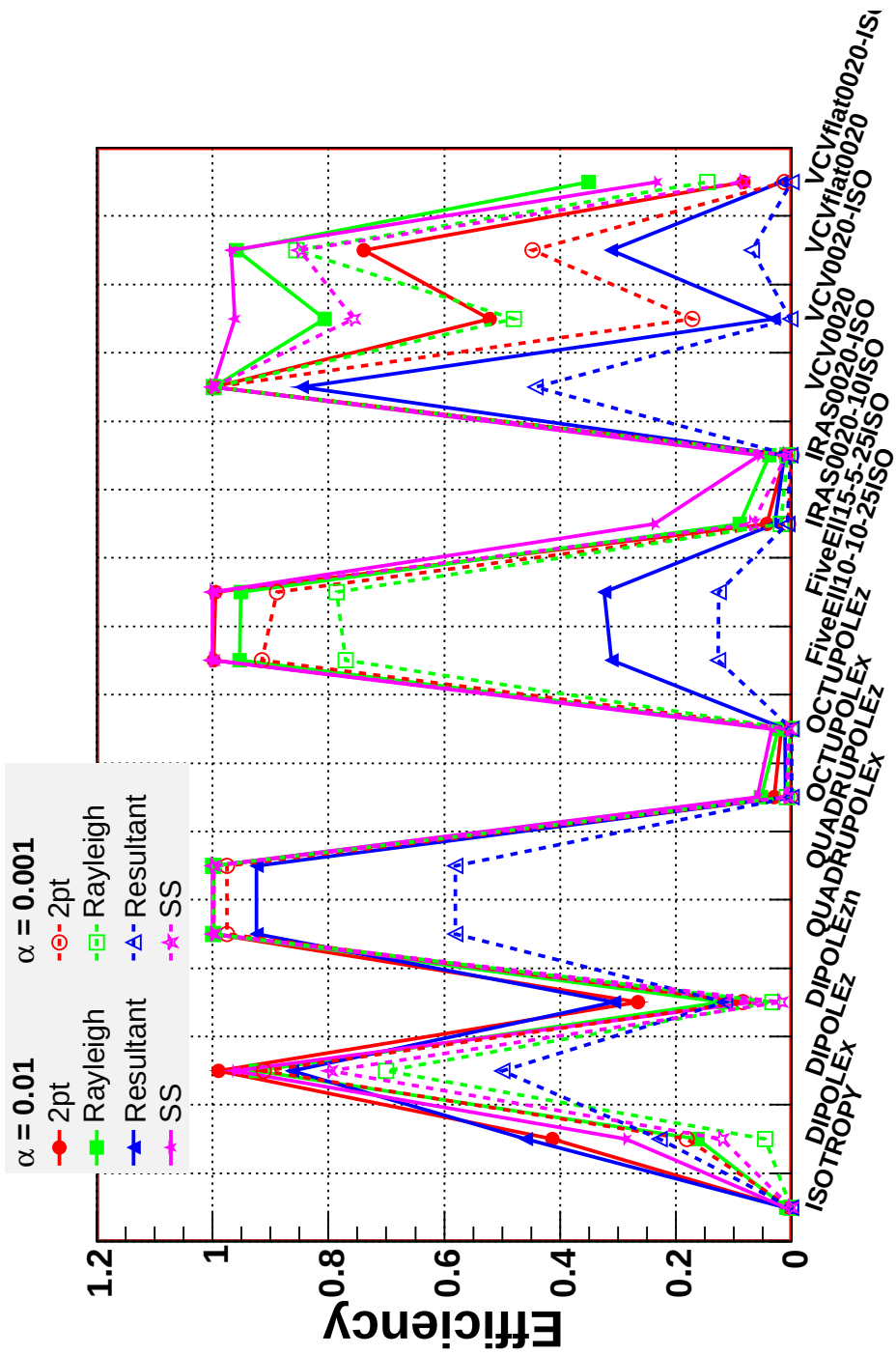


Figure B.27: The efficiency β is the probability of rejecting the isotropy hypothesis (in favor of the mock signal) given that the isotropy hypothesis is false. Each ensemble contains 10^4 with 60 events per sky, see Table B.4 for a description of the mock signal types. See FIGs. B.31, B.32, B.33 and B.34 for examples of the distribution of Σ_P used to compute β.

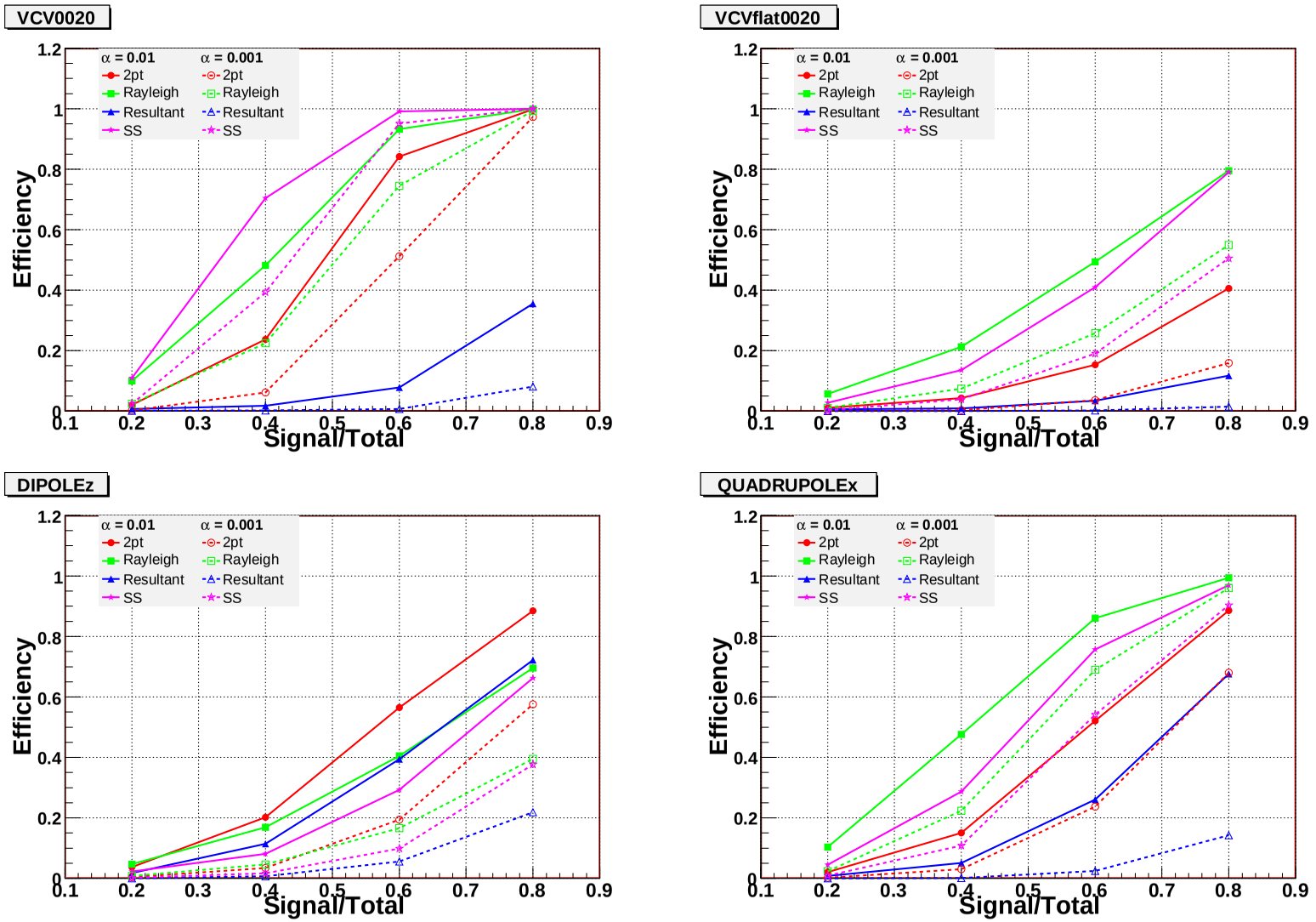


Figure B.28: The efficiency β is the probability of rejecting the isotropy hypothesis (in favor of the mock signal) given that the isotropy hypothesis is false. These plots are similar to Figure B.27 except that each file contains 10^4 skies, with 60 events per sky, and that the four mock signal types, DIPOLEz, QUADRUPOLEx, VCV0020 and VCVflat0020, have been diluted with isotropic(background) events such that the ratios of signal events to total events are 20%, 40%, 60% and 80% signal.

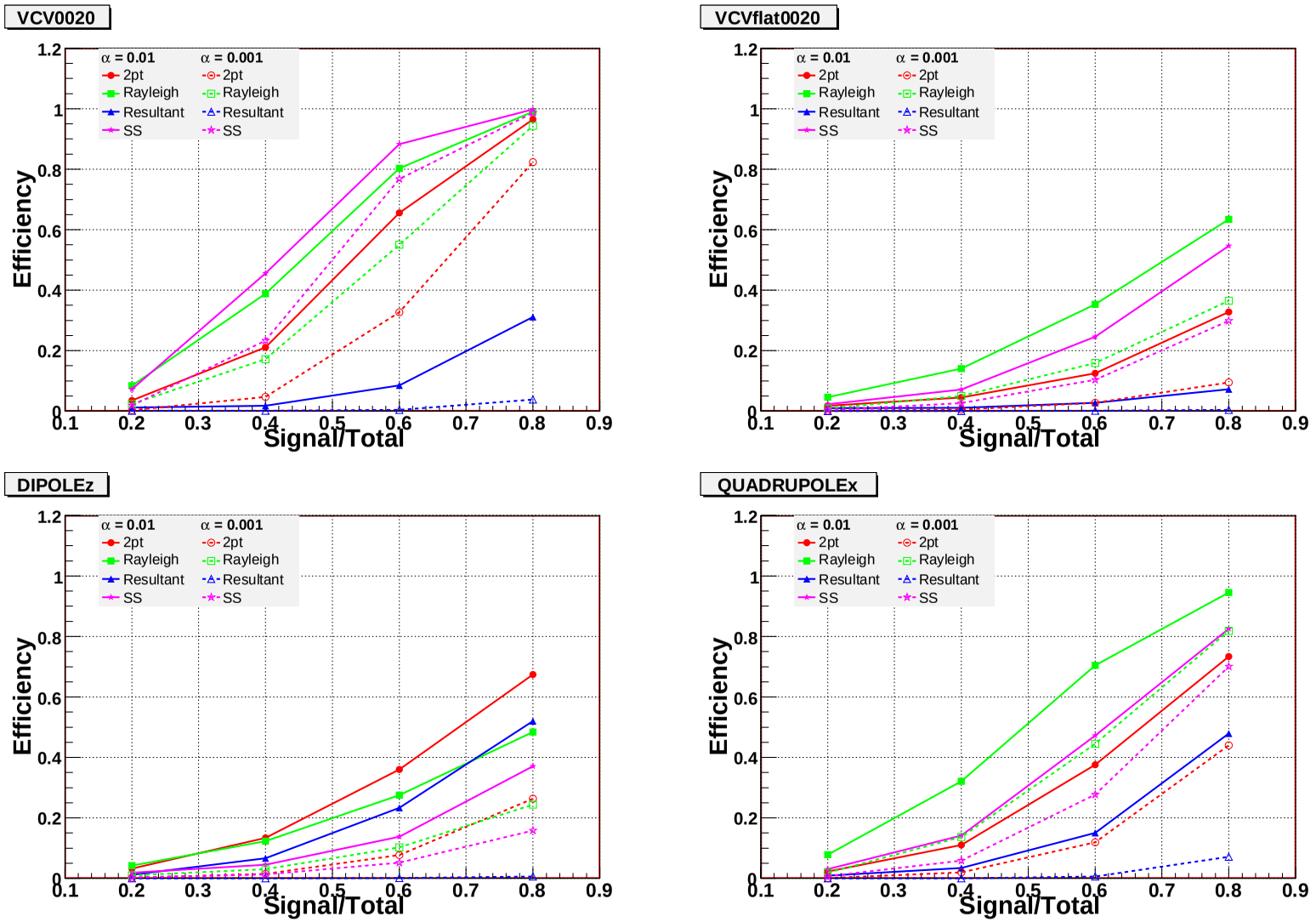


Figure B.29: The efficiency β is the probability of rejecting the isotropy hypothesis (in favor of the mock signal) given that the isotropy hypothesis is false. These plots are similar to Figure B.27 except that each file contains 10^4 skies, with 40 events per sky, and that the four mock signal types, DIPOLEz, QUADRUPOLEx, VCV0020 and VCVflat0020, have been diluted with isotropic(background) events such that the ratios of signal events to total events are 20%, 40%, 60% and 80% signal.

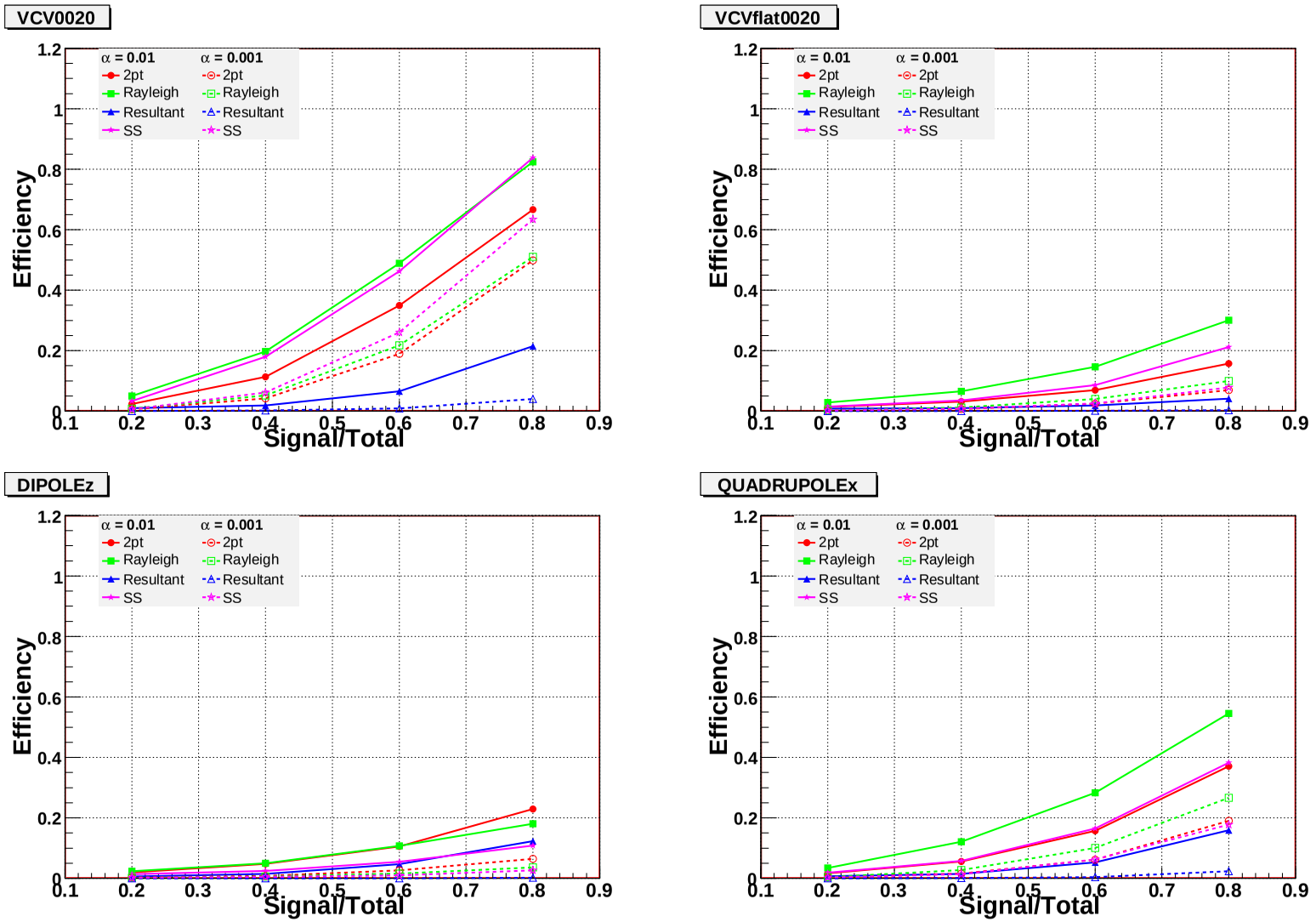


Figure B.30: The efficiency β is the probability of rejecting the isotropy hypothesis (in favor of the mock signal) given that the isotropy hypothesis is false. These plots are similar to Figure B.27 except that each file contains 10^4 skies, with 20 events per sky, and that the four mock signal types, DIPOLEz, QUADRUPOLEx, VCV0020 and VCVflat0020, have been diluted with isotropic(background) events such that the ratios of signal events to total events are 20%, 40%, 60% and 80% signal.

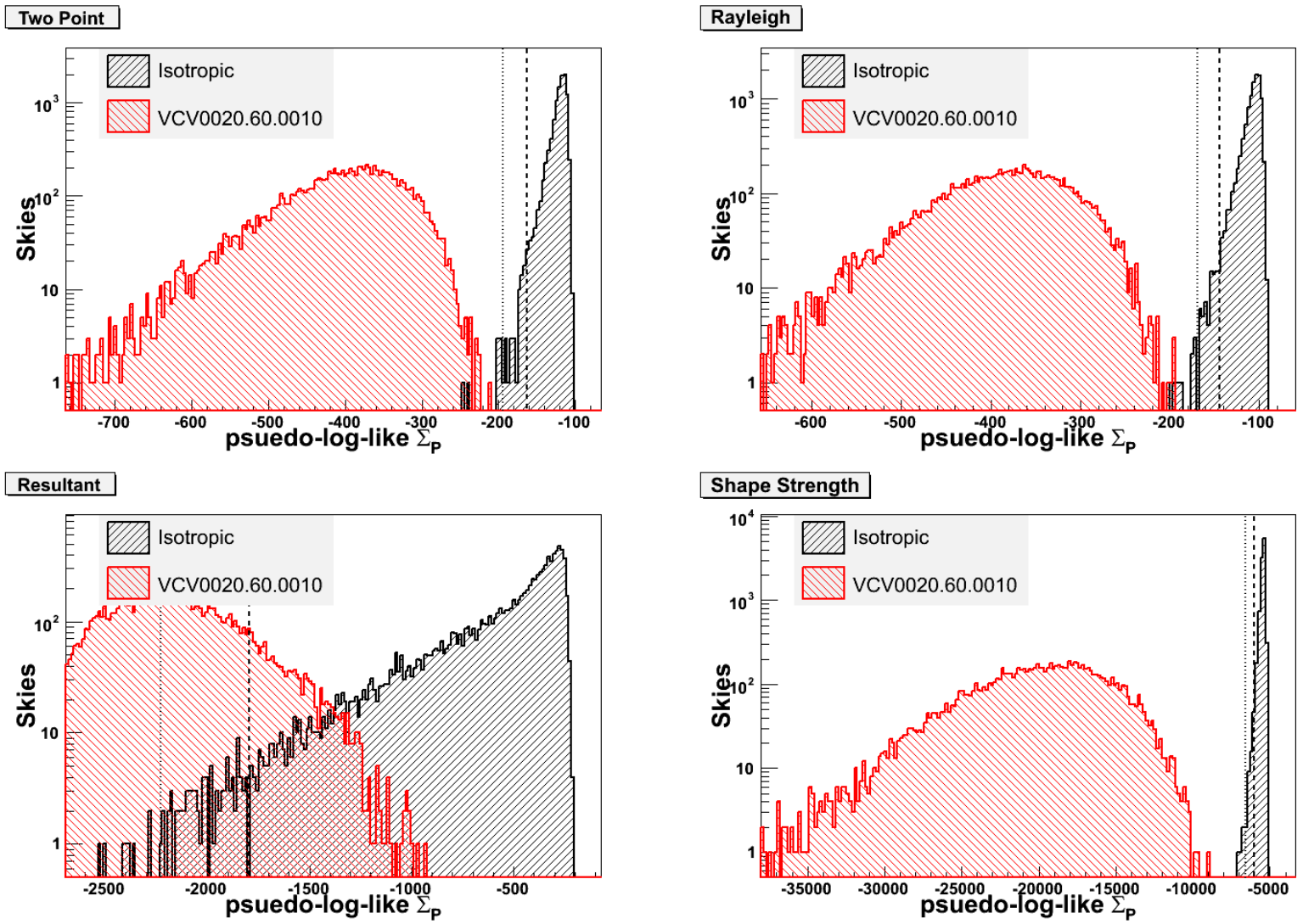


Figure B.31: Histograms of each metric's pseudo-log-likelihood Σ_P for two ensembles of MC skies; ISOTROPY (black, positive slope hatching) and VCV0020 (red, negative slope hatching), see Table B.4. The long and short dashed vertical lines define two signal regions and correspond to ${\Sigma_P}^{\alpha_{01}}$ and ${\Sigma_P}^{\alpha_{001}}$ respectively. One(one-tenth) percent of the isotropic likelihoods are less than ${\Sigma_P}^{\alpha_{01}}({\Sigma_P}^{\alpha_{001}})$. The efficiencies calculated using these histograms are plotted above the VCV0020 label in Figure B.27.

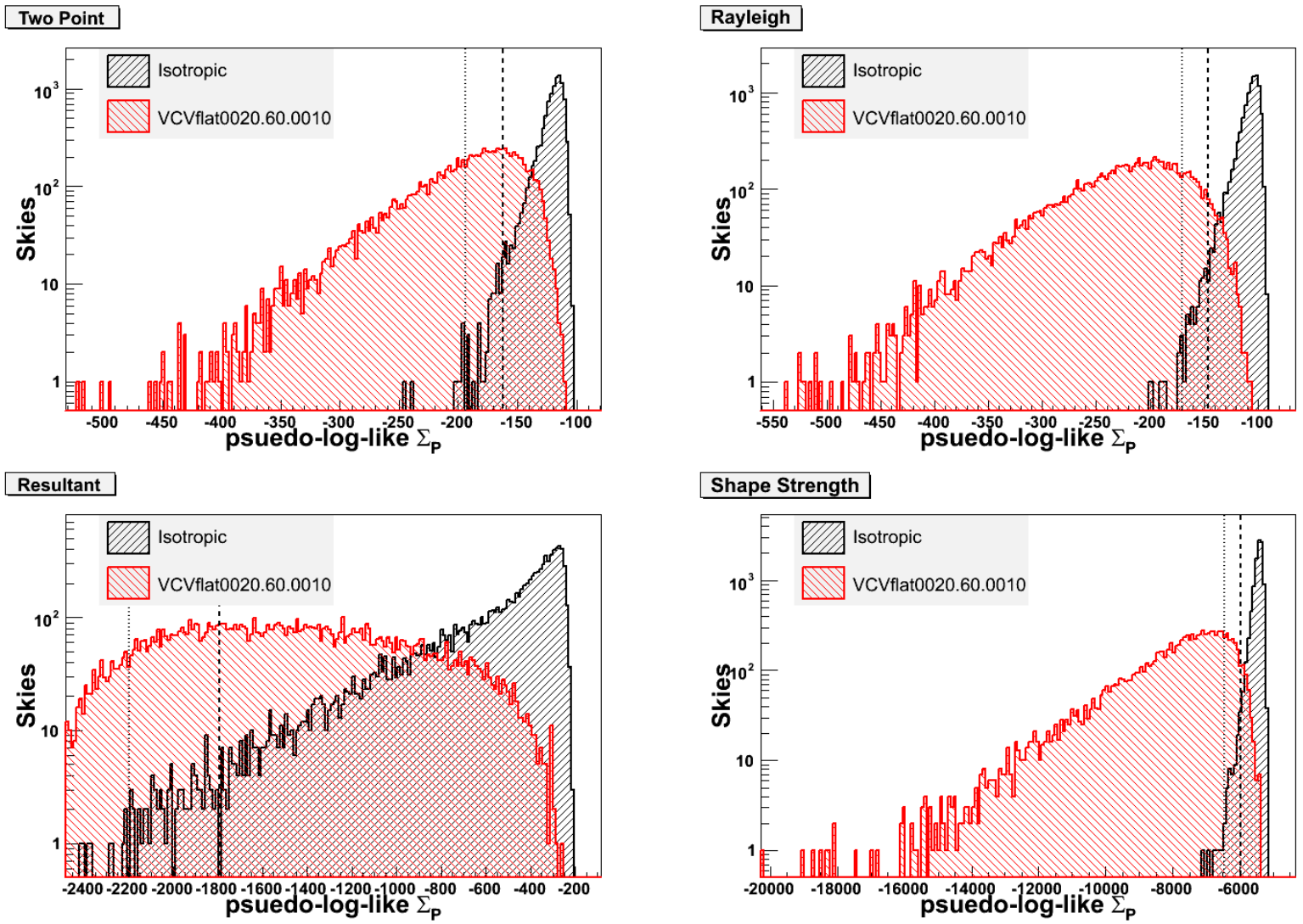


Figure B.32: Histograms of each metric's pseudo-log-likelihood Σ_P for two ensembles of MC skies; ISOTROPY (black, positive slope hatching) and VCVflat0020 (red, negative slope hatching), see Table B.4. The long and short dashed vertical lines define two signal regions and correspond to ${\Sigma_P}^{\alpha_{01}}$ and ${\Sigma_P}^{\alpha_{001}}$ respectively. One(one-tenth) percent of the isotropic likelihoods are less than ${\Sigma_P}^{\alpha_{01}}({\Sigma_P}^{\alpha_{001}})$. The efficiencies calculated using these histograms are plotted above the VCVflat0020 label in Figure B.27.

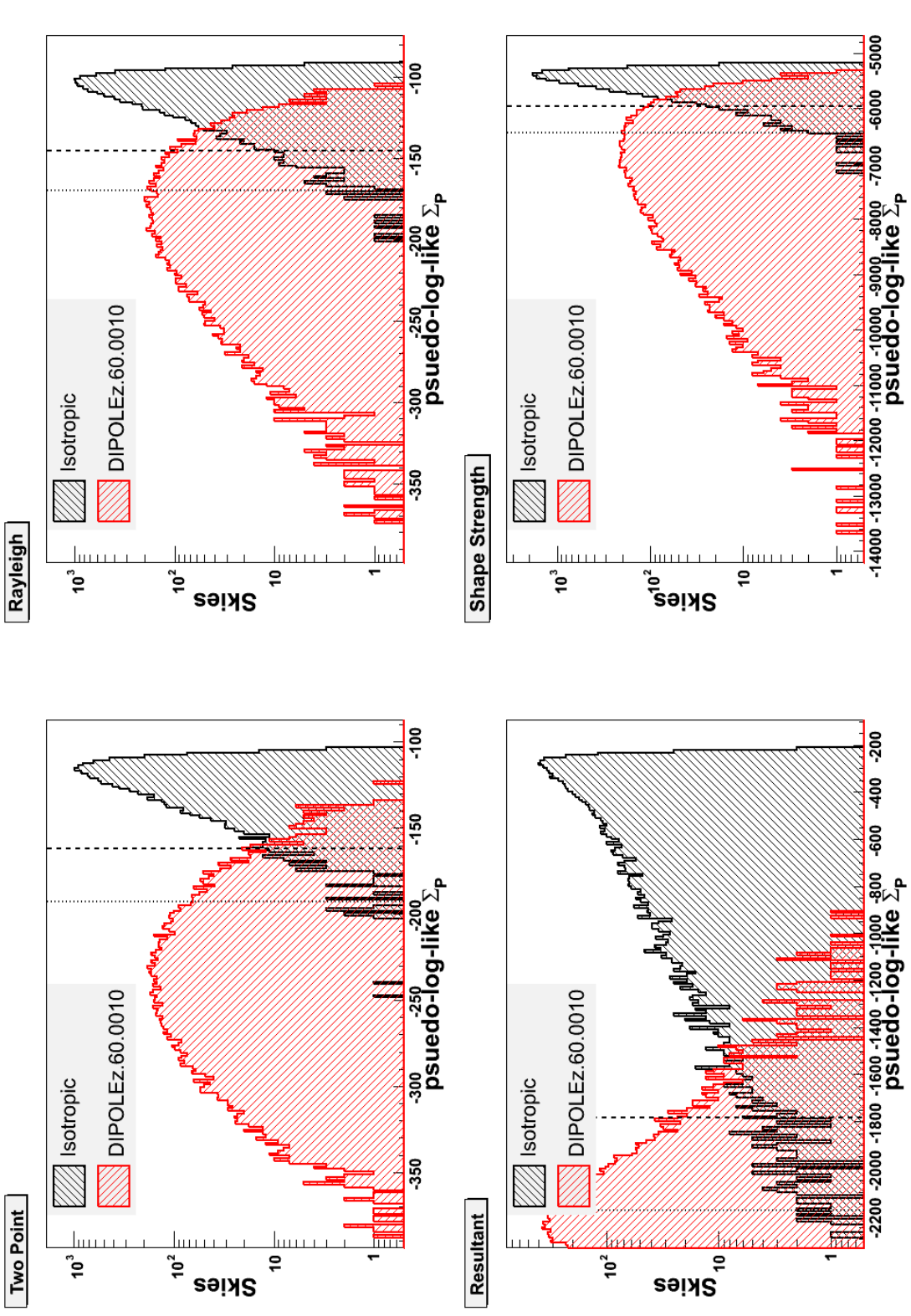


Figure B.33: *Bottom right panel:* Histograms of the **SS** metric pseudo-log-likelihood Σ_P for two ensembles of MC skies; ISOTROPY (black, positive slope hatching) and DIPOLEz (red, negative slope hatching), see Table B.4. The long and short dashed vertical lines define two signal regions and correspond to ${\Sigma_P}^{\alpha_{01}} = -5940$ and ${\Sigma_P}^{\alpha_{001}} = -6416$ respectively. One(one-tenth) percent of the isotropic likelihoods are less than ${\Sigma_P}^{\alpha_{01}}({\Sigma_P}^{\alpha_{001}})$. However, 96%(80%) of the mock signal skies have a Σ_P value less than these values, see the solid magenta star above the DIPOLEz label in Figure B.27. Thus the efficiency of the SS method for detecting the dipole anisotropy is 96%(80%) at the 1%(0.1%) confidence level. *Other panels:* The distribution of Σ_P for the other three metrics.

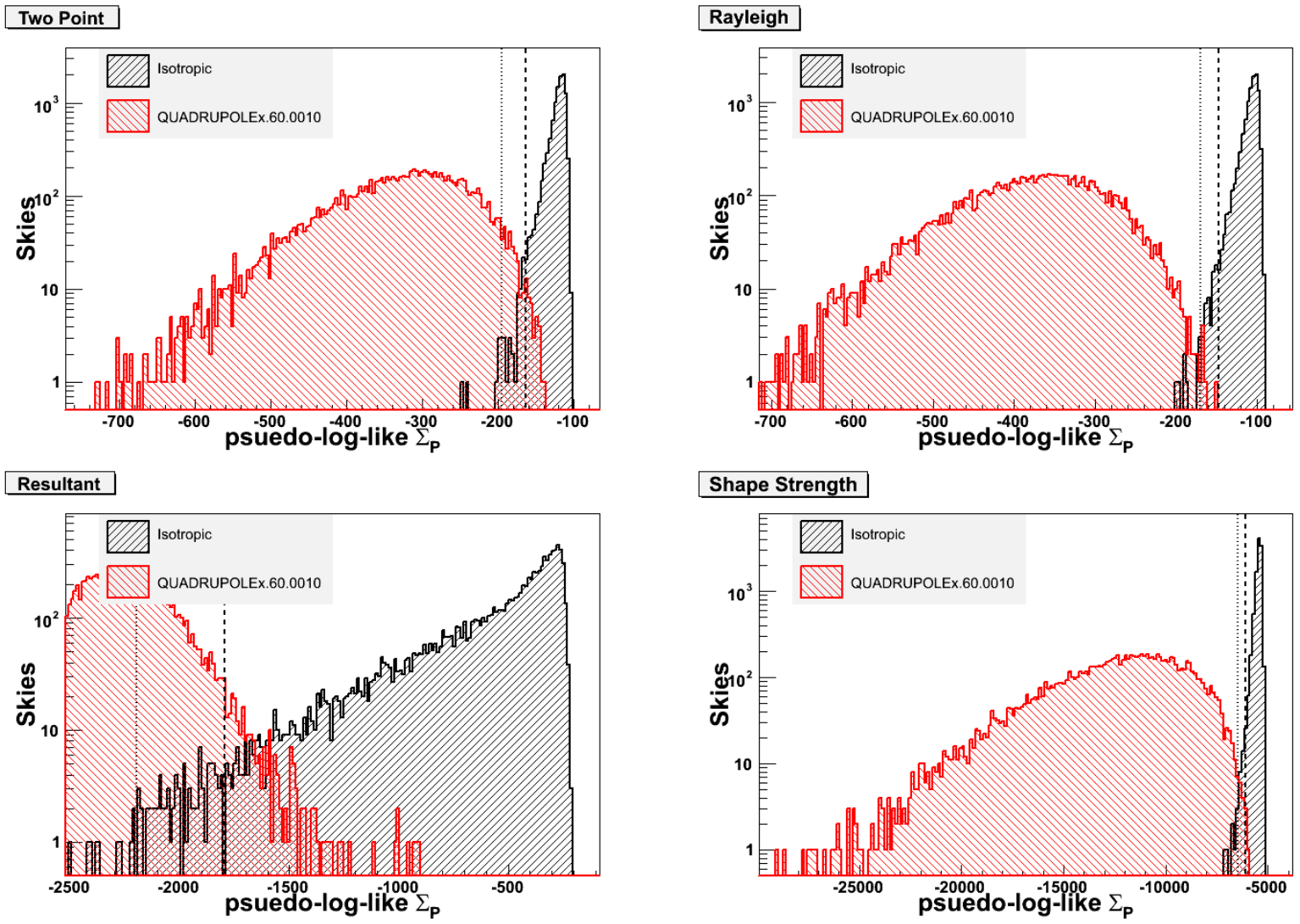


Figure B.34: Histograms of each metric's pseudo-log-likelihood Σ_P for two ensembles of MC skies; ISOTROPY (black, positive slope hatching) and QUADRAPOLEx (red, negative slope hatching), see Table B.4. The long and short dashed vertical lines define two signal regions and correspond to ${\Sigma_P}^{\alpha_{01}}$ and ${\Sigma_P}^{\alpha_{001}}$ respectively. One(one-tenth) percent of the isotropic likelihoods are less than ${\Sigma_P}^{\alpha_{01}}({\Sigma_P}^{\alpha_{001}})$. The efficiencies calculated using these histograms are plotted above the QUADRAPOLEx label in Figure B.27.

B.4 Post August 2007 Data Analysis

B.4.1 Section Overview

This section was originally published as [101].

In this note we apply the methods we have previously investigated[97, 98, 99, 100] to data collected after the cutoff date for the *Science* publication[8, 9] and through Oct. 31, 2008. We present possible interpretations of the collected data. While there are examples of anisotropy in PAO data, the strongest conclusion that can be reached to date is that we need more data, as the anisotropy tests in the current data set are subject to large statistical fluctuations.

In previous GAP notes we have described four anisotropy search techniques[97, 99], studied their sensitivity to various diluted toy signal types[98] and applied them to PAO events arriving before August 31, 2007[100]. In this note we add fourteen more months of data; up to October 31, 2008. Each anisotropy search method (Two Point-2pt, modified 'Chicago Rayleigh'-Rayleigh, Resultant-R^2, and the Shape Strength-SS) yields a p-value for rejecting isotropy. The reporting and interpretation of these values for various subsets of the data is the primary focus of this note.

B.4.2 Data

We use CDAS Herald v4r62f[111] data with arrival dates between 1/1/04 and 31/8/08 (dd/mm/yy). For the event energy we use `$39` $\equiv$ CIC Energy with latest FD calibration. We make awk style quality cuts on the Herald, `$2>0 && $22 && (($26 && $36)||($23>1))`, in addition to restricting the zenith angle[7], `$3>0 && $3<60`. We also check that the data does not fall within a bad period as defined in [112].

[7] The 'standard' cuts.

The subsets of the Herald data studied in this note are illustrated in Figure B.35. The *inclusive* data consists of the 20, 30, ..., 80 most energetic events. We also study a collection of *exclusive* subsets based on the energy ordered events; 4 sets of 20, 2 sets of 30 and 2 sets of 40. Particular attention is paid to the events with energies exceeding 57EeV, of which there are 38. The 24 events with $E \geq 57$EeV arriving before 31/8/07 are called the "Pre" set and the 27 events arriving after 27/5/06 are called the "Post" set. [8] Dividing the 60 most energetic events in half chronologically yields two more subsets of the Herald data (here denoted "1^{st} 30" and "2^{nd} 30"). This gives 17 subsets in total and they are (of course) correlated.

B.4.3 Results

Following the type of analysis done in [100] we calculate the p-values for the data subsets described above.

Inclusive by Energy

Figure B.37 shows the p-values for the inclusive data sets. Some trends are clear. All 4 methods have their p-values drop from 20 to 30 events and drop again from 30 to the 40 highest energy events. Furthermore, all four methods also have their p-values rise from 60 to 70 and rise again from 70 to 80 highest energy events.

[8] The "Pre" and "Post" data sets include common events are in this sense not exclusive. However, they are subsets of the data which are related to the three stages of data collection which make up the current data set. There is the first set of scanned data, which is known to be biased as it is based off a scan, the second set of data which is known to anisotropic as it was the basis of the Science (1 %) result and the third set of data which is everything after. The first and second set are combined to make the 'pre' data set and the second and third are combined to make the 'post' data set. This is done because each by itself has too few events to have any chance to show a significance, while combined these sets are large enough to show a possible anisotropy. The penalty of this is a correlation which limits the interpretation.

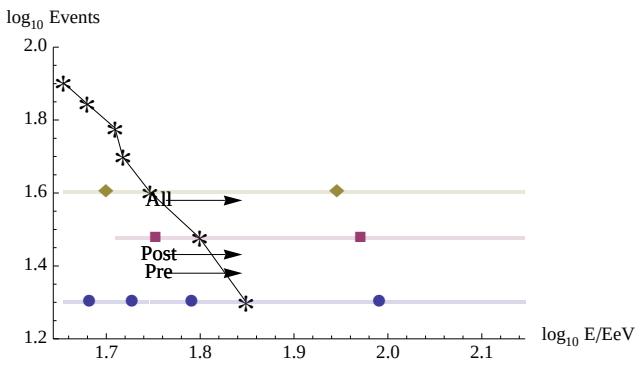


Figure B.35: Schematic of the subsets of the data studied in this note. The asterisks show the minimum energy and number of events for the *inclusive* data. The *exclusive* sets are represented by the bins; blue circles for the 4 sets of 20, red squares for the 2 sets of 30 and yellow diamonds for the 2 sets of 40. The words show a time cut; "All" shows the number of events greater than 57EeV between 1/1/04 and 31/8/08, "Pre" those between 1/1/04 and 31/8/07 and "Post" those between 27/5/06 and 31/8/08.

If we apply a 1% test (as in [8, 9]) there is some evidence of anisotropy for events with a minimum energy of a little over 50 EeV. In particular there are 5 p-values that are smaller than 1%: the SS method for 40, 50 and 60 events and the Rayleigh method for 40 and 60 events (the Rayleigh with 50 events is 1.2%). However, it should be remembered that each of these data sets are correlated with each other as well as including events that were found in the Science paper. From the work with mock MC maps done in[97, 98], we expect that the SS method and Rayleigh method have the best sensitivity to a variety of signals.

Exclusive by Energy

Figure B.38 shows the p-values for the exclusive (by energy) data subsets. Applying a 1% test to these subsets yields two sets where there is evidence of anisotropy; the 40 most energetic events with the Rayleigh and S-S method.

Exclusive by Time

Figure B.39 shows the p-values for the exclusive (by time) data subsets. Applying a 1% test to these subsets yields 5 subsets where there is evidence of anisotropy; "All," "Pre" and "1^{st} 30" using the S-S method, and "Pre" and "1^{st} 30" using the Rayleigh method.

It should also be noted that all of these include the events which were scanned during the initial search and are therefore biased. The 'Post' set (27 events) and 2^{nd} 30 events include no scanned data and are not consistent with a 1% claim in any method. There is over all very little independent evidence of anisotropy in these data sets.

B.4.4 Interpretations

Isotropy + Signal

The most striking feature apparent in comparing of the exclusive and inclusive data sets is that two exclusive data sets can have large p-values but when combined (in an inclusive manner) the p-value decreases dramatically. For example, the top 60 events have nominally strong evidence of anisotropy with a very small p-value while the individual 30 events sets do not. Given the previous work with the mock maps this is not a total surprise[98] and the studies here give credence to idea that a real signal could behave this way. It is important to note that since multiple combinations of events have been examined, the p-values should be subject to an unknown penalty factor.

A simple study was done to examine this naive p-value puzzle and the results are shown in Figure B.36. The "1^{st} 30" events (p_{ss}−-value= 0.5%) were combined with 30 events from an isotropic distribution and this was simulated 1000 times. Using the

S-S method it is found that $\sim 7\%$ of the generated p-values are as small or smaller than the value obtained from the 60 most energetic events ("1^{st} 30" + "2^{nd} 30" gives $p = 0.3\%$). The other three methods have $\sim 3\%$ of the generated p-values as small or smaller than the value obtained from the 60 most energetic events. This suggests that it is entirely possible that the 1^{st} 30 events (which includes scanned data) and an isotropic set of 30 events could produce a set of 60 events with a smaller p-value then the initial 30 events. There is no strong evidence of anisotropy in any data which does not include energy scanned data. This issue can be resolved only with additional data.

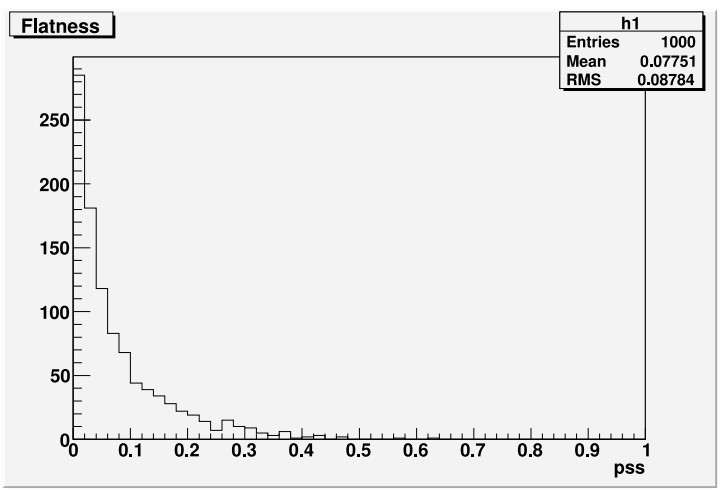


Figure B.36: A histogram of the shape strength test p-values for 1000 skies of 60 events that consist of the "1^{st} 30" events in addition to 30 events drawn from an isotropic distribution. Approximately 7% of the generated p-values are as small or smaller than the value obtained from the 60 most energetic events ($p = 0.3\%$). This is evidence that the data is consistent with the hypothesis that the "1^{st} 30" events ($p = 0.5\%$) have a signal, "the 2^{nd} 30" ($p = 36\%$) events do not and yet the combination still shows a stronger signal than the "1^{st} 30" events alone.

Mock Signals

Given that we have access to toy models of different types of signal, an interesting question is how consistent is our data with these models? This is addressed in Table B.4 using four diluted mock signal types: a dipole (DIPOLEz), a quadrapole (QUADRUPOLEx) a toy signal generated using the redshift weighted VCV catalog (VCV0020) and un-weighted VCV catalog (VCVflat0020). These four toy signal types are considered with 4 dilution factors r: 20%, 40%, 60% and 80% signal fractions. We use 10^4 skies with 60 events per sky and the data which these maps are compared against are the 60 highest energy 60. We choose 60 events because the 60 most energetic Herald events have a p-value smaller than the most energetic 20 or 40. [9]

We judge the toy signal types based on the fraction of mock MC maps that have smaller p-values than the data (are more anisotropic then the data). For example, if 99.0% of the mock maps from a particular signal have a smaller p-value than the data then this would suggest that the data did not come from this distribution. It should be noted that these models have many assumptions – magnetic field distributions, source redshift distributions and source strength distributions to name a few – and so this only shows a 'proof of principle' that it is possible to eliminate models in this manner. The numbers reported are not an true limit.

B.4.5 Conclusions

Using the shape-strength method the events with $E \geq 57$EeV arriving between 1/1/04 and 31/10/08 have a p-value of 0.4%, consistent with the 1% result published

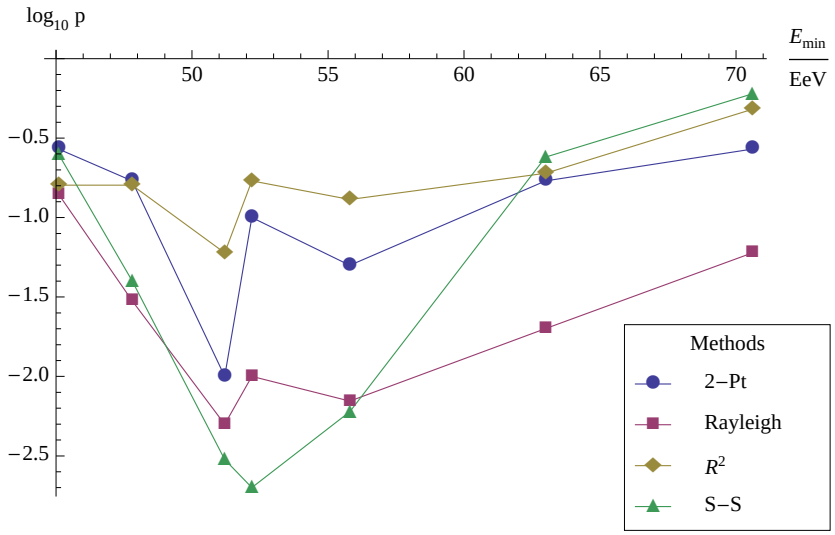
[9] Thus, the results presented should have a penalty factor included as this is a biased manner in which to select events. The question on what selection should be chosen is not simple and will not be dealt with here.

	Fraction of $p_{MC} \leq p_{data}$			
Mock Type	$r = 20\%$	$r = 40\%$	$r = 60\%$	$r = 80\%$
DIPOLEz	0.0114	0.0509	0.2030	0.5502
QUADRUPOLEx	0.0244	0.2020	0.6548	0.9608
VCV0020	0.0820	0.6744	0.9895	1.00
VCVflat0020	0.0151	0.0963	0.3530	0.7274

Table B.4: This table shows a comparison of certain models (with varying signal fraction r) to the 60 most energetic events. The VCV0020 model seems to be ruled out at a fraction of 80% signal (all mock files have a smaller p-value) and even 60% fractions seem unlikely (98.9 % smaller). None of the other models seem ruled out, although the quadrapole 80% is unlikely.

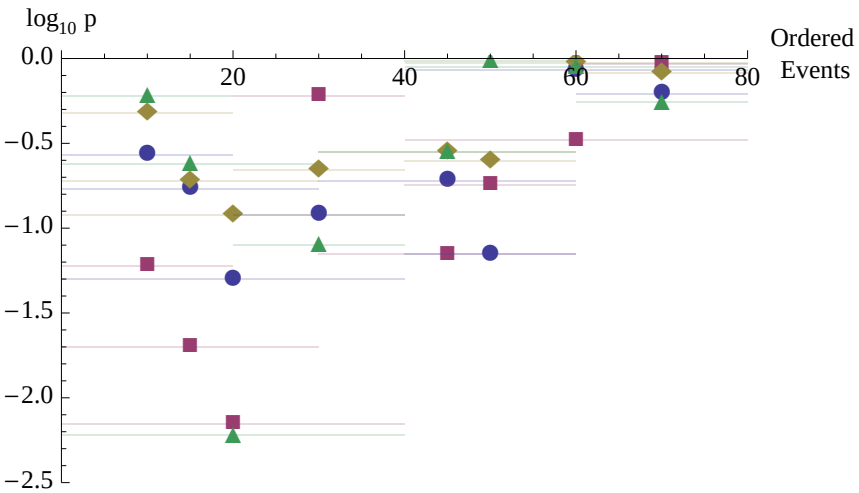
by Auger[8, 9]. When the scanned events are removed from this sample, however, the p-value increases to 22% which is certainly not consistent with the claim of anisotropy. Scanning the exclusive subsets in Figure B.37 a minimum p-value (0.2%) is found for the 50 most energetic events ($E \geq 52.2$EeV) using the S-S method. For the 60 most energetic events we get $p = 0.3\%$. The 30 highest energy events give $p = 24\%$ and the next 30 by energy give $p = 28\%$. Dividing the 60 chronologically the 1^{st} 30 events give $p = 0.5\%$ and the 2^{nd} 30 events give $p = 36\%$. When combining the "1^{st} 30" events with 30 events from an isotropic distribution (and simulating 1000 times) it is found that $\sim 7\%$ of the generated p-values are as small or smaller than the value obtained from the 60 most energetic events.

Taken as a whole, the evidence in this note suggests that data sets as small as 30 events are sensitive to noise and have unstable values and this conclusion is only made much stronger by the toy signal studies[98]. Furthermore, in order to find a (real) physical signal, a larger data set is needed reduce the importance of these fluctuations.



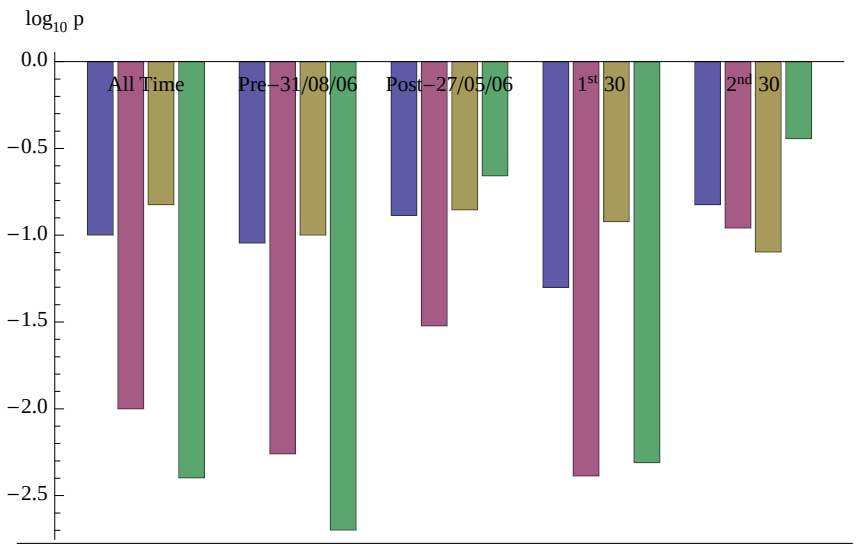
		p-value (%)			
Data Set	E_{min}/EeV	2-Pt	Rayleigh	R^2	S-S
Top 20	70.6	59.605	6.485	47.885	27.170
Top 30	63.0	16.965	2.155	18.720	24.135
Top 40	55.8	4.990	0.745	12.865	0.570
Top 50	52.2	9.865	1.185	16.720	0.160
Top 60	51.2	3.260	0.570	6.250	0.345
Top 70	47.8	16.635	3.345	16.375	3.735
Top 80	45.1	27.290	14.350	16.430	25.920

Figure B.37: The p-values of the four methods applied to the 20, 30, ..., 80 highest energy events for data through October 31, 2008. In the plot the 20 highest energy events are on the right and the 80 highest energy events are on the left. The right most column of the table is the energy in EeV of the lowest energy event in the subset. There appears to be a minimum in the probability between 40 and 60 events. It is possible with less then 40 events the statistics are simply not sufficient to see a signal. On the other hand at 70 events the probability goes up, this could be evidence of the signal vanishing because it becomes isotropic at lower energies. It should be noted that all of these data sets are correlated with each other.



			p-value (%)			
Total Events	Data Set	ΔE/EeV	2-Pt	Rayleigh	R^2	S-S
20	1-20	144-68.4	27.170	6.485	47.885	59.605
	21-40	68.4-55.8	12.530	6.425	22.420	8.340
	41-60	55.6-51.2	6.760	18.175	25.175	97.460
	61-80	51.2-45.1	62.525	93.365	81.715	55.180
30	1-30	144-62.4	16.965	2.155	18.720	24.135
	31-60	62.4-51.2	18.530	6.510	28.485	28.340
40	1-40	144-55.6	4.990	0.745	12.865	0.570
	41-80	55.6-45.1	85.050	32.735	94.770	89.165

Figure B.38: The *p*-values of the four methods applied to the exclusive by energy data subsets. The plot labels are the same as those in Figure B.37. The first row of the table, for example, shows the numerical values obtained for the 20 most energetic events, the second row shows the values for the next 20 events (by energy) and so on.



		p-value (%)			
Data Set	E_{min}/EeV	2-Pt	Rayleigh	R^2	S-S
All	57.6	9.645	1.400	15.735	0.445
Pre	57.6	9.250	0.550	10.335	0.185
Post	57.6	12.560	3.820	14.105	22.010
1^{st} 30	51.2	5.290	0.405	12.415	0.495
2^{nd} 30	51.3	15.275	11.540	7.670	35.850

Figure B.39: The p-values of the four methods applied to the exclusive by time data subsets. "All" is for all of the38 events with $E \geq 57$EeV, "Pre" is for the 24 events before 31/8/07 and "Post" is for the 27 events after 27/5/06 (note that "Pre" and "Post" are not independent sets). The 1^{st} 30 events are the first 30 events chronologically of 60 most energetic events. The 2^{nd} 30 events are the last 30 events chronologically of 60 most energetic events.

References

[1] Cosmic Ray wiki, June 2009. en.wikipedia.org/wiki/Cosmic_ray.

[2] Photo of Victor Hess in a balloon. pamela.physik.uni-siegen.de/pamela/bitmaps/viktorhess_001.jpg.

[3] John Linsley, Livio Scarsi, and Bruno Rossi. Extremely energetic cosmic-ray event. *Phys. Rev. Lett.*, 6(9):485–487, May 1961.

[4] Photo of John Linsley at Volcano Ranch. www.auger.org/photos/photo_archive/images/volcano_ranch.72dpi.jpg.

[5] J. Linsley. Evidence For A Primary Cosmic-Ray Particle With Energy 10^{20}-Ev. *Phys. Rev. Lett.*, 10:146, 1963.

[6] B.E. Smith. The Mass Composition of Cosmic Rays Above 1 EeV Inferred Using the Spread in Arrival Times of Air Shower Particles. *GAP Notes – A Thesis*, 2008. GAP2008-161.

[7] J. D. Hague, B. R. Becker, M. S. Gold, and J. A. J. Matthews. Power Laws and the Cosmic Ray Energy Spectrum. *Astropart. Phys.*, 27:455–464, 2007.

[8] The Pierre Auger Collaboration. Correlation of the highest energy cosmic rays with nearby extragalactic objects. *Science*, 318:938–943, 2007. arXiv:0711.2256v1 [astro-ph].

[9] The Pierre Auger Collaboration. Correlation of the highest-energy cosmic rays with the positions of nearby active galactic nuclei. *Astropart. Phys.*, 29:188–204, 2008. arXiv:0712.2843v2 [astro-ph].

[10] The Pierre Auger Collaboration. Upper limit on the cosmic-ray photon fraction at EeV energies from the Pierre Auger Observatory. 2009. `arXiv:0903.1127v1 [astro-ph.HE]`.

[11] M. Lemoine et al. Physics and Astrophysics of Ultra-High-Energy Cosmic Rays. *Lecture Notes in Physics*, 576, 2001.

[12] F. Schussler. Measurement of the Energy Spectrum of Ultra-High Energy Cosmic Rays using Hybrid Data of the Pierre Auger Observatory. *GAP Notes – A Thesis*, 2008. GAP2008-155.

[13] Kenneth Greisen. End to the cosmic-ray spectrum? *Phys. Rev. Lett.*, 16(17):748–750, Apr 1966.

[14] G. T. Zatsepin and V. A. Kuzmin. Upper limit of the spectrum of cosmic rays. *JETP Lett.*, 4:78–80, 1966.

[15] A. Penzias and R. Wilson. A Measurement of Excess Antenna Temperature at 4080 Mc/s. *Astrophysical Journal*, 142:419–421, 1965.

[16] F. Aharonian and J. Cronin. *Phys. Rev. D*, 50:1892, 1994.

[17] S. Lee et al. Extragalactic Magnetic Field and the Highest Energy Cosmic Rays. *Astrophysical Journal*, 455:L21+, 1995.

[18] K. Dolag et al. Constrained simulations of the magnetic eld in the local Universe and the propagation of ultrahigh energy cosmic rays. *Journal of Cosmology and Astroparticle Physics*, 1:9, 2005.

[19] N. Hayashida et al. The anisotropy of cosmic ray arrival directions around 10**18-eV. *Astropart. Phys.*, 10:303–311, 1999.

[20] T. Gaisser and G. Hillas. *Proc. 15th International Cosmic Ray Conference Plovdiv*, 8:353, 1977.

[21] M. Denia. Energy calibration of the Pierre Auger Observatory. Measurement of the spectrum of Ultra-High Energy Cosmic Rays. *GAP Notes – A Thesis*, 2008. GAP2008-167.

[22] D. Allard et al. *29th Int. Cosmic Ray Conf.*, 2005. astro-ph/0510320.

[23] D. Veberic and M. Roth. O line reference manual: Sd reconstruction. 2008. Included in the Offline distribution.

[24] D. Veberic and M. Horvat. On shower-front start-time variance. *GAP Notes*, 2007. GAP-2007-057.

[25] M. Ave et al. Reconstruction accuracy of the surface detector array of the pierre auger observatory. *Proc. ICRC 2007*, 2007. http://arxiv.org/abs/0709.2125v1.

[26] A. M. Hillas. *Proc. 12th ICRC*, 3:1001, 1971.

[27] D. Newton et al. *Astropart. Phys.*, 26:414, 2007.

[28] J. Hersil. *Phys. Rev. Lett.*, 6:22, 1961.

[29] A. Cordero et al. Proposal for the optical system of the fluorescence detector of the auger project. *GAP Notes*, (039), 1996. GAP-1996-039.

[30] M. Ambrosio. *Nucl. Instrum. Meth.*, A478:125, 2002.

[31] A. N. Bunner. Cosmic ray detection by atmospheric fluorescence. 6, 1967. Ph.D. Thesis, Cornell Iniversity.

[32] A. Cordero et al. Shower prole reconstruction from fluorescence and cerenkov light. *GAP Notes*, (010), 2006. GAP-2006-010.

[33] M. Prouza et al. Star tracking using background data of fd telescopes. *GAP Notes*, (041), 2005. GAP-2005-041.

[34] F. Kakimoto et al. A measurement of the air uorescence yield. *Nucl. Instr. Meth.*, page 527533, 1996.

[35] M. Ave et al. Temperature and humidity dependence of air fluorescence yield measured by airfly. *Proc. 5th Fluorescence Workshop, El Escorial - Madrid.*, 2007.

[36] The Pierre Auger Collaboration. Observation of the suppression of the flux of cosmic rays above 4×10^{19}eV. *Phys. Rev. Lett.*, 101:061101, 2008. arXiv:0806.4302v1.

[37] M. E. J. Newman. Power Laws, Pareto distributions and Zipf's Law. *Contemporary Physics*, 46:323–351, 2005.

[38] A. Clauset, M. Young, and K.S. Gleditsch. On the Frequency of Severe Terrorist Events, 2007. arXiv:physics/0606007v3 [physics.soc-ph].

[39] John N. Bahcall and Eli Waxman. Has the GZK cutoff been discovered? *Phys. Lett.*, B556:1–6, 2003. arXiv:hep-ph/0206217v5.

[40] M. Takeda et al. Extension of the cosmic-ray energy spectrum beyond the predicted Greisen-Zatsepin-Kuzmin cutoff. *Phys. Rev. Lett.*, 81:1163–1166, 1998. arXiv:astro-ph/9807193v1.

[41] Paul M. Mantsch. The Pierre Auger Observatory: Progress and first results. 2005. arXiv:astro-ph/0604114v1.

[42] Tokonatsu Yamamoto. The first scientific results from the Pierre Auger observatory. *AIP Conf. Proc.*, 842:1016–1018, 2006. arXiv:astro-ph/0601035v1.

[43] M. L. Goldstein, S. A. Morris, and G. G. Yen. Problems with fitting to the power-law distribution. *The European Physical Journal B - Condensed Matter and Complex Systems*, 41(2):255–258, 2004.

[44] L. W. Howell. Statistical Properties of Maximum Likelihood Estimators of Power Law Spectra Information. *NASA/TP-2002-212020/REV1, Marshall Space Flight Center*, 2002. http://www.sti.nasa.gov.

[45] V. Pisarenko, D. Sornette, and M. Rodkin. Deviations of the Distributions of Seiesmic Energies from the Gutenberg-Richter Law. *Computational Seismology*, 35:138–159, 2004. arXiv:physics/0312020v1 [physics.data-an].

[46] V. Pisarenko and D. Sornette. New statistic for financial return distributions: power-law or exponential? *Physica A*, 366:387–400, 20046. arXiv:physics/0312020v1 [physics.data-an].

[47] T. Yamamoto et al. The UHECR spectrum measured at the Pierre Auger Observatory and its astrophysical implications. 2007. arXiv:0707.2638v1.

[48] R. Abbasi et al. Observation of the GZK cutoff by the HiRes experiment. *Phys. Rev. Lett.*, 100:101101, 2008. arXiv:astro-ph/0703099v2 [astro-ph].

[49] S. Eidelman et al. (Particle Data Group). Review of Particle Physics. *Phys. Lett. B*, 592:1+, 2004. statrpp.pdf.

[50] A. Clauset, C.R. Shalizi, and E.J. Newman. Power-law Distributions in Emperical Data, 2007. arXiv:0706.1062v1 [physics.data-an].

[51] Eric Armengaud, Gunter Sigl, Tristan Beau, and Francesco Miniati. CRPropa: A numerical tool for the propagation of UHE cosmic rays, gamma-rays and neutrinos. *Astropart. Phys.*, 28:463–471, 2007. astro-ph/0603675.

[52] J. D. Hague, Bernard Raymond Becker, Michael S. Gold, J. A. J. Matthews, and J. Urbar. Statistical Methods for Investigating the Cosmic Ray Energy Spectrum. *ICRC-07 Abs.Num.1217*, 2007.

[53] I. Zehavi et. al. On Departures From a Power Law in the Galaxy Correlation Function. *arXiv*, 2005. arXiv:astro-ph/0508499v1.

[54] M. Schroedter et. al. A Very High Energy Gamma-Ray Spectrum of 1ES 2344+514. *arXiv*, 2005. arXiv:astro-ph/0508499v1.

[55] V. S. Berezinsky and S. I. Grigorieva. *Astro. and Astrophys.*, 199(1), 1988.

[56] F. James and M. Roos. Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations. *Comput. Phys. Commun.*, 10:343–367, 1975. We use the CERN root-systems v5.1 implementation available at http://www.root.org.

[57] Quang H. Vuong. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, 57(2):307–333, 1989.

[58] Auger Observatory. $\overline{\text{Off}}\underline{\text{line}}$ Observer, July 2008. http://augerobserver.fzk.de/index.php?id=16.

[59] Auger Collaboration. 29th ICRC. *Pune, India*, 10:115, 2005. www.auger.org/technical_info/pdfs/icrc2005/icrc_2005_124.pdf.

[60] Antoine Letessier-Selvon. A question. Auger Spectrum listserve ¡auger_spectrum@fnal.gov¿, June 2008.

[61] Antoine Letessier-Selvon, Email to AAARG list server, May 28, 2007.

[62] C. Covault and A. Letessier-Selvon. A prescription. *GAP Notes*, (010), 2006.

[63] B. Connolly et al. Comments on the prescription. *GAP Notes*, (093), 2007.

[64] J. D. Hague, B. R. Becker, M. S. Gold, and J. A. J. Matthews. A Three-Point Cosmic Ray Anisotropy Method. 2009. `arXiv:0905.4488v2 [astro-ph.IM]`.

[65] The Pierre Auger Collaboration. Upper limit on the cosmic-ray photon flux above 10^{19} eV using the surface detector of the Pierre Auger Observatory. *Astropart. Phys.*, 29:243–256, 2008. arXiv:astro-ph/0602153v2.

[66] The Pierre Auger Collaboration. An upper limit to the photon fraction in cosmic rays above 10^{19}-eV from the Pierre Auger observatory. *Astropart. Phys.*, 27:155–168, 2007. arXiv:astro-ph/0606619v2.

[67] S. W. Barwick et al. Constraints on cosmic neutrino fluxes from the ANITA experiment. *Phys. Rev. Lett.*, 96:171101, 2006. arXiv:astro-ph/0512265v2.

[68] P. Gorham et al. New Limits on the Ultra-high Energy Cosmic Neutrino Flux from the ANITA Experiment. 2008. arXiv:0812.2715v1 [astro-ph].

[69] M.-P. Véron-Cetty and P. Véron. A catalogue of quasars and active nuclei: 12th edition. *Astron. Astrophys.*, 455:773–777, 2006.

[70] Will Saunders et al. The PSCz Catalogue. *Mon. Not. Roy. Astron. Soc.*, 317:55, 2000. astro-ph/0001117.

[71] N. Hayashida et al. Possible clustering of the most energetic cosmic rays within a limited space angle observed by the akeno giant air shower array. *Phys. Rev. Lett.*, 77(6):1000–1003, Aug 1996.

[72] M. Takeda et al. Small-scale anisotropy of cosmic rays above 10^{19}-eV observed with the Akeno Giant Air Shower Array. *Astrophys. J.*, 522:225–237, 1999. arXiv:astro-ph/9902239v2.

[73] Michael Kachelriess and D. V. Semikoz. Clustering of ultra-high energy cosmic ray arrival directions on medium scales. *Astropart. Phys.*, 26:10–15, 2006. arXiv:astro-ph/0512498v2.

[74] Ronnie Jansson and Glennys R. Farrar. Maximum Likelihood Method for Cross Correlations with Astrophysical Sources. *JCAP*, 0806:017, 2008. arXiv:0711.0177v2 [astro-ph].

[75] Silvia Mollerach. Studies of clustering in the arrival directions of cosmic rays detected at the Pierre Auger Observatory above 10 EeV. 2007. arXiv:0706.1749v1 [astro-ph].

[76] Silvia Mollerach and the Pierre Auger Collaboration. Search for clustering of ultra high energy cosmic rays from the Pierre Auger Observatory. 2009. Contribution to the Proceedings of the CRIS 2008 Conference. `arXiv:0901.4699v1 [astro-ph.HE]`.

[77] Hylke B. J. Koers and Peter Tinyakov. Testing large-scale (an)isotropy of ultra-high energy cosmic rays. 2008. arXiv:0812.0860v1 [astro-ph].

[78] Diego Harari, Silvia Mollerach, and Esteban Roulet. Astrophysical magnetic field reconstruction and spectroscopy with ultra high energy cosmic rays. *JHEP*, 07:006, 2002. arXiv:astro-ph/0205484v1.

[79] Diego Harari, Silvia Mollerach, and Esteban Roulet. Detecting filaments in the ultra-high energy cosmic ray distribution. *Astropart. Phys.*, 25:412–418, 2006. arXiv:astro-ph/0602153v2.

[80] The isotropic expectation is typically computed by considering an ensemble of 10^4 to 10^5 isotropic skies with the same number of events per sky and weighted by the Auger exposure[9].

[81] The tail (large n_{exp}) of the true distribution in each bin falls somewhat less steeply than a Poisson tail. This approximation is made to save computational time and complexity, and results in slightly conservative p-values.

[82] Alessandro Cuoco, Steen Hannestad, Troels Haugboelle, Michael Kachelriess, and Pasquale D. Serpico. Clustering properties of ultrahigh energy cosmic rays and the search for their astrophysical sources. *Astrophys. J.*, 676:807, 2008. arXiv:0709.2712v2 [astro-ph].

[83] N.I. Fisher, T. Lewis, and B.J.J. Embleton. *Statistical Analysis of Spherical Data*. Cambridge University Press, 1987. ISBN-13: 9780521456999.

[84] N.H. Woodcock. Specification of fabric shapes using an eigenvalue method. *Geological Society of America Bulletin*, 88(9):1231–1236, 1977.

[85] M.A. Woodcock, N.H. Naylor. Randomness testing in three-dimensional orientation data. *Journal of Structural Geology*, 5(5):539–548, 1983.

[86] www.hpc.unm.edu.

[87] John Kent. The Fisher-Bingham distribution on the sphere. *J. Royal. Stat. Soc.*, 44:71–80, 1982. See also the Kent Distribution.

[88] J.D. Hague. Correlation of the highest energy cosmic rays with nearby extragalactic objects in pierre auger observatory data. *Proc. 31 International Cosmic Ray Conference (ICRC)*, (0143), 2009. `arXiv:0906.2347v1 [astro-ph.HE]`, page 6.

[89] J. Aublin. Discriminating potential astrophysical sources of the highest energy cosmic rays with the pierre auger observatory. *Proc. 31 International Cosmic Ray Conference (ICRC)*, 2009.

[90] F. Schussler. Measurement of the cosmic ray energy spectrum above 10^{18} ev with the pierre auger observatory. *Proc. 31 International Cosmic Ray Conference (ICRC)*, 2009.

[91] C. Giulio. Energy calibration of data recorded with the surface detectors of the pierre auger observatory. *Proc. 31 International Cosmic Ray Conference (ICRC)*, 2009.

[92] C. Bonifazi. The angular resolution of the pierre auger observatory. *Nuclear Physics B (Proc. Suppl.)*, 190:20–25, 2009.

[93] A. Wald. *Sequential Analysis*. 2009. John Wiley and Sons, New York.

[94] S. BenZvi et al. The angular resolution of the pierre auger observatory. *The Astrophysical Journal*, 687:1035–1042, 2008. arXiv:0711.3937v2 [astro-ph].

[95] NASA/IPAC Extragalactic Database. http://nedwww.ipac.caltech.edu/.

[96] W.T. Eadie. *Statistical Methods in Experimental Physics*. 1971. North-Holland, Amsterdam.

[97] B. Becker et al. Study of anisotropy metrics. *GAP Notes*, (084), 2008. GAP2008_084.

[98] B. Becker et al. Study of anisotropy metrics ii: Mock signal detection efficiency. *GAP Notes*, (124), 2008. GAP2008_124.

[99] B. Becker et al. Study of anisotropy metrics iii: Summary of methods. *GAP Notes*, (142), 2008. GAP2008_142.

[100] B. Becker et al. Study of anisotropy metrics iv: Application to cdas data through august 2007. *GAP Notes*, (143), 2008. GAP2008_143.

[101] B. Becker et al. Study of anisotropy metrics v: Post august 2007 data and analysis. *GAP Notes*, (014), 2009. GAP2009_014.

[102] PAO. Data Challenge. https://www.auger.unam.mx/AugerWiki/stage_1_CI_Paper_on_Anisotropy.

[103] B. Ave et al. Anisotropy of the highest-energy auger events. *GAP Notes*, (076), 2007.

[104] "Some studies regarding the 2 point rayleigh estimator", D. Harari, S. Mollerach and E. Roulet, GAP-Note-2008-035. "Search for clustering of ultra high energy cosmic rays from the Pierre Auger Observatory", Maximo Ave, Piera Ghia, Aurelio Grillo, Silvia Mollerach, Gap-Note-2008-015. "Searching for Threadlike Multiplets", Pierre Billoir, GAP-Note-2006-071.

[105] B. Becker et. al. Study of Anisotropy Metrics. *GAP Notes*, (084), 2008. http://www.auger.org/admin/GAP_Notes/GAP2008/GAP2008_081.pdf.

[106] M. Ave et. al. Anisotropy of the Highest-Energy Auger Events. *GAP Notes*, (076), 2007. http://www.auger.org/admin/GAP_Notes/GAP2007/GAP2007_076.pdf.

[107] M. Ave et. al. Search for clustering of ultra high energy cosmic rays from the Pierre Auger Observatory. *GAP Notes*, (015), 2008. http://www.auger.org/admin/GAP_Notes/GAP2008/GAP2008_015.pdf.

[108] D. Harari et. al. Some studies regarding the 2 point Rayleigh estimator. *GAP Notes*, (035), 2008. http://www.auger.org/admin/GAP_Notes/GAP2008/GAP2008_035.pdf.

[109] P. Billoir. Searching for Threadlike Multiplets. *GAP Notes*, (071), 2006. http://www.auger.org/admin/GAP_Notes/GAP2006/GAP2006_071.pdf.

[110] PAO. Data Challenge Maps. https://www.auger.unam.mx/AugerWiki/MockMaps.

[111] http://auger.colostate.edu/private/herald/.

[112] http://ipnweb.in2p3.fr/ auger/AugerProtected/Documents/AcceptanceWork/BadPeriods_01