A SEARCH FOR ANISOTROPY IN THE ARRIVAL DIRECTIONS OF ULTRA HIGH ENERGY COSMIC RAYS OBSERVED BY THE HIGH RESOLUTION FLY'S EYE DETECTOR

by

Benjamin Taylor Stokes

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Physics

The University of Utah

August 2006

THE UNIVERSITY OF UTAH GRADUATE SCHOOL

SUPERVISORY COMMITTEE APPROVAL

of a dissertation submitted by

Benjamin Taylor Stokes

This dissertation has been read by each member of the following supervisory committee and by majority vote has been found to be satisfactory.

	Chair: Charles C. H. Jui
	Pierre V. Sokolsky
	Benjamin C. Bromley
	Brian T. Saam
	Don H. Tucker

THE UNIVERSITY OF UTAH GRADUATE SCHOOL

FINAL READING APPROVAL

To the Graduate Council of the University of Utah:

I have read the dissertation of Benjamin Taylor Stokes in its final form and have found that (1) its format, citations, and bibliographic style are consistent and acceptable; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the Supervisory Committee and is ready for submission to The Graduate School.

Date

Charles C. H. Jui
Chair: Supervisory Committee

Approved for the Major Department

Pierre V, Sokolsky
Chair/Director

Approved for the Graduate Council

David S. Chapman
Dean of The Graduate School

ABSTRACT

Although the existence of cosmic rays with energies extending well above 10^{19} eV has been confirmed, their origin remains one of the most important questions in astro-particle physics today. The High Resolution Fly's Eye Detector (HiRes) detects Ultra High Energy Cosmic Rays (UHECRs) by employing the air-fluorescence technique to observe Extensive Air Showers (EAS) in the atmosphere over Dugway, Utah. It has been collecting data since May 1997 during which time over 1500 events with energies greater than $10^{18.5}$ eV have been observed in monocular mode. These events are characterized by arrival directions with asymmetric uncertainties, which are explored in detail for this study. Multiple methods are developed and utilized to search for anisotropies in the arrival directions. A primary emphasis is placed upon previous reported observations that suggested small-scale clustering and global dipole biases. Additionally a new method for searching for anisotropies is explored. While all conclusions are statistically limited in their applicability, the sensitivities are shown to be compatible with prior experiments. However, all evidence to date on the arrival directions of the UHECRs observed by HiRes in monocular mode is consistent with an isotropic distribution.

I would like to dedicate this work to my daughter Phoebe and all of the many other purveyors of sunshine whom I have been fortunate enough to know. You lend meaning to the most irrational thing of all: existence itself.

I would also like to dedicate this work to the memory of my grandfathers:
Allan Carlyle Speirs (1915-2001)
and
Lee Albert Stokes (1911-2005)
May the fields of Elysium suit you well.

CONTENTS

ACKNOWLEDGMENTS

First and foremost, I must thank my advisor Prof. Charles C.H. Jui. While many advisors fill the role of a paternal figure, Charlie was much more of a big brother. He encountered me at point in my life when I had nearly succumbed to self-doubt. He enabled me to believe in myself again. Charlie is the creative genesis behind many of the ideas contained herein. His influence over my development as a scientist cannot be overstated.

I would also like to thank the members of the HiRes Collaboration. Our project administrator, John N. Matthews, has been of immeasurable help in ensuring that this project has continued to collect data and as the proof-reader-in-chief of my publications. I must also thank the other members of the anisotropy group: Jose Bellido, John Belz, Bruce Dawson, Chad Finley, Malina Schindel, and Stefan Westerhoff. Their criticism and advice has been of great value. Furthermore, without the pioneering efforts of Tareq AbuZayyad, this work would have not been possible. Tareq's support as both a colleague and as a friend have been priceless. Finally I must tip my hat to our technical support staff, Mike Cassidy, Gary McDonough, Jeremy Smith, and Stan Thomas. Their often thankless toiling is what keeps this experiment alive.

I also must credit Prof. Glennys Farrar of New York University for many stimulating conversations. Her invaluable insights have led to the refinement of many of my analyses.

I cannot forget the impact of the many great teachers that I was fortunate enough to have in the years leading up to graduate school. I must credit the staff of Timpview High School in Provo, Utah for providing me with a secondary education that was superlative in every way. I also cannot forget the impact of the two professors, Robert Helbling and Gerald Davey, who corroborated to teach my

honors core sequence as an undergraduate student. They taught me what it meant to be truly educated.

My studies have kept me up many long nights and required me to travel for extended periods of time. I want to thank my daughter, Phoebe, for her understanding in this matter. I also cannot the overstate the contribution of Phoebe's mother, Jennie. First as my wife and then as my best friend, no one has done more to support my ambitions.

I wish to thank my father and mother. They always encouraged my sense of scientific curiosity for as long as I can remember. And while this certainly had unforeseen consequences, I will always be appreciative of the intellectual upbringing that they provided me with.

I also want to thank all of the many other colleagues, family members, and friends who have helped keep my chin up over the past seven years. Some of them I haven't seen for a long time, but I have not forgotten them. I will attempt to name just a few (in alphabetical order): Rasha Abbasi, Rob Atkins, Doug Bergman, Kristin Cantrell, Rachel Costello, Adam Everett, John Halpern, Petra Hüntemeyer, Jon Jensen, Ben Jones, Daniel Maslak, Jeremy Meyer, Steffanie Moore, Ruadhan O'Flanagan, Harrison Pope, Elizabeth Rennie, Patrick Rennie, Pierre Sokolsky, Wayne Springer, Clark Stacey, Esther Stokes, Sarah Stokes, Gordon Thomson, Todd VanderVeen, and Andreas Zech. Without the joy that they all added to my life I doubt I would have made it this far.

This research was funded by US NSF grants PHY 9322298, PHY 9321949, PHY 9974537, PHY 0098826, PHY 0071069, PHY 0140688, PHY 0307098, by the DOE grant FG03-92ER40732, and by the Australian Research Council. I gratefully acknowledge the contributions from the University of the Utah Center for High Performance Computing. The cooperation of Colonels E. Fisher and G. Harter, the US Army and the Dugway Proving Ground staff is appreciated.

CHAPTER 1

INTRODUCTION

The observation of cosmic rays with energies in excess of 10^{17} eV, or Ultra-High Energy Cosmic Rays (UHECRs), has now spanned over 40 years. Over that period, many source models have been proposed to explain the origin of these remarkably energetic particles. Yet there has been no conclusive evidence that would vindicate any of these various source models. We know no more about the origins of UHECRs now than we did at the inception of this field of research in 1963 [1].

This work will focus on UHECRs with measure energies above $10^{18.5}$ eV. The estimated integrated flux of UHECRs above this energy is $\sim$10 events km^{-2}sr^{-1}yr^{-1}. That is, we could expect $\sim$10 UHECRs to land within a 1 km^2 surface in one year's time. The rarity of these particles makes it impractical to measure them directly. In this work, we rely upon the air-fluorescence technique. The air-fluorescence technique observes UHECRs by using the Earth's atmosphere as a giant calorimeter. UHECRs interact with the atmosphere to form Extensive Air Showers (EAS). These showers can be observed by ground-based optical apparatuses that are designed to detect to brief flash of scintillation light that occurs when an EAS passes through the atmosphere. This method allows one to observe UHECRs over a much larger volume than would be practical with a direct measurement.

The High Resolution Fly's Eye Experiment (HiRes) [2, 3] consists of two "eyes", HiRes-I and HiRes-II, situated 12 km apart at Dugway, Utah as is shown in Figure 1.1. The sites have multiple "mirror" units each of which consists of a spherical mirror and a cluster of photo-multiplier tubes (PMTs) which are placed at the focal plane. By simultaneously observing a UHECR from both sites, one can make "stereo" measurements of air showers. An observation that is only made

Figure 1.1. Location of the High Resolution Fly's Eye Experiment relative to Salt Lake City, Utah (the location of the University of Utah).

from one site is referred to as a "monocular" measurement. For this study, we will focus solely upon monocular measurements made by the HiRes-I detector. We have chosen to do so because HiRes-I has accumulated a much larger exposure than HiRes-II thus providing us with larger statistics for our investigations.

This work will focus on the search for anisotropies in the distribution of arrival directions of cosmic rays observed by HiRes-I. Several different potential anisotropies will be considered. Because of the irregular nature of sky coverage for air-fluorescence detectors and the highly asymmetric angular resolution (see Chapter 6) of monocular observations, it is necessary to develop some novel techniques in order to perform the requisite analysis.

The data set that we consider is the events that were included in the HiRes-I monocular spectrum measurement [2, 3]. This set contains 1526 events with measured energies greater than $10^{18.5}$ eV and 52 event with measured energies greater than $10^{19.5}$ eV observed between May 1997 and February 2003 . This data set represents a cumulative exposure of ~ 3000 km^2·sr·yr at 5×10^{19} eV.

1.1 Organization

Below is a synopsis of the contents of this work beginning with a general description of UHECR physics and the HiRes-I detector and then continuing with a description of the anisotropy analysis that was performed.

Chapter 2 provides a description general description of the physics of UHECRs and of the EAS. Potential sources of UHECRs are explored as is their propagation through the galactic and extra-galactic medium. The known properties of EASs are explained as they relate to the air-fluorescence technique.

Chapter 3 gives a physical description of the HiRes-I detector. The general layout of the HiRes-I detector is explained and a description of the data acquisition apparati is provided. The electronic and atmospheric calibration methods are also explained.

Chapter 4 describes the Monte Carlo technique used to simulate showers for the HiRes-I detector. All of the anisotropy studies in this work are rely upon expected exposure estimates that are made by creating simulated data sets that consist of events created the by the simulation routine.

Chapter 5 provides a description of the data processing chain and monocular event reconstruction technique. Specifically, the profile constraint reconstruction method is explained.

Chapter 6 explores detector resolution issues that are pertinent to anisotropy analyses. These include systematic shifts in energy and angular resolution in the reconstruction and the overall angular resolution of the reconstructed events.

Chapter 7 documents the procedure for quantifying the HiRes-I exposure with particular attention to the simulation of data sets that accurately recreate the expected HiRes-I exposure. Specific issues include seasonal variations and accounting individual detector unit (mirror) ontimes.

Chapter 8 presents an analysis that searches for dipole enhancements in the distributions of arrival directions for the HiRes-I UHECR event set. This analysis found no evidence of dipole effects oriented towards the Galactic Center, Centaurus A, or M87. This analysis was published in a HiRes Collaboration paper [4].

Chapter 9 presents a search for small scale anisotropy (i.e. autocorrelation) in the HiRes-I events with estimated energies above $10^{19.5}$ eV. This analysis shows that the arrival directions of the observed events provide no evidence of autocorrelation. This analysis was published in a HiRes Collaboration paper [5].

Chapter 10 explains and demonstrates a novel method involving a fractal dimensionality technique to search for anisotropies in the absence of an *a priori*

preferred source model. This technique is then demonstrated upon *simulated* data sets of different sizes with different source models to show explore its effectiveness in detecting arrival direction anisotropies. This technique was published in an independent paper [6].

Chapter 11 demonstrates the application of the technique explained in chapter 10 to the dipole analysis performed in chapter 8. It is shown that while the fractal technique is not as sensitive as the direct measurement, it provides estimates that are consistent with the results reported in Chapter 8. This analysis was published in a HiRes Collaboration paper [4].

The Appendix is a study that reassesses the significance of the small-scale anisotropy reported by the Akeno Giant Air Shower Array (AGASA). This analysis shows that the chance probability of the AGASA observation is cannot be lower than $\sim 10^{-1}$ due to the limited sensitivity of the AGASA experiment.

CHAPTER 2

THE PHYSICS OF ULTRA-HIGH ENERGY COSMIC RAYS AND EXTENSIVE AIR SHOWERS

UHECRs have been a subject of interest in experimental high energy astrophysics since the first observations were made by Volcano Ranch 40 years ago [7]. In that time we have learned only a little about their underlying nature. This is partly due to their extreme rarity. For UHECRs with energies above $10^{18.5}$ eV we can optimistically expect a particle incidence of $\sim 10^{-17}\ cm^{-2}sr^{-1}sec^{-1}$. Even by using indirect methods for observing these particles, it is very difficult to accumulate the statistics necessary to make a thorough study of these particles.

With the HiRes experiment, some significant progress has been made, both in the measurement of the chemical composition [8] and of the energy spectrum [2, 3] including evidence for the Greisen-Zatsepin-Kuz'min (GZK) suppression near 6×10^{19} eV [9, 10] (see Section 2.3.2. However, we have made less progress in answering the astrophysical questions: where do UHECRs come from and how are they produced. There are a few candidate sources for UHECRs, but as yet there is no evidence that UHECRs are originating at these source and there is no widely accepted model of how these sources accelerate elementary particles to the UHE regime.

Because we cannot directly observe UHECRs, it is very important that we have a thorough understanding of the extensive air showers (EAS) generated by UHECRs in the atmosphere. Using detailed knowledge of the characteristics of EAS, we can estimate the energy, composition, and arrival directions of the precursor UHECRs.

2.1 Current Knowledge: The Energy Spectrum and Primary Composition

In Figure 2.1, we see the differential cosmic ray energy spectrum for twelve decades of energy, including measurements in the UHECR regime. Because this spectrum is steeply falling, it is customary to multiply it by $E^{2.7}$ in order to clearly show its features as is done in Figure 2.2. In the past 40 years, UHECR energy spectrum measurements have been published by several collaborations including Volcano Ranch [7], Yakutsk [13], Haverah Park [14], Fly's Eye [15], the Akeno Giant Air Shower Array (AGASA) [16], the HiRes-MIA Prototype [17] and HiRes Mono [2, 3]. Fly's Eye and HiRes have both also published their findings on UHECR composition [18, 17, 8].

While the various experiments are not in complete agreement, it can be concluded that the differential flux, $\frac{d\Phi}{dE}$, follows a power-law spectrum:

$$\frac{d\Phi}{dE} \propto E^{-\alpha}, \tag{2.1}$$

where $\alpha \simeq 2.7$ for $E > 10^{18.5}$ eV.

This spectral index seems to remain approximately the same as one increases in energy until one reaches an energy of $E > 10^{19.8}$ eV at which point there simply is not enough statistics from all of the experiments performed to date to make a definitive determination. In Figure 2.3, one can see the most recently reported differential spectrum observations of AGASA and HiRes monocular. While these results are certainly not completely compatible, they can made to agree at all but the highest energies by applying an energy correction factor of $\sim$ 1.3 to either experiment. At the highest energies, the reported HiRes spectrum is compatible with the GZK suppression (see Section 2.3.2 while the AGASA spectrum is not.

Primary composition is determined by comparing the atmospheric penetration summarized by the depth of shower maximum, $x_{\max}$, of the observed showers with that predicted by CORSIKA [19] simulations of different primary particles. Composition measurements performed by Fly's Eye and HiRes have shown that the observed showers with $E > 10^{18.5}$ eV have mean $x_{\max}$ values consistent with a primary composition mostly of protons with the admixture of a heavier nuclear

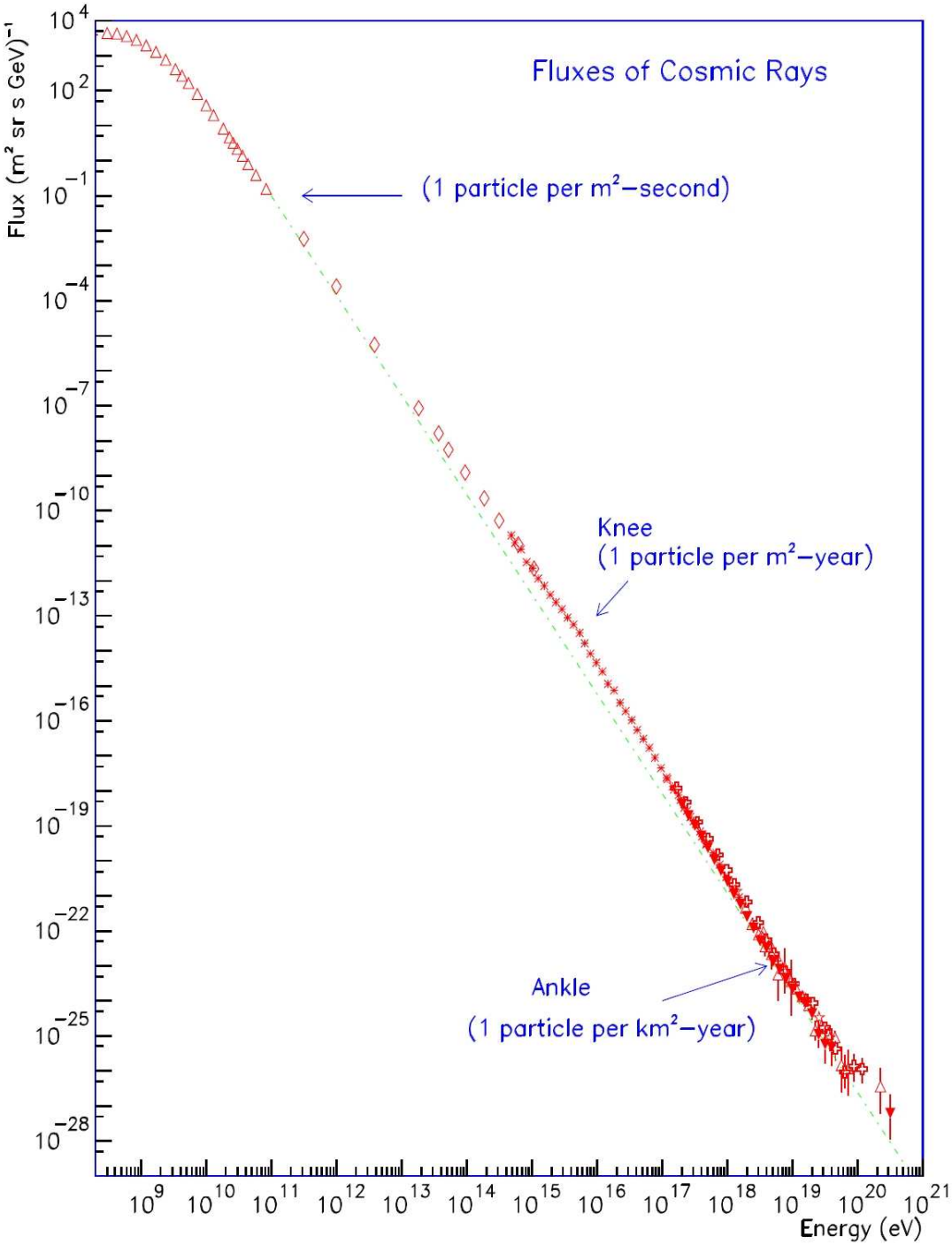


Figure 2.1. The differential cosmic ray energy spectrum [11]. The spectral index is ~ -2.7 up to the knee and ~ -3.1 between the knee and the ankle.

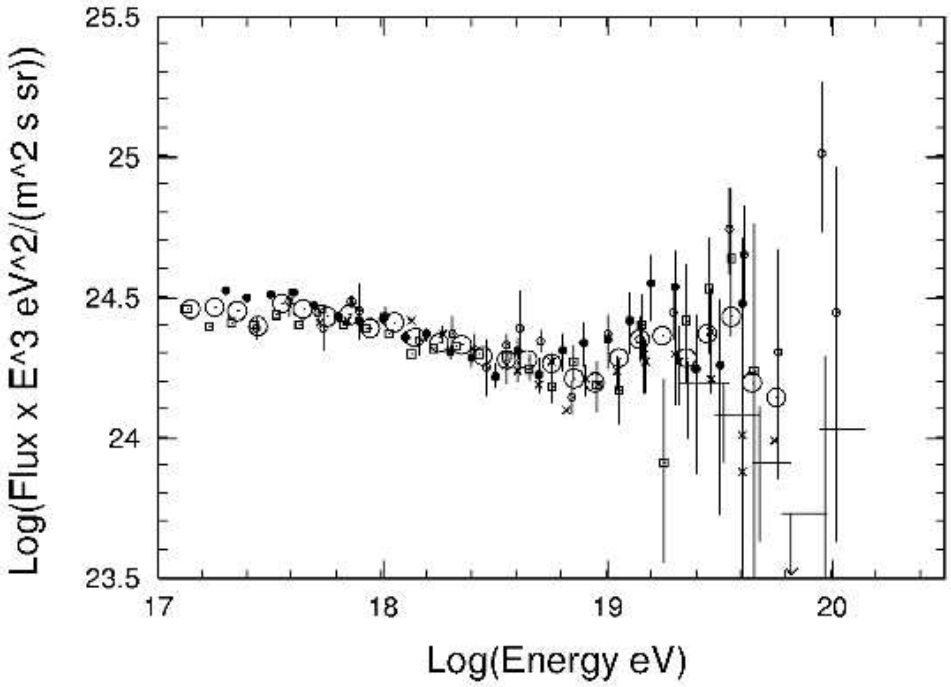


Figure 2.2. The differential cosmic ray energy spectrum multiplied by $E^{2.7}$ [12]. The data was taken from several experiments including: Akeno (□), JACEE (▼), Proton-4 (▲), Tien Shan (▽), MSU (△), Tibet (+), HEGRA (×), CASA-MIA (⧫), Fly's Eye (stereo (○), mono (◊), and Haverah Park (■). The high energy data points indicated by ▼ are from AGASA. See [12] for references to original papers.

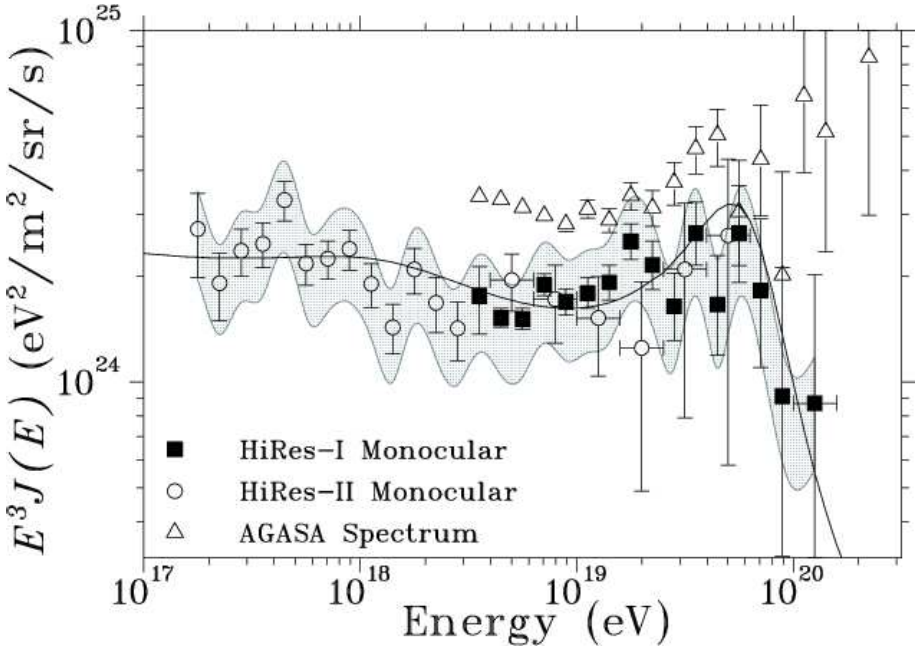


Figure 2.3. The reported UHECR differential spectrum observations of AGASA [16] and HiRes mono [2, 3].

component toward the lower energies. In Figure 2.4, we can see the $x_{\max}$ values for showers observed in the stereo mode by HiRes with a comparison to the results of the HiRes-MIA prototype results and the predictions of CORSIKA with the QGSJet01 and SIBYLL 2.1 hadronic interaction models.

2.2 The Origins of UHECRs

UHECRs with $E > 10^{18.5}$ eV are believed to be extra-galactic in origin. This is because there is no known mechanism in our galaxy for producing particles at these extreme energies. The most likely scenario is that UHECRs are produced in some of the most violent places in the universe. We therefore hope that the study of UHECRs will eventually lead to a better understanding of cosmology and other highly energetic astrophysical phenomena.

2.2.1 Production Mechanisms

UHECRs are most likely produced in environments where shock fronts and strong magnetic fields conspire to trap charged particles (i.e. protons and atomic

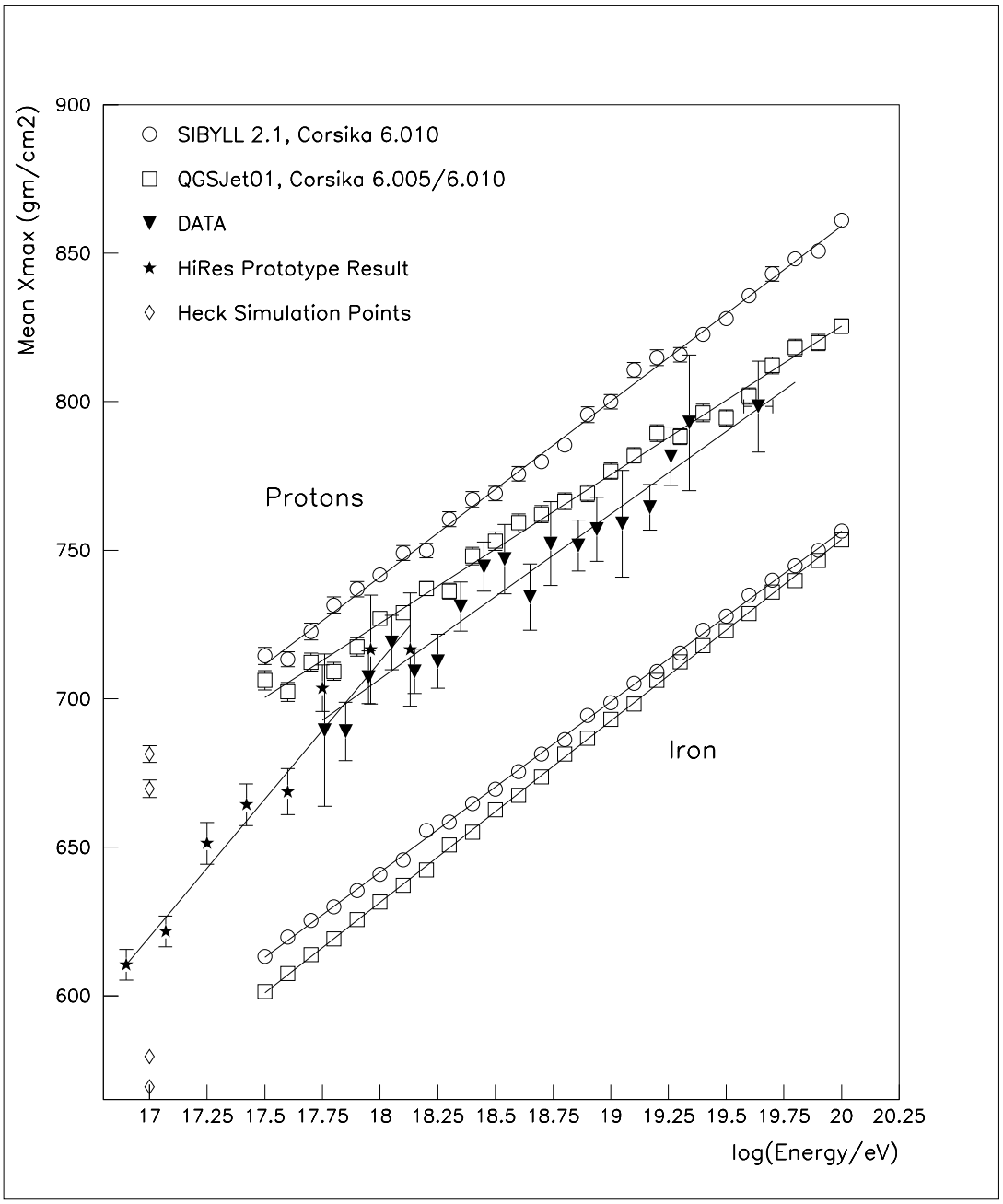


Figure 2.4. The atmospheric penetration, $x_{\max}$, of showers observed in the stereo mode by HiRes with comparison to the HiRes-MIA prototype results and the CORSIKA predictions using the QGSJet01 and SIBYLL 2.1 hadronic interaction models.

nuclei) and accelerate them over extended periods of time. The primary limitation to this model is that there are relatively few observed astrophysical phenomena that possess large enough magnetic fields over sufficient distance scales to achieve the energies of UHECRs. In Figure 2.5, we can see the relationship between source size and the observed magnetic field for a variety of potential UHECR sources. When one considers that the current evidence points to a primary composition that is

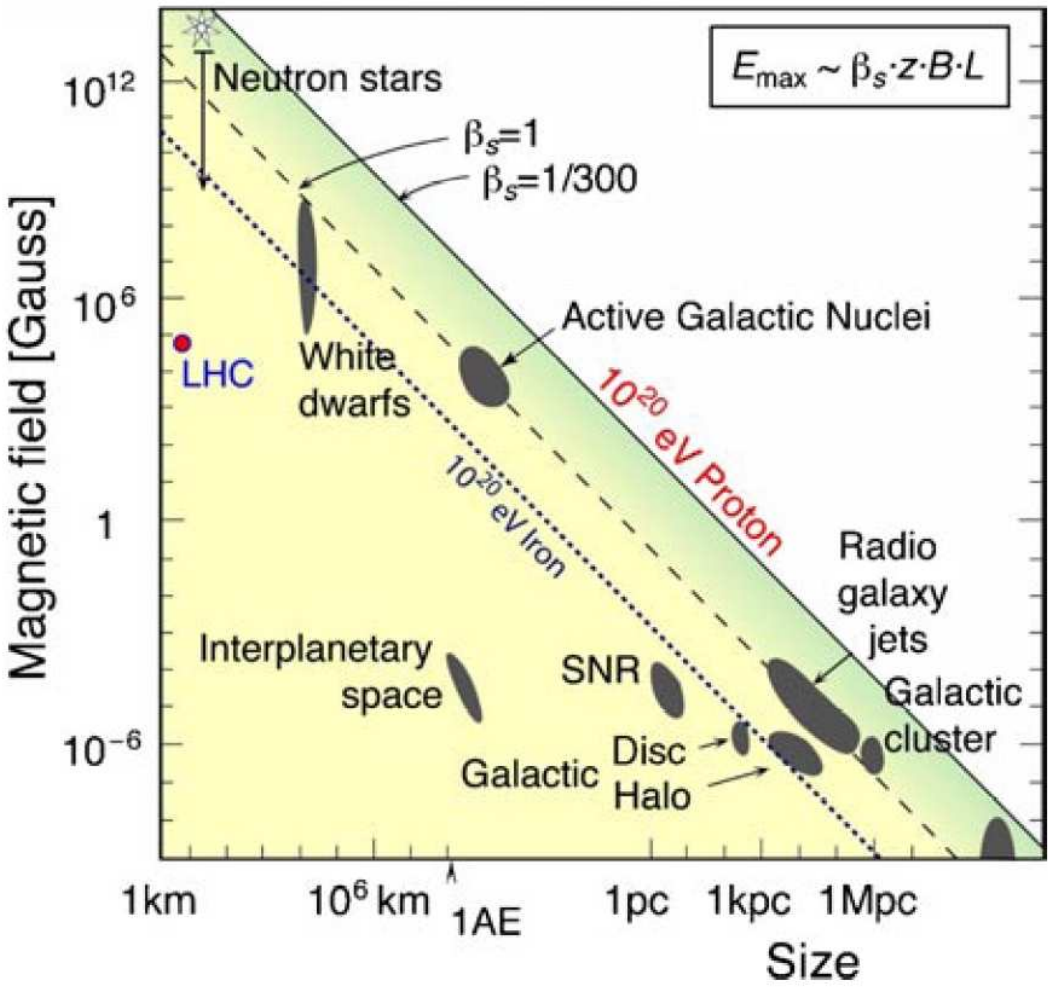


Figure 2.5. The Hillas Diagram. In this diagram we can see the combination of magnetic field strength and source size necessary to accelerate protons (solid line) and iron nuclei (dashed line) to an energy of 10^{20} eV with a shock front assumed to velocity $\beta \simeq 1$. The top of the shaded area represents the size and magnetic field requirements necessary to accelerate a proton to and energy of 10^{20} eV with a shock front velocity of $\beta = 1/300$.

almost entirely proton, it becomes clear that our most likely sources are active galactic nuclei (AGN) and radio galaxy lobes.

Alternatively, it is also possible that UHECRs are not accelerated at all but are instead the decay product of some, yet unknown, super-massive objects. Several such "top-down" models have been proposed including:

1. The decay of large quasi-stable super-symmetric particles that reside in the galactic halo [20]
2. Topological defects such as "cosmic strings" [21]
3. Magnetic monopoles that are easily accelerated by the magnetic fields in the interstellar medium and then initiate EAS when they encounter the Earth's atmosphere [21]

However, there is no evidence that any of these objects do in fact exist and even if they did, these theories do no provide predictions that are consistent with the observations that have been made.

2.2.2 Candidate Sources

Within the galaxy, there is very little likelihood that any known astrophysical phenomena could produce the highest energy UHECRs. There are only two objects that have the possibility of possessing large enough magnetic fields over sufficient distance scales. First, the galactic center, Sagittarius A which is $\sim$8 kpc in distance. Sagittarius A is a super-massive black hole with a mass of $3.5 \times 10^6 M_\odot$ [22]. It is is currently quiescent, but in the past is has undoubtedly had periods of extreme activity. The second potential source is Cygnus X-3. Cygnus X-3 is a stellar black hole that is part of binary system $\sim$2.3 kpc away in the Sagittarius Arm of the Milky Way. Cygnus X-3 is extremely loud in the X-ray bandwidth and could potentially be producing some UHECRs.

Outside of the galaxy, there are two likely sources within 20 Mpc. The first is Centaurus A, which is $\sim$5 Mpc in distance. Centaurus A is a very loud radio galaxy with lobes that extend out $\sim$50 kpc from its center. The other potential source is

M87. M87 is a massive elliptical galaxy with a extraordinarily active AGN which is $\sim$15 Mpc in distance.

On a cosmological scale, one can also consider several extremely energetic astrophysical phenomena that typically occur at a distance scale of at least 500 Mpc. All of these objects almost certainly possess large enough magnetic fields over sufficient distances to accelerate particles into the GZK regime. These include gamma-ray bursts (GRBs), quasars, and blazars. However, even if one disregards other complicating factor, the cosmological distances to these objects, seemingly preclude them as source candidates for the highest energy UHECRs because of the GZK threshold.

2.3 Propagation

2.3.1 Magnetic Deflection

Because cosmic rays are thought to be charged particles, they are expected to be subject to deflection in the galactic and extragalactic magnetic fields. The Larmour radius, R_L, for a particle with charge Ze can be approximated as follows:

$$R_L(kpc) \simeq \frac{1}{Z}\left(\frac{E}{1\ EeV}\right)\left(\frac{B}{1\ \mu G}\right)^{-1}, \tag{2.2}$$

where E is the energy of the particle and B is the magnetic field of the propagating medium. For protons in a field comparable to the galactic field ($B \sim 2\ \mu G$) $R_L(kpc) \simeq E(EeV)/2$. This implies that a proton must have an energy of $\sim 4\times10^{19}$ in order to propagate within the galaxy. Extra-galactic magnetic fields are believed to be ($B \sim 1\ nG$) and randomly oriented [23]. This would lead to deflection that is approximately three orders of magnitude less than that of the galactic magnetic field, albeit over much greater distances.

2.3.2 The GZK Threshold

In 1966, Greisen, Zatsepin, and Kuz'min independently concluded that UHECRs with energies above a threshold would interact with the Cosmic Microwave

Background (CMB) and consequently be degraded in energy [9, 10]. For a proton this is dominated by the Delta resonance:

$$p + \gamma_{2.7\mathrm{K}} \longrightarrow \Delta^*(1232\mathrm{MeV}) \longrightarrow p + \pi^0. \tag{2.3}$$

This means that above the so-called GZK threshold, the universe becomes opaque to UHECRs. For protons, the nominal threshold is $\sim 6 \times 10^{19}$ eV. Above that energy, the distance which UHECRs can travel is limited to $\sim$ 50 Mpc. In Figure 2.6, the estimated energy degradation for UHECRs in the super-GZK regime can be seen. This degradation seems to necessitate that UHECRs with energies on the order of 10^{20} eV must originate in our local part of the universe. This drastically curtails the number of potential sources. However, the original super-GZK events observed the Fly's Eye [15] and AGASA [25] do no correlate with local violent objects. Alternative explanations include:

1. Violation of Special Relativity (Lorentz Invariance) [26]

2. Stronger than anticipated extra-galactic magnetic fields, which would deflect arrival directions away from local sources [27]

3. Exotic super-massive objects that initiate "top-down" decays within the galactic halo [20]

2.4 The Extensive Air Shower

Once a UHECR is generated, it must travel across the vast expanse of space before it finally arrives at the point where it can be observed by detectors here on Earth. The EAS is the cascade of billions of particles that are generated when a UHECR interacts with the terrestrial atmosphere. The EAS can be observed in two principal ways; either by ground arrays that sample the remnant particles in the shower (usually past shower maximum) as they strike the ground, or by optical detectors that record either the cherenkov light or the air fluorescence of an EAS as it develops in the atmosphere. HiRes relies upon the observation of air-fluorescence, so we will focus upon the properties of the EAS that lead to the fluorescence measurement.

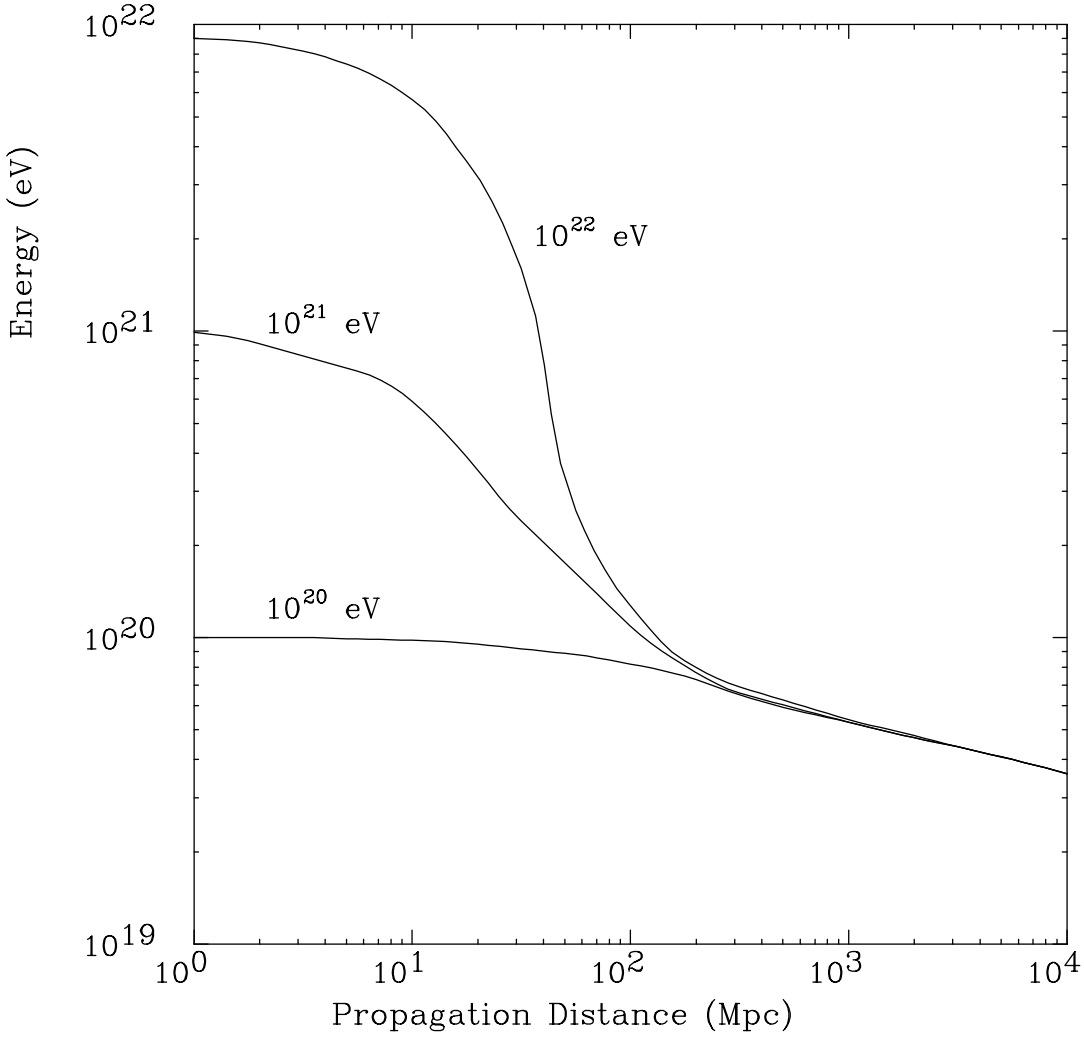


Figure 2.6. Energy of a proton as a function of propagation distance [24].

2.4.1 EAS Development

A hadron induced shower can be thought of as a series of electromagnetic cascades that are all induced by a hadronic core that remains very close to the shower axis. As the EAS develops, the hadronic core produces π^0 particles which then electromagnetically decay initiating the electromagnetic cascades. The longitudinal profile an EAS can be modeled using the Gaisser-Hillas parameterization [28] which is based upon simulations of electromagnetic showers. The Gaisser-Hillas profile is as follows:

$$N(x) = N_{\max}\left(\frac{x - x_\circ}{x_{\max} - x_\circ}\right)^{(x_{\max}-x_\circ)/\lambda} \exp\left[\frac{x_{\max} - x}{\lambda}\right], \tag{2.4}$$

where x is the shower depth in gm/cm^2, $x_\circ$ is nominally the depth of first interaction, $x_{\max}$ is the depth of shower maximum, N and $N_{\max}$ are the total number of particles at depth x and $x_{\max}$ respectively, and λ is a constant of value 70 gm/cm^2 for hadronic showers.

To make an accurate simulation of an air shower, one must also consider the latitudinal profile of the shower. This is approximated by the Nishimura-Kamata-Greisen (NKG) formula [29]:

$$\rho(r) = \frac{N}{r^2} f\left(s, \frac{r}{r_M}\right) \tag{2.5}$$

where N denotes the total number of electrons, s denotes the shower age, r_M denotes the Moliere radius for multiple scattering and f denotes the Nishimura-Kamata function:

$$f\left(s, \frac{r}{r_M}\right) = \left(\frac{r}{r_M}\right)^{s-2}\left(1 + \frac{r}{r_M}\right)^{s-4.5} \frac{\Gamma(4.5 - s)}{2\pi\Gamma(s)\Gamma(4.5 - 2s)} \tag{2.6}$$

2.5 Light Production

The final consideration is the light production generated by the EAS. This is what is in actuality observed by the air-fluorescence technique. There are two components: Cherenkov emission and air-fluorescence emission.

Cherenkov emission is narrowly confined to near the axis of the EAS. It produces a very strong beam of light with angular distribution:

$$\frac{dN_p}{d\Omega} \propto \frac{e^{-\theta/\theta_\circ}}{\sin\theta}, \tag{2.7}$$

where $\theta_\circ = 0.83E_t^{-0.67}$, E_t is the lower threshold energy given by:

$$E_t = mc^2/\sqrt{2(n-1)}, \tag{2.8}$$

and n is the index of refraction of the atmosphere.

Fluorescent emission is produced when the electromagnetic component of the shower excites molecular nitrogen in the atmosphere. When this occurs, primarily ultraviolet light is emitted isotropically by the molecular nitrogen. In Figure 2.7, we

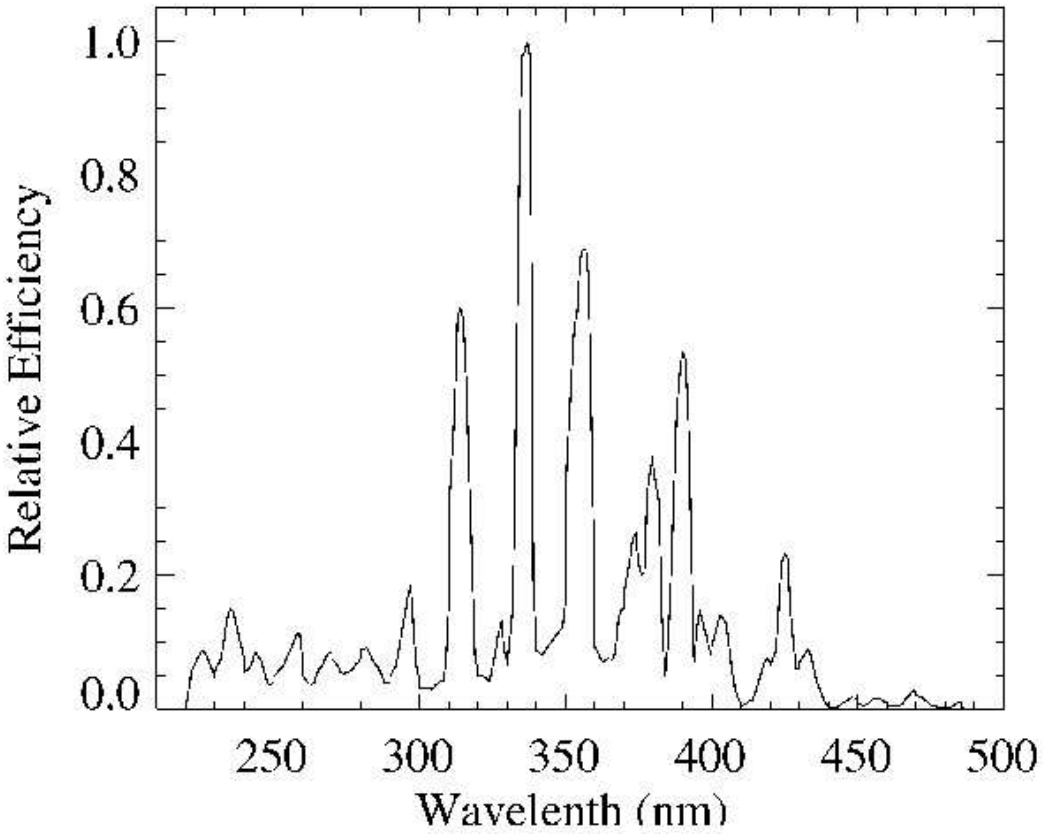


Figure 2.7. The fluorescence spectrum of nitrogen by Bunner [30].

can see the measured spectrum of nitrogen as a function of wavelength as measured by Bunner in 1968 [30]. More recent, yet unpublished measurements by the FLASH collaboration [31] seem to confirm Bunner's and suggest that the air-fluorescence is largely independent of atmospheric pressure.

The relative contributions of shower light resulting from Cherenkov and nitrogen fluorescence emission is dependent upon the orientation of the shower axis with regards to the detector. It is important that one distinguishes between the two different types of emission because only the fluorescent component will be utilized in determining shower energy.

CHAPTER 3

THE HIGH RESOLUTION FLY'S EYE DETECTOR AT LITTLE GRANITE MOUNTAIN

The High Resolution Fly's Eye Experiment is located at the United States Army's Dugway Proving Ground, about 100 km southwest of Salt Lake City, Utah in the western United States. The detector consists of two separate sites with a 12.6 km separation. The first site, HiRes-I, is located at Little Granite Mountain (a.k.a. "Five Mile Hill"). HiRes-I began operation in May 1997. The second site, HiRes-II, is located at Camels Back Ridge. HiRes-II began operation in June 1999. Since this work focuses almost exclusively on the monocular observations of HiRes-I, using stereo observations only for calibration purposes, only HiRes-I will be discussed in detail.

3.1 Detector Layout

During the history of its operation, HiRes-I has consisted of anywhere from 14 to 21 detector units in operation. Each detector unit views a section of sky 16° wide in azimuth covering elevation angles from 3° to 17°. With the original 21 detector units, this led to an azimuthal coverage of 336°, although the current configuration of only 20 detector units, since 2002, covers 320° in azimuth. A maximum elevation angle of 17° was chosen because this optimized the ability of the detector to observe the highest energy showers at extreme distances ($\sim 30\ km$) and the 3° lower edge raises the field of view above artificial light sources associated with military facilities on the ground.

The individual detector units are housed in pairs in shelters constructed at Little Granite Mountain in the early 1990's. Most shelters consist of steel sheds with a large garage-style doors that opens at night when the detector units are in operation and close during daylight hours to shield the mirrors and photo-tube units from direct sunlight. For a time, one detector unit was also housed in a shelter that was a modified grain silo.

Each detector unit is connected to the HiRes-I central facility via Ethernet. The central facility consists of a data acquisition (DAQ) computer, a GPS based Central-Timing crate and a YaG laser used for nightly detector calibration. A T1 line provides network connectivity to the outside world. In Figure 3.1, a schematic representation of the HiRes-I detector is shown.

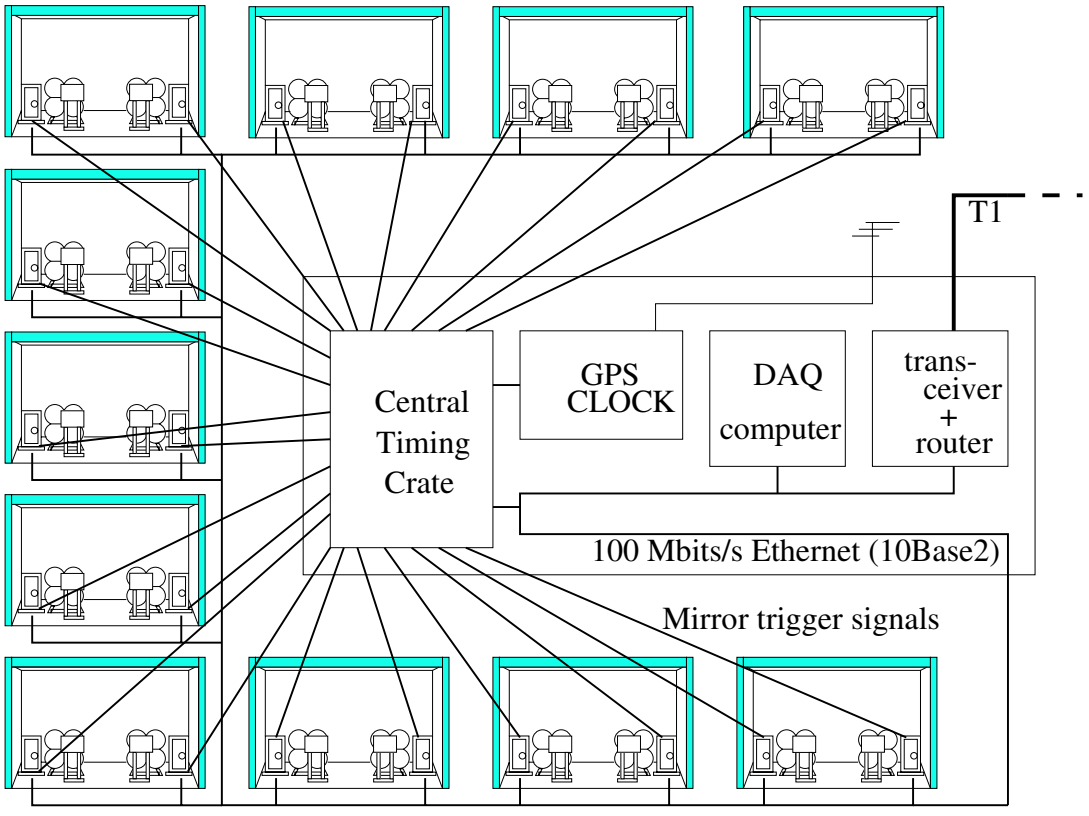


Figure 3.1. A schematic representation of the HiRes-I detector.

3.2 The Individual Detector Unit

The individual HiRes-I detector units consist of three primary components: a 4.18 m^2 spherical mirror, a photo-multiplier tube (PMT) cluster, and corresponding VME crate electronics. The data acquisition system employs a "sample and hold" (S/H) electronics system that was first implemented in the HiRes prototype in 1993. The most challenging design criterion for this system is robustness, as the electronics is operated essentially in an open environment with ambient temperatures ranging from $-20°$ to $50°$. Between its original prototype deployment and its reconfiguration in 1997 to form the HiRes-I detector, this system has required no significant modification.

The light collection component for each detector unit is a spherical mirror. Each mirror has an area of 4.18 m^2. However, due to the obscuration of the PMT clusters, the effective area of each mirror is reduced to 3.72 m^2. For the purpose of this analysis, the reflectivity coefficient of the mirror over the pertinent frequency band (300 - 400 nm) is estimated to be 0.8. In Figure 3.2, we can see the physical positioning of the HiRes mirror.

For each mirror, there is a corresponding PMT cluster. The PMT cluster is placed on the mirror axis at a distance of 0.485 times the radius of curvature of the mirror. This separation was chosen both to minimize the optical spot size, r_m, and its off-axis variation, $dr_m/d\theta$. Each PMT cluster is composed of 256 PMTs arranged in the hexagonal honeycomb configuration. In order to improve the signal to noise ratio, the tubes are covered with a UV-transmitting glass filter. The system has f/1 optics and with a radius of curvature of 4.73 m, each individual tube views an approximately $1°$ cone of sky with an entire PMT cluster viewing a section of sky that is approximately $16°$ in azimuth and $14°$ in elevation. Each tube has its own pre-amplifier and high-voltage (HV) divider and is plugged into a single-piece back-plane that provides high and low voltages (LV) as well as signal output to the mirror electronics crate. This is shown schematically in Figure 3.3.

Figure 3.2. The placement of a HiRes mirror with respect to the PMT cluster.

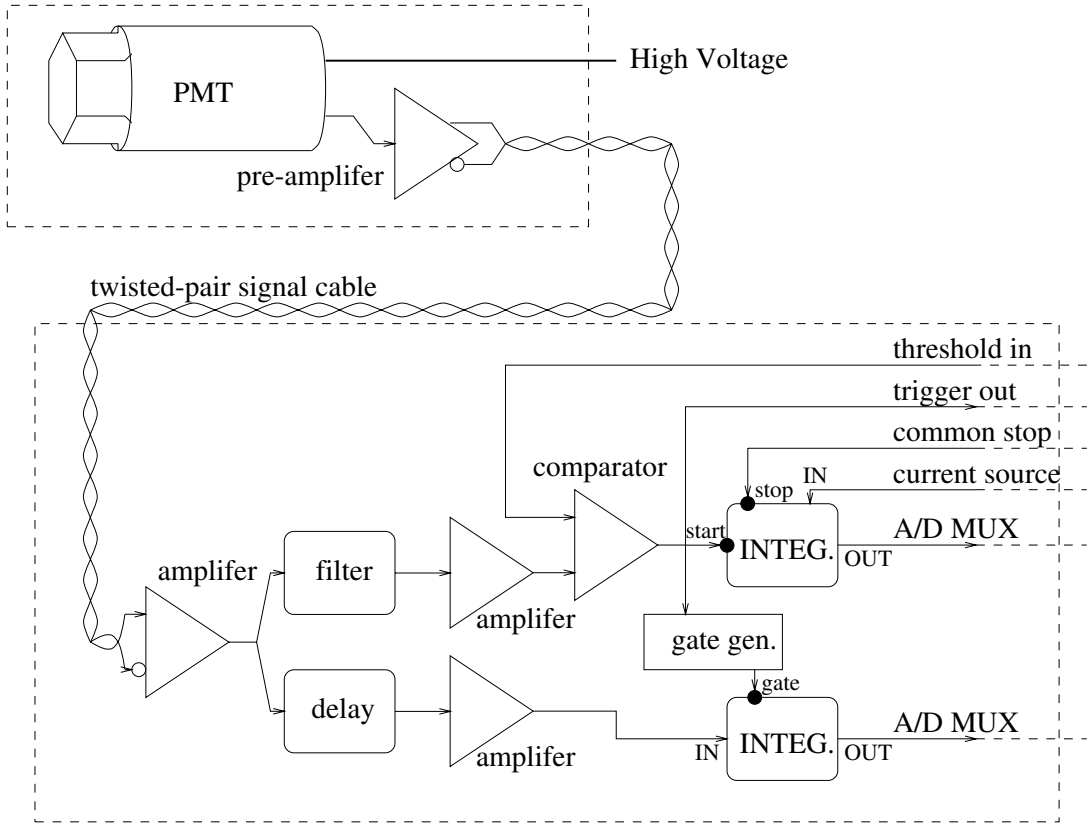


Figure 3.3. A schematic representation of an individual HiRes-I PMT channel.

The readout electronics is deployed in a double-height (6V) VME crate housed in a standard rack which also contains the requisite HV and LV power supplies. Each VME crate contains:

1. A CPU board
2. A Programmable Pulse Generator (PPG) board
3. A trigger logic board
4. 16 readout or "ommatidial" boards (OMBs)
5. A multifunction control/monitor board

In Figure 3.4, we see the schematic overview of readout electronics. The PPG

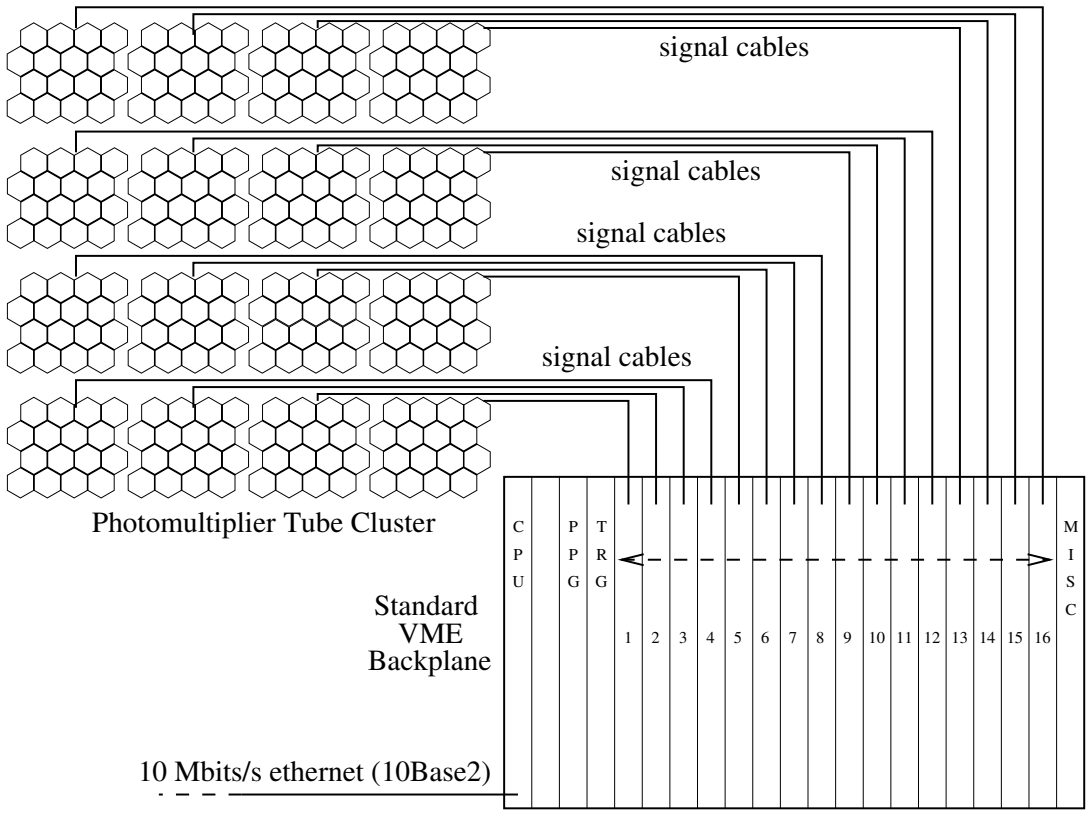


Figure 3.4. A schematic overview of the HiRes-I readout electronics.

board is both as a diagnostic tool and for nightly electronics calibration. The multifunction board provides miscellaneous functions such as HV and LV readout, temperature monitoring, and garage door control. The ommatidial boards provide secondary signal processing functions for the cluster including amplification, integration, and digitization of the PMT signals which then allows each OMB to form subcluster level triggers. The OMB subcluster trigger is shown schematically in Figure 3.5. This information from the individual OMBs is then provided to trigger logic board which forms mirror-wide triggers as is shown in Figure 3.6. Mirror-wide triggers are then transmitted by the CPU board to the DAQ computer in the central facility.

3.3 The Central Facility

The central facility is a complex of interconnected trailers at Little Granite Mountain that houses the electronic DAQ equipment, repair facilities, and rudimen-

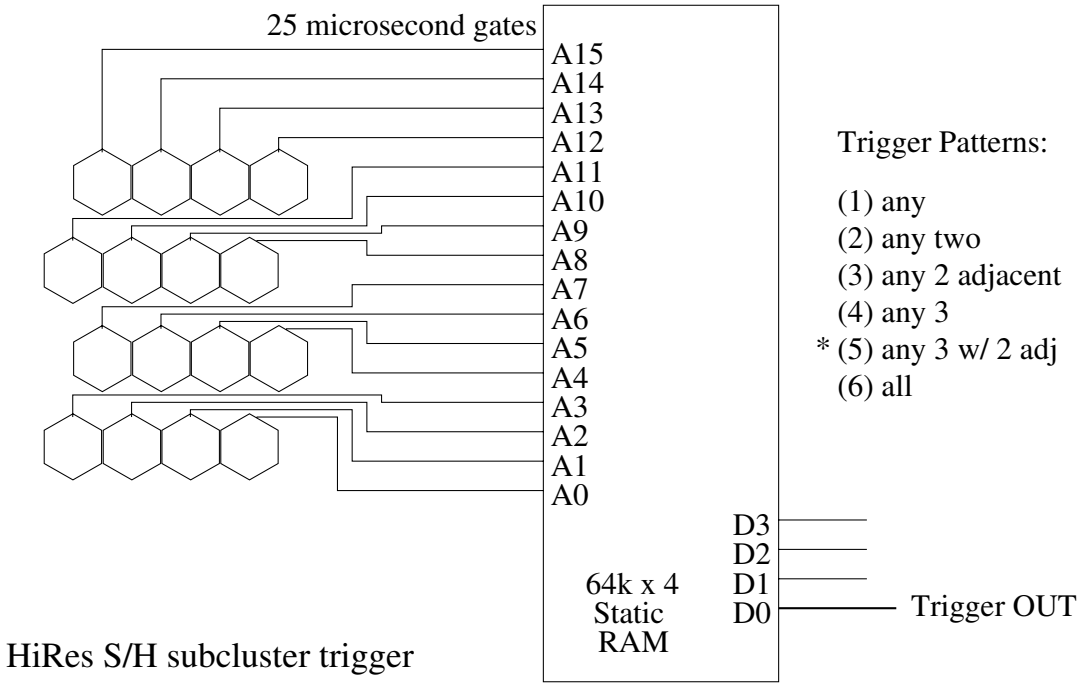


Figure 3.5. The HiRes-I OMB subcluster triggering system.

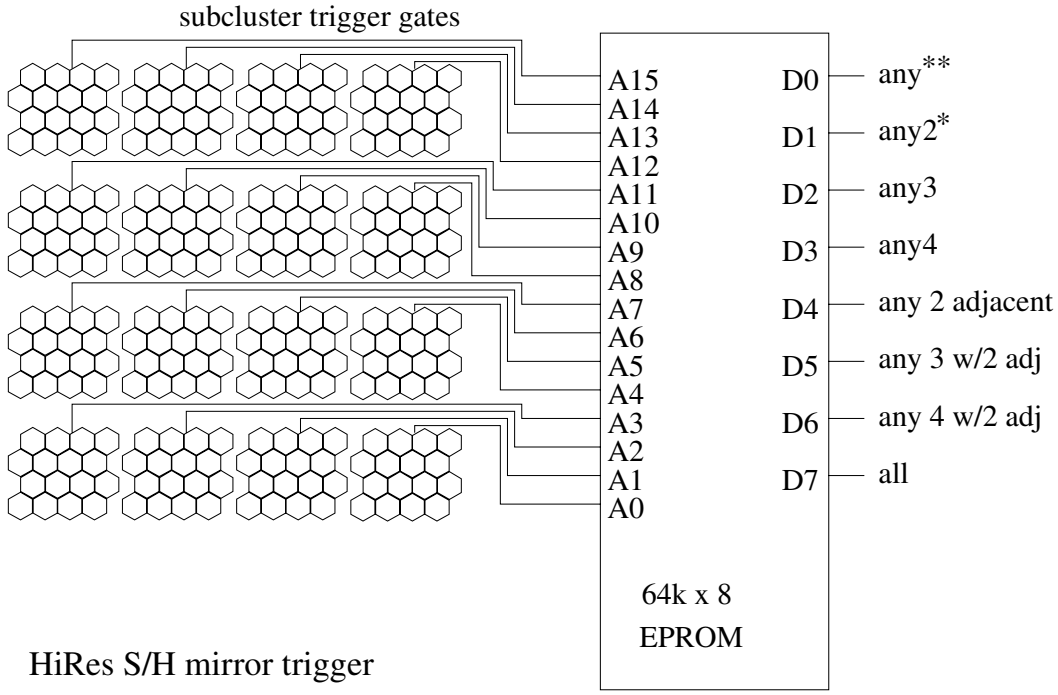


Figure 3.6. The HiRes-I mirror-wide triggering system.

tary living quarters for detector operators. Over the operation history of HiRes-I, the primary DAQ computer was first a DEC Alpha and is now a commodity Linux box with an Intel CPU. In conjunction with the central DAQ computer, there is a double-height VME crate that applies time stamps to the mirror triggers received by the central DAQ computer. The central timing crate assigns a time to the event which is based upon an absolute time provided by GPS. A second Linux box is also present in the central facility which provides online diagnostic tools for detector operation and quick online reconstruction of the observed events. A separate trailer at the central facility contains a YaG laser that is employed to illuminate the separate clusters by way of optical fibers for off-line calibration purposes.

3.4 Remote Operations

Currently, the HiRes-I detector is configured so that it can be operated remotely from collaborating institutions. Modifications for remote operations include the introduction of fail-safe curtains attached to light sensors. These curtains deploy in the event that a door fails to close to provide a physical barrier between the mirror and the cluster from sunlight. Further modifications include the installation of VNC software in the central facility computers to allow for the operation of these computers from remote locations, remote power controllers that allow us to cycle the power on any piece of essential equipment over a modem. The modem is also needed for emergency operation in the event of network outages.

3.5 Detector Calibration

The calibration of HiRes-I can be divided into to three primary constituents: electronics response, PMT response, and atmospheric clarity. For the sake of brevity, we will only consider the calibration efforts employed in the analysis of the HiRes-I data set as it now stands.

The electronics calibration of HiRes-I is primarily concerned with two quantities: pedestal values and the response of the individual pixel to a graduated signal. The

pedestals are measured at the beginning and end of each night by taking a series of random triggers of the individual clusters with the garage doors closed. By considering the fluctuation in the individual tube read-outs, the inherent electronic noise of the individual tubes can be calculated. The response of the individual pixel to a graduated signal is measured by using the PPG to provide a series of pulses of various heights and widths.

PMT response is calibrated using a highly stable roving xenon flasher (RXF). The RXF luminosity is known to an accuracy of $\pm 10\%$. Periodically during detector down times, the RXF is manually transported to each mirror and used to illuminate the cluster with a known amount of light with a series of neutral density filters. The response of each PMT can be established as a function of the number of photons observed. These calibration data sets are then introduced into the offline processing scheme with each data part employing the calibration data set that is chronologically closest to it.

Atmospheric calibration is accomplished mainly with the use of a steerable YaG laser situated at HiRes-II. Each hour, this laser fires a set program of shots into the night sky to comprise full coverage in both elevation and azimuth. By reconstructing the HiRes-I observation of these laser shots, it is possible to establish the night-to-night values of atmospheric attenuation. However, because this laser was not in use for more than half of the operational time life HiRes-I, an average value of atmospheric attenuation is employed for the purpose HiRes-I monocular data analysis. The specific details of the atmospheric model will be discussed in the next chapter.

CHAPTER 4

EVENT SIMULATION

Event simulation plays in critical role in all aspects of HiRes data analysis. A comprehensive simulation code has developed corrabatively by the collaboration for this purpose. The ability to realistically simulate air showers allows us to calculate aperture, model the resolution of event reconstruction, and to produce large libraries of simulated data sets. In order to create simulated events that are representative of the actual data, all of the processes that affect the observation of actual events must be taken into account. These include shower development, light production, light propagation, detector optics, and detector electronics.

4.1 Event Geometry and Shower Development

Each event can be described using three parameters: energy, composition, and geometry. For the purpose of this study, a two component mix of composition was used: protons and iron nuclei. In accordance with the composition and spectral measurements previously made by Fly's Eye stereo [15, 18], proton events were simulated with a differential energy spectral index of $\gamma = -2.6$ and iron events with $\gamma = -3.5$. The proton and iron fluxes were set to be equal at an energy of $E = 10^{18.4}\ eV$. Event geometry was set assuming uniform distributions on $\cos\theta_{sh}$ and ϕ_{sh} where θ_{sh} is the zenith angle and ϕ_{sh} is the azimuthal angle of the arrival direction of the event. A value for the impact parameter, R_p, is then chosen on a uniform distribution of R_p^2 for a selected minimum and maximum value of R_p. For the highest energy simulations (where the aperture is expected to be larger), we used: $3.0\ km \leq R_p \geq 70.0\ km$.

If, for a given event, the above procedure produces a shower that is observable from the HiRes-I detector, then the simulation is allowed to proceed. The longitudi-

nal development of the shower follows the Gaisser-Hillas (G-H) parameterization of equation 2.4 with the three free parameters: $x_\circ$, x_{max}, and N_{max} selected according to physical distributions. The depth of first interaction, $x_\circ$, is chosen on the probability density,

$$p(x_\circ) = e^{-x_\circ/\overline{x_\circ}}, \tag{4.1}$$

where $\overline{x_\circ} = 70\ gm/cm^2$ for protons and $\overline{x_\circ} = 15\ gm/cm^2$ for iron. The depth of shower maximum is then calculated using the following:

$$x_{max} - x_0 = 725.0 + 55.0 \times (\log(E) - 18) - \overline{x_0} \ \text{(proton)} \tag{4.2}$$

$$x_{max} - x_0 = 650.0 + 55.0 \times (\log(E) - 18) - \overline{x_0} \ \text{(iron)} \tag{4.3}$$

The shower size at x_{max}, N_{max}, is calculated by [32]:

$$N_{max} = E/1.3 \times 10^9, \tag{4.4}$$

with E being the energy of the primary particle in eV.

Once all of the descriptive characteristics of the shower have determined, the shower track is divided in 1000 angular segments. For each segment, the shower depth, size, age, and the Moliere multiple scattering radius are calculated. The shower size is derived from the G-H parameterization. The age parameter is determined by the formula:

$$s = 3/(1 + 2x_{max}/x). \tag{4.5}$$

The Moliere radius (which determines the lateral width of the shower) is given by:

$$r_m = X_\circ \times E_s/E_c, \tag{4.6}$$

where $X_\circ = 37.1\ gm/cm^2$ is the radiation length for electrons in air, $E_s = 21\ MeV$ and the critical energy, $E_c = 97\ MeV$ [33].

4.2 Light Production

The light produced by the EAS consists of two components: beamed cherenkov radiation and isotropic fluorescence emission. While energy reconstruction depends

primarily on the fluorescence light, it is necessary to model accurately the cherenkov light contamination in the observed EAS so that this component can be subtracted out.

Cherenkov light is produced when charged particle, primarily electrons in an EAS, travel through a given medium at speed greater than the local speed of light. For the purpose of our simulation, the number of cherenkov photons produced per unit length can be written:

$$\frac{dN_\gamma}{dl} = 4\pi\alpha\delta F(E_t)\int \frac{d\nu}{c}, \tag{4.7}$$

where α is the fine structure constant, $\delta = n-1$ with n being the index of refraction of air, ν is the frequency of the emitted radiation, and $F(E_t)$ is the fraction of the electrons in the shower with energies greater than $E_t = mc^2/\sqrt{2delta}$ given by [34]:

$$F(E_t) = \frac{34.8}{(40.4 + E_t)(1 + 10^{-4}E_t)^2} \tag{4.8}$$

Furthermore, the cherenkov radiation has an angular dependence:

$$\frac{dN_\gamma}{dld\Omega} = \frac{dN_\gamma}{dl}\frac{e^{-\theta/\theta_0}}{2\pi\sin\theta} \tag{4.9}$$

with $\theta_0 = 0.83\,E_t^{-0.67}$ [35].

On the other hand, the fluorescence light emitted by the shower is approximately proportional to the number of electrons present in the shower at a given point in the shower track. The amount of photons emitted per unit length per steradian is given by:

$$\frac{d^2N_\gamma}{dld\Omega} = \frac{YN_e}{4\pi}, \tag{4.10}$$

where Y is the fluorescent yield in photons/electron/m and N_e is the number of electrons in the shower.

Several attempts at experimentally measuring the fluorescent yield of high energy electrons have been made over the years. The integral value of Y used in this analysis is based upon the recent reported by Kakimoto et al. [36] in conjunction with the spectrum reported by Bunner [30] with the light generated in 1 nm bins. These measurements provides an estimate of fluorescent as a function of energy and temperature with a systematic uncertainty of $\sim 10\%$.

4.3 Light Propagation

Now that we have established the means of light production in the EAS, we consider the processes by which that light impinges on the HiRes-I detector. There are three atmospheric mechanisms that we must consider: scattering by air molecules, scattering by aerosol contaminants suspended in the atmosphere, and absorption by ozone molecules. The effects of these mechanisms can be put into two categories: scattering of directional cherenkov light into the field of view and effective attenuation of isotropic fluorescent light by scattering/absorbing it out of the field of view.

4.3.1 Rayleigh Scattering

We first consider scattering due to molecular air. This process is commonly referred to as Rayleigh scattering. The number of photons scattered per unit length is given by [37]:

$$\frac{dN_\gamma}{dl} = -\rho \frac{N_\gamma}{x_R} \left(\frac{400}{\lambda}\right)^4 \tag{4.11}$$

with ρ being the atmospheric density and $x_R = 2970 gm/cm^2$ [33] being the mean free path for scattering at 400 nm. The angular distribution is given by:

$$\frac{d^2N_\gamma}{dl d\Omega} = \frac{dN_\gamma}{dl} \frac{3}{16\pi} \left(1 + \cos^2\theta\right) \tag{4.12}$$

By using equation 4.12 we can now calculate the contribution of Rayleigh-scattered cherenkov light to the total light observed for each segment of the angular track. However, we must also take into account the effective attenuation of the observed light between the track and the point of observation. The transmission coefficient for light traveling through molecular air for a path length of $\Delta x (gm/cm^2)$ is:

$$T_R = \exp\left[-\frac{\Delta x}{x_R}\left(\frac{400}{\lambda}\right)^4\right] \tag{4.13}$$

4.3.2 Aerosol Scattering

Aerosol scattering is treated approximately the same way as Rayleigh scattering. However, because the precise nature and concentration of the aerosols is variable,

an entirely analytical treatment is not possible. In general, the number of photons scattered by aerosol per unit length per steradian is given by:

$$\frac{d^2N_\gamma}{dld\Omega} = -\frac{N_\gamma}{L_M(\lambda)}\rho_a(h) \times \varphi(\theta). \tag{4.14}$$

There are a number of free parameters in this equation that we must either determine from prior observations or with our own atmospheric calibration. First, $\rho_a(h)$ is the reduced aerosol density at height h above the ground. For the purpose of our analysis, we assume an exponential aerosol profile:

$$\rho_a(h) = e^{-(h-h_m)/H_a}, \tag{4.15}$$

where H_a is the scale height of the aerosols and is set to 1.0 km from laser measurements. This assumes an aerosol structure comparable to the U.S. Standard Desert Atmosphere.

The scattering phase function used, $\varphi(\theta)$, is based upon the Longtin desert atmosphere aerosol model at a wind speed of 10 m/sec for 550 nm light. This is shown in Figure 4.1 [33].

Finally, data from the laser calibration system is used to set the average extinction length, $\lambda_m(334\ nm) = 25\ km$. The wavelength dependence of λ_m used is based on the Etterman extinction model, with the fractional difference from 334 nm shown in Figure 4.2 [33].

By considering each segment of the track, we now calculate the contribution of the aerosol-scattered cherenkov light to the total observed light using equation 4.14. We also must find the aerosol attenuation between the track and the point of observation. For a given reduced slant depth Δs m (the integral of ρ_a from the track to the point of observation) the transmission factor is:

$$T_A(\lambda) = e^{-\Delta s/L_M(\lambda)} \tag{4.16}$$

4.3.3 Ozone Absorption

We also consider the effect of ozone absorption. This is a small effect in comparison to molecular and aerosol scattering but is included for completeness

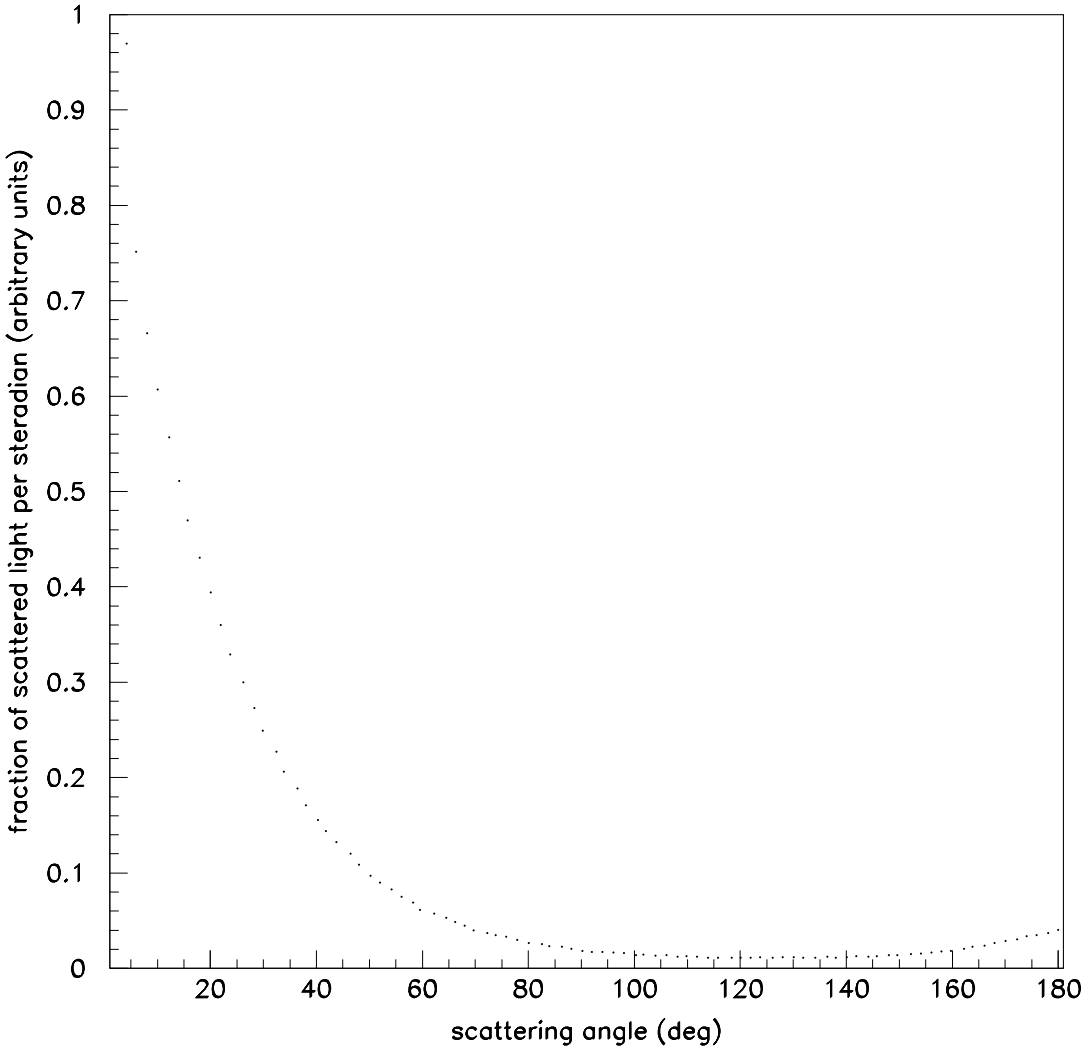


Figure 4.1. The Longtin phase function.

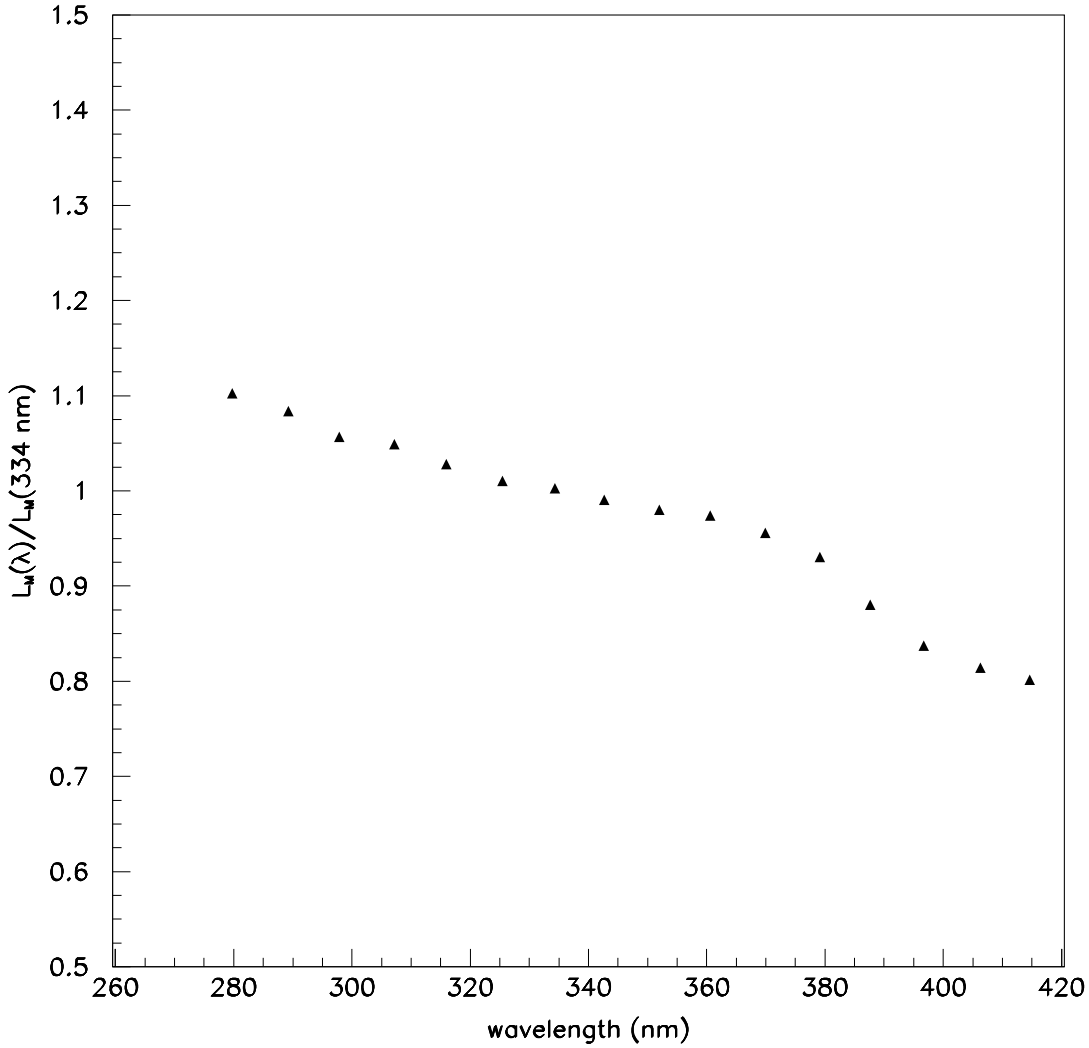


Figure 4.2. The Etterman model of the wavelength dependence of the aerosol extinction length.

The transmission coefficient due to ozone is:

$$T_{O_3} = exp(-\Delta x_{O_3} A_{O_3}(\lambda)) \tag{4.17}$$

There are two empirically determined parameters in equation 4.17. The ozone attenuation coefficient, $A_{O_3}(\lambda)$, is shown in Figure 4.3 [33]. The integrated ozone density, Δx_{O_3} relies upon a working knowledge of the ozone density profile as a function of altitude. This is shown in Figure 4.4 [33].

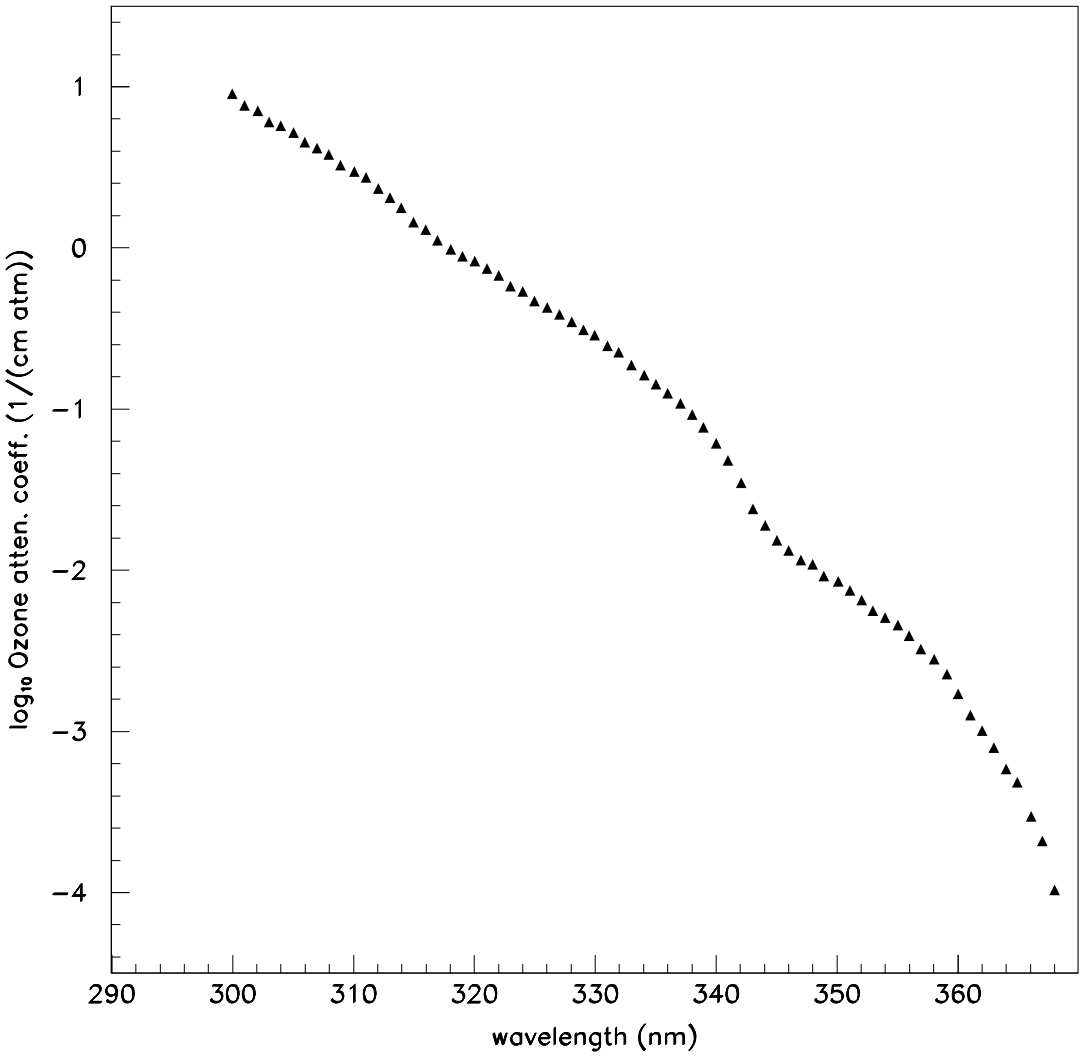


Figure 4.3. The ozone attenuation coefficient as a function of wavelength.

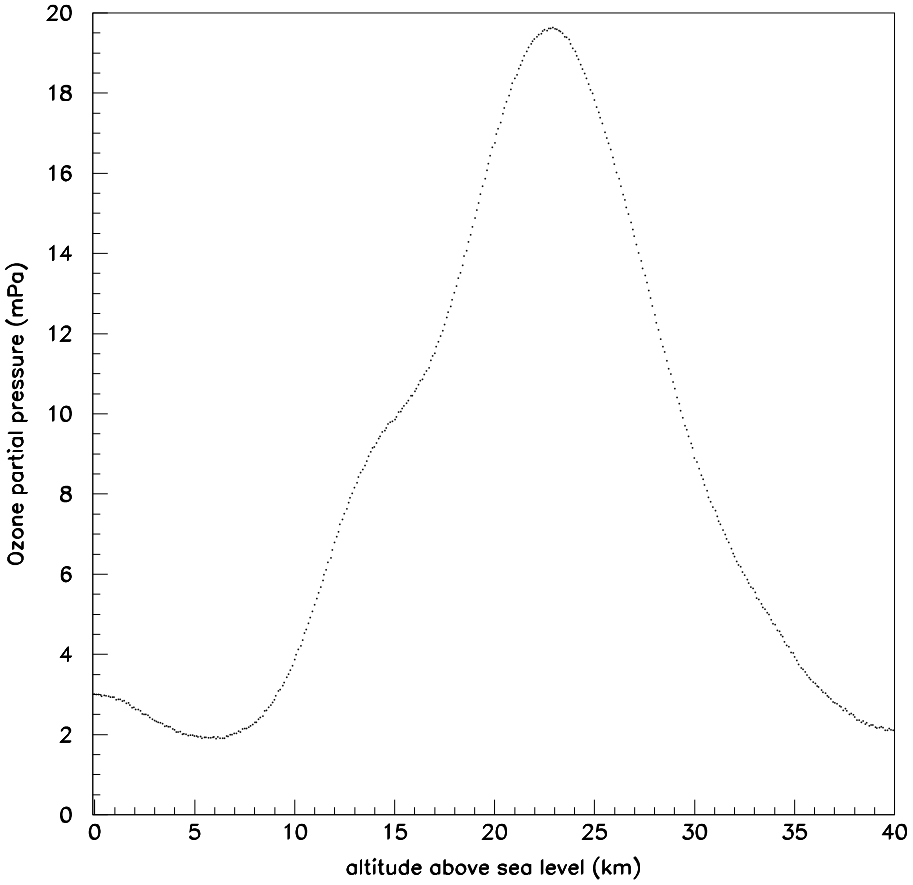


Figure 4.4. Ozone concentration as a function of altitude.

4.4 Detector Optics

We now turn to the task of simulating the amount of light that reaches the HiRes-I detector from a given EAS. This is done in two steps: For each track segment, an analytical calculation is done to establish the expectation value for the number of photons that reach the PMTs at HiRes-I. The actual number of photons is then fluctuated according to the Poisson distribution corresponding to the calculated mean. Then, each individual photon is ray-traced in order to take into account the clover-leaf shape of the individual mirrors and the obscuration of the PMT cluster.

The analytic calculation relies upon the formulae given in the previous two sections. First, the number of photons are calculated using by multiplying the expressions of $d^2N_\gamma/dld\Omega$ by the length of a given track segment $\delta\gamma$. Then, number of photo-electrons collected by the track segment are found by a disk-shaped mirror are calculated by:

$$\begin{aligned} N_{pe} &= \sum_{\lambda=300}^{420} \left(\frac{dN_\gamma}{d\Omega}(\lambda,\theta_e)\Big|_{scin} + \frac{dN_\gamma}{d\Omega}(\lambda,\theta_e)\Big|_{ascat} + \frac{dN_\gamma}{d\Omega}(\lambda,\theta_e)\Big|_{rscat} \right) \\ &\times \; T_R(\lambda)\, T_A(\lambda)\, T_{O_3}(\lambda)\, T_{UV}(\lambda)\, R_m\, QE(\lambda)\, \delta\Omega. \end{aligned} \tag{4.18}$$

For this calculation, λ is summed over 1 nm steps and θ_e is the light emission angle. The labels $scin$, $ascat$, and $rscat$ refer to light provided by scintillation, aerosol scattering of cherenkov light, and Rayleigh scattering of cherenkov light. The transmission coefficients of the Rayleigh scattering, aerosol scattering, and ozone absorption are taken into account by T_R, T_A, and T_{O_3}, respectively. Detailed detector characteristics are taken into account by T_{UV}, the UV filter transmission coefficient; R_m, the mirror reflectivity; QE, the PMT quantum efficiency; and $\delta\Omega$ which is the solid angle subtended at the center of the track segment by the disc shaped mirror placed at the observation point.

Once the expected number of photons per track segment is established, we then use a Monte Carlo implementation of ray tracing to numerically simulate the actual response of the HiRes-I detector optics to an EAS. For each track segment, the mean number of photo-electrons found in equation 4.18 are taken and Poisson

fluctuated. For each photo-electron, a longitudinal position in the track segment is randomly chosen and a lateral position is chosen according to the NKG distribution in equation 2.5. The photo-electron is then ray-traced to the PMT in order to take into account the obscuration of the PMT cluster and the actual clover shape of the real mirrors. At the PMT, the position of the photo-electron is fluctuated on a Gaussian with $\sigma = 0.25\ cm$ in order to take into account the mirror imperfection. Finally, this position is tested to see if it falls into the acceptance of a PMT. If the photo-electron falls within the PMT acceptance, then it is weighted by the measured tube response profile and added to the observed signal for that PMT.

4.5 Detector Electronics

The simulation of the HiRes-I detector electronics is primarily concerned with two issues:

1. The modeling of the interaction between the S/H electronics and the individual tube signal

2. The application of the detector trigger requirements

The time interval for a shower observation consists of three parts:

1. The actual shower transit time, T_e, across the mirror

2. A pre-shower interval, $T_s = 25\ \mu s$, that takes into account tube noise triggers formed prior to the transit of the shower

3. A post-shower interval, $T_h = 50\ \mu s$, that allows tubes to finish triggering once the shower has crossed the field of view

Thus the total time interval that must be considered is: $T_{tot} = T_e + T_s + T_h$.

In our simulation, the T_{tot} interval is divided into 20 ns time bins. Each tube in the PMT cluster is assigned its own set of bins. Each bin for each tube is then filled with photo-electron contribution from the shower. Additionally, sky noise fluctuations are added to the individual bins from a Poisson distribution with a

mean value of 40 photo-electrons per μs. The mean value is suppressed to reflect the AC coupling of the tube anode signal.

Once all of the bins are filled with the expected photo-electrons, each tube is scanned for one or more triggers during the interval T_{tot}, where the signal processing through the pre-amplifier, amplifier, and low-pass filters are emulated digitally. We also simulate the splitting and delays of signal strength through the separate trigger and integration/digitization channels (see Figure 3.3). Signals in the trigger channel exceeding the preset threshold of 700 mV in any 20 ns bin initiate an integration in the delayed channel of 5.6 ns, duplicating the functionality of the actual electronic circuitry on the OMB.

The final step is to sort all of the recorded tube triggers in ascending chronological order and to search for an event trigger. This is done by first searching for subcluster triggers (see Figure 3.5). A subcluster trigger requires that three tubes within a subcluster (two of which are adjacent) trigger within a 6 μs time interval. Once all of the subcluster triggers are recorded, they are then scanned to see if any two coincide in a 25 μs time interval, which would then create an mirror-level trigger. If no mirror-level trigger is found in the list of tube triggers, it is assumed that the event would not have generated a trigger in the HiRes-I detector.

CHAPTER 5

DATA PROCESSING AND EVENT RECONSTRUCTION

In the course of its operational history, the HiRes-I detector has collected on the order of 10^8 event triggers. The vast majority of these triggers are due to electronic and sky noise. Additionally, flashers and lasers produce event triggers as part of the event calibration system. The data processing chain filters out these artificial triggers and then selects events whose geometries are suitable for reconstruction.

5.1 PASS0 and PASS1: Event Preparation

The first step in the data processing chain is to pair together event triggers and times. For each trigger an event packet is recorded by the main data acquisition (DAQ) computer. Simultaneously , the central timing crate records a GPS time for each event packet as part of a time packet sent to the DAQ computer once per second. The program *hma* then pairs event packets with their respective times to form the PASS0 stage in the data processing chain.

The next step, PASS1, involve using the program *hpass1* to apply event calibration. Raw TDC values are converted to PMT trigger times measured relative to the time of the mirror trigger. QDC values are converted to numbers of photo-electrons employing the values from the routine roving xenon flasher (RXF) calibration and pedestals from the nightly electronics calibration.

5.2 PASS2: The Rayleigh Filter

Once we have the processed events, we can filter out noise triggers to produce the PASS2 stage data. The first step in this part is to search for track-like events using a Rayleigh filter. Basically, this method employs the formula:

$$plog = \frac{R^2}{N \ln 10}, \tag{5.1}$$

where R is the length of the vector and N is the number of steps. *Plog* is the negative of the base-10 logarithm of the probability that a given vector was the result of a random walk process

In practice, the hit tubes are first ordered by the time that they triggered. A *step* is then added for each nearest neighbor pair (tube within 1.5°) triggered beginning with the earlier tube and ending with the later tube. All of the steps are added vectorially and the resulting value of *plog* is then calculated according to equation 5.1. Events with $plog \geq 2.0$ (i.e., with a 1% probability that they are due to a random process) are kept. This initial filter reduces the total number of events by $\sim 95\%$ with the vast majority of the remaining events being due to the atmospheric calibration systems.

5.3 PASS3: The Plane Fit

Figure 5.1 shows the geometry of a shower relative to the HiRes-I detector. In order to reconstruct the shower trajectory, we first identify the shower-detector (SD) plane by minimizing the function:

$$\chi^2 = \sum_i \frac{[(\hat{n} \cdot \hat{n}_i)]^2 \cdot w_i}{\sigma_i^2}, \tag{5.2}$$

where the sum is over triggered tubes, $\hat{n}$ is the plane normal, $\hat{n}_i$ is the tube viewing direction unit vector and w_i is the number of photo-electrons seen by tube i. An angular error of $\sigma = 1.0°$ is assumed for all tubes. By initially selecting only tubes that occur in clusters of three or more, we can obtain a rough value for $\hat{n}$. We then use that value to reject tubes that are either spatially or temporally uncorrelated to the shower track. This process is iterated until either there are no more rejected tubes or less than three tubes remain.

The event rejection criteria during PASS3 include:

1. Failure of the plane fit (i.e., less than three remaining tubes)

2. Angular track length of less than 6°

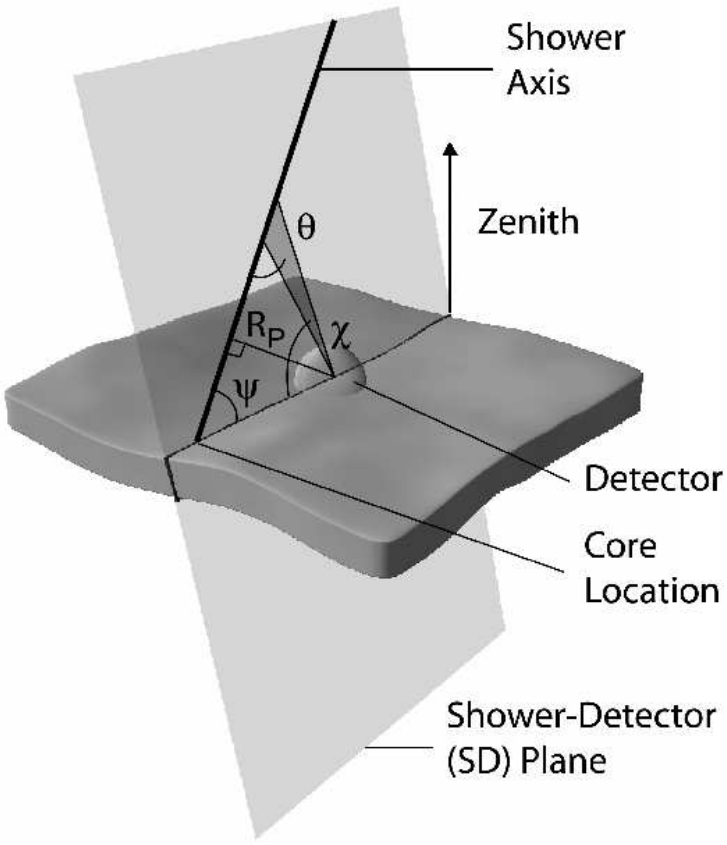


Figure 5.1. Shower geometry relative to detector.

3. Average number of photons per good tube less than 75

4. Track width RMS greater than 1°

5. Angular speed greater than $5.73°/\mu s$ which corresponding to a distance to a vertical shower, otherwise know as a "pseudo-distance" of 3 km.

Further filtering is done in PASS3a and PASS3b. PASS3a employs a "correlation cut" where we calculate the correlation coefficient, $\rho(x,y) = Cov(x,y)/(\sigma_x\sigma_y)$ where the two variables in question are the tube trigger times and the χ_i values from Figure 5.1. PASS3b tightens two of the cuts from PASS3. Specifically, the angular speed cut is lowered to $3.33°/\mu s$ (corresponding to a pseudo-distance of 5 km) and the minimum average number of photons per good tube is increased to 200. These two additional cuts are specifically applied for monocular reconstruction using the profile-constrained fit (PCF) described below.

5.4 PASS4 Profile-Constrained Reconstruction

While the original Fly's Eye experiment was able to reconstruct shower geometry purely with a timing fit of tube signals, simulations showed that this was not the case for HiRes-I. Because a single-ring detector observes much shorter track lengths than a detector with full sky coverage, R_p values are frequently underestimated by as much as 75% in simulations when a pure timing fit is employed. This necessitates the implementation of a profile-constrained reconstruction routine, where, in addition to timing, the expected shower profile is used as a constraint. This is done by fitting both the tube timing and the tube signal amplitude information in parallel. By varying the profile to which the tube signal amplitude is fit, the best *combined* χ^2-fit, of both the profile and timing information, is then chosen as the optimal solution.

First, we consider the timing fit. In Figure 5.2 we can see the relationship between shower viewing angles, χ_i, and shower viewing times, t_i. The two are related by the following equation:

$$t_i = t_0 + \frac{R_p}{c}\tan\left(\frac{\pi - \psi - \chi_i}{2}\right), \tag{5.3}$$

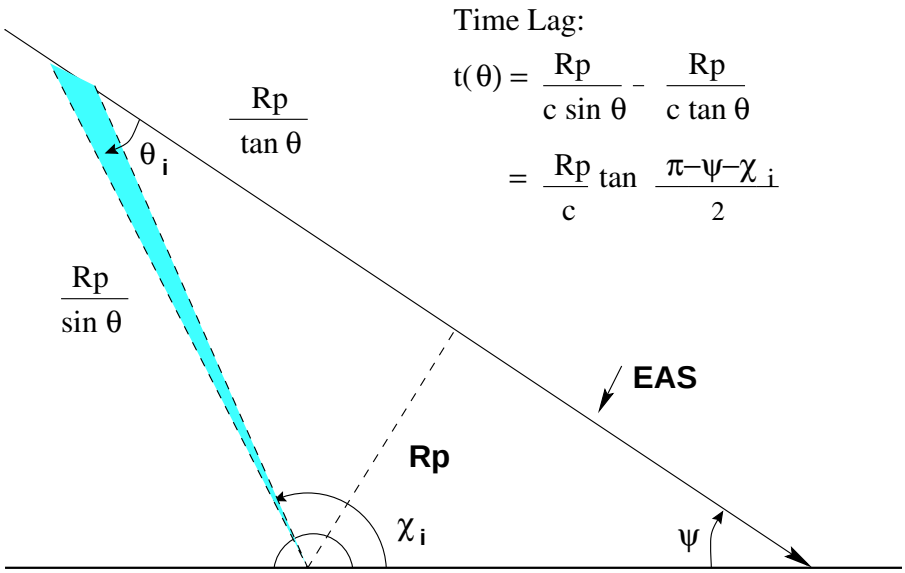


Figure 5.2. Track geometry in the Shower-Detector plane.

where t_0 is the time at which the shower passes through the point of closest approach, R_p, the impact parameter and ψ is the shower in-plane angle shown in Figure 5.2. Once this relationship is known, we can then reconstruct an observed shower track by minimizing the χ^2 function:

$$\chi^2_{tim} = \sum_i \frac{1}{\sigma_i^2} \left\{ t_i - \left(t_0 + \frac{R_p}{c} \tan\left(\frac{\pi - \psi - \chi_i}{2} \right) \right) \right\}^2, \tag{5.4}$$

where $\sigma_i = (500/\sqrt{S_i})\ ns$ is the estimated uncertainty in tube signal times with S_i being the tube signal amplitude in photo-electrons.

While the timing fit is sufficient when considering the longer angular track lengths ($\geq 40°$)) provided by the full-sky coverage of the Fly's Eye detector, the vast majority of HiRes-I events have track lengths of less then 15°. For these events, the corresponding $t_i(\chi_i)$ span is so short that equation 5.3 is nearly linear. Previous simulation studies showed that for a timing fit, $\sigma_{R_p} \sim 1/L^{5/2}$ [38], where L is the angular track length. This rapid deterioration of resolution for shorter track lengths forces us to place additional constraints on our χ^2 minimization.

The shower profile fit for a given shower geometry is based on:

$$\chi^2_{pfl} = \sum_i \frac{1}{\sigma_i^2}(S_i^{(m)} - S_i^{(p)})^2, \tag{5.5}$$

where $S_i^{(m)}$ is the observed tube signal strength (in photo-electrons), $S_i^{(p)}$, is the signal strength predicted by simulation, and $\sigma_i^2 = S_i^{(m)} + 200$ which is the uncertainty of the tube signal strength added in quadrature to the ambient sky noise measured at HiRes-I. The $S_i^{(p)}$ values are determined for a given geometry by the Monte Carlo simulation described in Chapter 4. The profile χ^2 is minimized with respect to the Gaisser-Hillas profile parameters x_{max} and N_{max} as seen in equation 2.4.

The actual reconstruction is done for a series of trial x_{max} values that include 650, 685, 720, 755, 790, 825, and 860 gm/cm^2. Since 10^{18} eV proton shower has an average x_{max} of 725 gm/cm^2 and a 10^{20} eV proton shower has an average x_{max} of 835 gm/cm^2, these trial values cover the range of x_{max} expected for hadronic showers.

For each x_{max} value, the shower profile fit is made while varying ψ over the range $[0, 180^\circ]$ with the R_p value specified by the relation in equation 5.3. Once χ^2_{pfl} is minimized with respect to ψ, χ^2_{tim} is calculated for the specified geometry. Both χ^2 values are then normalized per degree of freedom and a combined $\chi^2_{com} = \chi^2_{pfl} + \chi^2_{tim}$ is calculated for the given value of x_{max}.

Next, the best fits corresponding to the seven x_{max} values are compared and the one with the lowest value of χ^2_{com} is chosen. Additional quality cuts are then imposed, which include:

1. The angular track length is required to exceed 7.9°
2. The depth for the highest elevation hit must be less than 1000 gm/cm^2.
3. The value of ψ must be less than 120°.
4. No more than one angular bin can be contaminated with ¿25% Cherenkov light.

In Figure 5.3, the best profile fit for the first chronological event in the HiRes-I data set is shown.

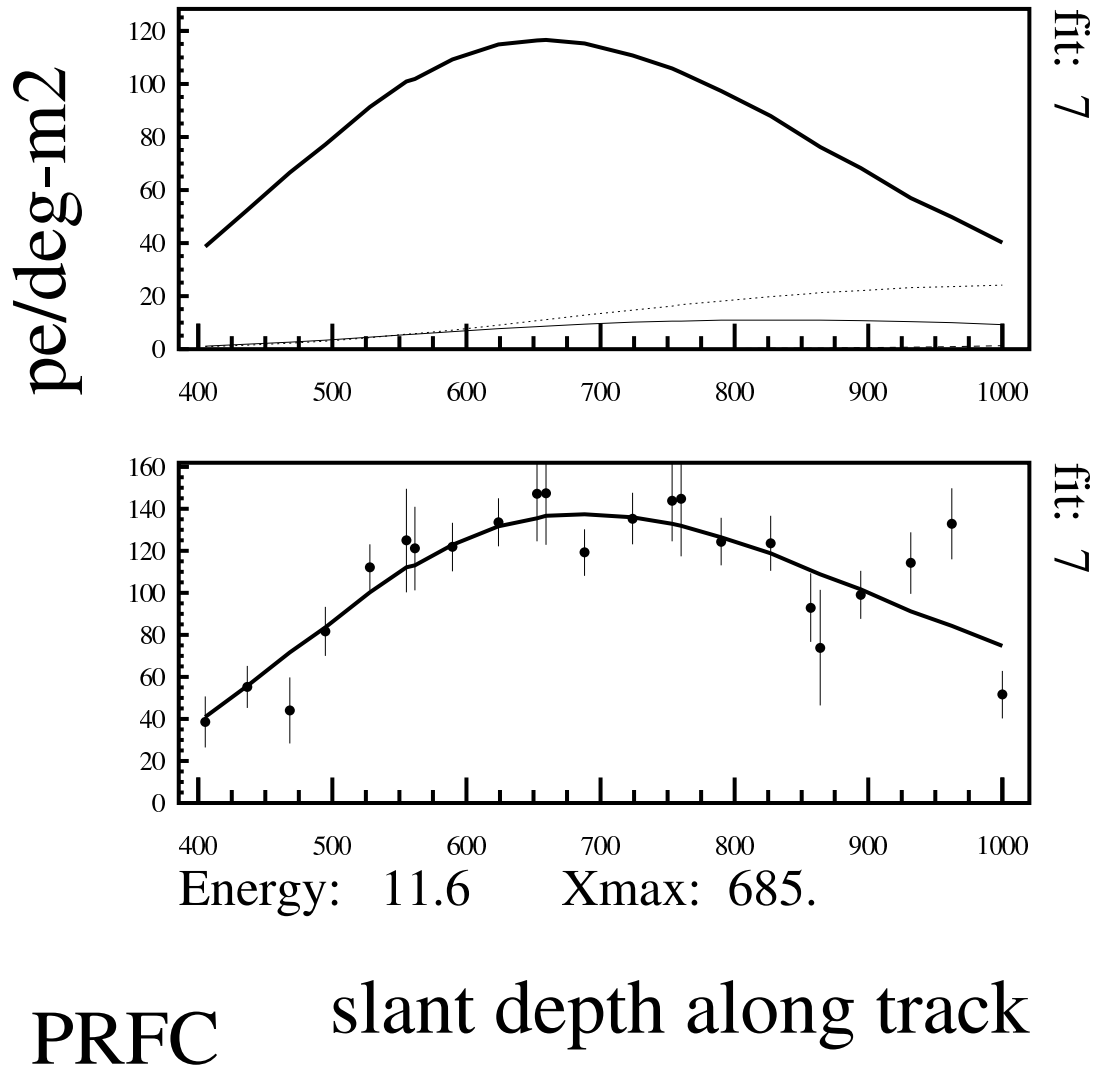


Figure 5.3. The profile-constrained geometry fit for the first event in the HiRes-I data set. This shower was observed on May 29, 1997 and has a reconstructed energy of 11.6×10^{19} eV.

CHAPTER 6

RECONSTRUCTION RESOLUTION

In order to properly utilize the HiRes-1 monocular data, we must carefully consider resolution of the profile-constraint reconstruction technique. This method leads to broad, energy-dependent resolution in both energy estimation and the determination of the arrival angle (ψ) within the plane of reconstruction. Furthermore, Monte Carlo studies indicate that there are energy-dependent systematic shifts in the central values of both the estimated energy and ψ when one compares the profile constraint fit results with input Monte Carlo.

6.1 Energy Resolution

For the purpose of our anisotropy studies, we are only considering events whose reconstructed energies are above $10^{18.5}\ eV$. This is the data set that was used in the published energy spectrum analysis [2, 3]. We first need to correct for the energy-dependent systematic *shift* in energy that is observed in Monte Carlo studies. The fractional energy shift can be shown to have the following relationship (see Figure 6.1):

$$\frac{\Delta E}{E} = 0.34556 e^{-\frac{\log_{10} E}{1.0206}} - 0.099633 \tag{6.1}$$

where E is the energy estimated by the profile constraint fit in EeV. For events with energies of $10^{18.5}\ eV$, this leads to a positive correction of $\sim 11\%$. For a data set with a differential spectrum proportional to E^{-3}, it leads to an $\sim 25\%$ increase in the total number of events over $10^{18.5}\ eV$. The energy resolution is obtained by studying the reconstruction of simulated events. In Figure 6.2, we see that the energy resolution at $10^{18.5}$ eV is approximately 26%, improving to 16% at $10^{19.5}$ eV.

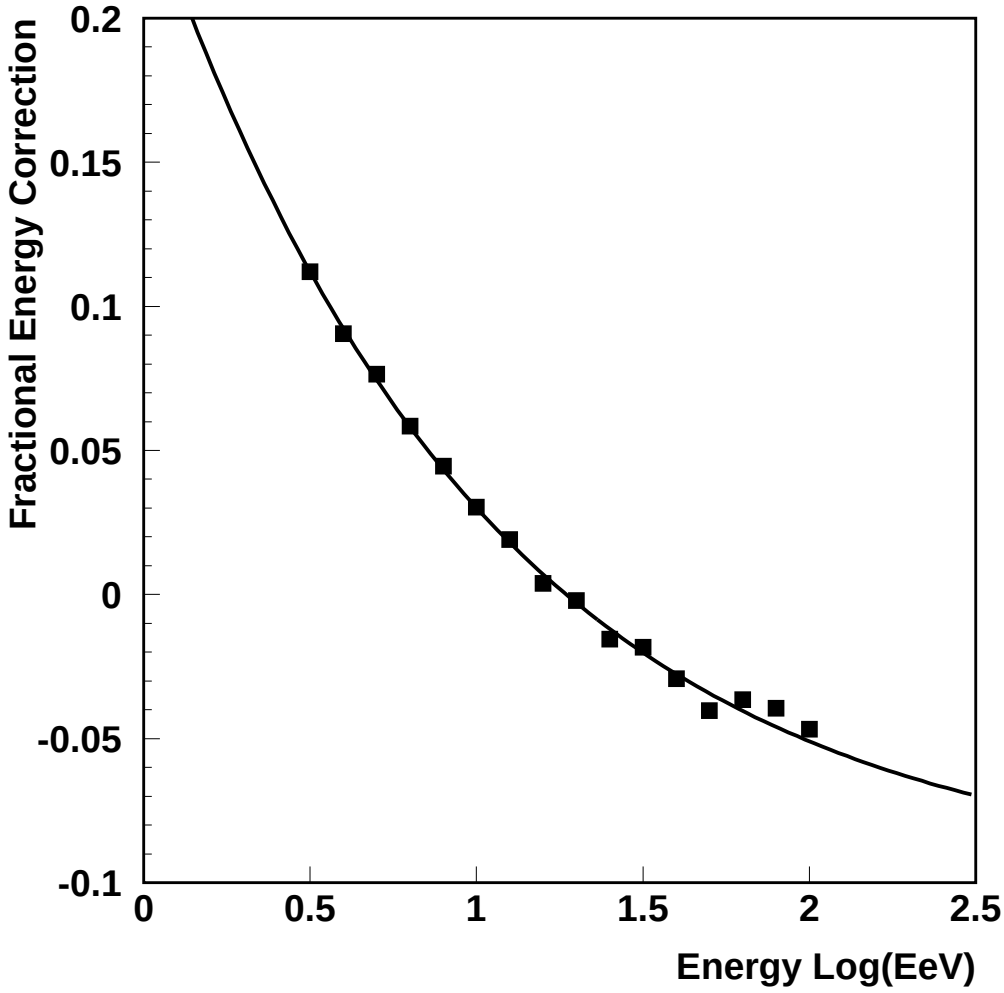


Figure 6.1. The mean fractional shift in energy between the Monte Carlo input and reconstructed energy estimated by the profile constraint fit.

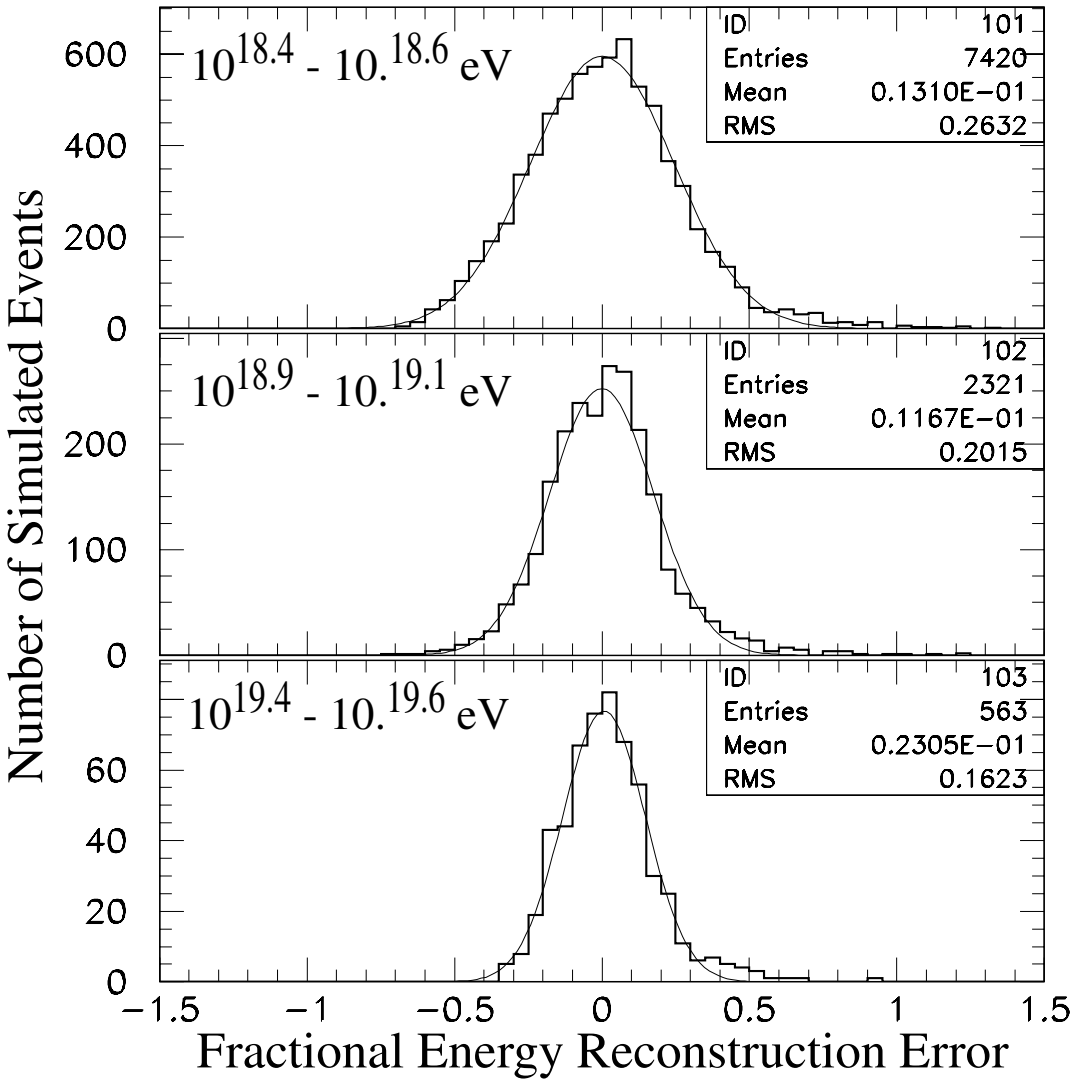


Figure 6.2. The fractional energy resolution observed by applying the reconstruction technique to simulated events.

6.2 Verifying the Energy Resolution and Correction Estimations

We now can perform a series of checks to verify that the simulation is accurately capturing the energy resolution of the events. First, we compare the geometrical distribution of events in the observed data and the simulation. This is done by considering the distribution of impact parameters, R_p (Figure 6.3) and zenith angles, θ (Figure 6.4).

As a further check, we consider the subset of events that were successfully reconstructed in both the HiRes-I monocular data set and in the HiRes stereo data set. Because the stereo events have a much better constrained geometry, we can use the stereo reconstruction to study monocular reconstruction in the same way we use simulated events. In Figure 6.5, we see that monocular reconstruction (with the correction applied) does not have a systematic bias when compared with stereo geometry. In Figure 6.6, we see that the distribution of fractional energy reconstruction errors is the same for simulation and the comparison of the energy values obtained via the monocular and stereo techniques.

6.3 Angular resolution

For a monocular air fluorescence detector, angular resolution consists of two components, the plane of reconstruction and the angle ψ within the plane of reconstruction. Figure 5.1 illustrates how this geometry would appear with a particular plane of reconstruction and a particular value for ψ. We can see that we should be able to determine the plane of reconstruction very accurately. However, the value of ψ is more difficult to determine accurately because it is highly dependent on the precise results of the profile constraint fit. The profile constraint fit determination of ψ is also subject to a systematic shift in its central value. This shift in can be characterized by its relationship to the reconstructed energy estimated by the profile constraint fit (see Figure 6.7):

$$\psi_{\mathrm{SHIFT}} = 18.353^{\circ} e^{-\frac{\log_{10} E}{0.66080}} - 2.7311^{\circ} \tag{6.2}$$

While the χ^2 fit itself can provide an error estimate for the fitted value of ψ,

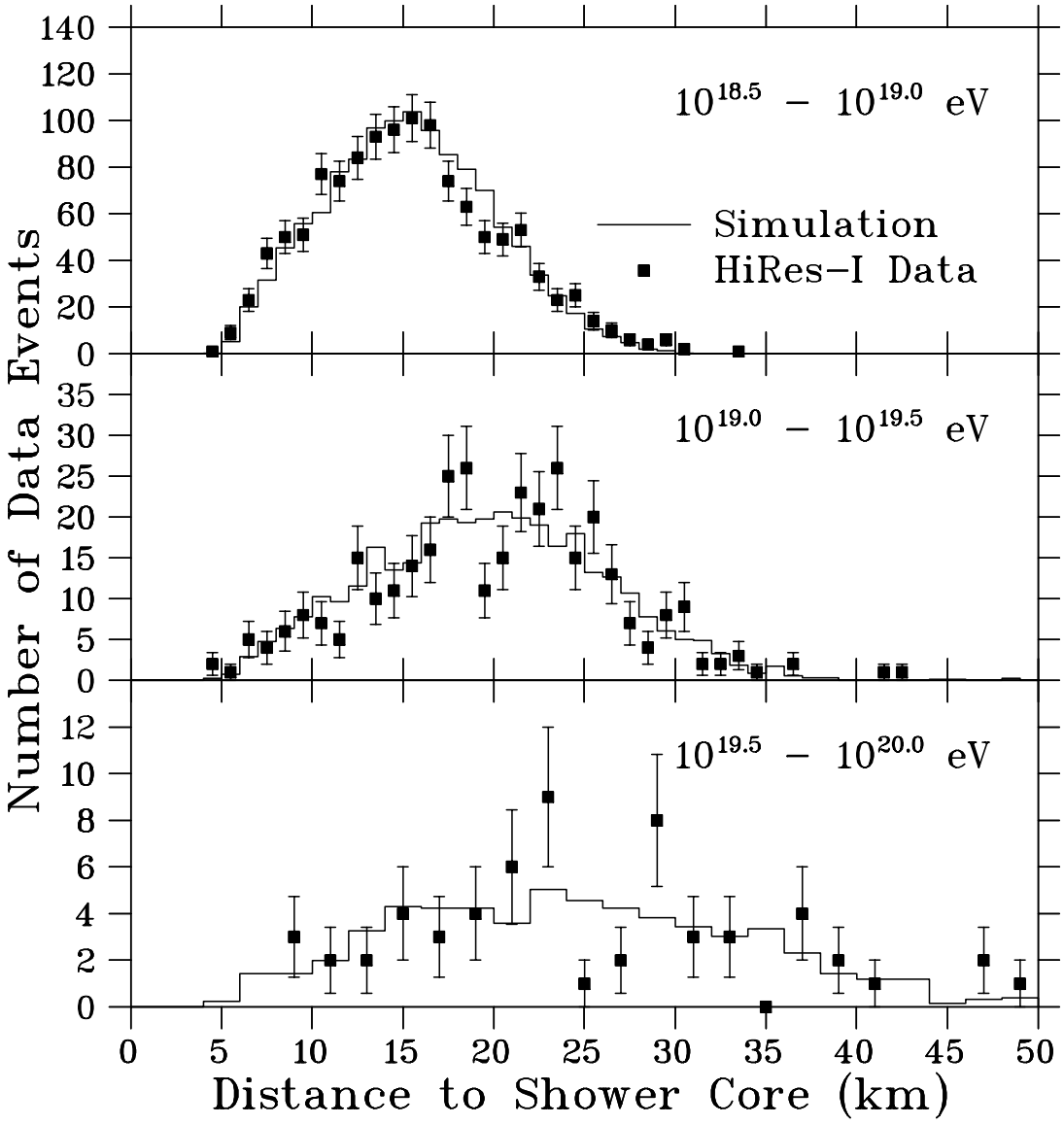


Figure 6.3. The distribution of R_p for real versus simulated events at different energies. In each case, the solid line histogram corresponds to distribution of R_p values for the simulation. The data points correspond to the specified angle distribution of the real data with Gaussian uncertainties assumed for each bin.

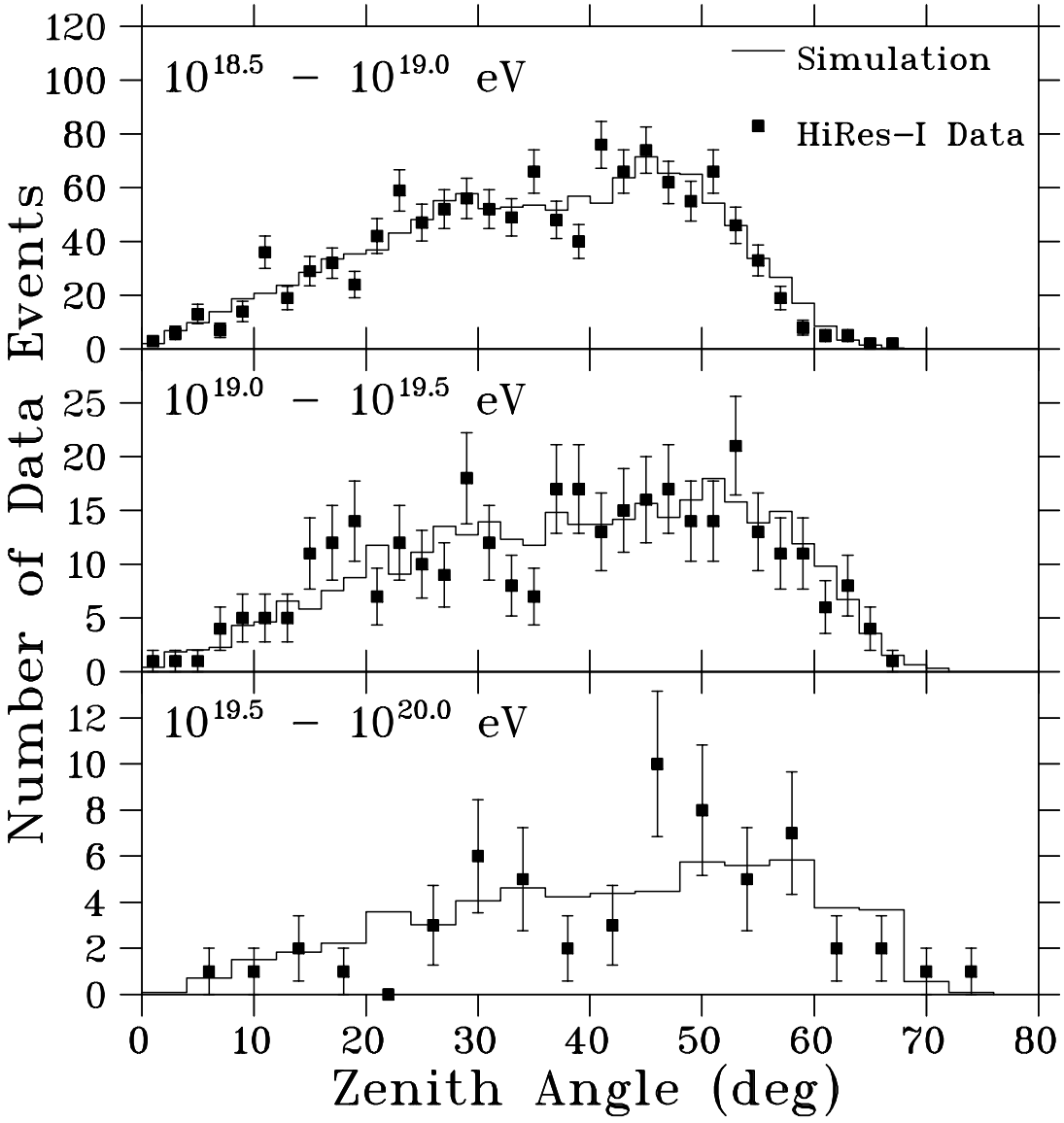


Figure 6.4. The distribution of zenith angles for real versus simulated events at different energies. In each case, the solid line histogram corresponds to distribution of zenith angle values for the simulation. The data points correspond to the specified angle distribution of the real data with Gaussian uncertainties assumed for each bin.

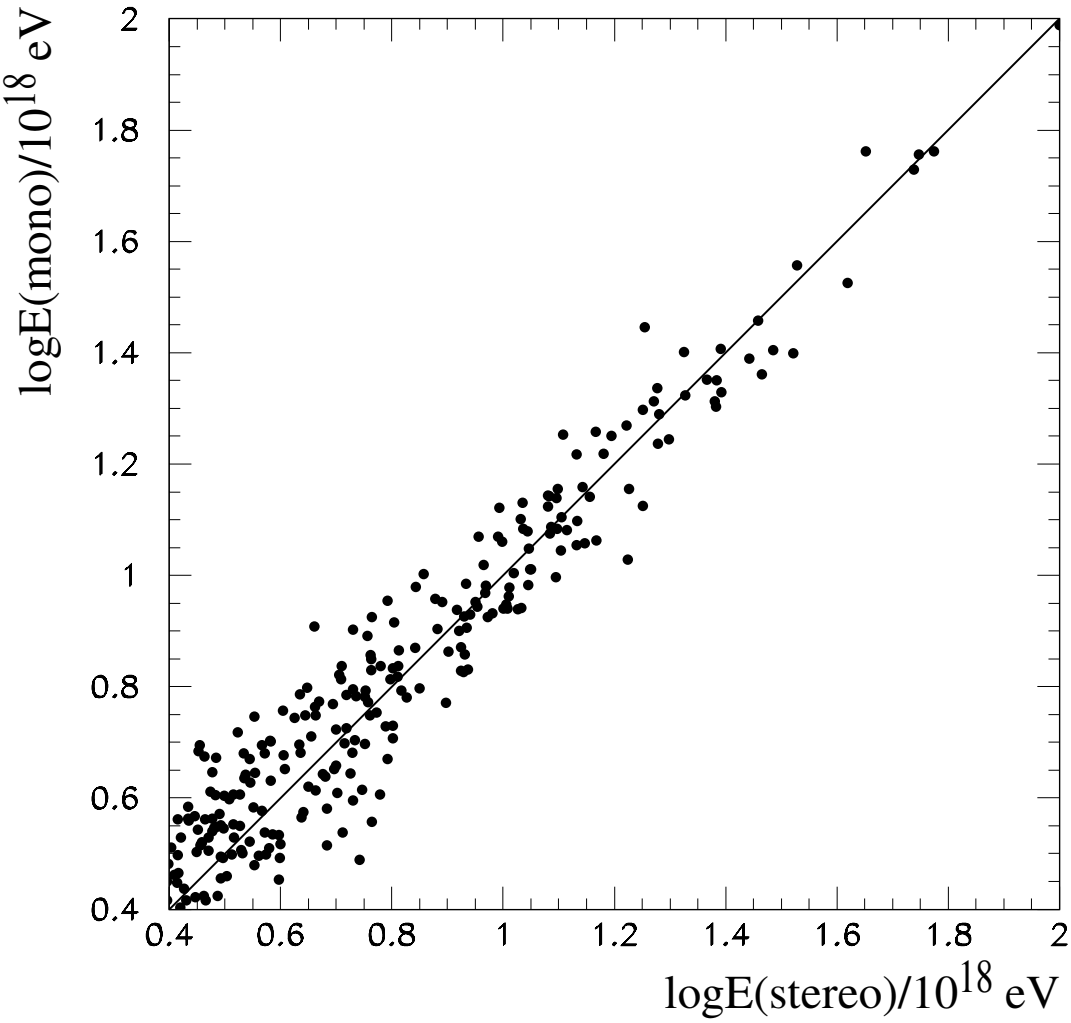


Figure 6.5. The comparison of the estimated energies from the HiRes-I monocular reconstruction and the HiRes stereo reconstruction.

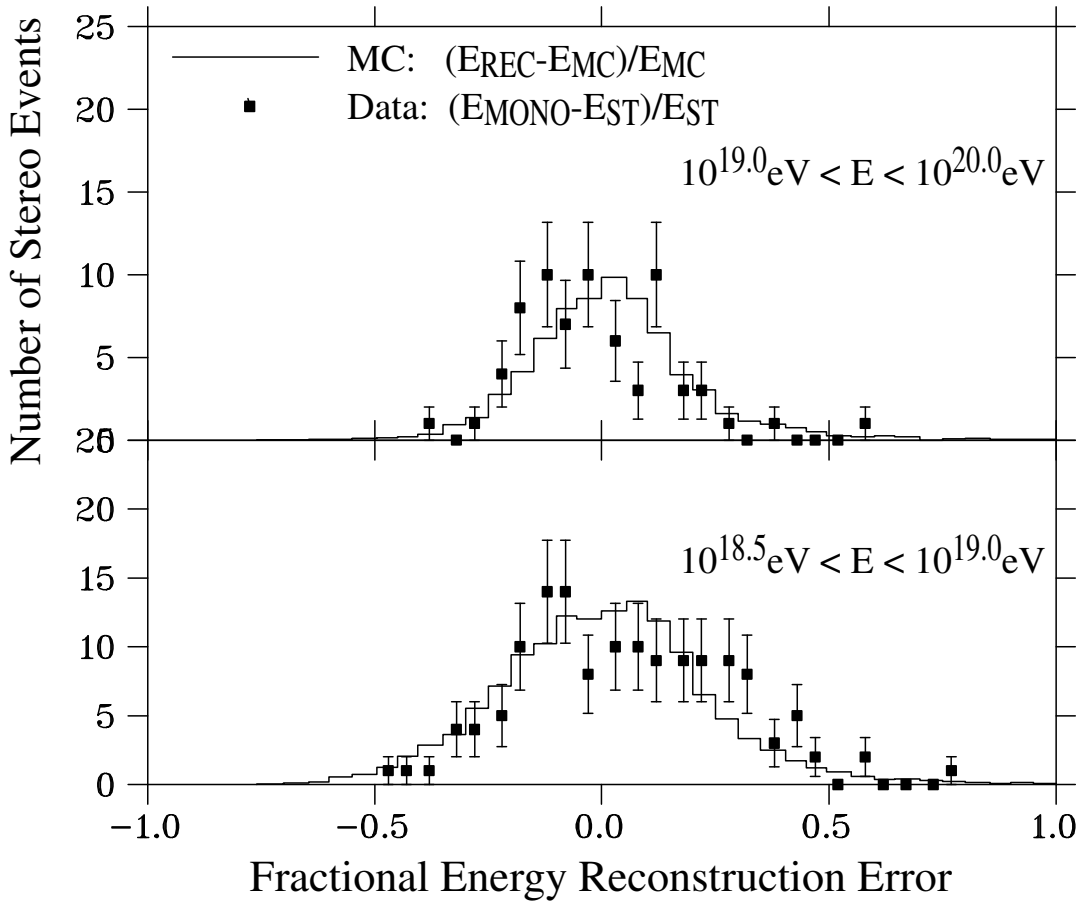


Figure 6.6. The distribution of fractional energy reconstruction error values for real versus simulated events at different energy ranges. In each case, the solid line histogram corresponds to the fractional energy error with simulated events. The data points correspond to the fractional energy reconstruction error of the real data constructed with the monocular and stereo techniques with Gaussian uncertainties assumed for each bin.

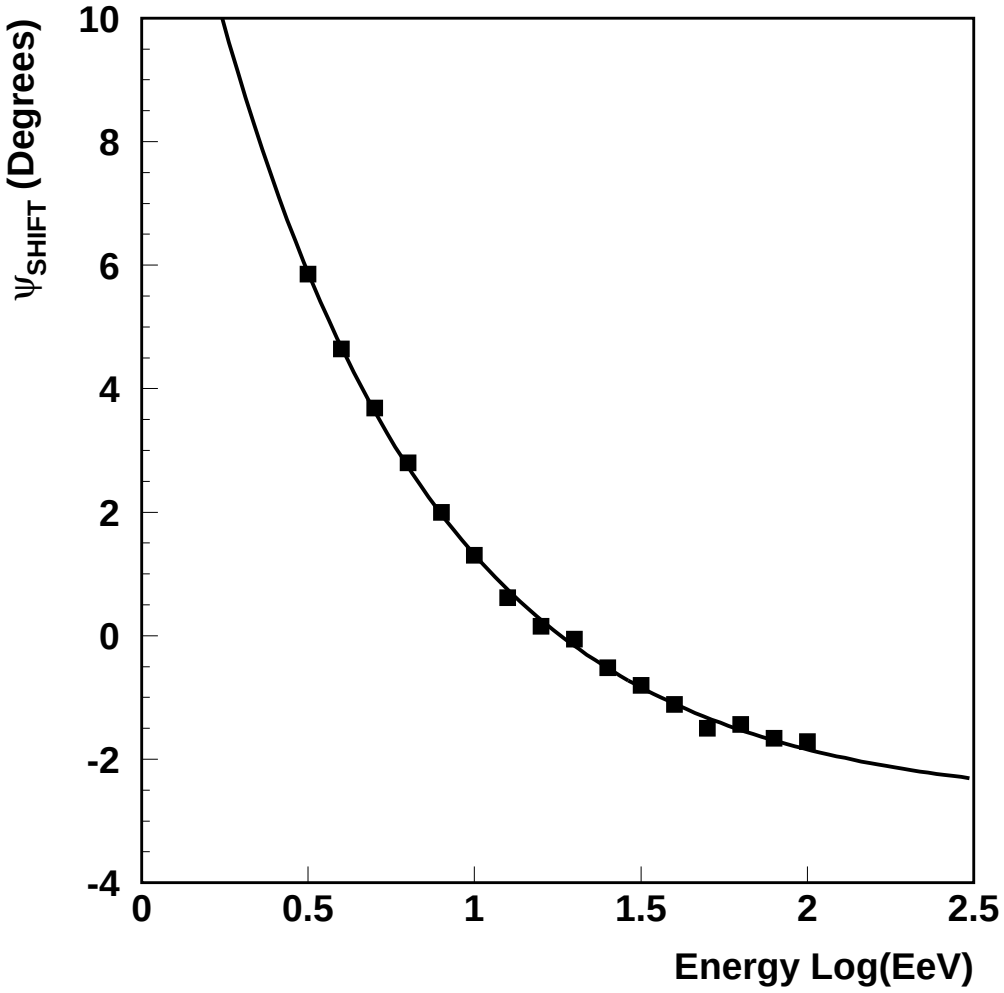


Figure 6.7. The systematic shift in the central value of ψ between the Monte Carlo input and profile constraint fit results.

the fact that systematic shift exists suggests that this error estimate is, by itself, not reliable. This can be confirmed by comparing the input ψ values of simulated events with the results of the profile constraint fit. The quantity that we are in fact interested in is: $|\psi_{\mathrm{MC}} - \psi_{\mathrm{PCF}}|/\Delta\psi_{\chi^2}$. The distribution for this quantity should be characterized by:

$$\left\langle \frac{|\psi_{\mathrm{MC}} - \psi_{\mathrm{PCF}}|}{\Delta\psi_{\chi^2}} \right\rangle = \frac{1}{\sqrt{2\pi}\sigma^2} \int_{-\infty}^{\infty} |x| e^{-\frac{x^2}{2\sigma^2}} dx = \sqrt{\frac{2}{\pi}} = 0.79788 \tag{6.3}$$

and

$$\left(\frac{|\psi_{\mathrm{MC}} - \psi_{\mathrm{PCF}}|}{\Delta\psi_{\chi^2}} \right)_{\mathrm{RMS}} = \left(\frac{1}{\sqrt{2\pi}\sigma^2} \int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{2\sigma^2}} dx \right)^{\frac{1}{2}} = 1 \tag{6.4}$$

where $\Delta\psi_{\chi^2}$ is the angular resolution predicted by the χ^2 fit. However, if we refer to the actual distribution of $|\psi_{\mathrm{MC}} - \psi_{\mathrm{PCF}}|/\Delta\psi_{\chi^2}$ in Figure 6.8 we in fact have a mean value of 2.221 ± 0.006 with an RMS of 3.115.

Customarily, the field of high-energy particle physics uses the resolution function (the distribution of the difference between reconstructed and input values from simulated data) to establish a measure of the average statistical uncertainty of a measurement. So a possible alternative would be to see if we can parameterize the angular resolution as a function of the reconstructed energy provided by the profile constraint fit. This can be done with the following relationship (see Figure 6.9):

$$\Delta\psi = 15.2^\circ e^{-\frac{\log_{10} E}{0.69085}} + 3.4^\circ \tag{6.5}$$

We verify that this parameterization is providing the proper errors by once again considering the distribution of $\frac{|\psi_{\mathrm{MC}} - \psi_{\mathrm{PCF}}|}{\Delta\psi}$, this time for with $\Delta\psi = \Delta\psi_{\mathrm{EP}}$, the value of $\Delta\psi$ predicted by the energy parameterization. The resulting distribution pictured in Figure 6.10 has a mean of 0.7786 ± 0.0019 and an RMS of 1.0435, values that conform well with theoretical values predicted in equations (6.3) and (6.4).

Before we abandon the χ^2 angular resolution predictions altogether, there is one more question that we should ask: Does the profile constraint fit behave the same on both simulated and real events? There is some reason to think that it would not.

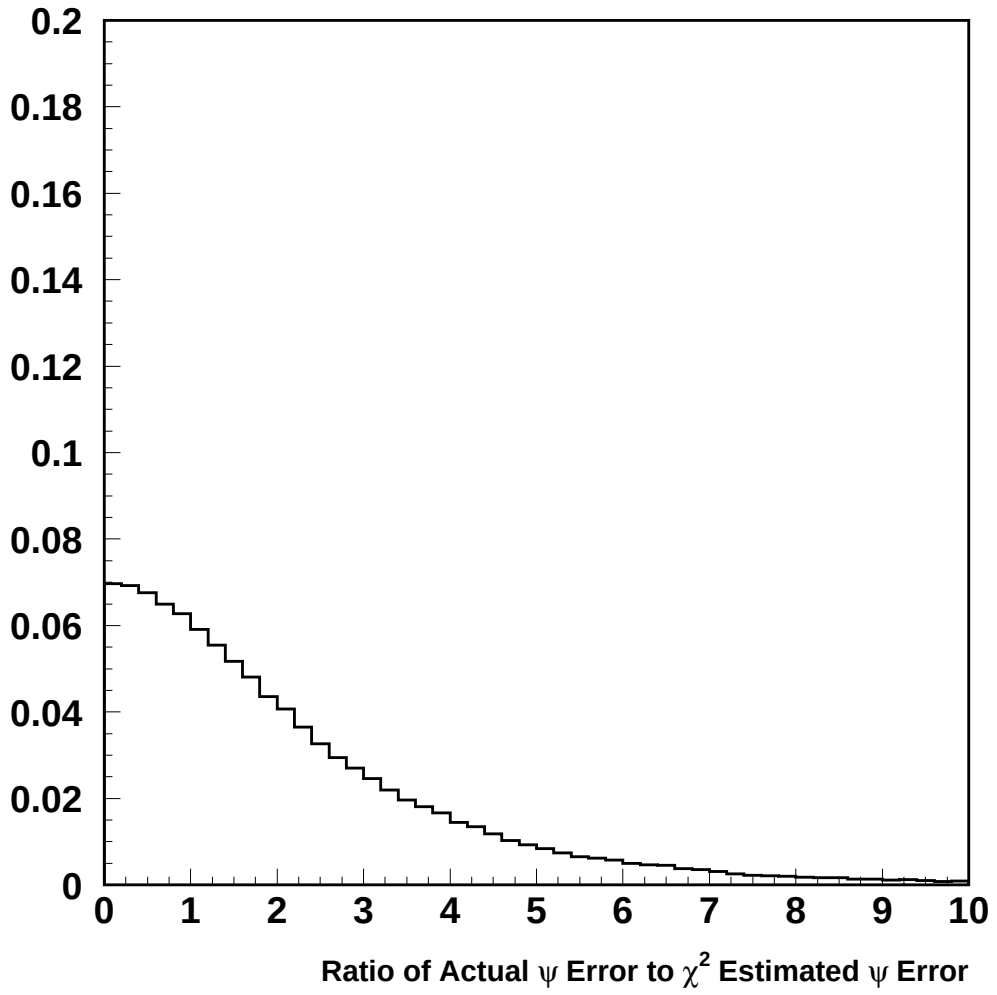


Figure 6.8. Comparison of the actual ψ error and the predicted ψ error for 135,186 simulated events using the χ^2 fit to predict the angular resolution ($\mu = 2.221 \pm 0.006$ and $RMS = 3.115$).

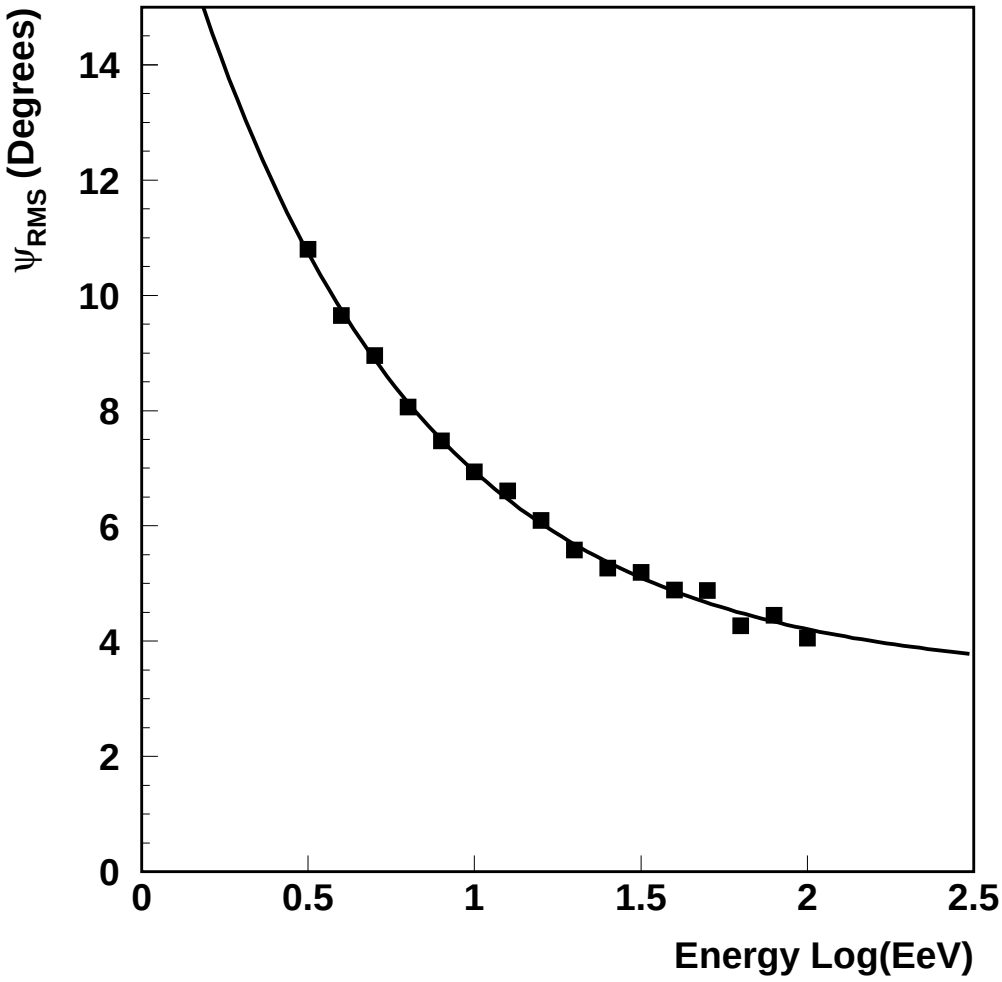


Figure 6.9. The the resolution of ψ determined from comparisons between Monte Carlo inputs and profile constraint fit results.

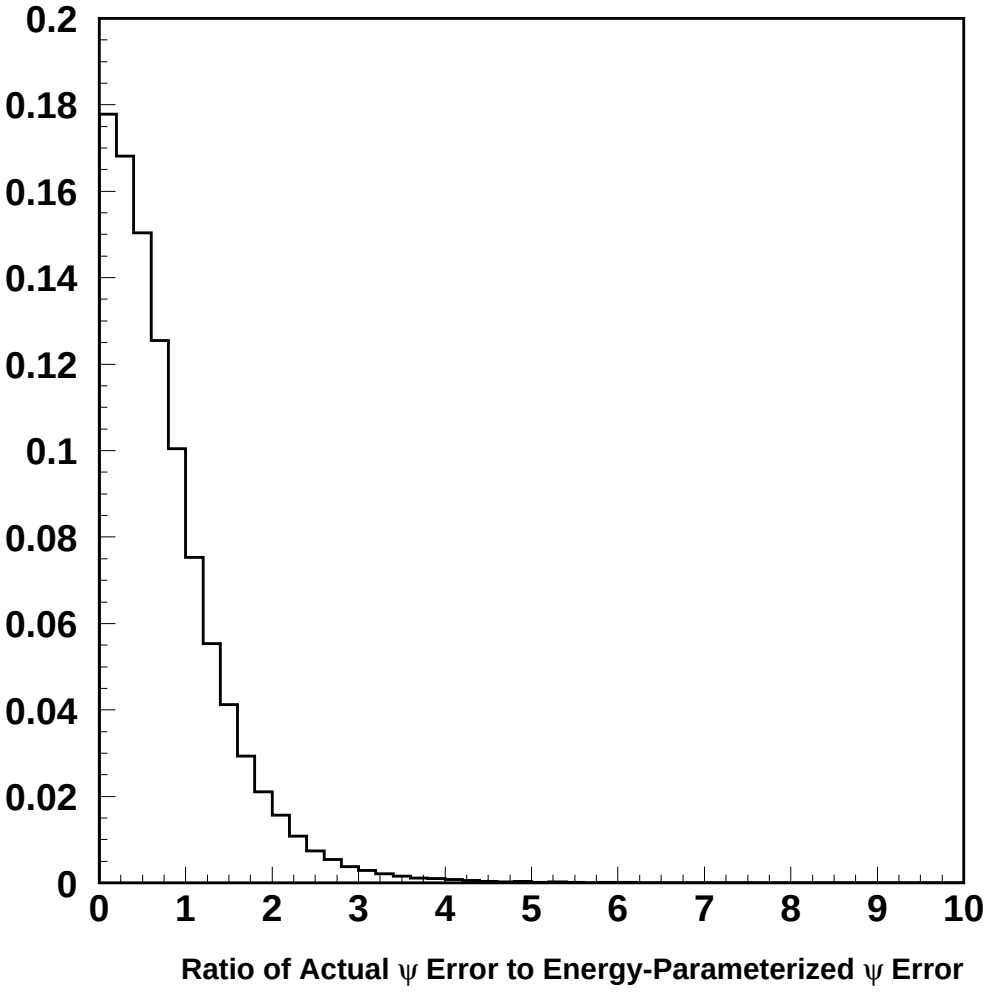


Figure 6.10. Comparison of the actual ψ error and the predicted ψ error for 135,186 simulated events predicting the angular resolution via an energy parameterization ($\mu = 0.7786 \pm 0.0019$ and $RMS = 1.0435$).

Calibration uncertainties will inevitably lead to rougher profiles with real events and it would impractical to develop an event simulation that actually took into account *all* of the fluctuations of a real particle shower. A simple check that we can provide would be to look at the *ratio* of $\Delta\psi$ predicted values from the χ^2-fit and from the energy parameterization for both simulated and *real* events. From Figures 6.11 and 6.12, we can see the distribution of ratios is noticeably broader for the real data. A possible solution to this is to put in a smearing constant of $s_\circ = 1.210$ for the Monte Carlo, that is $\Delta\psi = s_\circ \Delta\psi_\circ$ for the energy parameterized value of $\Delta\psi$. As we can see in Figure 6.13, the addition of this smearing constant produces a distribution of ratios that is very similar to the distribution of ratios for the simulated data. If we now multiply the smearing constant into equation (6.5), we will get the corrected expression for $\Delta\psi$:

$$\Delta\psi = s_\circ \Delta\psi_\circ = 18.4^\circ e^{-\frac{\log_{10} E}{0.69085}} + 4.1^\circ \tag{6.6}$$

For the actual analysis described in later chapters, equation (6.6) is used fro the parameterization of ψ resolution. Correspondingly, the reconstruction error for simulated events are additionally smeared by the factor $s_\circ$.

We also must take into account the error in the determination of the plane of reconstruction, Δn. The plane of reconstruction is not determined by a separate fit that is applied earlier in the data processing chain. We again use the comparison between Monte Carlo input and the reconstructed output to parameterize the error in plane determination. However, instead of parameterizing the error in plane determination with the reconstructed energy provided by the profile constrain fit, we will use the angular track length, $\Delta\chi$ (in degrees) of the event. The actual parameterization is as follows (see Figure 6.14):

$$\Delta n = 72.889^\circ e^{-\frac{\Delta\chi}{1.9595}} + 0.30908^\circ \tag{6.7}$$

A comparison between the results of the plane fitter applied to events with and without the RXF calibration, one will see a 12% difference in the events that

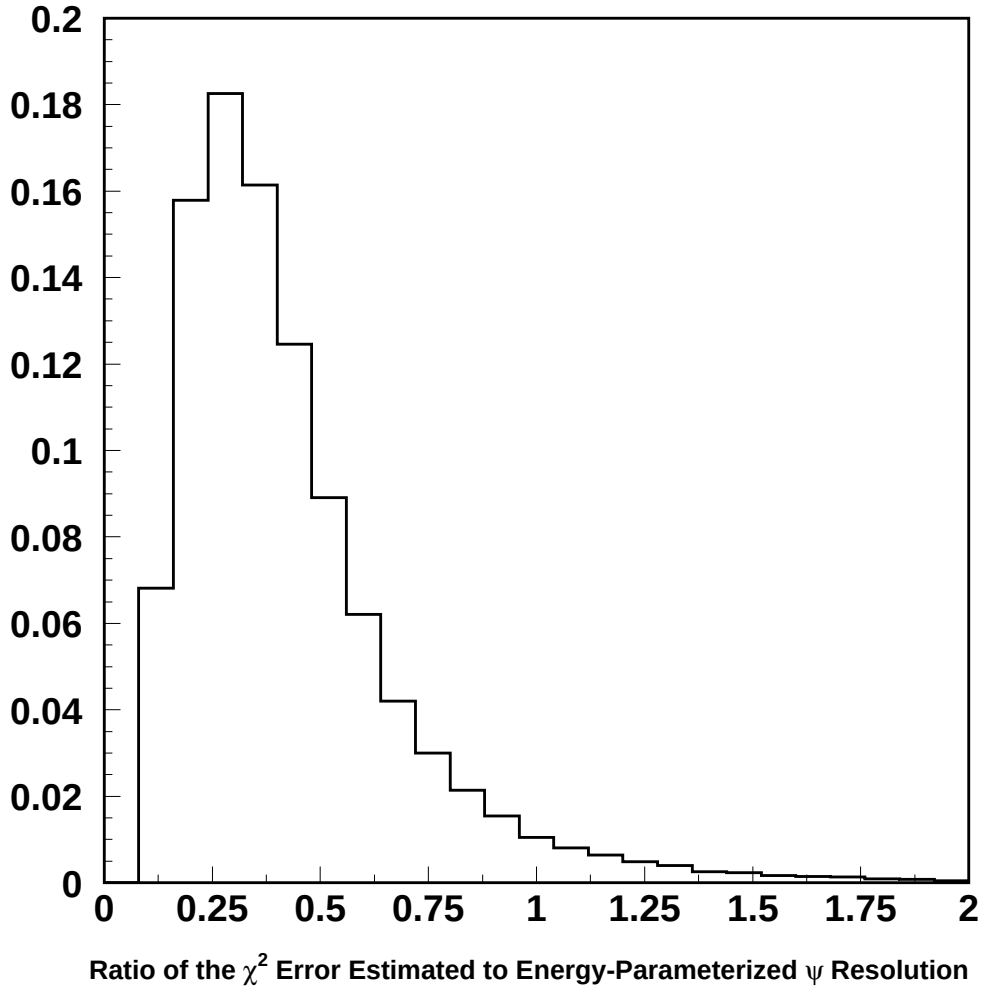


Figure 6.11. Distribution of $\frac{\Delta\psi_{\chi^2}}{\Delta\psi_{\mathrm{EP}}}$ where $\Delta\psi_{\chi^2}$ is the ψ resolution predicted by the χ^2-fit and $\Delta\psi_{\mathrm{EP}}$ is the ψ resolution predicted by energy parameterization method for the simulated data ($\mu = 0.4334$ and $RMS = 0.5184$).

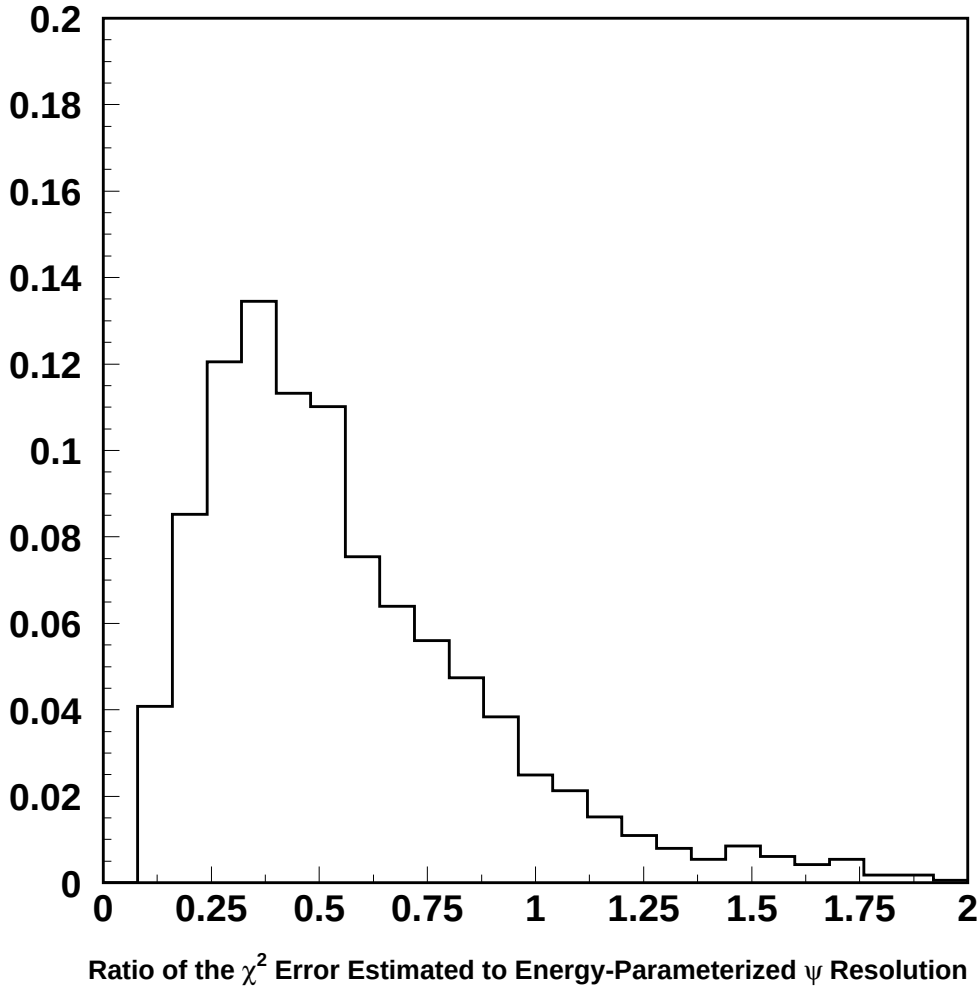


Figure 6.12. Distribution of $\frac{\Delta\psi_{\chi^2}}{\Delta\psi_{\mathrm{EP}}}$ where $\Delta\psi_{\chi^2}$ is the ψ resolution predicted by the χ^2-fit and $\Delta\psi_{\mathrm{EP}}$ is the ψ resolution predicted by energy parameterization method for the real data ($\mu = 0.5760$ and $RMS = 0.6849$).

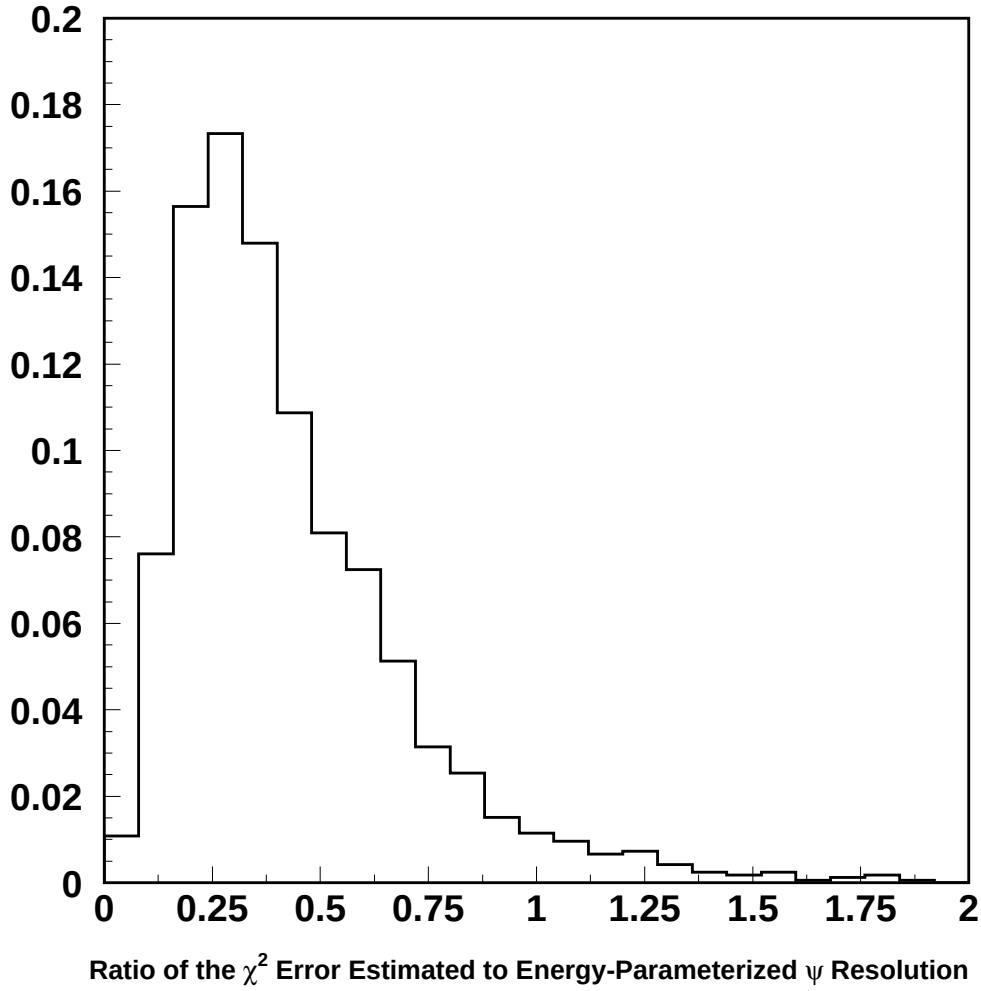


Figure 6.13. Distribution of $\frac{\Delta\psi_{\chi^2}}{\Delta\psi_{\rm EP}}$ where $\Delta\psi_{\chi^2}$ is the ψ resolution predicted by the χ^2-fit and $\Delta\psi_{\rm EP}$ is the ψ resolution predicted by energy parameterization method for real data with $\Delta\psi = s_\circ\Delta\psi_\circ$ ($\mu = 0.4339$ and $RMS = 0.5156$).

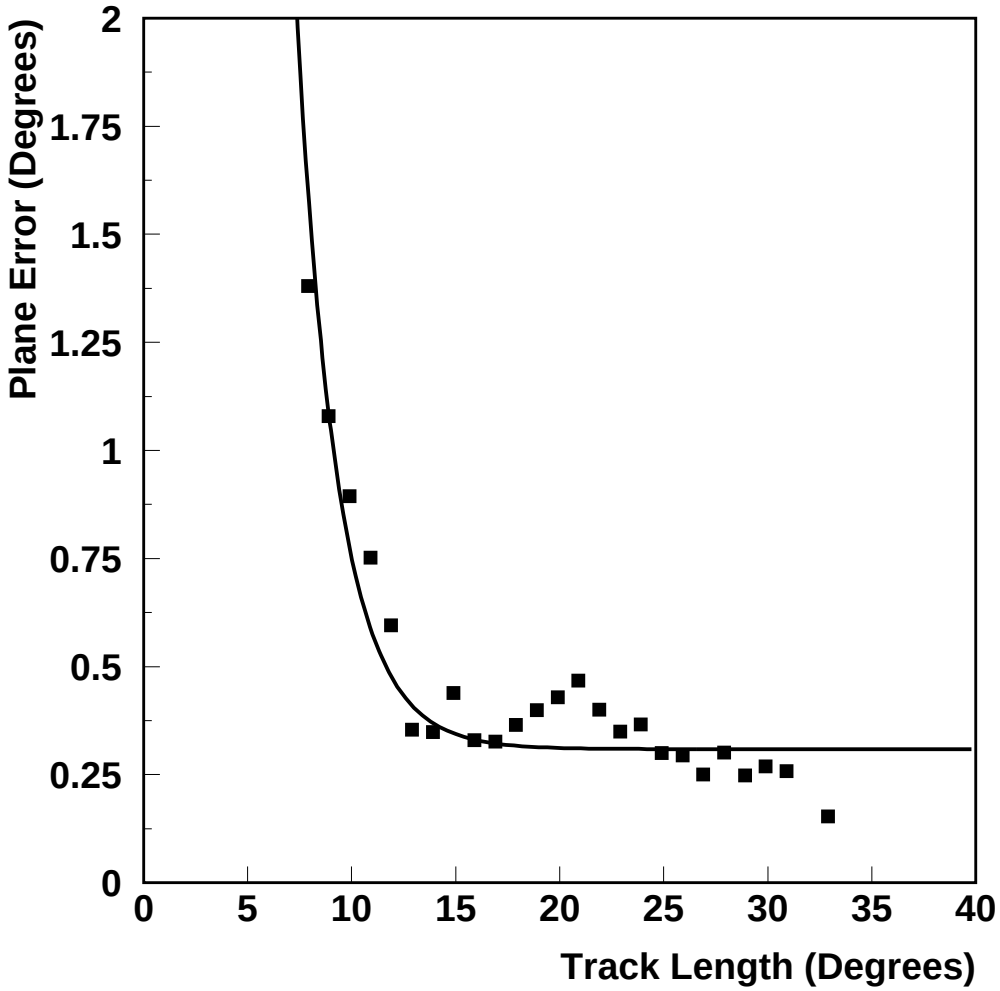


Figure 6.14. The the resolution of the plane of reconstruction determined from comparisons between Monte Carlo inputs and profile constraint fit results.

are accepted. This suggests that calibration uncertainties can have a substantial impact on the determination of the plane of reconstruction. Hence we will apply the smearing constant, $s_\circ$ to equation (6.7), which yields:

$$\Delta n = s_\circ \Delta n_\circ = 88.2^\circ e^{-\frac{\Delta\chi}{1.9595}} + 0.37^\circ \tag{6.8}$$

6.4 Using Stereo Data to Further Refine the Angular Resolution and Establish the Angular Resolution Systematic Uncertainty

One way to check the efficacy of the angular resolution parameterization is by considering events that were successfully reconstructed by both the HiRes-1 monocular profile constraint routine and by the stereo reconstruction routine that considers both HiRes-1 and HiRes-2 data. In stereo mode, the shower detector planes of the two detectors are intersected, thus the geometry is much more precisely known and the total angular resolution is of order 0.6°, a number that is largely correlated to σ_{plane} and thus is negligible when added in quadrature to the larger term, σ_ψ. This allows us to perform a comparison of the angular resolution estimated through simulations to the observed angular resolution values of actual data.

We analyzed $\sim$ 200 actual events from the HiRes-I monocular data set which were also seen in stereo whose energies were estimated to be above $10^{18.5}$ eV by the profile constraint reconstruction routine and were successfully reconstructed by both mono stereo reconstruction routines. For each event, the opening angle between the nominal arrival directions estimated by mono and stereo reconstructions were histogrammed. We then considered the full library of simulated events. For each simulated event, we sampled the error space five times, both with and without the smearing constant, $s_\circ$, and then histogrammed the resulting opening angles to the nominal arrival direction estimated by the profile constraint reconstruction routine. The results shown in Figure 6.15 seem to validate the angular resolution parameterization and the application of the smearing coefficient to the simulated events. In Figure 6.15c 6.15d, we show the distribution of angular errors for real

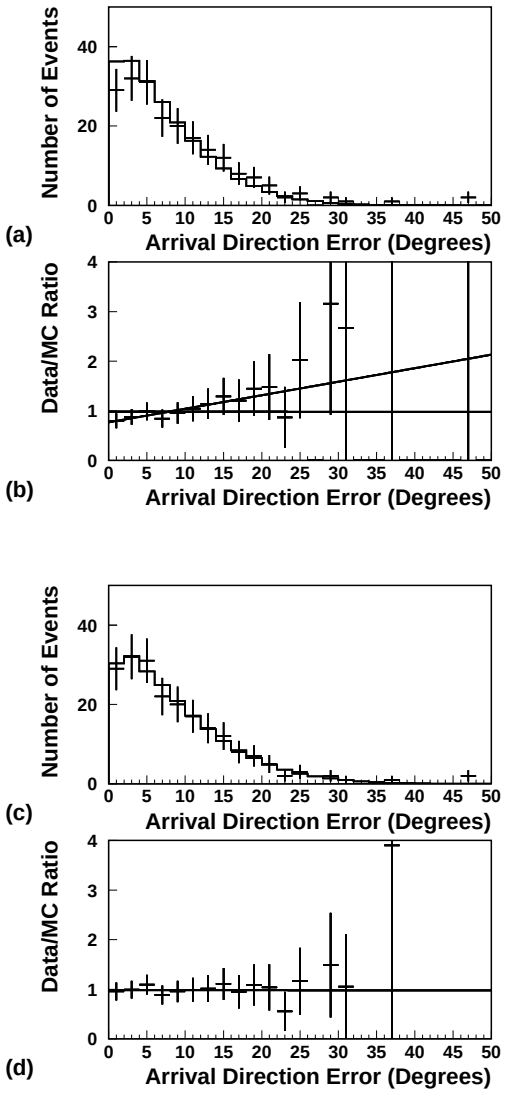


Figure 6.15. Arrival direction error comparison between real data (mono vs. stereo) and simulated data for events with estimated energies above $10^{18.5}$ eV. The solid line histogram corresponds to the arrival direction error distribution of the monocular reconstructed Monte Carlo simulated data. The crosses correspond to the arrival directions error distribution observed for actual data by comparing the arrival directions estimated by the monocular and stereo reconstructions. —(a and b): without the application of the smearing constant, $s_\circ$; (c and d): with the application of the smearing constant, $s_\circ$. In each case, the solid line histogram corresponds to the estimated angular resolution distribution of the entire library of simulated events. In the case where $s_\circ$ has been applied, the solid line in the ratio component corresponds to the fit $y = ax + b$ where $a = 0.000 \pm 0.011$ and $b = 0.98 \pm 0.11$.

and simulated data. The uncertainty in the slope of the ratio (Figure 6.15d) leads to an 7.5% uncertainty in the angular resolution.

CHAPTER 7

MODELING THE HIRES EXPOSURE

In order to calculate the significance or to set meaningful limits on anisotropies seen within the HiRes-I monocular data set, it is important to simulate data sets that possess the same exposure characteristics as the actual data set. This presents a unique challenge. First of all, the aperture of the HiRes detector is dependent upon both energy and arrival direction. Secondly, the HiRes-1 exposure is asymmetric in seasonal coverage, night to night weather conditions and individual mirror ontimes.

For the purpose of creating simulated data sets, we generated a library of simulated events which were subsequently reconstructed using the profile-constraint reconstruction routine. This library of events possesses the spectrum and composition that were obtained by the Fly's Eye Stereo experiment [15, 18]. A total of $\sim 1.3 \times 10^5$ simulated events were successfully reconstructed with energies above $10^{18.5}\ eV$.

Once a library of simulated events were created, we then turned to the task of creating simulated data sets of equal exposure to the HiRes-1 monocular data set. In general, the apertures of air-fluorescence detectors are complicated: We need to assign times to individual events that accurately reflect that distribution of observation times seen in the actual data. The main considerations include: detector ontimes, mirror ontimes, and weather conditions that can lead to asymmetric sky coverage.

The detector ontimes are calculated by a program in the hires_soft library called "HSUM." HSUM is run on the HiRes-1 data each month after the global weather cuts are applied to exclude any data where the weather is so uniformly bad that it would preclude any viable events observations. For each data part, HSUM notes

(among other things) the start and end times of the HiRes-1 detector (there are generally several data segments for each night of observation). We then randomly assign a time from the recorded observation periods to each prospective individual simulated event for a simulated event set.

It is also necessary to apply a mirror-by-mirror correction. A detector start record does not ensure that every individual mirror was alive or that its view was unobstructed by doors, curtains, clouds, or excessive sky noise. On a minute to minute basis, each mirror provides a status packet that provides critical information on that mirror's operational status. Two pieces of information are of particular interest to us: the event rate and the mean threshold of the tubes in the individual mirror. We will only consider a particular minute of data for an individual mirror to be valid for the sake of aperture calculation if a specific set of criteria are met:

1. The particular minute of data in question must have at least two event triggers. This excludes minutes where the view of a mirror could potentially be blocked by a door or curtain.

2. The minute of data cannot have more 1000 event triggers. This excludes data where the aperture of a mirror is being severely compromised by excessive sky noise or problems with the cluster electronics.

3. The mean threshold, Θ_{i} for the mirror for a given minute is constrained by the following condition:

$$\Theta_i \geq < \Theta > -2\sqrt{< \Theta^2 > - < \Theta >^2} \tag{7.1}$$

   where $< \Theta >$ and $< \Theta^2 >$ are taken for a given mirror over two year periods for *all* data before the manual weather cut is applied.

The last cut allows us to exclude data for which the sky noise is abnormally low due to the presence of clouds in the field of view. The data are considered in two year increments in order to account for the two competing factors. First, one has annual climatological cycles that can vary quite a bit from year to year. One also

has to consider the rapid expansion of the Wasatch population corridor 80 km away, which is leading to ever-higher amounts of ambient sky noise. The distributions of minute-to-minute mean thresholds are shown for two different mirrors for a two year period in Figures 7.1 and 7.2. While this cut excludes $\sim 15\%$ of *all* data, it only excludes $< 5\%$ of the data that remains *after* the weather cut is applied.

Next, we apply the correction to our simulated data set by only accepting events when the mirrors in the Monte Carlo event and the randomly selected event time coincide during a minute where the mirrors are retained by the above cuts. In Figures 7.3 and 7.4 we see the results of this mirror-by-mirror correction by comparing the sidereal time distributions of real and simulated data sets before and after the correction is applied. A definite improvement is seen in Figure 7.4 in the correlation between the sidereal time distributions of the real and simulated data, especially at the sidereal times where the event count is at its maximum and minimum. As an additional check, we also compare the distribution of azimuth and zenith angle distributions for real and simulated data sets after the above corrections were applied. The results are shown in Figures 7.5 and 7.6. In both figures, we find excellent agreement between data and simulation. Note however, that these correspond to the shower directions in terrestrial coordinates and do not introduce an *a priori* bias in the equatorial coordinates in the subsequent analysis.

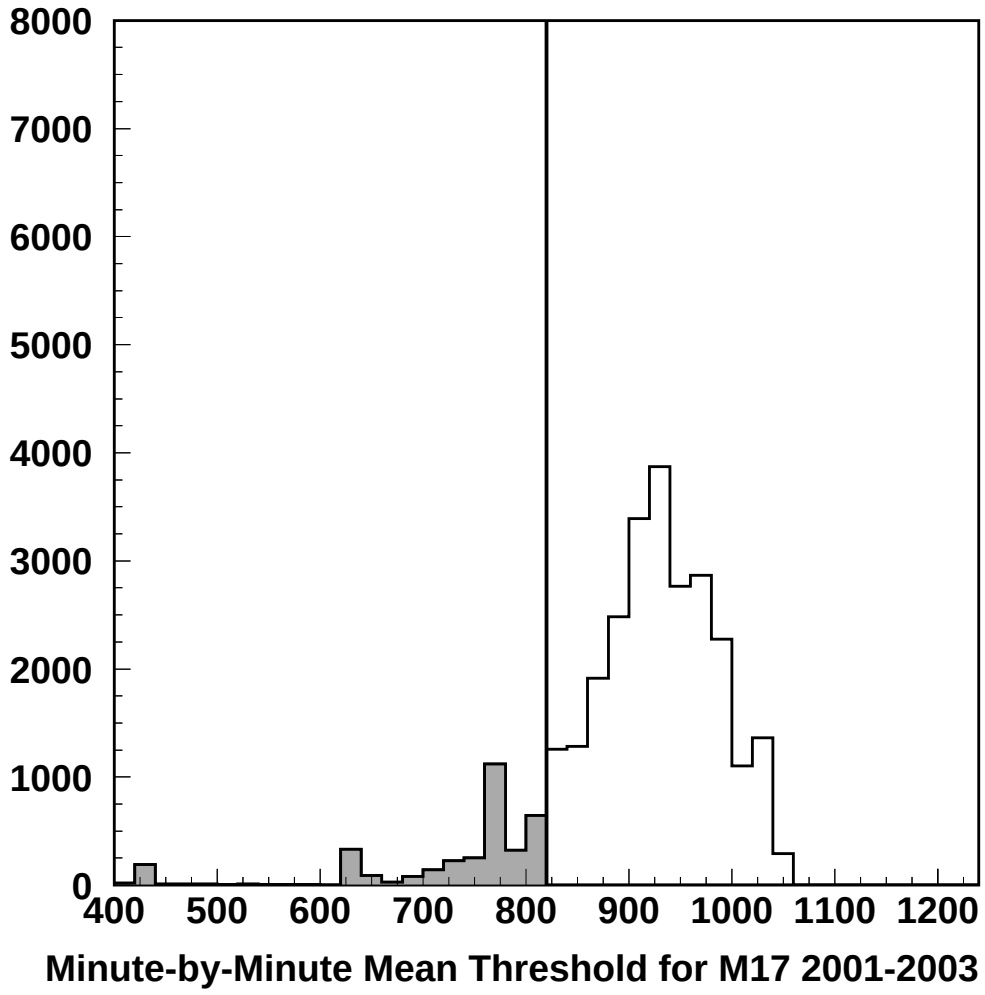


Figure 7.1. Distributions of the minute-to-minute mean threshold for Mirror 17 in 2001-2003. The shaded area shows the cut that was applied in order to determine whether a given minute would would be considered in HiRes-1 aperture estimation when the mirror-by-mirror correction was applied.

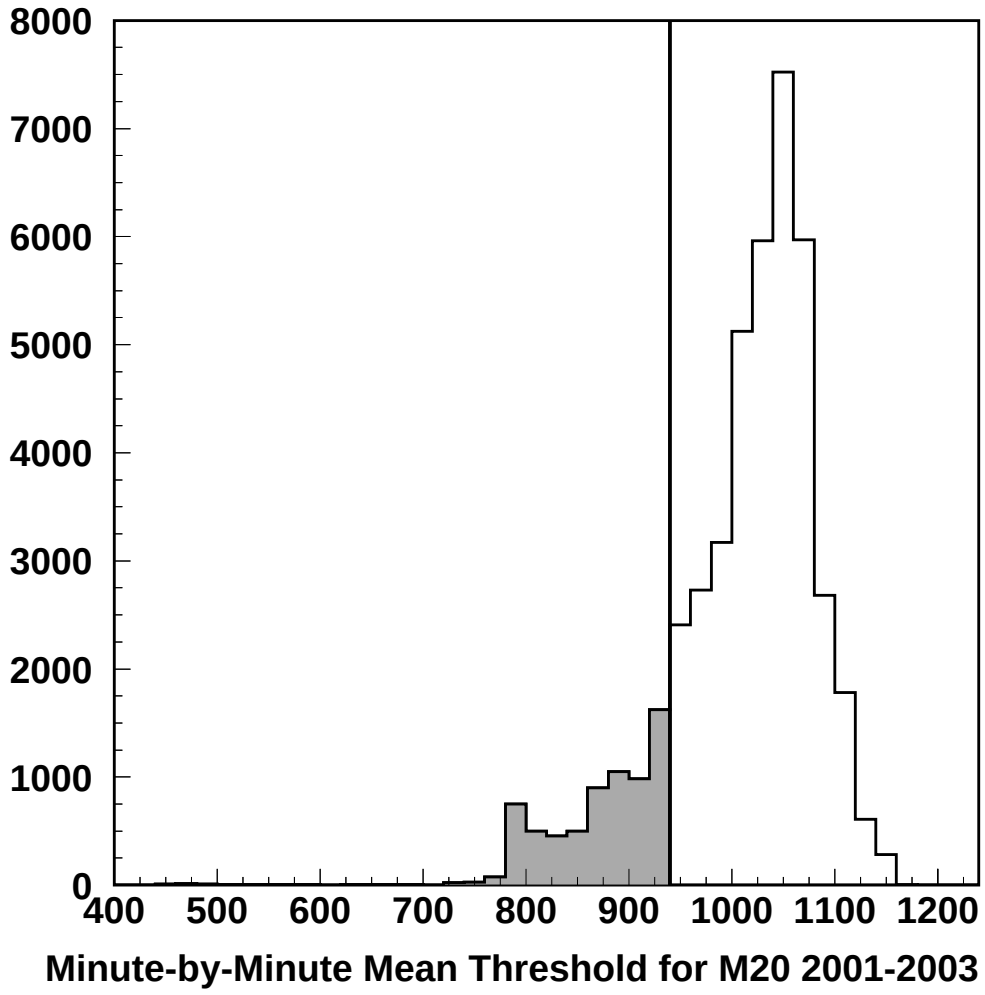


Figure 7.2. Distributions of the minute-to-minute mean threshold for Mirror 20 in 2001-2003. The shaded area shows the cut that was applied in order to determine whether a given minute would would be considered in HiRes-1 aperture estimation when the mirror-by-mirror correction was applied.

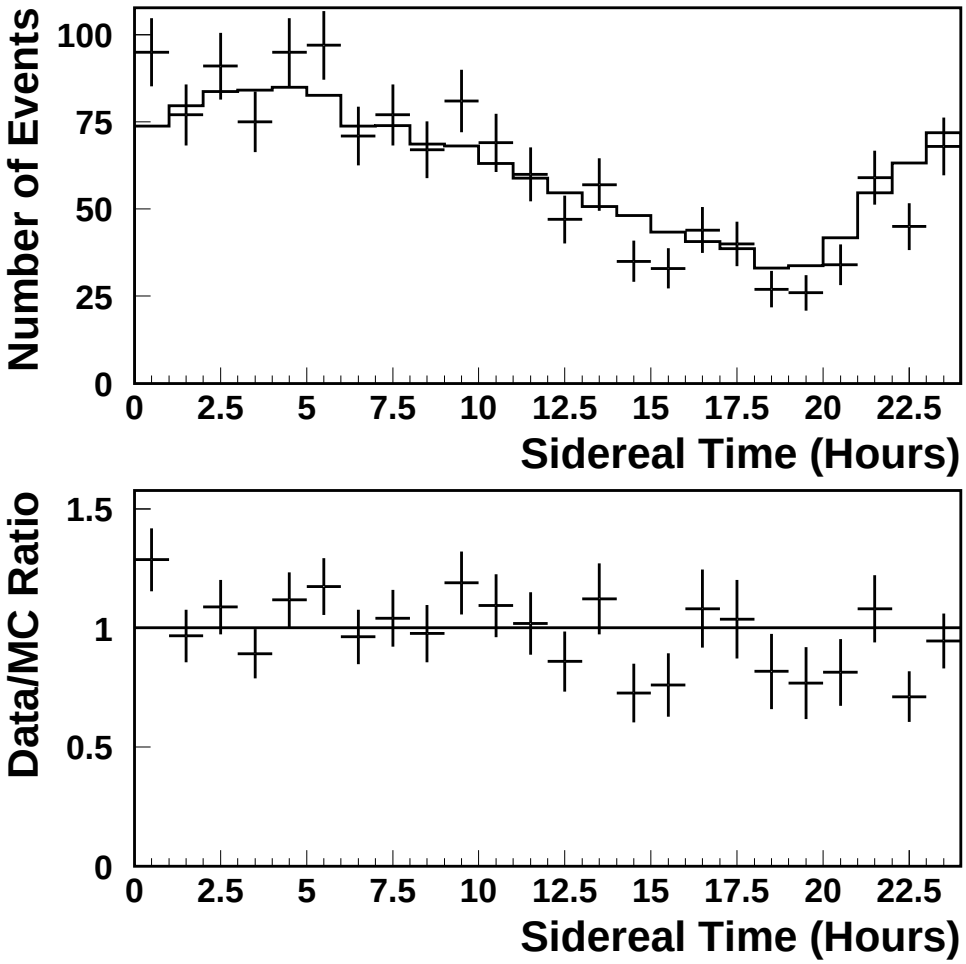


Figure 7.3. Sidereal time distribution comparisons between the ~ 1500 real data events and 10^6 simulated event/time pairings *before* the mirror-by-mirror correction was applied ($\chi^2/d.f. = 1.5$). The solid line histogram corresponds to the sidereal time distribution of simulated event/time pairings. The crosses correspond to the sidereal time distribution of the real data with Gaussian uncertainties assumed for each bin.

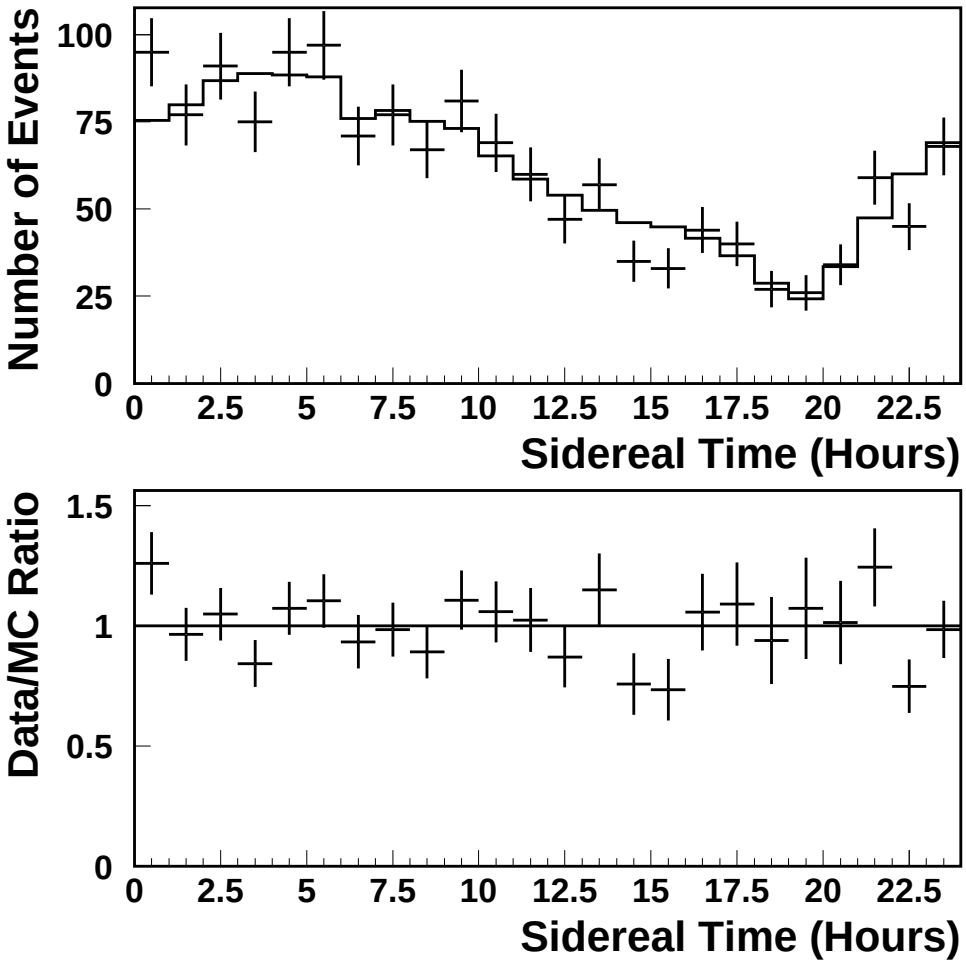


Figure 7.4. Sidereal time distribution comparisons between the ~ 1500 real data events and 10^6 simulated event/time pairings *after* the mirror-by-mirror correction was applied ($\chi^2/d.f. = 1.2$). The solid line histogram corresponds to the sidereal time distribution of simulated event/time pairings. The crosses correspond to the sidereal time distribution of the real data with Gaussian uncertainties assumed for each bin.

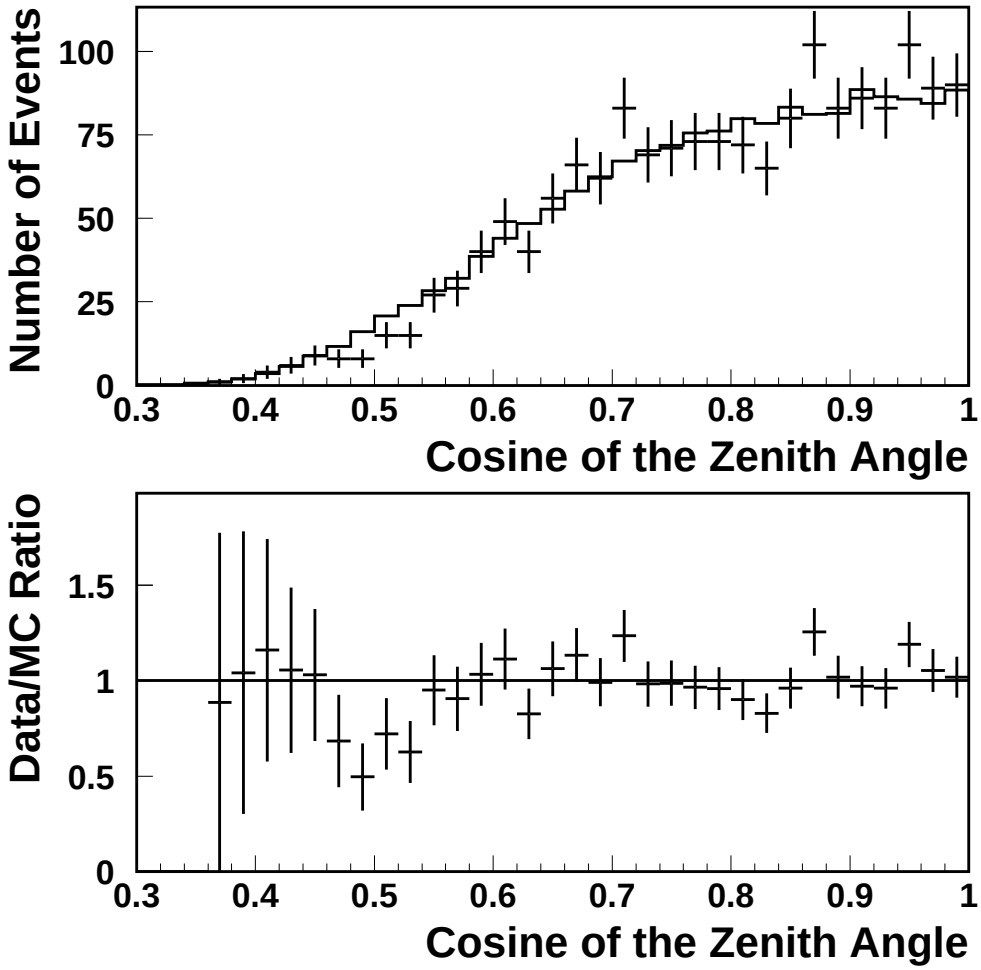


Figure 7.5. Distribution of the cosines of the zenith angles ($\chi^2/d.f. = 0.9$). The solid line histogram corresponds to the specified angle distribution of 10^6 simulated event/time pairings. The crosses correspond to the specified angle distribution of the real data with Gaussian uncertainties assumed for each bin.

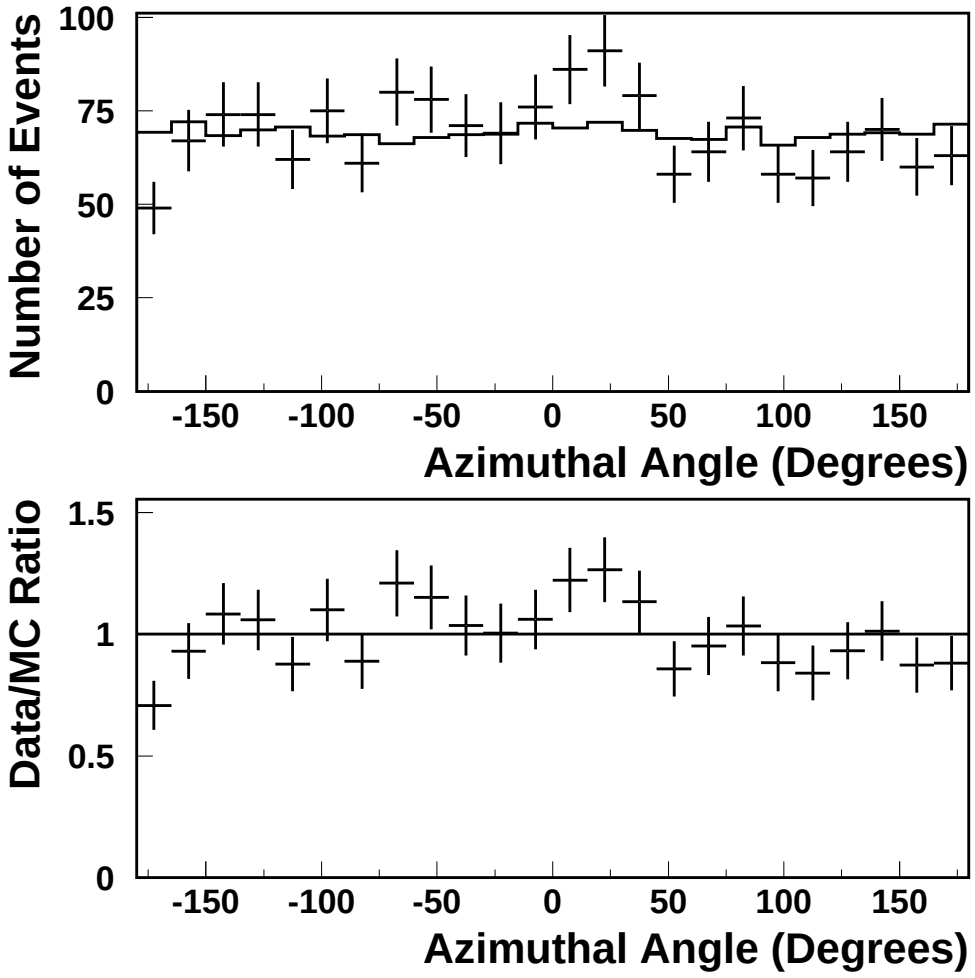


Figure 7.6. Distribution of the azimuth angles ($\chi^2/d.f. = 1.31$). The solid line histogram corresponds to the specified angle distribution of 10^6 simulated event/time pairings. The crosses correspond to the specified angle distribution of the real data with Gaussian uncertainties assumed for each bin.

CHAPTER 8

SEARCH FOR GLOBAL DIPOLE ENHANCEMENTS

In the past five years, theoretical models have been suggested that would potentially produce dipole distributions oriented towards M87 [39] or Centaurus A [40, 41]. In addition, the Akeno Giant Air Shower Array (AGASA) has reported findings suggesting a 4% dipole-like enhancement oriented towards the Galactic Center present in its events with energies around 10^{18} eV [42]. This result seemed to be corroborated by findings published by the Fly's Eye experiment in 1999 that suggested the possibility of an enhancement in the galactic plane also at energies around 10^{18} eV [43], and in addition by a re-analysis of data from the SUGAR array that was published in 2001 [44] that showed an enhancement in the general vicinity of the Galactic Center.

However, both AGASA and Fly's Eye are subject to a limiting factor; they are both located too far north in latitude to directly observe the Galactic Center itself. The re-analysis of SUGAR data actually demonstrated an excess that was offset from the Galactic Center by 7.5° and was more consistent with a point source than a global dipole effect [44]. While the current High Resolution Fly's Eye (HiRes) experiment is subject to a similar limitation in sky coverage as the AGASA and Fly's Eye experiments, we will show that, by properly estimating the HiRes aperture and angular resolution, we can effectively exclude these dipole source models to a certain degree of sensitivity. However, we are not able to completely exclude the findings of AGASA or the theoretical predictions mentioned above.

Our methods for detecting the presence of a dipole source model will be based upon comparisons between the real data and a large quantity of events generated

by our Monte Carlo simulation program. The simulated data possess the same aperture and exposure as the actual HiRes-I monocular data set as was discussed in the previous chapter. In order to measure the presence of a dipole effect in our event sample, we use first a conventional binning technique that considers the event counts for the full range of opening angles from the center of each proposed dipole distribution. We then show how the asymmetric angular resolution of a monocular air fluorescence detector can be accommodated in this method. We ascertain the 90% confidence interval for a dipole source model for each of the three dipoles considered by comparing our real data with large numbers of similar-sized simulated data sets. We then consider the effects of systematic uncertainties on our measurements.

8.1 The Dipole Function

A dipole source model can be described, as first proposed by Farrar and Piran [40]:

$$n = \frac{1}{2} + \frac{\alpha}{2}\cos\theta, \tag{8.1}$$

where n is the relative density of cosmic rays in a given direction, θ is the opening angle between that direction and the global maximum of the distribution, and α is the customary anisotropy amplitude [45]:

$$\alpha = \frac{n_{\max} - n_{\min}}{n_{\max} + n_{\min}}. \tag{8.2}$$

The cases of $\alpha = 1$ and $\alpha = -1$ correspond to 100% dipole distributions in the direction of the center and anti-center of the dipole source model, respectively. The case of $\alpha = 0$ corresponds to an isotropic source model.

A simple scheme for measuring α consists of constructing a *dipole function* in the following manner:

1. The opening angle is measured between the arrival direction of an event and the center of the proposed dipole source model.

2. The cosine of the opening angle is then histogrammed.

3. The preceding steps are repeated until all of the events are considered.

4. The resulting curve produced by the histogram is the dipole function.

The dipole function has two variable parameters: the bin width, $\Delta(\cos\theta)$, and the total number of counts in all of the bins. At first glance, it would seem that the total bin count is fixed upon the total number of events, but we will show that this is not necessarily the case when we consider how to accommodate angular resolution.

In the simplest case of a sample that contains a very large number of events with a constant exposure and aperture over the entire sky, the dipole function will be proportional to equation 8.1. We propose two simple ways that one can quantify the dipole function for this sample; the most obvious way is to consider its slope. We can see by referring to equation 8.1 that this is equal to $\frac{\alpha}{2}$. A second way of quantifying α is to consider the mean cosine value, $<\cos\theta>$ for the dipole function:

$$<\cos\theta>= \frac{1}{2}\int_{-1}^{1} \cos\theta(1+\alpha\cos\theta)\, d(\cos\theta) = \frac{1}{3}\alpha. \tag{8.3}$$

Both methods of quantification produce values that are dependent upon α. While the dependence of $<\cos\theta>$ is linear in α for the case of homogeneous full-sky coverage, we will find that this is not necessarily the case when considering the cumulative exposure of a ground-based air fluorescence detector.

8.2 Calculating the Dipole Function for the HiRes-I Monocular Data

As a first order measurement, we construct the dipole function for a source model with a maximum value at the Galactic Center. For now, we only consider the nominal arrival directions of the events in our data sample. For this demonstration, we set the bin width of the dipole function to $\Delta(\cos\theta) = 0.04$. This provides us with a mean bin count, $\langle n\rangle = 30.52$. Figure 8.1a shows the resulting dipole function. However, in order to estimate the value of α, we first normalize our dipole function with respect to aperture and exposure. This is done by considering 10^7 pairs of simulated events and event times that correspond to the actual HiRes-I observation periods. By constructing a dipole function for this simulated set, we

then estimate the normalization factor for each $\cos\theta$ bin in the dipole function. The result is shown in Figure 8.1b. The dipole function is then normalized and a χ^2-fit performed to determine its slope, m, and y-intercept, b. The normalized dipole function is pictured in Figure 8.1c with the best linear fit applied. The scaling constant, α, is then estimated by the quotient, $m/\langle n \rangle$. The result for the galactic dipole source model is then: $\alpha = -0.010 \pm 0.055$.

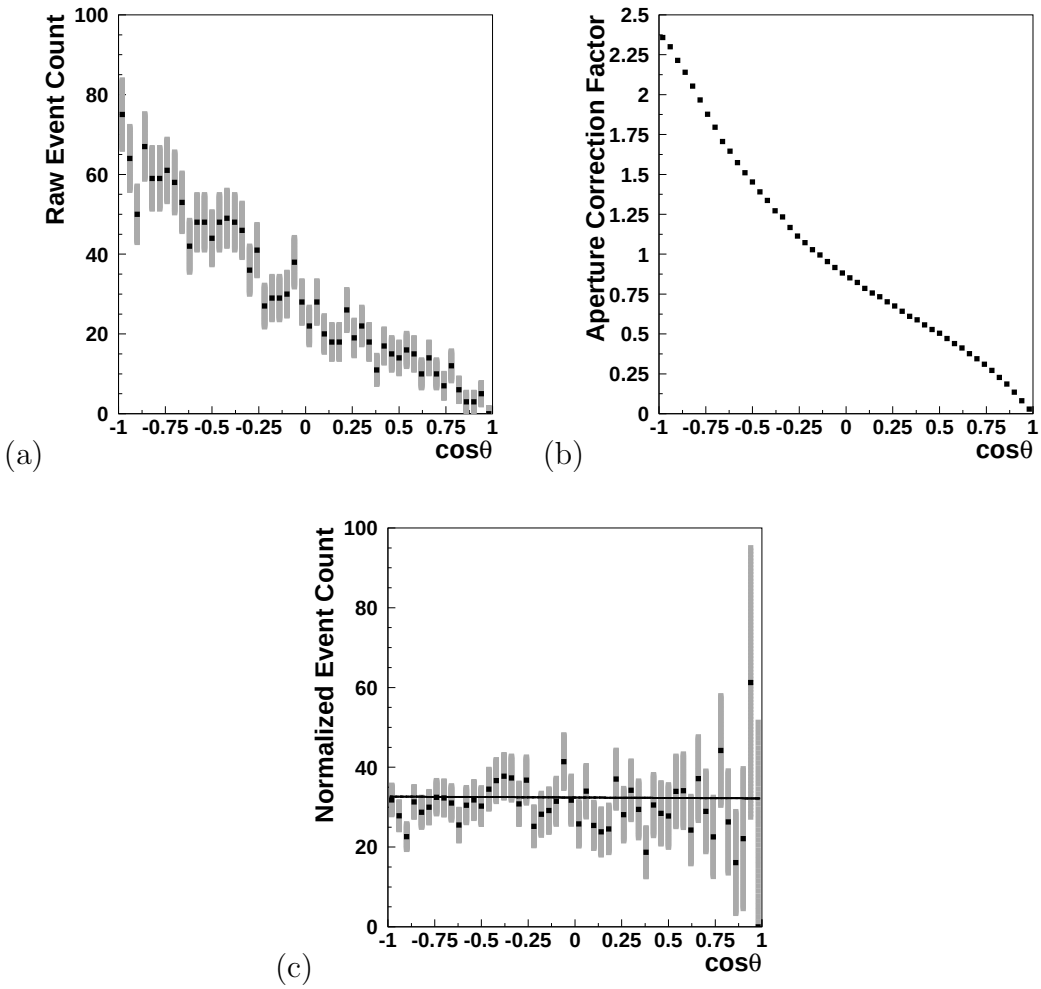


Figure 8.1. The dipole function for the nominal arrival directions of the HiRes-I data set—(a) the number of counts in each $\cos\theta$ bin; (b) the aperture/exposure normalization factor for each bin; (c) the normalized bin count with the χ^2-fit to a line (all uncertainties are Poisson).

The same method was employed to calculate α in the cases of Centaurus A and M87. For Centaurus A, we obtained a result of: $\alpha = -0.035 \pm 0.060$. For M87, we found $\alpha = -0.005 \pm 0.045$.

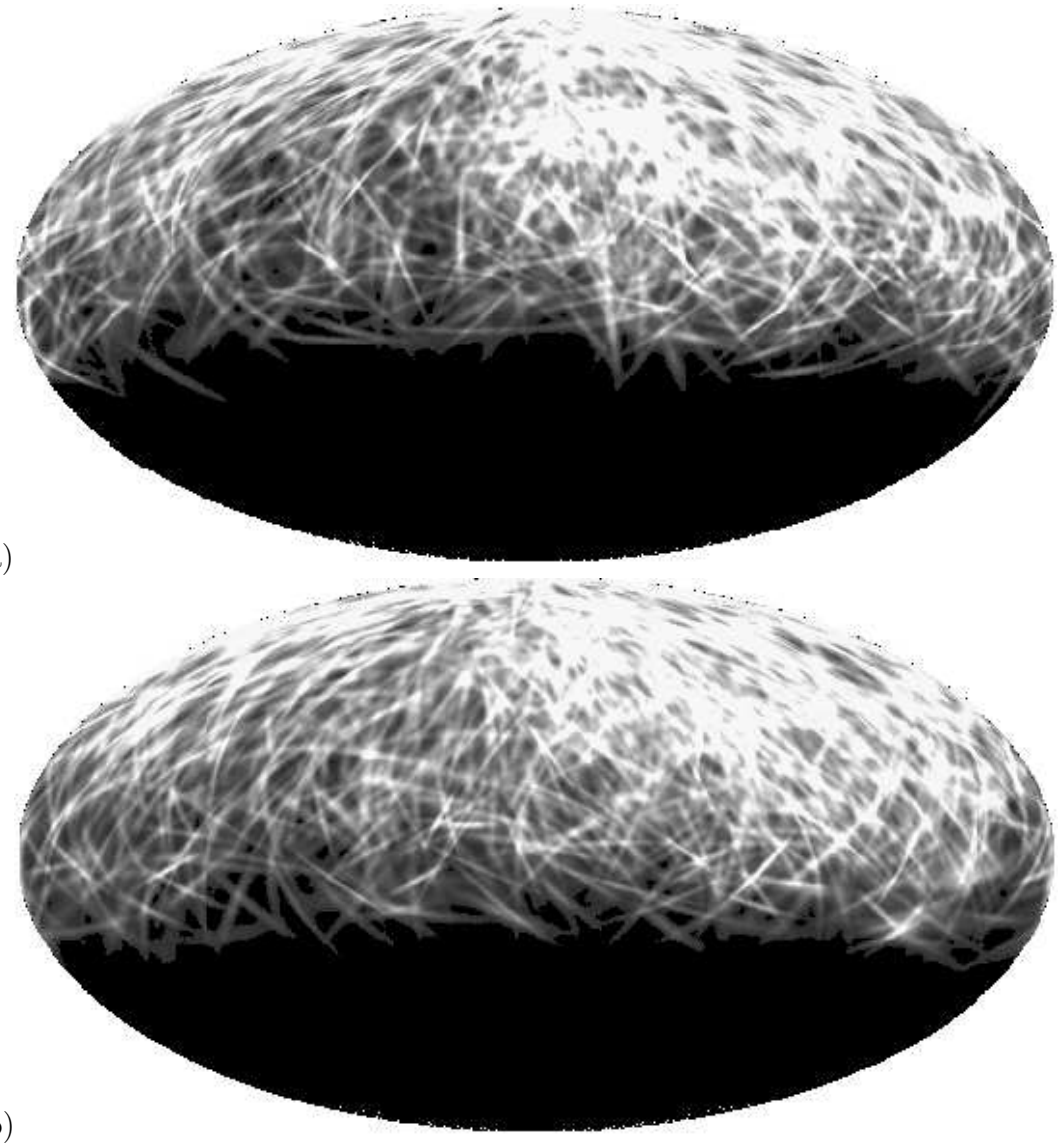
8.3 Incorporating Angular Resolution into the Dipole Function

The analysis described in the previous section does not take into account the experimental angular resolution. Accommodating the angular resolution is important to the overall integrity of this analysis because the HiRes-I monocular data contains very asymmetric errors in arrival direction determination.

In order to accommodate the HiRes-I monocular angular resolution, it is necessary to revise the method we use to construct the dipole function. Instead of considering each event as a single arrival direction, we will consider each event to be an elliptical, two-dimensional Gaussian distribution of N points with the two Gaussian parameters, σ_1 and σ_2, being defined by the parameters that describe the angular resolution, corresponding to equations 6.6 and 6.8. Figures 8.2 and 8.3 show how entire sets of events with these error parameters appear when projected on a density plot using a Hammer-Aitoff projection and equatorial coordinates.

In order to account for angular resolution in the construction of the dipole function, we add an additional step. Instead of simply calculating the opening angle between the arrival direction of the event and the center of the dipole for the preferred arrival direction, we do so separately for each of the N points in the Gaussian distribution that describes each event's arrival direction. By choosing a sufficiently large value for N and a sufficiently small bin width, $\Delta(\cos\theta)$, we can then construct the dipole function as a smooth curve. Examples of the dipole function are shown in Figures 8.4 and 8.5 for each of the event sets in Figures 8.2 and 8.3.

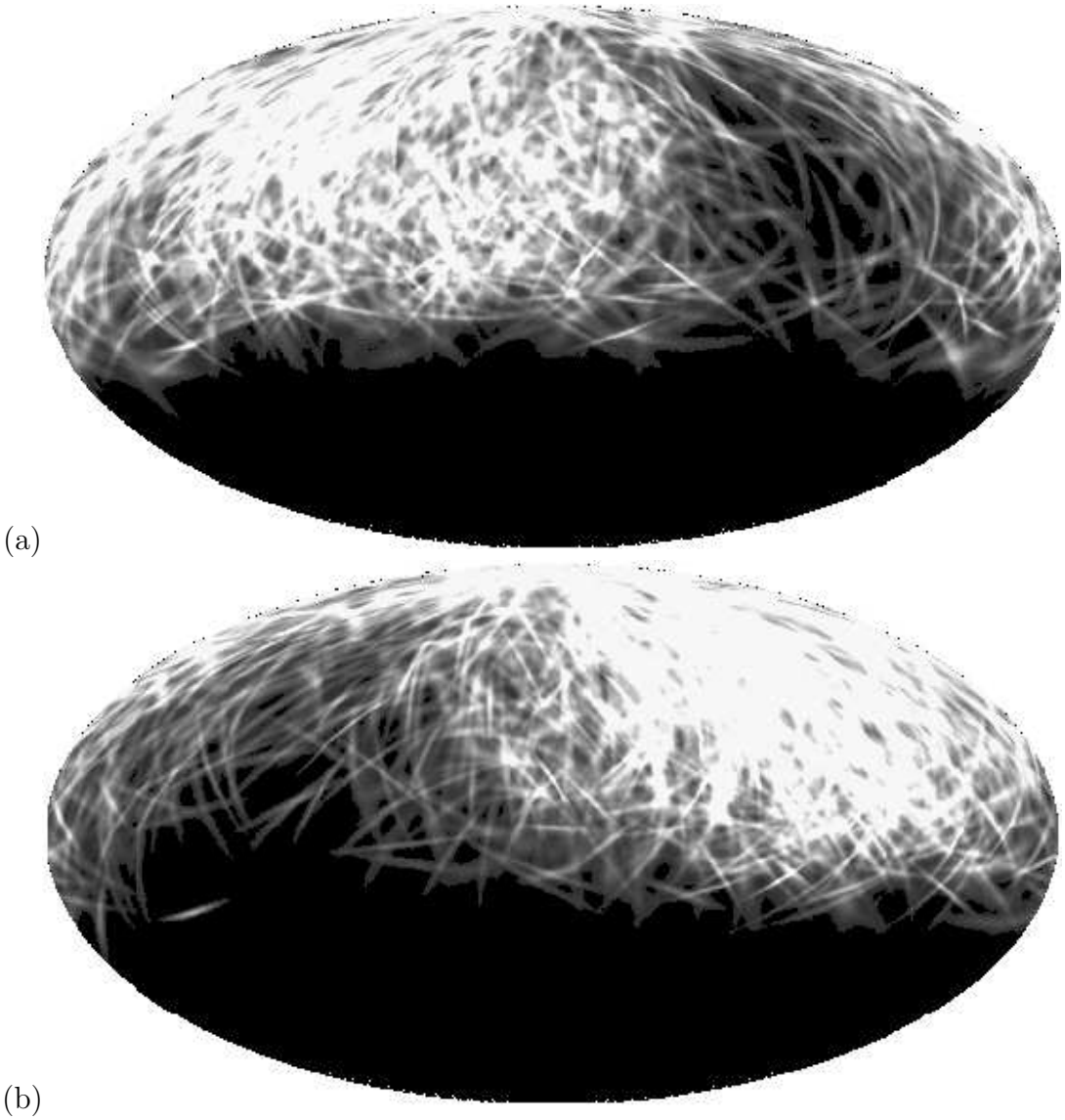
The next logical step would be to attempt to normalize the dipole function of the real data with respect to aperture and exposure and then to calculate the slope, m, and the y-intercept, b. However, this program would run into a major complication. Because the Gaussian distributions that are used to approximate



(a)

(b)

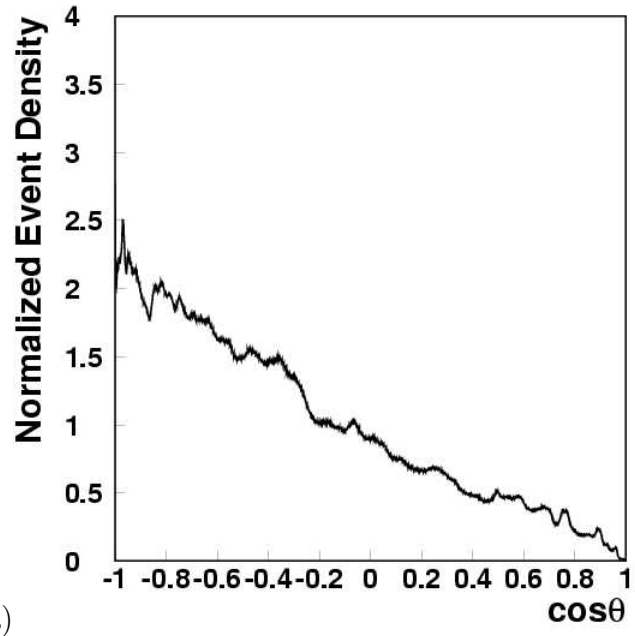
Figure 8.2. Density plots of event arrival directions with the angular resolution parameters of the Hires-1 monocular data on a Hammer-Aitoff projection with equatorial coordinates (right ascension right to left)— (a) HiRes-I monocular data set; (b) simulated data set with an isotropic source model; In each case, the lighter regions correspond to a higher density of event arrival directions.



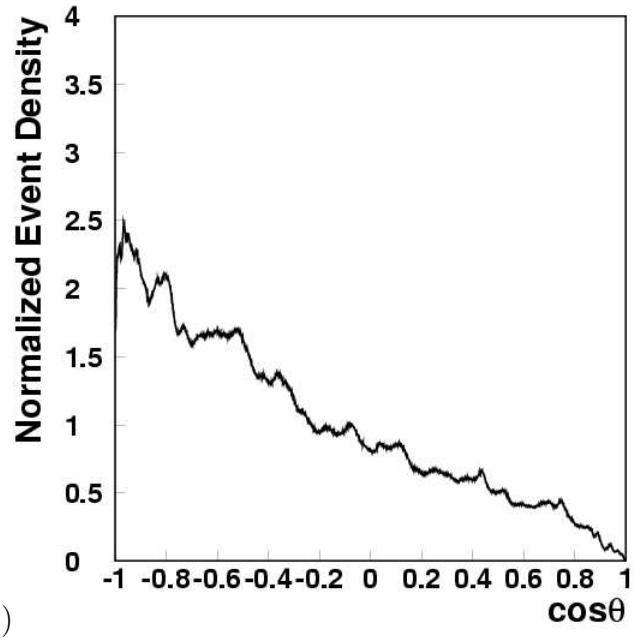
(a)

(b)

Figure 8.3. Density plots of event arrival directions with the angular resolution parameters of the Hires-1 monocular data on a Hammer-Aitoff projection with equatorial coordinates (right ascension right to left)— (a) simulated data set with a galactic dipole source model ($\alpha = 1$); (b) simulated data set with a galactic dipole source model ($\alpha = -1$). In each case, the lighter regions correspond to a higher density of event arrival directions.

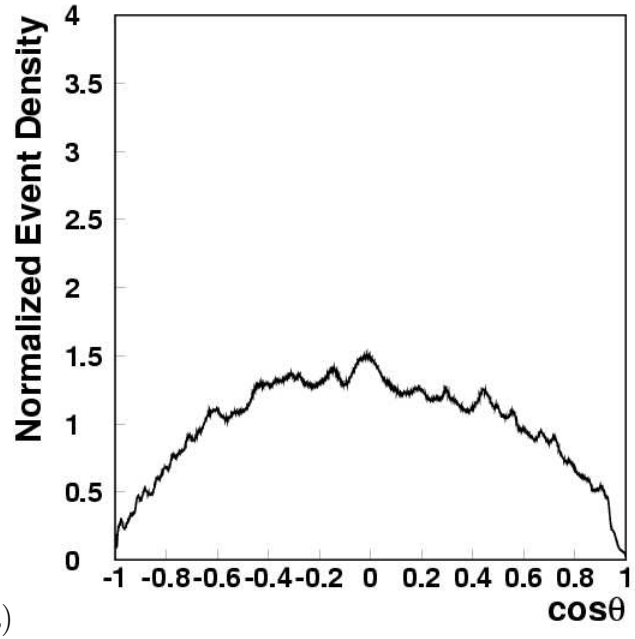


(a)

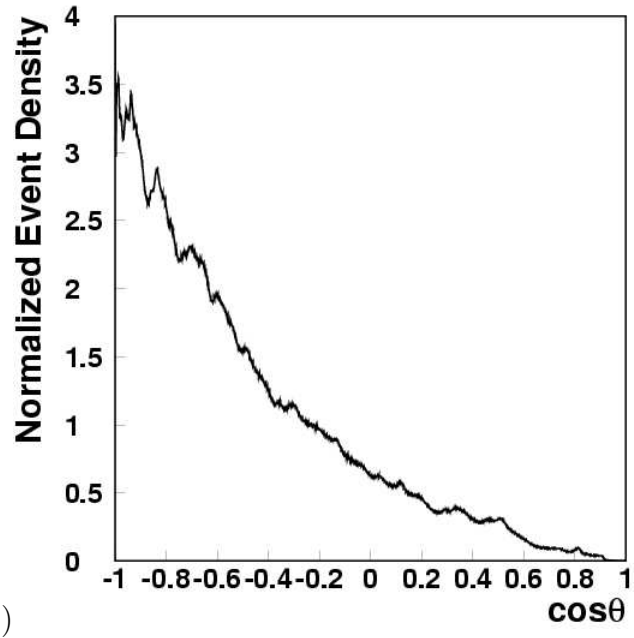


(b)

Figure 8.4. The dipole function, with angular resolution included, for a galactic dipole model for different event sets without correction for aperture and exposure—(a) HiRes-I monocular data set; (b) simulated data set with an isotropic source model.



(a)



(b)

Figure 8.5. The dipole function, with angular resolution included, for a galactic dipole model for different event sets without correction for aperture and exposure—(a) simulated data set with a galactic dipole source model ($\alpha = 1$); (b) simulated data set with a galactic dipole source model ($\alpha = -1$).

the individual event arrival directions can overlap into a large number of bins, the individual data points in the dipole function are highly correlated. This makes it impossible to apply either the χ^2-fit or a bootstrap method to estimate the error in the values of m and b—and thus the error in α—for the normalized dipole function. Another approach needs to be developed.

The method that we used is to compare the value of $<\cos\theta>$ for the dipole function of the real data sample with that of a large number of similar-sized simulated data samples with a discrete spectrum of α-values. We can then show how $<\cos\theta>$ varies with respect to α for different dipole source models.

For each of the three dipole source models considered we used the following procedure to measure the α parameter:

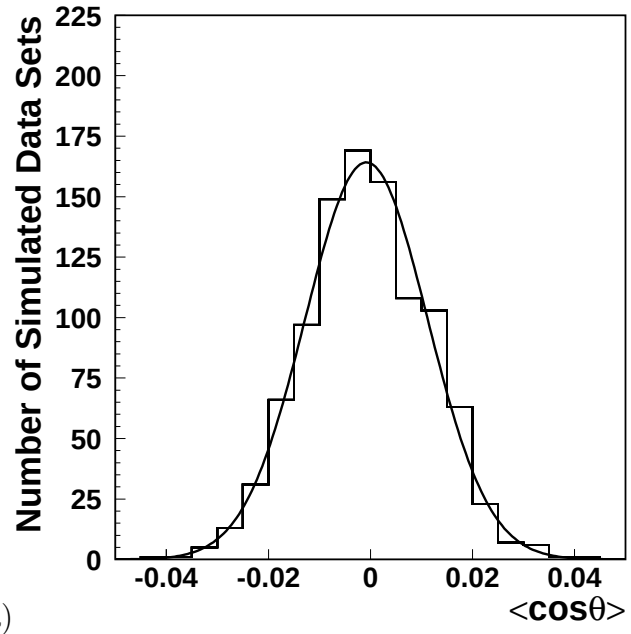
1. We calculated the value of $<\cos\theta>$ for the dipole function of the real data sample.

2. We created a total of 20,000 simulated data samples, 1000 each for 0.1 increments of α from -1.0 to 1.0, each with the same number of events as the actual data. In Figure 8.6 we can see that the distribution of $<\cos\theta>$ values for each α-value generated conforms well to a Gaussian distribution.

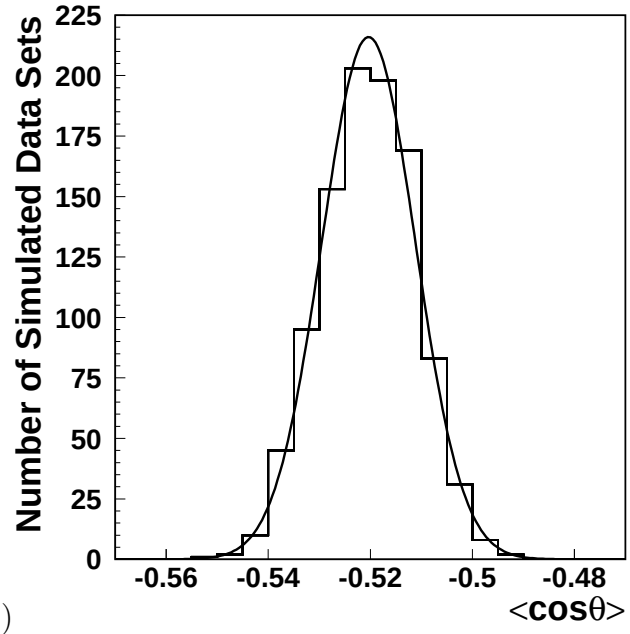
3. We constructed curves corresponding to the mean and standard deviation of $<\cos\theta>$ of the dipole function for each value of α.

4. We determined the preferred value of α and the 90% confidence interval of α for each dipole source model by referring to the intersections of the 90% confidence interval curves with the actual value of $<\cos\theta>$ for the dipole function of the real data.

The results for all three dipole source models are shown in Figures 8.7, 8.8, and 8.9. In each case, the nominal values of α and the 90% confidence levels only deviated marginally from the values obtained without considering angular resolution.

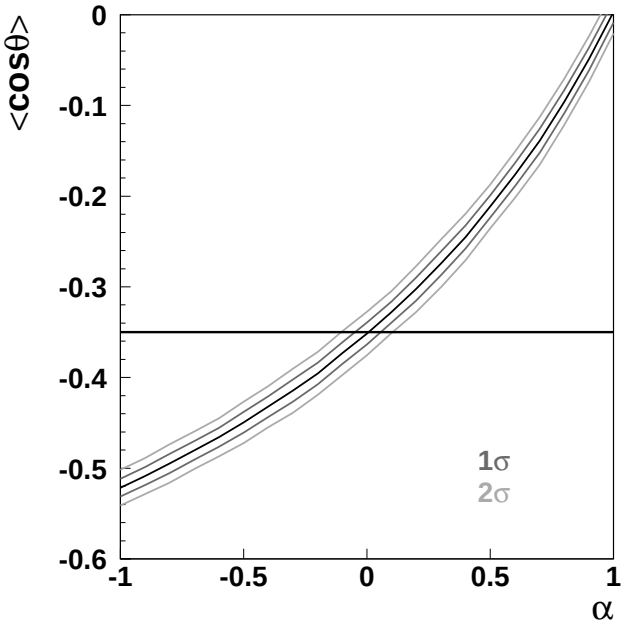


(a)

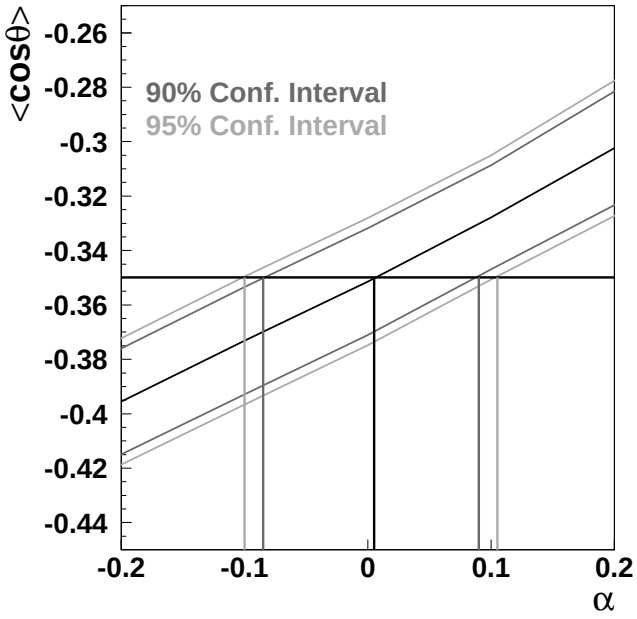


(b)

Figure 8.6. The distribution of <cos θ> values for the dipole functions of simulated data sets with a single α-value (without exposure correction)—(a) the galactic dipole source model with $\alpha = 1.0$; (b) the galactic dipole source model with $\alpha = -1.0$.

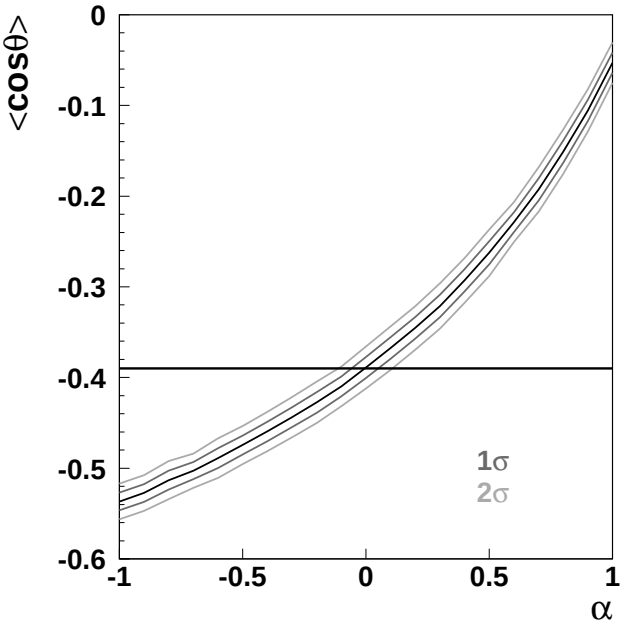


(a)

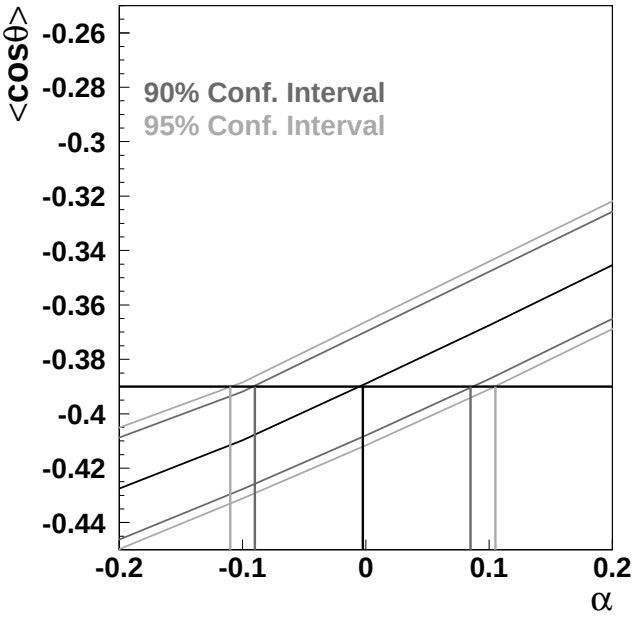


(b)

Figure 8.7. Estimations of the value of α for three different dipole source models. The curves demonstrate the dependence of $<\cos\theta>$ of the dipole functions upon α. The horizontal lines represent the value of $<\cos\theta>$ of the real data for the dipole functions of each dipole source model—(a) the galactic dipole source model for $\alpha = [-1.0, 1.0]$; (b) the critical region for the galactic dipole model: $\alpha = 0.005 \pm 0.055$ with a 90% confidence interval of: $[-0.085, 0.090]$.

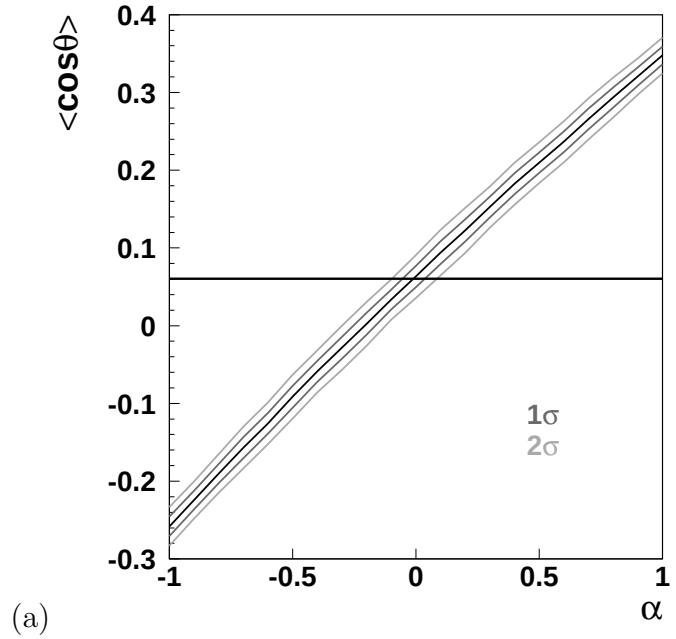


(a)

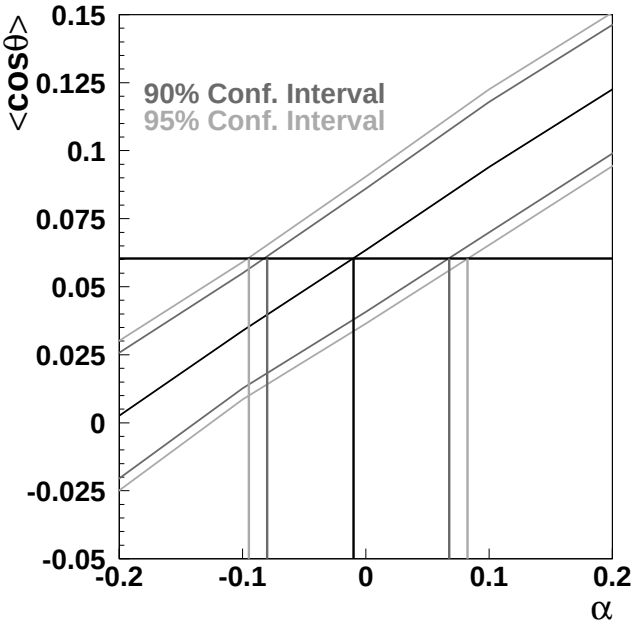


(b)

Figure 8.8. Estimations of the value of α for three different dipole source models. The curves demonstrate the dependence of $<\cos\theta>$ of the dipole functions upon α. The horizontal lines represent the value of $<\cos\theta>$ of the real data for the dipole functions of each dipole source model— (a) the Centaurus A dipole source model for $\alpha = [-1.0, 1.0]$; (b) the critical region for the Centaurus A dipole model: $\alpha = -0.005 \pm 0.065$ with a 90% confidence interval of: $[-0.090, 0.085]$;



(a)



(b)

Figure 8.9. Estimations of the value of α for three different dipole source models. The curves demonstrate the dependence of $<\cos\theta>$ of the dipole functions upon α. The horizontal lines represent the value of $<\cos\theta>$ of the real data for the dipole functions of each dipole source model— (a) the M87 dipole source model for $\alpha = [-1.0, 1.0]$; (b) the critical region for the M87 dipole model: $\alpha = -0.010 \pm 0.045$ with a 90% confidence interval of: $[-0.080, 0.070]$.

8.4 Systematic Uncertainty in the Estimation of α

There are two principal potential sources of systematic error in the determination of α with HiRes-I monocular data. The first lies in the estimation of the angular resolution. If the error in arrival direction estimation was being underestimated or overestimated, it could lead to an improper evaluation of the confidence intervals for α. In order to study the effect of angular resolution on our determination of α, we repeated our analysis of the galactic dipole model twice. In the first case, we increased the estimated angular resolution parameters for both the real and simulated data sets by 33%. In the second case, we decreased the angular resolution parameters for both types of data sets by 25%. In both cases, the width of the 90% confidence interval for α changed by less than 0.010 and the nominal value of α remained unchanged. The results suggest that the determination of α is largely independent of the angular resolution—at least for the plausible range of values that one could adopt for the angular resolution parameters.

The second issue of concern is the uncertainty in the determination of atmospheric clarity. Because hourly atmospheric observations are not available for the entire HiRes-I monocular data set, we have relied upon the use of an average atmospheric profile for the reconstruction of our data [46]. Different atmospheric conditions can influence how the profile constraint reconstruction routine interprets an observed shower profile and thus can lead to slightly divergent determinations of an event's arrival direction. Unfortunately, we do not have large libraries of simulated data with differing atmospheric parameters used in the generation and reconstruction of events. However, we do have the real data reconstructed with a full range of atmospheric parameters. By considering the value of $<\cos\theta>$ over the 1σ error space of atmospheric parameters, we can establish the degree of systematic uncertainty that is contributed to the determination of α by atmospheric variability. We saw that in the most extreme case, the nominal value of α shifted by less than .01. There was no broadening in the 90% confidence interval.

8.5 Conclusion

We are able to place upper limits on the value of $|\alpha|$ for each of our three proposed dipole source models. However, these limits are not small enough to exclude the theoretical predictions [39, 40, 41]. Also, they do not exclude the findings of the AGASA collaboration in terms of the intensity of the dipole effect that they observed or in terms of the energy considered because the events in the dipole effect observed by the AGASA detector possessed energies below $10^{18.5}$ eV [42]. Furthermore, it should be noted that our findings apply only to each dipole direction individually and not to all three simultaneously. Since it appears that angular resolution has little impact on the measurement of α and we do not appear to be systematically limited, we conclude that the driving factor in making a better determination of α will simply be larger event samples. HiRes-I mono will continue to have the largest cumulative aperture of any single detector for the next three to five years, thus it will continue to serve as an ever more powerful tool for constraining dipole source models.

CHAPTER 9

SEARCH FOR SMALL-SCALE CLUSTERING

Another topic of interest in UHECR physics over the past decade has been the search for small scale anisotropy in event arrival directions. This term refers to statistically significant excesses occurring at the scale of $\leq 2.5^{\circ}$. The interest in this sort of anisotropy has largely been fueled by the observations of the Akeno Giant Air Shower Array (AGASA). In 1999 [47] and again in 2001 [48], the AGASA collaboration reported observing what eventually became seven clusters (six "doublets" and one "triplet") with estimated energies above $\sim 3.8 \times 10^{19}$ eV. Several attempts that have been made to ascertain the significance of these clusters returned chance probabilities ranging from 4×10^{-6} [49] to 0.08 [50].

By contrast, the monocular (and stereo) analyses that have been presented by the High Resolution Fly's Eye (HiRes) demonstrate that the level of autocorrelation observed in our sample is completely consistent with that expected from background coincidences [51, 52, 53]. Any analysis of HiRes monocular data needs to take into account that the angular resolution in monocular mode is highly asymmetric.

It is difficult to compare the results of the HiRes monocular and AGASA analyses. They are very different in the way that they measure autocorrelation. Differences in the published energy spectra of the two experiments suggest an energy scale difference of 30% [2, 54]. Additionally, the two experiments observe UHECRs in very different ways. The HiRes experiment has an energy-dependent aperture and an exposure with a seasonal variability [2]. These differences make it very difficult to anticipate what HiRes should see if the AGASA claim of autocorrelation

is based on actual source. In order to develop this sort of intuition, we apply the same analysis to both AGASA and HiRes data.

9.1 The HiRes-I Monocular Data

The data set that we consider consists of events that were included in the HiRes-I monocular spectrum measurement [2, 3]. This set contains 52 events observed between May 1997 and February 2003 with measured energies greater than $10^{19.5}$ eV. This collection of events is a subset of the data used in the dipole analysis of the previous chapter and are subject to the same event reconstruction and analysis.

For the range of estimated energies considered the angular resolutions are in the range, $\sigma_\psi = [4.9, 6.1]^\circ$ and $\sigma_{\rm plane}[0.4, 1.5]^\circ$. The arrival directions of these events along with their 1σ error ellipses are plotted in equatorial coordinates in Figure 9.1.

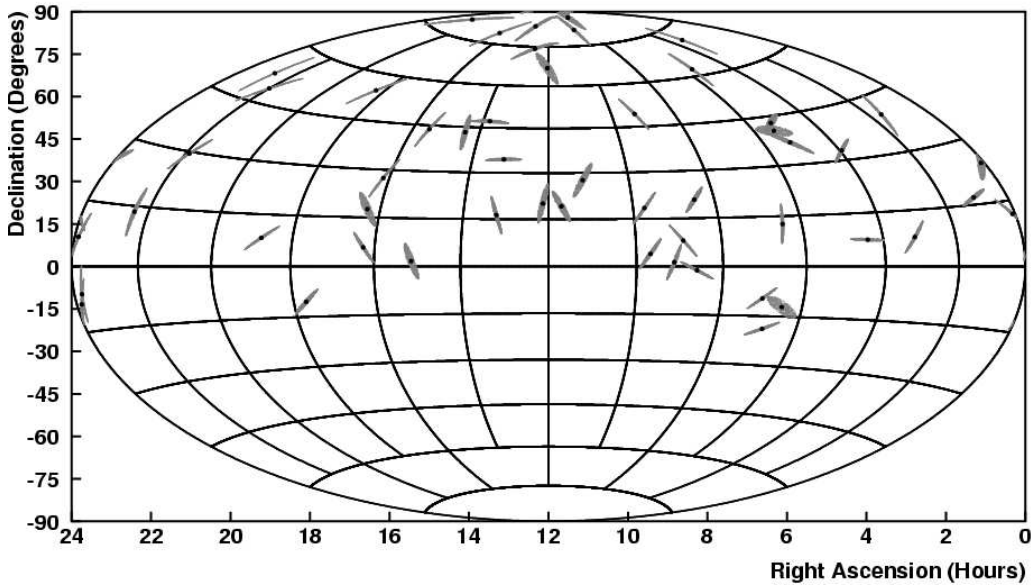


Figure 9.1. The arrival directions of the HiRes-I monocular with reconstructed energies above $10^{19.5}$ eV events and their 1σ angular resolution.

9.2 The Published AGASA Data

The AGASA data with energies above 40 EeV has been published up to the year 2000 [25] and all but one of these events used for this calculation has a measured energy greater than 4×10^{19} eV. The AGASA estimated angular errors [47] are shown in Figure 9.2. The AGASA angular errors (Figure 9.2) are fit to a two-component Gaussian distribution:

$$n = N_{\circ}(E_{\mathrm{EeV}})\left[0.33\Delta\theta e^{-(\Delta\theta)^2/2\sigma_1^2} + 0.67\Delta\theta e^{-(\Delta\theta)^2/2\sigma_2^2}\right] \tag{9.1}$$

where $\sigma_1 = 6.52^\circ - 2.16^\circ \log_{10} E_{\mathrm{EeV}}$, $\sigma_2 = 3.25^\circ - 1.22^\circ \log_{10} E_{\mathrm{EeV}}$, and $N_{\circ}(E)$ is a numerically determined normalization constant. Figure 9.3 shows the arrival directions of the published AGASA events plotted in equatorial coordinates with their 68% angular resolution.

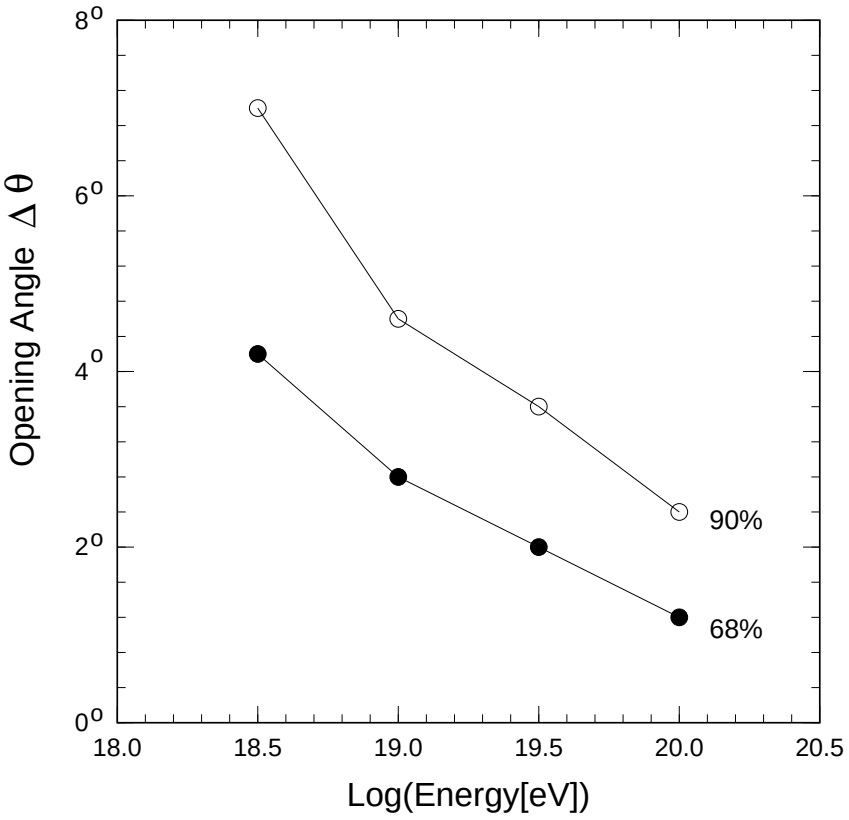


Figure 9.2. The AGASA angular resolution as a function of estimated energy [47].

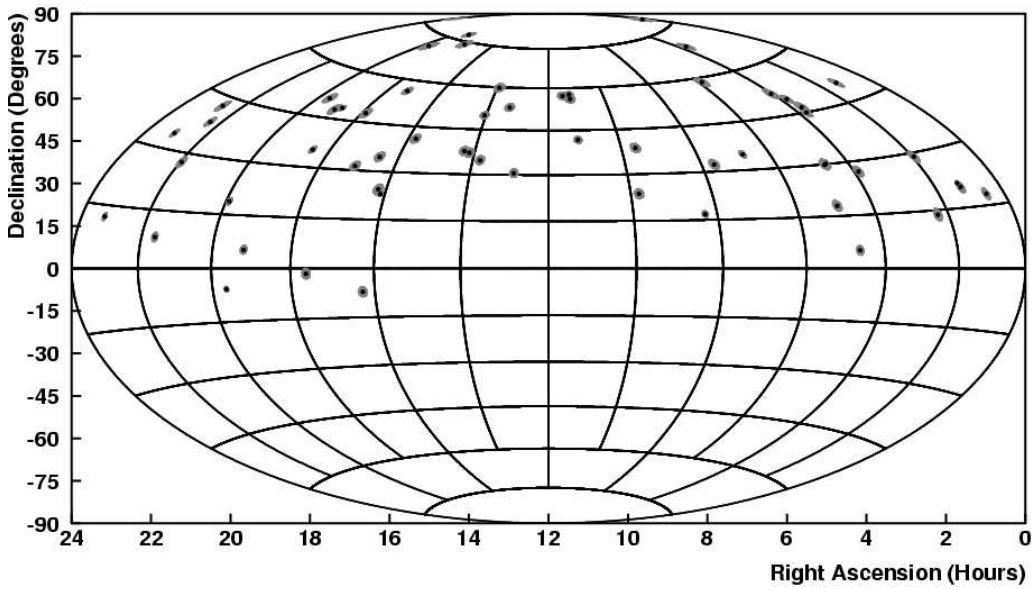


Figure 9.3. The arrival directions of the published AGASA events with their 68% angular resolution.

9.3 The Autocorrelation Function

We measure the degree of autocorrelation in both samples by means of an autocorrelation function. It is calculated as follows:

1. For each event, an arrival direction is sampled on a probabilistic basis from the error space defined by the angular resolution of the event.

2. The opening angle is measured between the arrival directions of a pair of events.

3. The cosine of the opening angle is then histogrammed.

4. The preceding steps are repeated until all possible pairs of the events are considered.

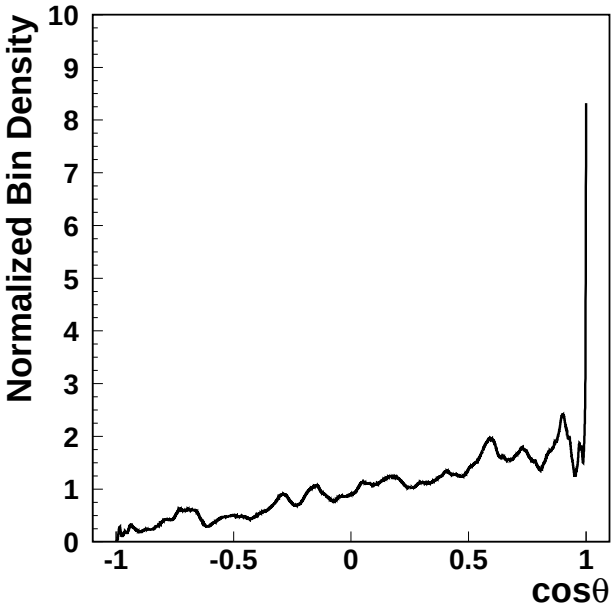
5. The preceding steps are repeated until the error space, in the arrival direction of each event, is thoroughly sampled.

6. The histogram is normalized and the resulting curve is the autocorrelation function.

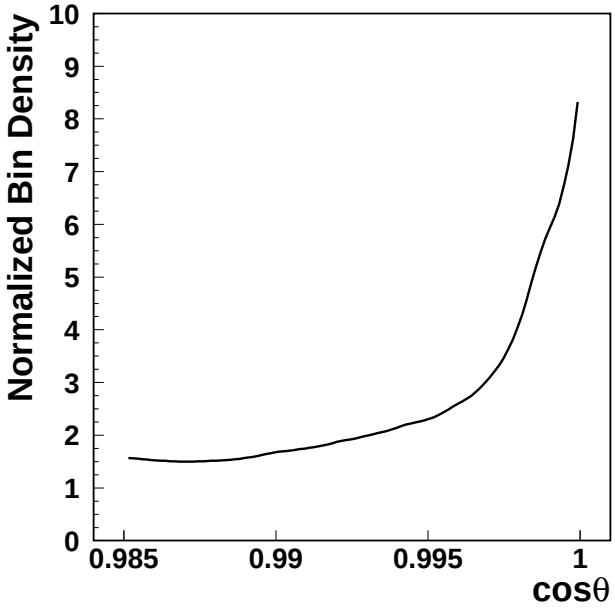
Figure 9.4a and 9.4b show an example of the autocorrelation function for a highly clustered set of simulated data. The sharper the peak at $\cos\theta_{\min}$ is, the more highly autocorrelated the data set is. There are many ways that one could quantify the degree of autocorrelation that a set possesses. The most obvious way is to look at the value of the bin which contains $\cos\theta_{\min}$. However, this method has some shortcomings. First of all, the value of the last bin is dependent upon the chosen bin width. Also, the value of the last bin is not stable unless the angular resolution is sampled at a level that is computationally unfeasible. Finally, the value of the last bin over a large number of similarly autocorrelated sets does *not* produce a Gaussian distribution (see Figure 9.5a),thus complicating the interpretation of the results of an analysis employing $\cos\theta_{\min}$ as an observable.

A more well-behaved measure of the autocorrelation of a specific set of data is the value of $<\cos\theta>$ for $\theta \leq 10^{\circ}$. This value is also a measure of the sharpness of the autocorrelation peak at $\cos\theta = 1$. However, this method of quantification does not depend on bin width and it does produce Gaussian distributions when applied to large numbers of sets with similar degrees of autocorrelation, as is demonstrated in Figure 9.5b. An additional advantage to this method is that by considering the continuous autocorrelation function over a specified interval, both the peak at the smallest values of θ and the corresponding statistical deficit in the autocorrelation function at slightly higher values of θ are taken into account. Thus we simultaneously measure both the positive and negative swings in the autocorrelation signal. The interval of $[0^{\circ}, 10^{\circ}]$ was chosen because it was found, in simulations, to optimize the autocorrelation signal for clusters resulting from point sources spread isotropically across the sky.

Using the description of the HiRes-I monocular angular resolution from Chapter 6, we then calculate the autocorrelation function via the method described above. In Figure 9.6, we show the result of this calculation. For this sample, we obtain $<\cos\theta>_{[0^{\circ},10^{\circ}]} = 0.99234$.

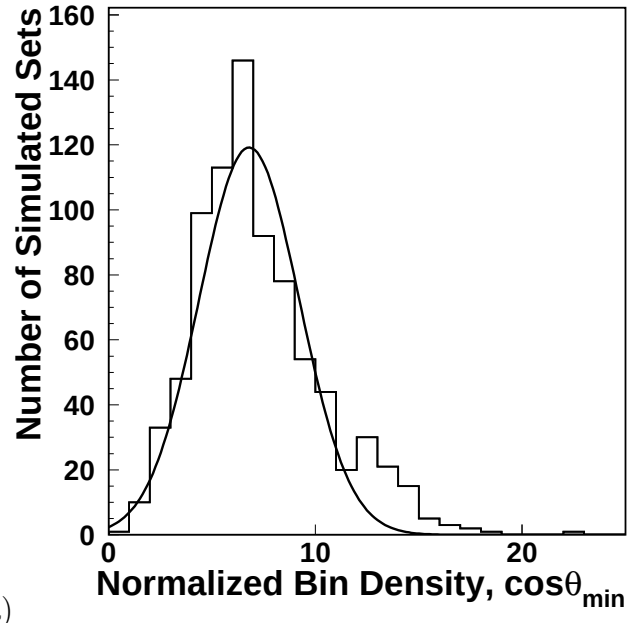


(a)

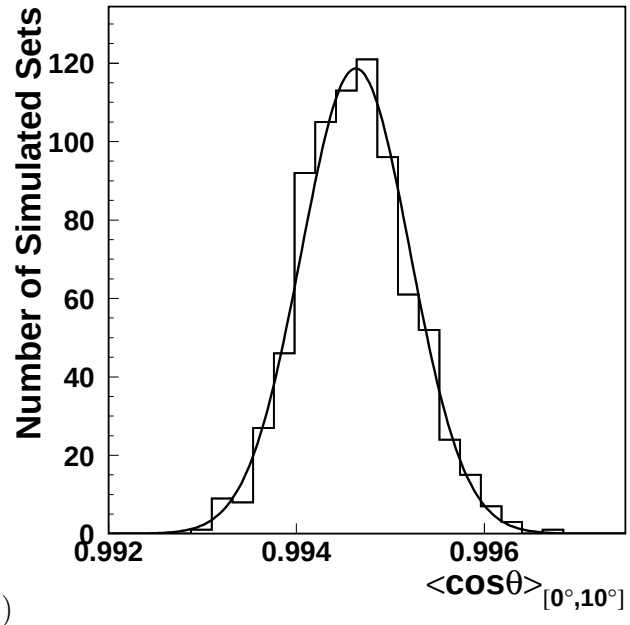


(b)

Figure 9.4. An example of the autocorrelation function for a simulated data set that contains ~ 10 clusters in a total of 60 events—(a) the full autocorrelation function for $\theta = [0°, 180°]$; (b) the critical region of the the autocorrelation function: $\theta = [0°, 10°]$.

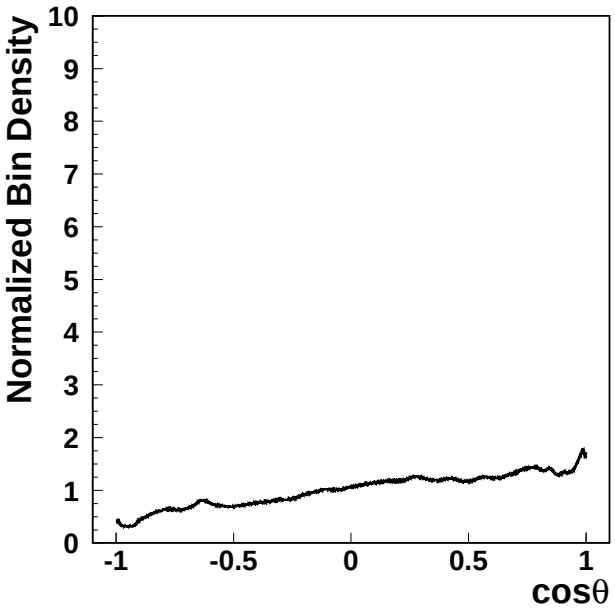


(a)

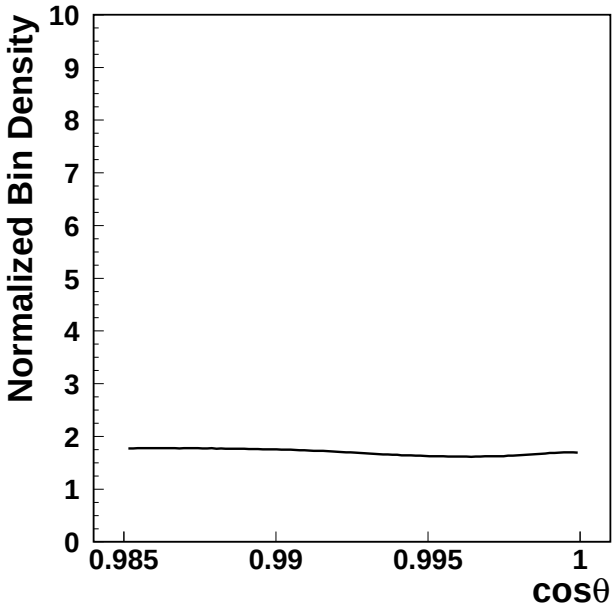


(b)

Figure 9.5. Distributions of normalized bin densities of $\cos\theta_{\min}$ and $<\cos\theta>_{[0^\circ,10^\circ]}$ values for a large number of simulated sets with the same level of clustering as in Figure 9.4—(a) Distribution of observed normalized bin densities of $\cos\theta_{\min}$, note that it is *not* Gaussian ($\chi^2/dof = 5.44$); (b) : $<\cos\theta>_{[0^\circ,10^\circ]}$ distribution ($\chi^2/dof = 1.09$).



(a)



(b)

Figure 9.6. The autocorrelation for the HiRes-I events above $10^{19.5}$ eV—(a) the full autocorrelation function for $\theta = [0°, 180°]$; (b) the critical region of the autocorrelation function: $<\cos\theta>_{[0°,10°]}= 0.99234$.

We also calculate the autocorrelation function for the published AGASA events. We show the result in Figure 9.7. For this sample, we obtain $<\cos\theta>_{[0^\circ,10^\circ]}=$ 0.99352.

9.4 Quantifying the Relative Sensitivity of HiRes-I and AGASA to Autocorrelation

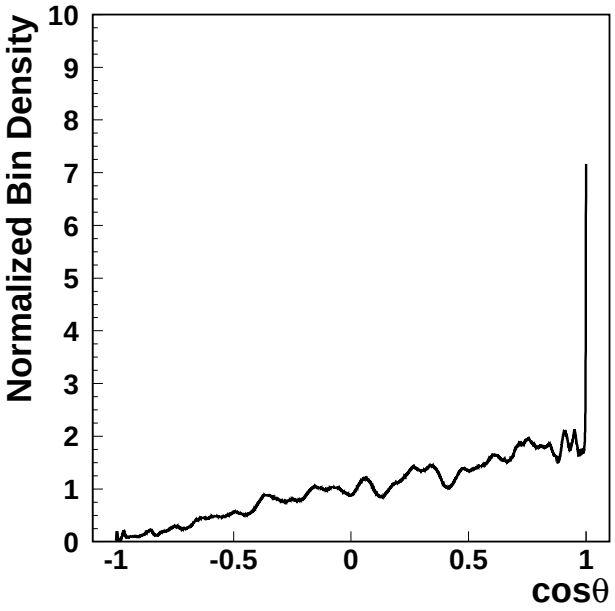
In order to quantify the relative sensitivity of the AGASA and HiRes-I data sets, we must first understand the exposures of both detectors. For HiRes-I, we assemble a library of approximately 8×10^4 simulated events with energies above $10^{19.5}$ eV. We then pair each event with randomly generated times during which the detector was operating. As described in Chapter 6, a mirror-by-mirror correction is applied where simulated events are rejected if the mirror(s) that would have observed the event in question was not operating at the time that event would have occurred. Once 10^7 pairings of simulated events and times are assembled, a surface plot is created of the event density on a bin by bin basis. The value of each bin is then normalized so that the mean value of all the bins in the *observable* sky δ (declination) $= [-30^\circ, 90^\circ]$ is 1. The resulting surface plot is shown in a Hammer-Aitoff projection in Figure 9.8. We have shown in Chapter 7 that this method produced zenith angle, azimuthal angle, and sidereal time distributions that were consistent with that observed in the actual data [4]. The highest exposure areas have a normalized relative exposure: $\rho_{\rm H}(\delta,\alpha) =\sim 2.5$.

For the AGASA detector, we refer to the distribution of event declinations presented in Uchiori et al. [55]. By following the lead of Evans et al. [56], we fit a normalized polynomial to this distribution:

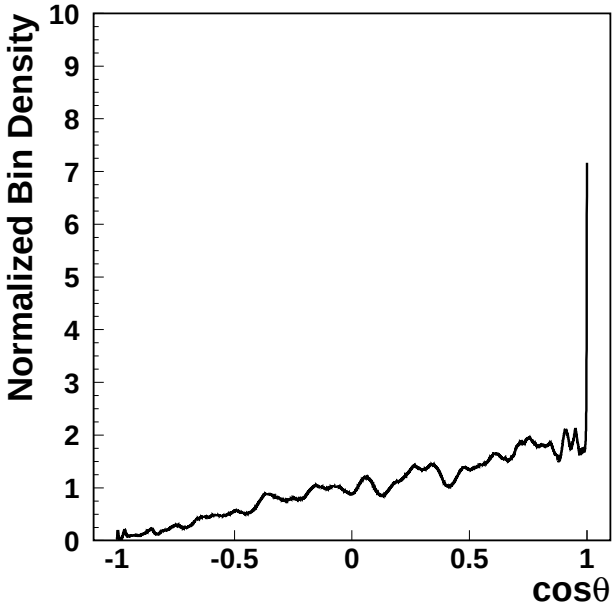
$$\begin{aligned} N(\delta) &= 0.323616 + 0.0361515\delta - 5.04019 \times 10^{-4}\delta^2 + \\ & \quad 5.539141 \times 10^{-7}\delta^3; \end{aligned} \tag{9.2}$$

where $N(\delta)$ holds for $\delta = [-8^\circ, 87.5^\circ]$the maximum value of $N(\delta)$ is 1. We also know that:

$$A_\circ \int_{-8^\circ}^{87.5^\circ} N(\delta)\, d\delta = \int_{-8^\circ}^{87.5^\circ} \rho_{\rm A}(\delta) \cos\delta \, d\delta, \tag{9.3}$$



(a)



(b)

Figure 9.7. The autocorrelation for the published AGASA events—(a) the full autocorrelation function for $\theta = [0°, 180°]$; (b) the critical region of the autocorrelation function: $<\cos\theta>_{[0°,10°]} = 0.99352$.

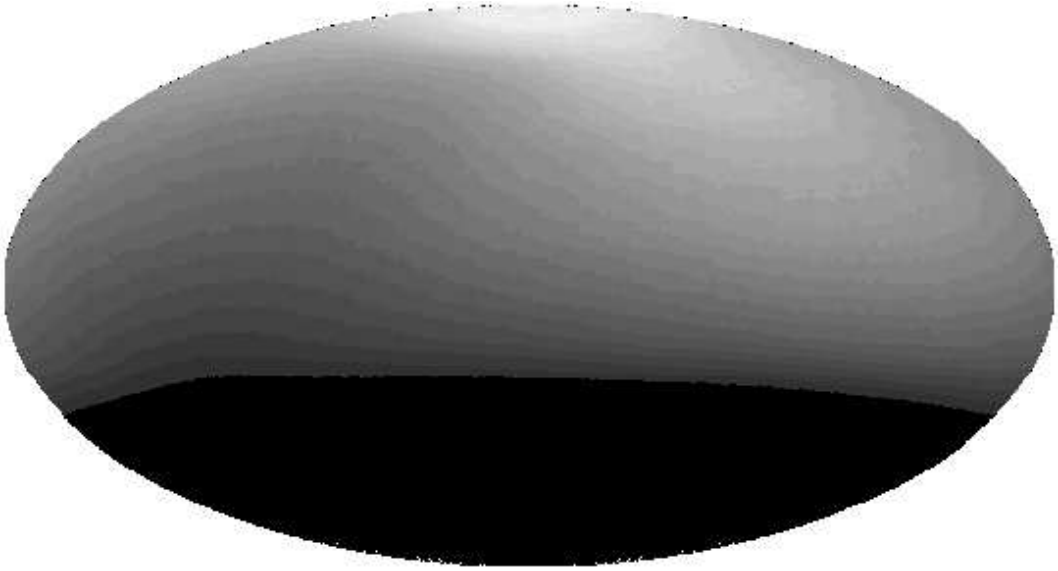


Figure 9.8. Hires-I estimated relative exposure, $\rho_{\rm H}(\delta, \alpha)$, for events above $10^{19.5}$ eV in equatorial coordinates (right ascension right to left). The lightest region corresponds to a normalized event density of 2.5. The observable sky extends from $\delta = -30^{\circ}$ to $\delta = 90^{\circ}$.

where $A_{\circ}$ is a numerically determined normalization constant. We then derive:

$$\rho_{\rm A}(\delta) = A_{\circ} N(\delta) \sec\delta; A_{\circ} = 1.0251. \tag{9.4}$$

The value of each bin is once again normalized so that the mean value of all the bins in the *observable* sky $\delta = [-8^{\circ}, 87.5^{\circ}]$ is 1. The resulting surface plot is shown in a Hammer-Aitoff projection of a equatorial coordinates in Figure 9.9. The highest exposure areas have $\rho_{\rm A}(\alpha) =\sim 1.6$. In Figure 9.10, we show the distribution of $< \cos\theta >_{[0^{\circ},10^{\circ}]}$ values for isotropic data sets with each of the two different exposure models (HiRes-I and AGASA). The AGASA data set manifests $\sim 10^{-3}$ chance probability above background. For the AGASA data, we also calculated the autocorrelation function without consideration to angular resolution and employed the more conventional θ_{min} observable. After varying the bin width for θ_{min} and accounting for the trials factor, we independently concluded that the chance probability is $\sim 10^{-3}$ for the optimal bin width, $\theta_{min} = [0^{\circ}, 2.5^{\circ}]$. We thus conclude that factoring angular resolution into our analysis and employing $<\cos\theta>_{[0^{\circ},10^{\circ}]}$ as an

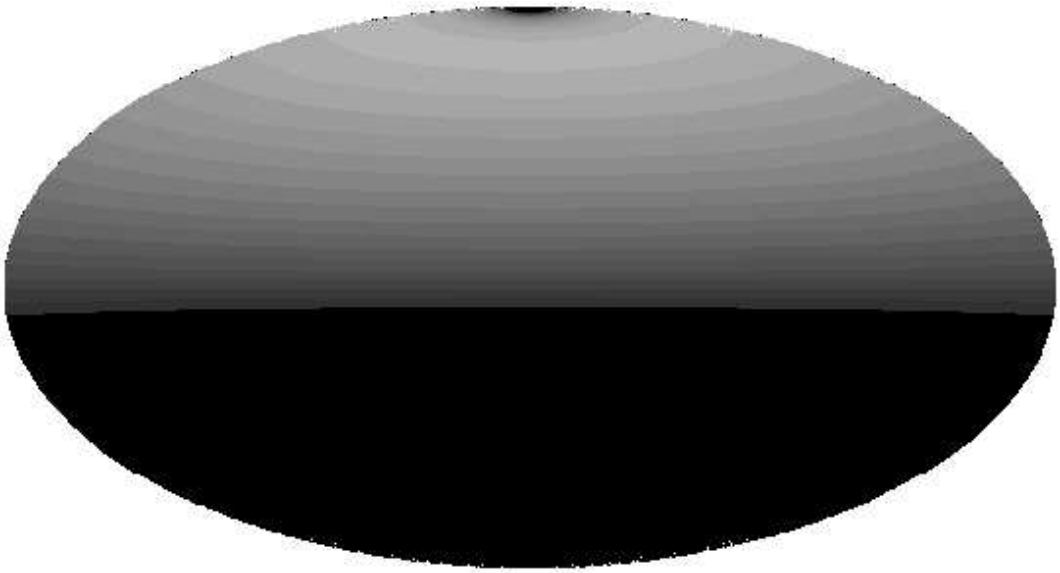


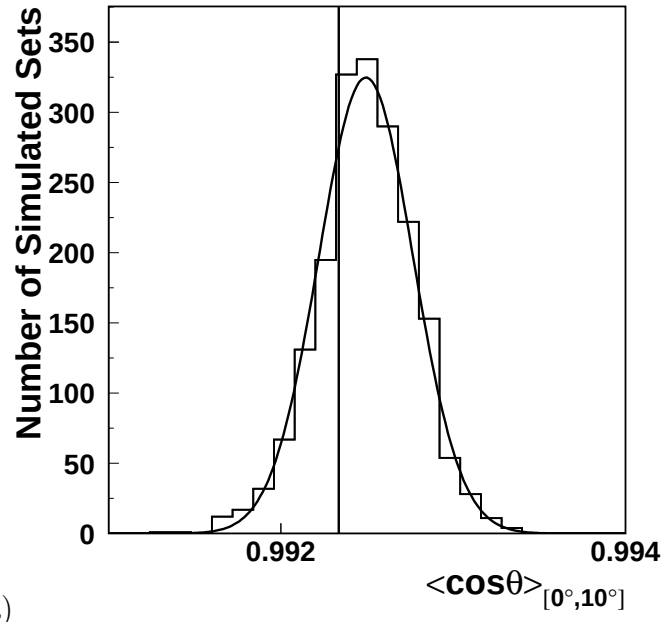
Figure 9.9. AGASA estimated relative exposure, $\rho_{\rm A}(\delta)$, for events above $10^{19.5}$ eV in equatorial coordinates (right ascension right to left). The lightest region corresponds to a normalized event density of ~ 1.6. The observable sky extends from $\delta = -8^\circ$ to $\delta = 87.5^\circ$.

observable in no way diminishes the sensitivity to autocorrelation in the reported AGASA data.

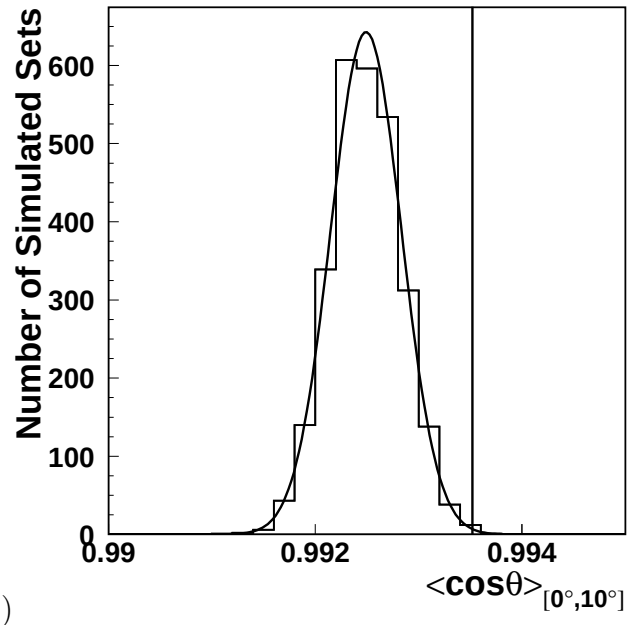
There are a few important differences between the exposure of the HiRes-I and AGASA detectors. First of all, the exposure of the HiRes-I detector is more asymmetric than the exposure of the AGASA detector. This is not only due to seasonal variations in the HiRes detector, but also due to its ability to constantly observe the region around $\delta = 90^\circ$ as a result of a higher zenith angle acceptance. This higher zenith acceptance also allows the HiRes detector to observe a greater region of the southern hemisphere. In general, while AGASA reports observations for 56.9% of the total sky, the HiRes-I detector reports observations for 75% of the total sky.

To simulate clustering we use the following prescription:

1. An event is chosen based upon the distribution in α and δ that is dictated by ρ. In the case of HiRes-I, this is simply done by selecting a simulated event from our library and then assigning it a time that is a known good-weather



(a)



(b)

Figure 9.10. Distribution of $<\cos\theta>_{[0^\circ,10^\circ]}$ values for simulated isotropic data sets—(a) HiRes-I; (b) AGASA. In each figure, the vertical line represents the the value of $<\cos\theta>_{[0^\circ,10^\circ]}$ for the observed data.

ontime for the mirror(s) that observed that event. In the case of the AGASA detector, this is done by selecting a random value for δ that conforms to the distribution in equation (9.4) and then assigning it a random value in α (right ascension) between 0h and 24h and sampling a value for the energy from the energies of the reported events.

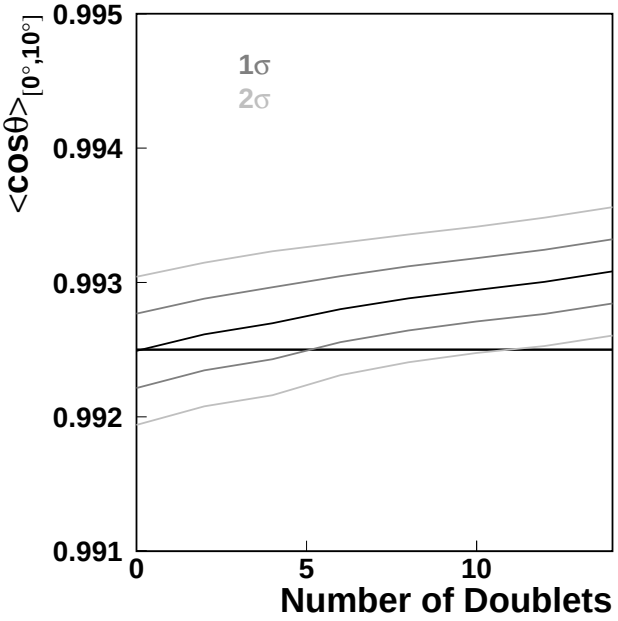
2. This event does not represent the source location itself, but is assumed to have arrived from the source location with some error. We construct a "true" source location by sampling the error space of this event.

3. For each additional event assigned to that source, a simulated event is selected with a "true" arrival direction that is the same as that of the initial event.

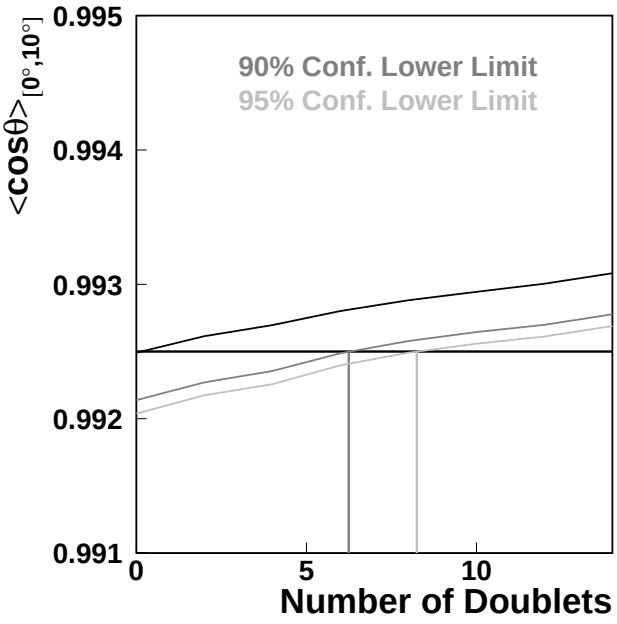
To study the relative sensitivity of AGASA and HiRes-I, we measure the value of $<\cos\theta>_{[0^\circ,10^\circ]}$ for multiple simulated sets with a variable number of doublets inserted. We then construct an interpolation of the mean value and standard deviation of $<\cos\theta>_{[0^\circ,10^\circ]}$ from a given number of observed doublets for each experiment. This will allow us to state the number of doublets required for each experiment in order for the 90% confidence limit of $<\cos\theta>_{[0^\circ,10^\circ]}$ to be above the background value of 0.99250. Figures 9.11 and 9.12 show the result of these simulations. In general, for the HiRes-I data set, the 90% confidence lower limit corresponds to the mean expected background signal with the inclusion of 6.25 doublets. For AGASA,the 90% confidence lower limit corresponds to the mean expected background signal with the inclusion of 5.5 doublets. This demonstrates that while AGASA has a slightly better ability to perceive autocorrelation, the sensitivity of the two experiments is comparable.

We now apply the actual Hires-I $<\cos\theta>_{[0^\circ,10^\circ]}$ to the sensitivity curve shown in Figures 9.11 and 9.12. In Figure 9.13 we can see the result of these simulations. The observed HiRes-I signal corresponds to the 90% confidence upper limit with the inclusion of only 3.5 doublets beyond random background coincidence.

If we repeat this analysis with first, a 7.5% reduction in the estimated angular resolution values and second, a 7.5% increase in the estimated angular resolution

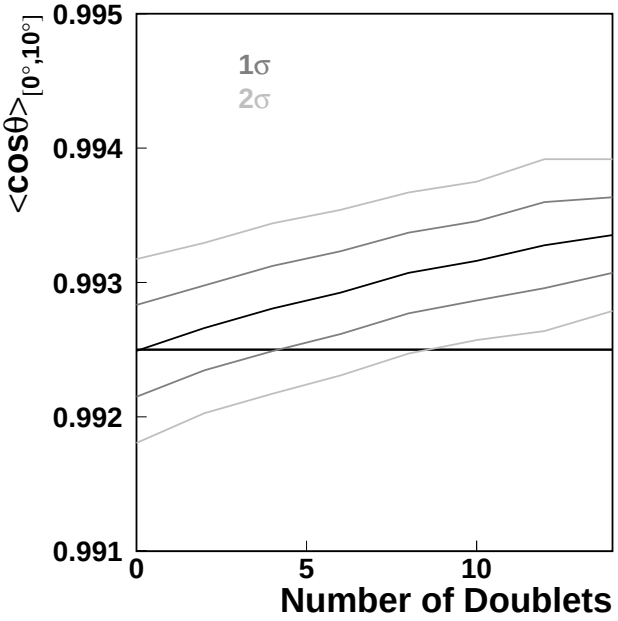


(a)

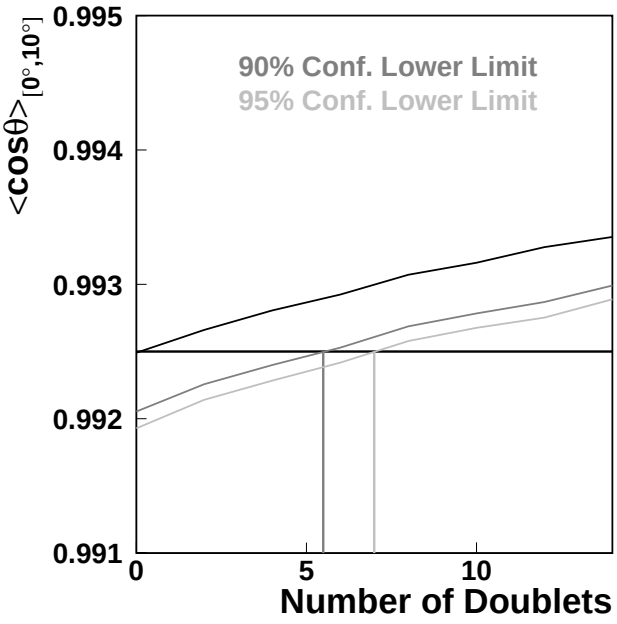


(b)

Figure 9.11. Relative sensitivity of HiRes-I and AGASA to doublets—(a) Simulations with the HiRes-I detector and 52 events; (b) 90% confidence above background: 6.25 doublets, 95% confidence above background: 8.25 doublets. In each figure, the horizontal line indicates the expected value of $<\cos\theta>_{[0°,10°]}$ for an isotropic background.

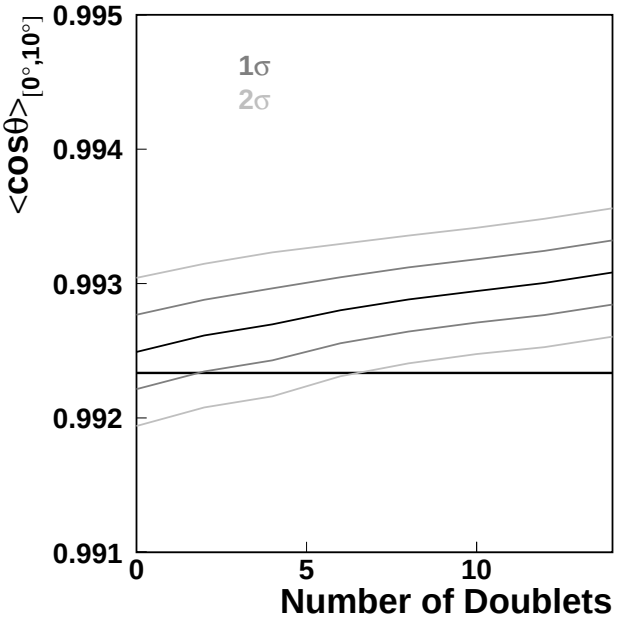


(a)

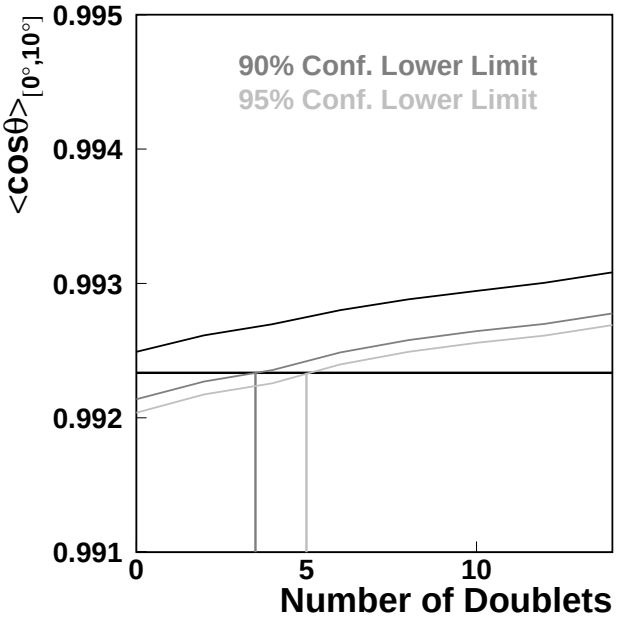


(b)

Figure 9.12. Relative sensitivity of HiRes-I and AGASA to doublets— (a) simulations with the AGASA detector and 59 events; (b) 90% confidence above background: 5.5 doublets, 95% confidence above background: 7.0 doublets. In each figure, the horizontal line indicates the expected value of $<\cos\theta>_{[0°,10°]}$ for an isotropic background.



(a)



(b)

Figure 9.13. Sensitivity of the HiRes-I monocular observations to doublets—(a) Simulations with the HiRes-I detector and 52 events; (b) 90% confidence above observed signal: 3.5 doublets, 95% confidence above observed signal: 5 doublets. In each plot, the horizontal line represents the value of $< \cos\theta >_{[0°,10°]}$ for the observed HiRes-I data.

values, we obtain a range for the 90% confidence upper limit of $[2.75, 4.0]$ doublets and a range for the 95% confidence upper limit of $[4.5, 5.5]$ doublets.

A final area of concern is the systematic uncertainty in the determination of atmospheric clarity. As before, we have relied upon the use of an average atmospheric profile for the reconstruction of our data [46]. While different atmospheric conditions have negligible impact on the determination of the arrival direction for events with measured energies this high, differing conditions can have an impact on energy estimation and thus the number of events that are included in our data set. Over the 1σ error space for our estimation of atmospheric conditions, the total number of events in our data set fluctuates on the interval $[41, 65]$. The value of the observable, $<\cos\theta>_{[0^\circ,10^\circ]}$, has a fluctuation on the interval $[0.99226, 0.99249]$ owing to addition and subtraction of events from the data set. Note that in neither case does the value of $<\cos\theta>_{[0^\circ,10^\circ]}$ exceed the mean value (0.99250) expected for a background set.

9.5 Conclusion

We conclude that the HiRes-I monocular detector sees no evidence of clustering in its highest energy events. Furthermore, the HiRes-I monocular data has an intrinsic sensitivity to global autocorrelation such that we can claim at the 90% confidence level that there can be no more than 3.5 doublets above that which would be expected by background coincidence in the HiRes-I monocular data set above $10^{19.5}$ eV. From this result, we can then derive, with a 90% confidence level, that no more than 13% of the observed HiRes-I events could be sharing common arrival directions. This data set is comparable to the sensitivity of the reported AGASA data set if one assumes that there is indeed a 30% energy scale difference between the two experiments. It should be emphasized that this conclusion pertains only to point sources of the sort claimed by the AGASA collaboration. Furthermore, because a measure of autocorrelation makes no assumption of the underlying astrophysical mechanism that results in clustering phenomena, we cannot claim that

the HiRes monocular analysis and the AGASA analysis are inconsistent beyond a specified confidence level.

CHAPTER 10

USING FRACTAL DIMENSIONALITY IN THE SEARCH FOR SOURCE MODELS

The previous two chapters have alluded to significant discrepancies in the anisotropy results reported by the three major cosmic ray experiments. In particular, AGASA has reported both a small excess toward the Galactic Center for events in the range 10^{18} to $10^{18.4}$ eV [42], and a clustering at small angular scales for events above 4×10^{19} eV [47]. The Fly's Eye did not see an excess towards the Galactic Center [57, 58], but did report a small enhancement along the Galactic Plane [43]. In 1995, it was reported by Stanev et al. [59] that the combined data of Haverah Park [60], Yakutsk [61], AGASA [62], and Volcano Ranch [63] showed an excess along the supergalactic plane with several potential point sources for events above 2×10^{19} eV. However, the HiRes experiment has reported that it saw no evidence of an enhancement towards the galactic center [4] or small-scale clustering [5].

These conflicting reports call for developing a more global way in which one could determine if a given sample possesses any statistically significant anisotropy. We will show that by considering the information dimension, D_{I}, of a given sample, one can simultaneously look for anisotropies at all angular scales greater than the angular resolution of the sample by considering the intrinsic heterogeneity of that particular data sample. This method is quite robust in that it can easily accommodate both asymmetric angular resolutions and irregular apertures. Furthermore, in the event that a sample is shown to be consistent with an isotropic distribution, this method can be used to place upper limits on possible source models.

An important advantage of this technique is that, regardless of the model being considered, only a single measurement of the data is actually taken. In this way,

there is no dilution of the statistical significance when the experimental data set is compared against a large number of models, as was the case for the reported AGASA excesses. In those studies, systematic scans of bin sizes were made to maximize the apparent significance of the observed excess. The fractal dimensionality method does not suffer from this shortcoming. For example, the observed $D_{\rm I}$ value can be compared to that generated for any number of *a priori* sources without incurring a statistical penalty whereas a conventional binned, background subtraction method would suffer from artificially enhanced statistical chance of finding a signal.

10.1 Calculating Fractal Dimensions of a Data Sample

Fractal dimensionality is a simple measure of the scaling symmetry of a structure. By measuring the fractal dimension of a data sample, one can examine its heterogeneity at different levels of magnification. There are several ways of exploiting this idea. From a computational perspective, the simplest is to use box-counting. For the most general case, the capacity dimension, $D_{\rm c}$ [64, 65], one partitions the sample space into equi-sized and equi-shaped "boxes" with edge size ϵ:

$$D_{\rm c} = \lim_{\epsilon \to 0^+} \frac{\log N(\epsilon)}{\log 1/\epsilon}. \tag{10.1}$$

Here, $N(\epsilon)$ is the minimum number of "boxes" with edge size ϵ necessary to cover one's sample.

The capacity dimension has a serious limitation: It only looks for the presence of the sample within the available space and does not consider variations in the density of the sample at a given point in space. In cases where the density may differ within the sample space, the appropriate alternative is to use the information dimension, $D_{\rm I}$ [66, 67]:

$$D_{\rm I} = -\sum_{i=1}^{N} \lim_{\epsilon \to 0^+} \frac{P_{\rm i}(\epsilon) \log P_{\rm i}(\epsilon)}{\log 1/\epsilon}, \tag{10.2}$$

where $P_{\rm i}(\epsilon)$ is the probability of finding a data point in the i-th box of edge size ϵ. This is a particularly suitable measurement when considering a data set consisting of UHECR arrival directions with finite angular resolution.

It should be noted that $D_{\rm C}$ and $D_{\rm I}$ are both particular cases of the q-dimension [68, 69],

$$D_q = \frac{1}{1-q} \lim_{\epsilon \to 0^+} \frac{\log I(q,\epsilon)}{\log 1/\epsilon}; \tag{10.3}$$

where

$$I(q,\epsilon) = \sum_{i=1}^{N} P_{\rm i}(\epsilon)^q. \tag{10.4}$$

We can then see that $D_{\rm C} = \lim_{q\to 0} D_q$ and that $D_{\rm I} = \lim_{q \to 1} D_q$.

10.2 Application to Arrival Direction Distributions for UHECRs

In principle, it is simple to calculate the information dimension for a given sample of events. However, two complications arise when considering a set of arrival directions of UHECRs. First, the event directions are not known with complete precision. This makes the determination of $D_{\rm I}$ as $\epsilon \to 0$ meaningless. Secondly, the determination of $D_{\rm I}$ requires that the sample space be divided into equi-sized and equi-shaped bins. For a spherical surface, this is simply not possible. Nevertheless, there are workable solutions for both of these problems.

10.2.1 Angular Resolution

We will consider a hypothetical monocular air-fluorescence detector. This detector is largely based upon the characteristics of the Hires-I detector [2, 3]. Our detector observes events with an angular resolution that is described by a highly asymmetric 2-d Gaussian. For a monocular air fluorescence detector, angular resolution consists of two components: σ_1, in the determination of the angle, ψ, within the plane of reconstruction, and σ_2, in the estimation of the plane of reconstruction itself. Figure 5.1 illustrates how this geometry would appear with a particular plane of reconstruction and a particular value for ψ. Intuitively, we can see that we should be able to determine the plane of reconstruction quite accurately.

However, the value of ψ is more difficult to determine because it is dependent on the precise results of the monocular reconstruction [2, 3].

The actual parameterizations of σ_1 and σ_2 assumed are as follows:

$$\sigma_1 = 20^\circ e^{-1.5 \log_{10} E_{\mathrm{EeV}}} + 4^\circ \tag{10.5}$$

and

$$\sigma_2 = 100^\circ e^{-0.5\Delta\chi} + 0.4^\circ. \tag{10.6}$$

Here, E_{EeV} is the primary energy of the shower in EeV. For the purpose of this study, the energy will be allowed to vary between $10^{18.5}$ eV and 10^{20} eV with a differential spectral index of -2.7. In this scenario, a shower with a primary energy of $10^{18.5}$ eV will have $\sigma_1 = 13.4^\circ$, while a shower with a primary energy of 10^{20} eV will have $\sigma_1 = 5.0^\circ$. This difference can be attributed to the fact that larger showers have better defined profiles and a better signal-to-noise ratio.

The factor, $\Delta\chi$, in equation 10.6 is the angular track length (in degrees) of the shower as observed by the detector, which is allowed to vary between 8° and 30°. A shower with an observed track length of 8° will have $\sigma_2 = 2.2^\circ$, while a shower with an observed track length of 30° will have $\sigma_2 = 0.4^\circ$; a longer track-length leads to a more accurate determination of the plane of reconstruction. The distribution of σ_2 values is largely independent of energy because while higher energy showers do lead to more longitudinal development, they are also brighter, which allows one to observe them at greater distances. These competing factors lead to $\Delta\chi$ distributions that are virtually identical across the observed spectrum.

In general, it should be noted the distributions of σ_1 and σ_2 values are relatively insensitive to the differential spectral index that is chosen. We ascertained this by considering two simulated data sets, one with a differential spectral index of -2.5 and one with a differential spectral index of -3.5. Even for a variation that was much larger than the accepted range of experimental values for the the UHECR spectrum [2, 3, 15, 54], the value of $\bar{\sigma_1}$ increased by only 11%. The value of $\bar{\sigma_2}$ remained unchanged. This can be explained by realizing for a steeply falling spectrum, the overwhelming majority of observed events in either case will occur

in the first half decade of the measurement. This is a very small effect compared to the expected statistical fluctuations that would occur between two consecutive sets of observations.

For the purpose of calculating $D_{\rm I}$, we can treat the arrival direction of each individual shower as a two-dimensional elliptical Gaussian distribution with the parameters σ_1, σ_2. The size of a bin's edge, ϵ, will be allowed to take on a series of values, $\Delta\theta$, which will be in an interval corresponding to a scale length of the sample or in the case of a smooth distribution, the smallest value that is computationally feasible. For finite events samples, we will use $\Delta\theta \simeq 0.5^\circ$. In the case of smooth distributions we will use a computationally limited value of $\Delta\theta = 1/6^\circ$. The number of points in each shower direction distribution, $N_{\rm Dist}$, will be determined by the mean value, $<n_{\rm i}>$, necessary to assure that the fractional Gaussian fluctuations of the count, $n_{\rm i}$, in each bin, do not on average, exceed a predetermined value (i.e. for 5% fluctuations, $N_{\rm Dist} \simeq 500$). For each value of ϵ, the probability, $P_{\rm i}(\epsilon)$, for the i-th bin will be:

$$P_{\rm i}(\epsilon) = \frac{n_{\rm i}}{N_{\rm Dist} \cdot N_{\rm Shower}}. \tag{10.7}$$

We calculate $D_{\rm I}$ for each value of ϵ:

$$D_{\rm I}(\epsilon) = -\sum_{i=1}^{N} \frac{P_{\rm i}(\epsilon) \log P_{\rm i}(\epsilon)}{\log 1/\epsilon}. \tag{10.8}$$

We then determine $D_{\rm I}$ to be $<D_{\rm I}(\epsilon)>$ over the specified interval of ϵ values.

10.2.2 Latitudinal Binning

For the purpose of calculating $D_{\rm I}$, it is necessary that all bins be equi-sized and equi-shaped as we vary the size of the side of the bins, ϵ. While it is impossible to achieve completely this criterion on the surface of a sphere, we will be able to do so approximately by adopting a latitudinal binning scheme.

Latitudinal binning is achieved by first dividing the sky into N_δ declinational (δ) bands where each band has a width

$$\Delta\theta = \frac{\pi}{N_\delta} \tag{10.9}$$

For each declinational band, the sky is then divided into $N_{\alpha,\delta}$ bins in right ascension (α) where:

$$N_{\alpha,\delta} = \left[\frac{2\pi \int_{\delta 1}^{\delta 2} \cos\delta \, d\delta}{(\Delta\theta)^2}\right] = \left[\frac{2(N_\delta)^2 \int_{\delta 1}^{\delta 2} \cos\delta \, d\delta}{\pi}\right]. \tag{10.10}$$

The solid angle, $\Delta\Omega_\delta$ of each bin (in steradians) is:

$$\Delta\Omega_\delta = \frac{2\pi \int_{\delta 1}^{\delta 2} \cos\delta \, d\delta}{N_{\alpha,\delta}}, \tag{10.11}$$

with a minimum value of $(\Delta\theta)^2$ (at the equator) and a maximum value of $\frac{\pi}{3}(\Delta\theta)^2$ (at the poles) regardless of the value of N_δ. This provides us with bins that are all almost the same area and nearly square-shaped (with the exception of three triangular bins at each pole). The total number of bins in the sky can be approximated by:

$$N_{sky} \simeq 4\pi(\frac{N_\delta}{\pi})^2 = \frac{4}{\pi}(N_\delta)^2. \tag{10.12}$$

In Figures 10.1 and 10.2, we visualize the latitudinal binning technique for a series of different N_δ values.

10.2.3 Application to the Calculation of $D_{\rm I}$

We can now apply the preceding machinery to the calculation of $D_{\rm I}$: First, we need to normalize the event count in each bin by its respective bin area, $\Delta\Omega_\delta$:

$$P_{\rm i}(\epsilon) = \frac{n_{\rm i}(\Delta\theta)^2}{N_{\rm Dist} N_{\rm Shower} \Delta\Omega_\delta} = \frac{n_{\rm i}\pi^2}{N_{\rm Dist} N_{\rm Shower} (N_\delta)^2 \Delta\Omega_\delta}. \tag{10.13}$$

If we then realize that $\epsilon = \frac{1}{N_\delta}$, we can obtain:

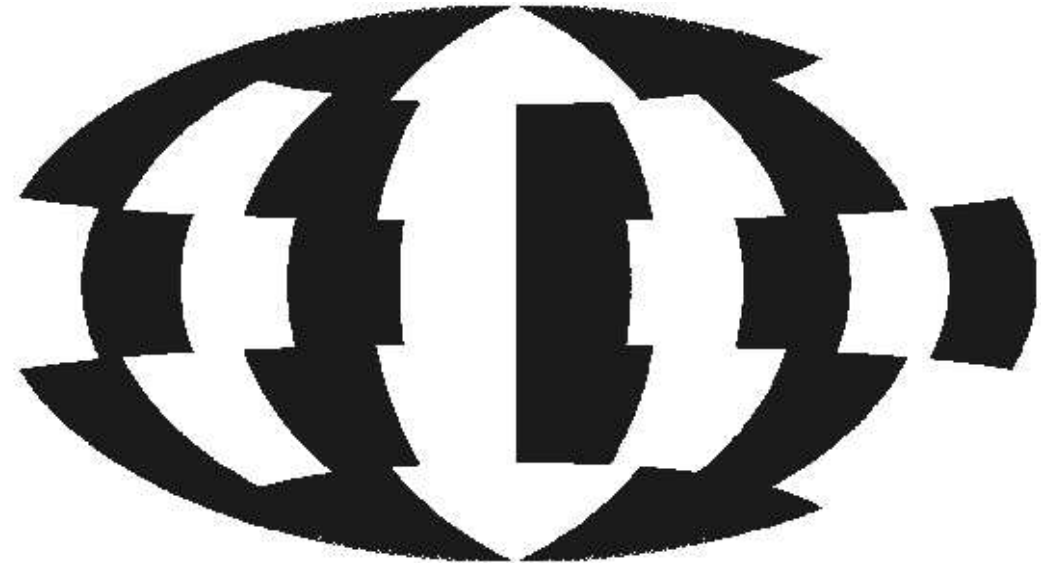
$$D_{\rm I}(N_\delta) = -\frac{1}{\log N_\delta} \sum_{i=1}^{N} P_{\rm i}(N_\delta) \log P_{\rm i}(N_\delta). \tag{10.14}$$

This expression is reminiscent of the the general formula for entropy from statistical mechanics:

$$S = -k \sum_r p_r \log p_r; \tag{10.15}$$

where p_r is the probability of a particle being the r-th state and k is the Boltzmann constant, which can be thought of as a scaling constant based upon the intrinsic scale of the given particle. The information dimension, $D_{\rm I}$, is an analogous measurement of the heterogeneity of a given data set.

(a)



(b)

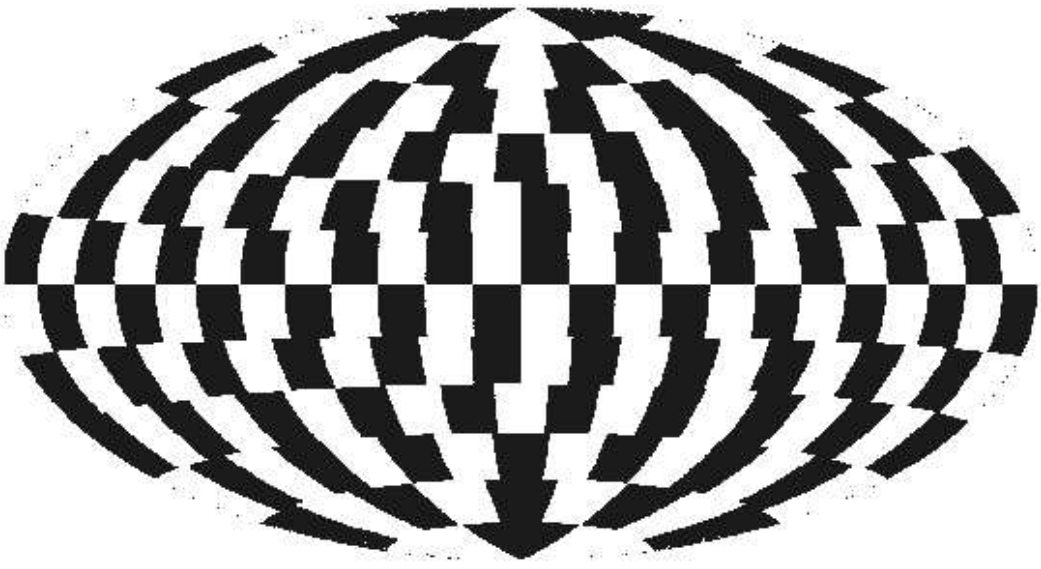
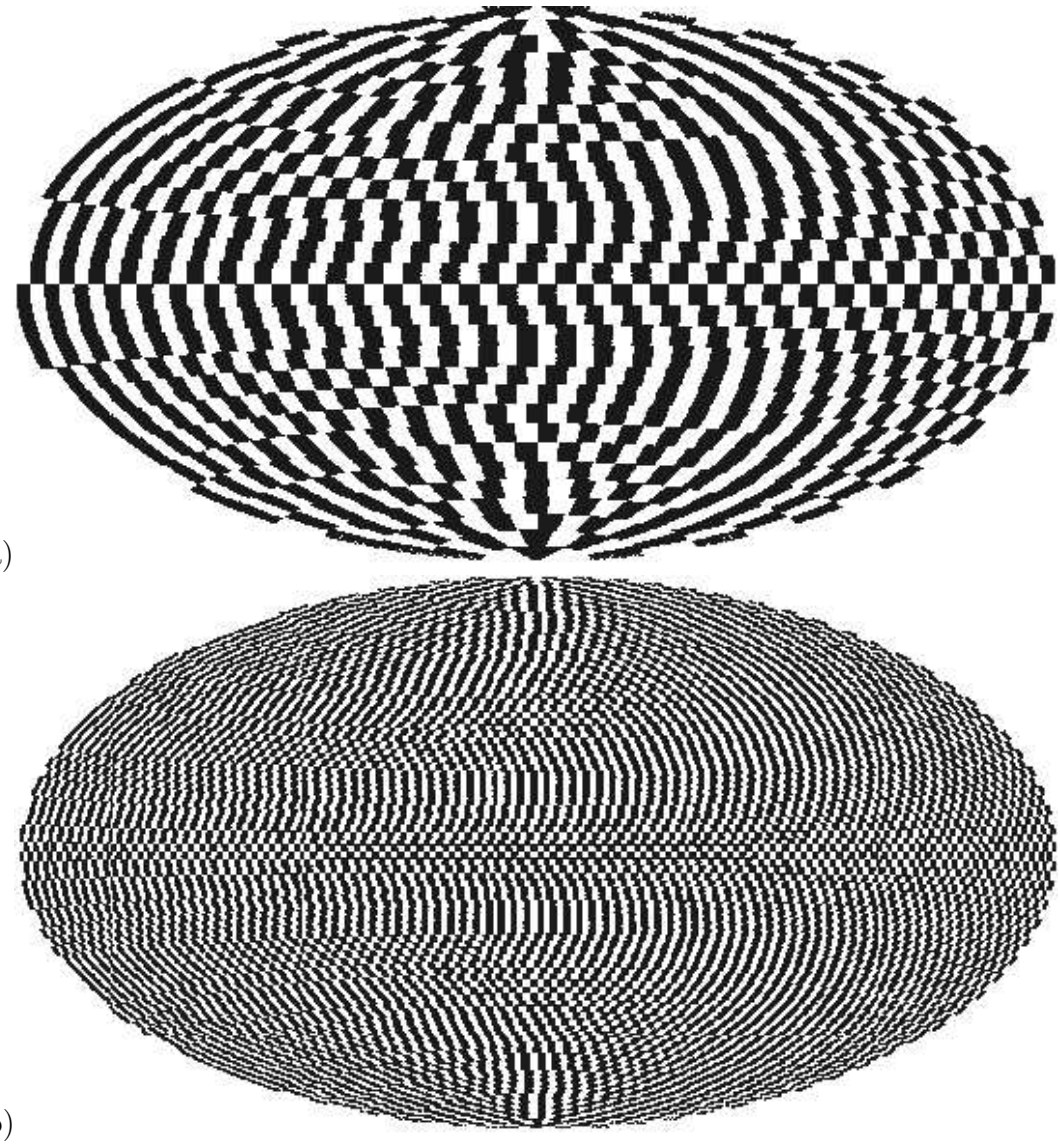


Figure 10.1. Hammer-Aitoff projection of latitudinal bins for different values of N_δ—(a): $N_\delta = 5$; (b): $N_\delta = 12$.



(a)

(b)

Figure 10.2. Hammer-Aitoff projection of latitudinal bins for different values of N_δ—(a): $N_\delta = 30$; (b): $N_\delta = 90$.

10.3 Calculating D_{I} for Exposures of Different Source Models

10.3.1 Exposure-Independent Source Descriptions

We began by examining four different source models independently of detector exposure. This allows us to calculate the value of the information dimension, D_{I}, without consideration to the detector aperture or statistical fluctuations from a finite event sample. The source models are: an isotropic source model, a dipole source model, a model with seven sources superimposed on an isotropic background, and a dark matter halo source model.

10.3.1.1 Isotropic Model

The first model that we will consider is an isotropic source model with distribution:

$$n_{\mathrm{isotropic}} = 1. \tag{10.16}$$

10.3.1.2 Dipole Model

The second is the Centaurus A dipole source model first proposed by Farrar and Piran [40]. This model has a distribution of arrival directions characterized by a scaling parameter, α, which can take on any value between -1 and $+1$ and by θ, which is the opening angle between a given event arrival direction and the center of the dipole distribution at Centaurus A. The overall distribution is:

$$n_{\mathrm{dipole}} = 1 + \alpha \cos\theta. \tag{10.17}$$

10.3.1.3 Discrete Source Model

The third model that we will consider is one with seven discrete sources superimposed on an isotropic background. For simplicity's sake, we will assume that all seven sources have an equal intensity indirectly determined by a parameter, F_s,

which will be defined as the fraction of the entire event sample which originates in the seven sources. We will define our source direction to be the centroids of the seven hypothetical point sources proposed by the AGASA collaboration [47]. The equatorial coordinates used for each point source are listed in Table 10.1

The arrival directions for the events from each source are assumed to be subjected to magnetic smearing. That is, in the course of traveling through space from the source to the point of observation, the velocity vector of the event is subject to bending in the galactic and extra-galactic magnetic fields. We assumed that this bending produces an apparent source that can be characterized by a Gaussian distribution:

$$P(\Delta\theta) = \frac{\Delta\theta}{\lambda^2} e^{-\frac{(\Delta\theta)^2}{2\lambda^2}}, \tag{10.18}$$

where $P(\Delta\theta)$ is the probability that an event will be observed with an opening angle, $\Delta\theta$, from the nominal direction of the apparent source. We will also assume that the arrival directions of the events from all the sources are subject to the same degree of magnetic smearing, parameterized by $\lambda_\circ = 1.5105 \cdot \lambda$. The parameter, $\lambda_\circ$, corresponds to the 68% confindence interval in $\Delta\theta$. For this simulation study, we will assume that $\lambda_\circ = 5°$. It should be noted that the nominal direction of the apparent source is not necessarily the direction of the *actual* source because the possibility exists that the path of the events in question traveled through large

Table 10.1. Coordinates used for the centers of seven discrete sources. These coordinates correspond to the centers of the seven clusters reported by the AGASA Collaboration [47].

Cluster	*Right Ascension*	*Declination*
C1	01h13m	20.6°
C2	11h17m	56.9°
C3	18h51m	48.2°
C4	04h38m	30.0°
C5	16h02m	23.3°
C6	14h11m	37.4°
C7	03h03m	55.5°

regions of homogeneous magnetic fields.

10.3.1.4 Dark Matter Halo Model

The fourth model that we will consider is a dark matter halo source model. Dark matter halos are characterized by a density profile that is assumed to take the Navarro-Frenk-White (NFW) form [70]:

$$\rho_{\mathrm{NFW}} = \frac{\rho_\circ}{r(1 + r/r_{\mathrm{s}})^2}, \tag{10.19}$$

where $\rho_\circ$ is a dark matter density parameter, r is the distance from the center of the halo, and r_{s} is a critical radius. For our source model, we will consider the contribution of only the four closest significant dark matter halos: the Milky Way, M31, LMC, and M33. We will assume that $\rho_\circ$ is the same for all four sources and that r_{s} scales with the cube root of the luminosity, $L^{\frac{1}{3}}$. Thus the Milky Way will have: $r_{\mathrm{s,MW}} = 10.0$ kpc, LMC will have $r_{\mathrm{s,LMC}} = 0.3 \cdot r_{\mathrm{s,MW}} = 3.0$ kpc, M31 will have $r_{\mathrm{s,M31}} = 1.5 \cdot r_{\mathrm{s,MW}} = 15.0$ kpc, and M33 will have $r_{\mathrm{s,M33}} = 0.4 \cdot r_{\mathrm{s,MW}} = 4.0$ kpc.

We now calculate the information dimension, D_{I}, for each of the our four models. Since these are smooth distributions with no statistical fluctuations, we will only use one, computationally limited value for the number of declinational bands for each model of $N_\delta = 1080$ (i.e. $\Delta\theta = \frac{1}{6}^\circ$), which implies:

$$P_{\mathrm{i}} = n_{\mathrm{i}} \left[\sum_{\mathrm{i}} n_{\mathrm{i}} \right]^{-1}. \tag{10.20}$$

Using equations 10.14 and 10.20 we can now calculate D_{I} for each of the four models. The results are in Table 10.2 column 1. Note that three of the four values of D_{I} exceed the analytical limit of 2 for a 2-D surface.

If we consider the analytic limit for the isotropic case, we have:

$$D_{\mathrm{I}} = -4 \frac{(N_\delta)^2}{\pi} \frac{P_{\mathrm{i}} \log P_{\mathrm{i}}}{\log N_\delta}. \tag{10.21}$$

If we then substitute in equation 10.20 we get:

$$D_{\mathrm{I}} = \frac{2 \log N_\delta + \log 4/\pi}{\log N_\delta}; \lim_{N_\delta \to \infty} D_{\mathrm{I}}(N_\delta) = 2. \tag{10.22}$$

The reason the we obtain values greater than 2 is because we are working with a finite number of elements on a surface where the total area does not equal $(N_\delta)^2$.

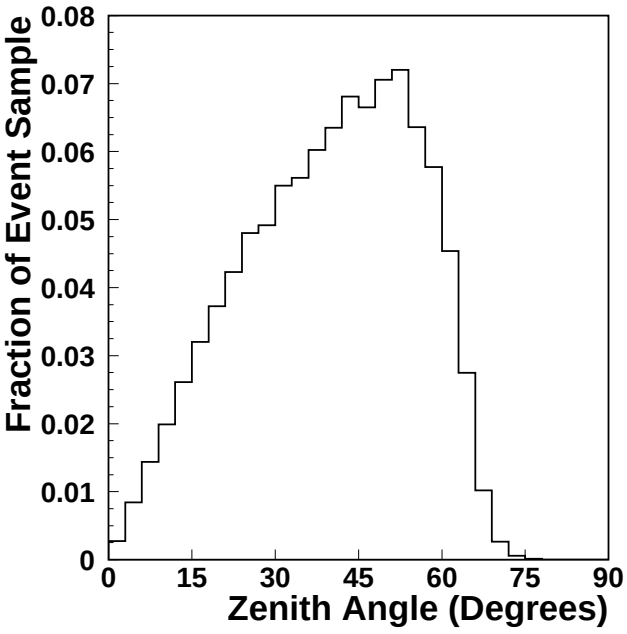
Table 10.2. The estimated values of D_{I} with $N_{\delta} = 1800$ for the four proposed source models for the entire sky independent of any real detector's aperture (1) and for the four source models superimposed on the estimated aperture of an air-fluorescence detector at 40° N. These values are mathematical descriptors of a data set that are whose number of significant digits are determined by how extensively each bin is sampled (in this case four digits). In the case of a real data set with a finite number of the observations, the number of significant digits is constrained by the fluctuations inherent to the sample size.

	1	**2**
SOURCE MODEL	D_{I} for Source Model **without** Detector Exposure	D_{I} for Source Model **with** Detector Exposure
Isotropic	2.035	1.967
Dipole Enhancement	2.007	1.945
Seven Source	2.033	1.946
Dark Matter Halo Model	1.999	1.978

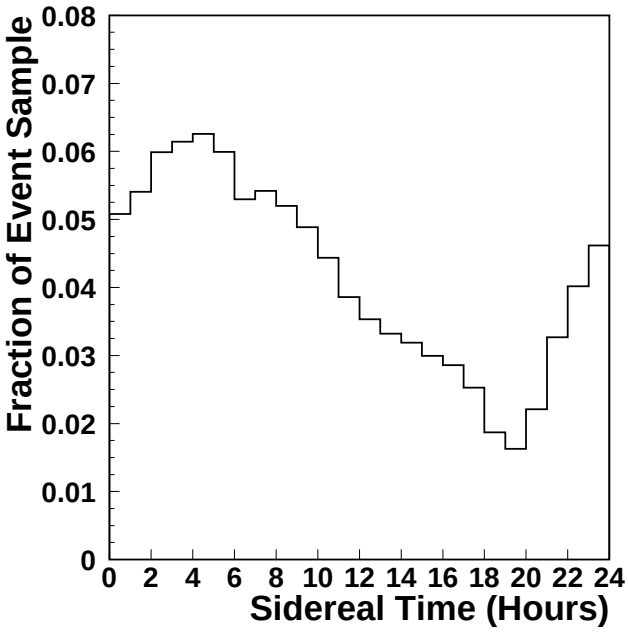
10.3.2 Exposure-Dependent Source Descriptions

For the purpose of this simulation study, we assume a hypothetical air-fluores[-]cence detector located at 40° N. This analysis assumes an isotropic distribution for the azimuthal component of the arrival directions and a zenith angle distribution that remains constant in time. This is what one would expect for a detector with 360° coverage with identical detector units and stable atmospheric conditions. The acceptance of our detector can thus be defined by two distributions: zenith angle and sidereal time which are shown in Figure 10.3. The zenith angle distribution is characterized by 100% acceptance until $\sim 50°$, at which point it drops off dramatically due to the lack of a well-defined profile to assist monocular reconstruction. The sidereal time distribution is the combination of the seasonal availability of dark, moonless sky at 40° N and seasonal climatic changes in a desert locale (the rainy season is assumed to extend from February to May which results in a loss of $\sim 30\%$ of exposure).

By defining acceptance this way, we can calculate the exposure of the detector. By also considering the finite, asymmetric angular resolution, we can then super-



(a)



(b)

Figure 10.3. The exposure characteristics for a monocular air fluorescence detector–(a): The distribution of zenith angles; (b): the distribution of sidereal times for a detector located at 40° N in a desert locale.

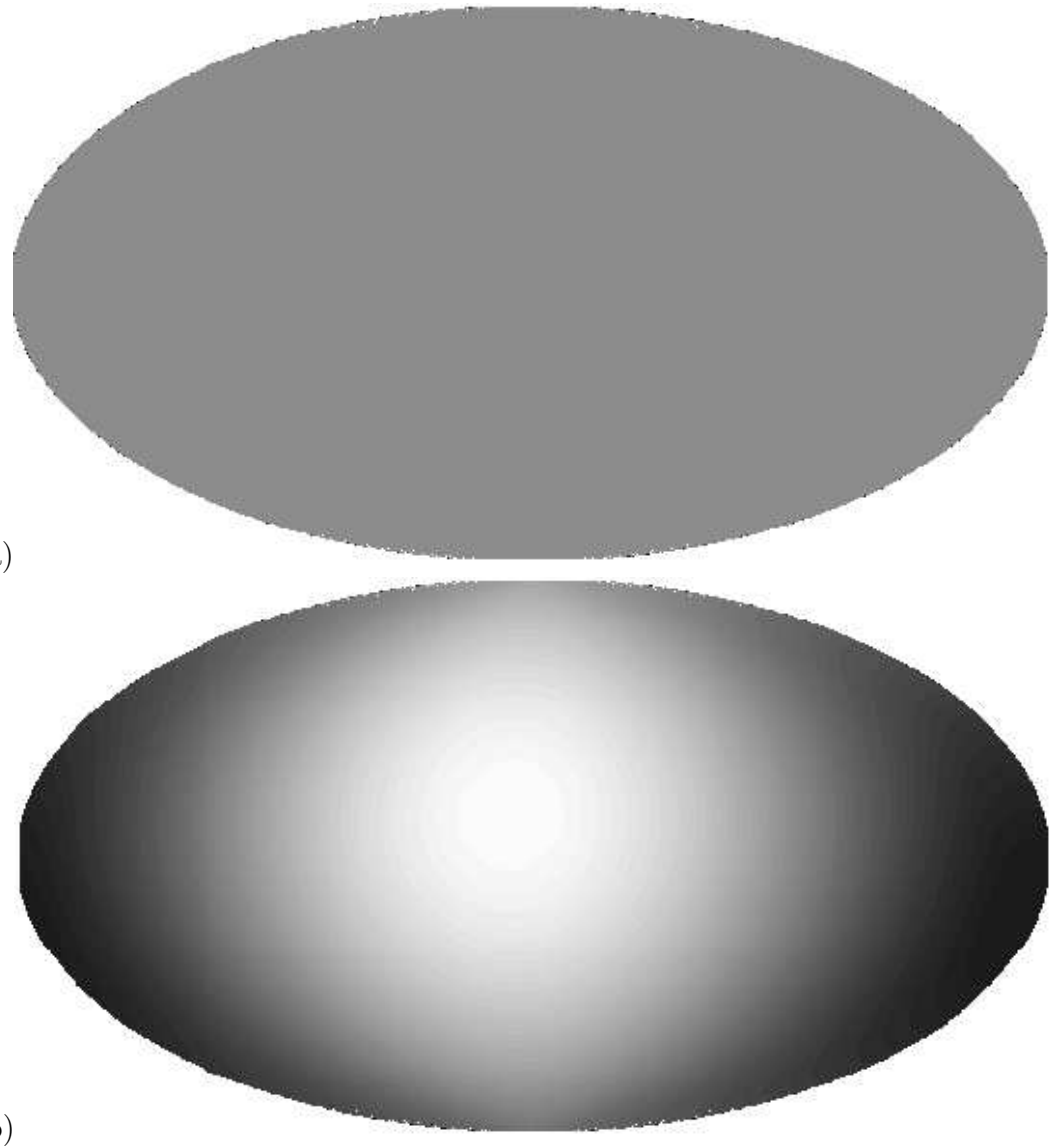
impose the detector exposure upon the various source models that we previously examined (Figures 10.4 and 10.5 and Table 10.2 column 1). We then obtain an effective detector response for the air fluorescence detector for each of our four source models. The results are shown in Figures 10.6 and 10.7. It should be emphasized that effective detector effect is due to the *combined* effect of asymmetric sky coverage and angular resolution smearing. A remarkable consequence of this combination is the possibility that point sources can take on an apparently asymmetric shapes due to preferential orientations of the plane of reconstruction for events arriving from a specific location in the physical sky.

We can now determine the value of D_{I} using the same method as before. The results are shown in Table 10.2 column 2. It is interesting to note that the dark matter halo source model now has a larger value for D_{I} than the isotropic source model. By looking at Figures 10.6 and 10.7, one can verify that the superposition of the detector exposure and source models actually yields a more uniform apparent distribution for the dark matter halo source model than it does for the isotropic source model.

10.4 Calculating D_{I} for Finite Event Samples

So far, we have only considered calculating D_{I} for smooth distributions. From an experimental standpoint, it is very difficult to collect enough data to obtain a smooth distribution. This is especially true for UHECRs. In order to make a measurement of D_{I}, we must first determine what value(s) we should assign to N_δ. A reasonable approach is to assign a scale length to our sample. Choosing $\Delta\theta = 0.5^\circ$, which approximately reflects the lowest value that can be obtained from σ_2 in equation 10.6, yields $N_\delta = 360$.

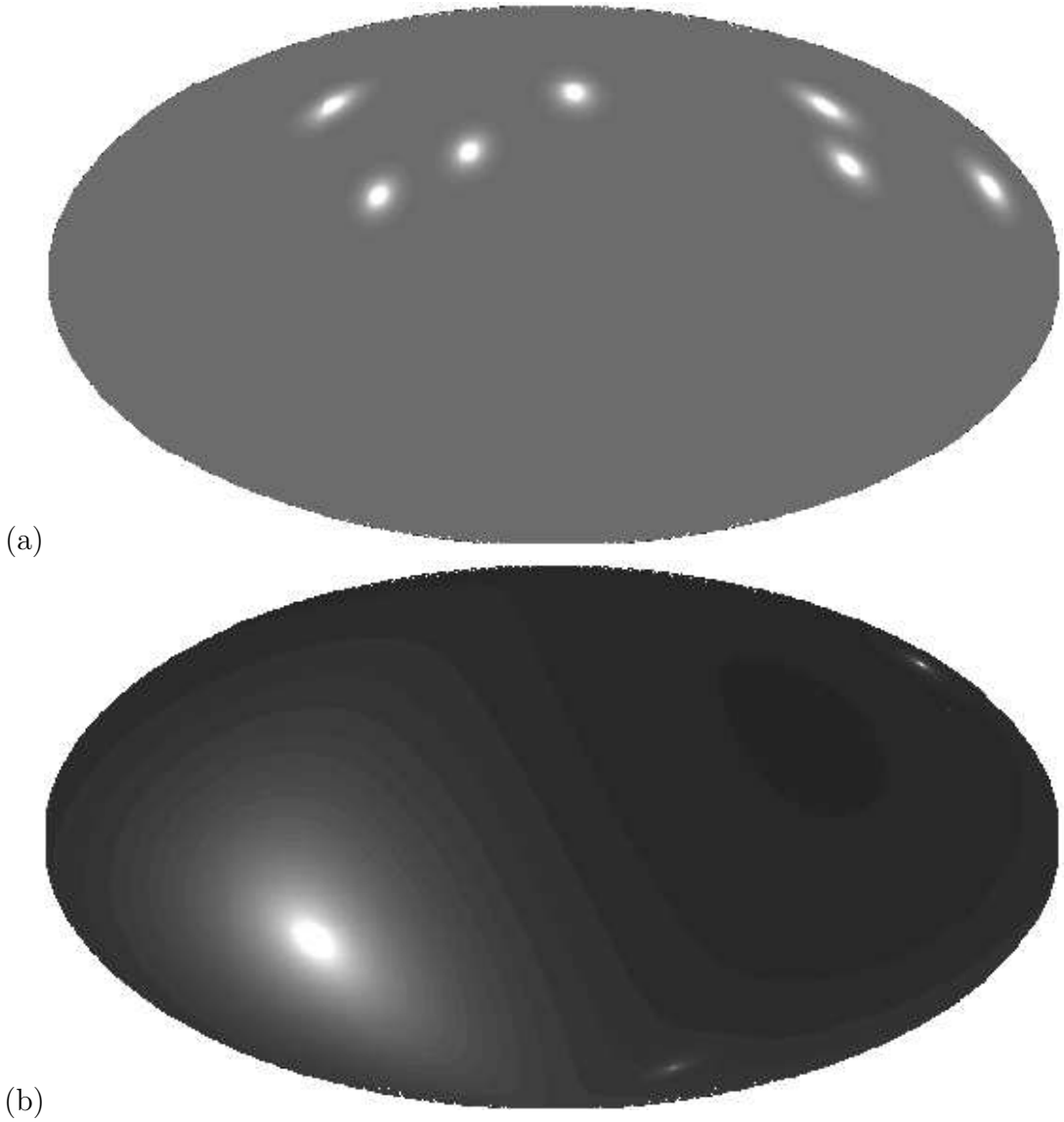
However, it can be beneficial to actually calculate D_{I} for a range of values around N_δ. In Figure 10.8a, we display the values $D_{\mathrm{I}}(N_\delta)$ over the range $N_\delta = [354, 360]$ for two separate simulated sets where $N_{\mathrm{Shower}} = 500$. These sets yield very similar values for D_{I} over the full range of values for N_δ. However, if we examine Figure 10.8b we can see that for the same two finite samples, the fractional



(a)

(b)

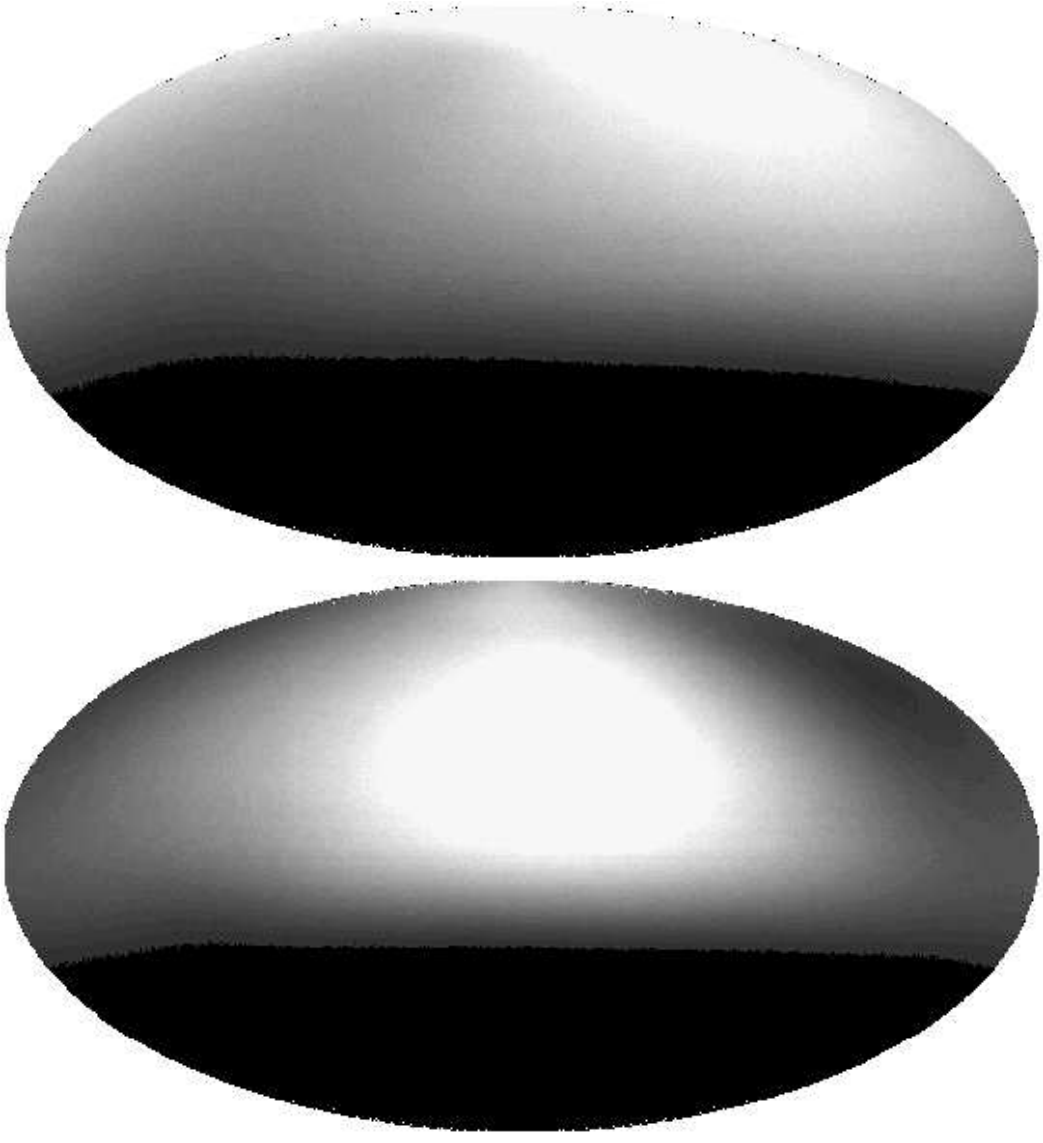
Figure 10.4. Density profiles for different source models—(a): isotropic model; (b): dipole enhancement model ($\alpha = 1.0$). All figures are shown in a Hammer-Aitoff projection of equatorial coordinates (right ascension right to left). The highest density in each panel corresponds to the lightest regions, the lowest density to the darkest regions.



(a)

(b)

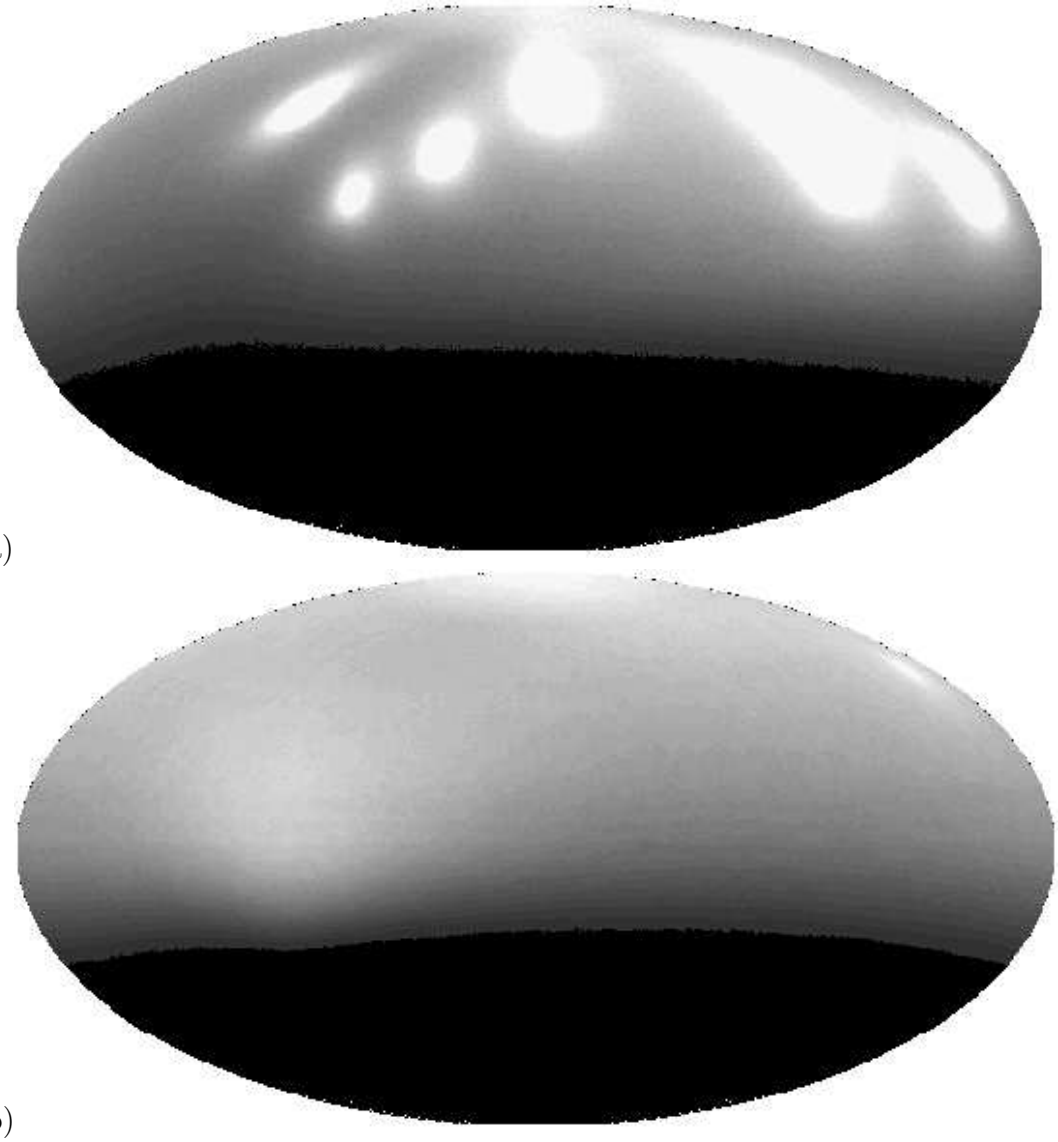
Figure 10.5. Density profiles for different source models—(a): seven source model ($F_s = 0.28$); (b): dark matter halo model ($r_s = 10$ kpc). All figures are shown in a Hammer-Aitoff projection of equatorial coordinates (right ascension right to left). The highest density in each panel corresponds to the lightest regions, the lowest density to the darkest regions.



(a)

(b)

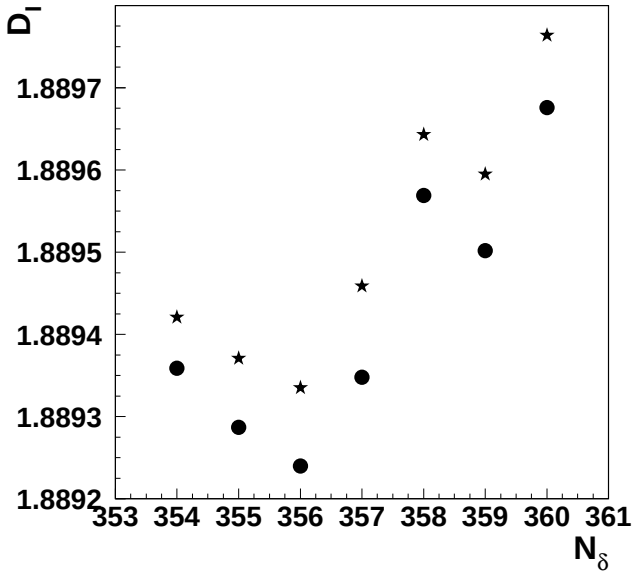
Figure 10.6. Effective detector response for the different source models—(a): isotropic model; (b): dipole enhancement model ($\alpha = 1.0$); All figures are shown in a Hammer-Aitoff projection of equatorial coordinates (right ascension right to left). The highest density in each panel corresponds to the lightest regions, the lowest density to the darkest regions.



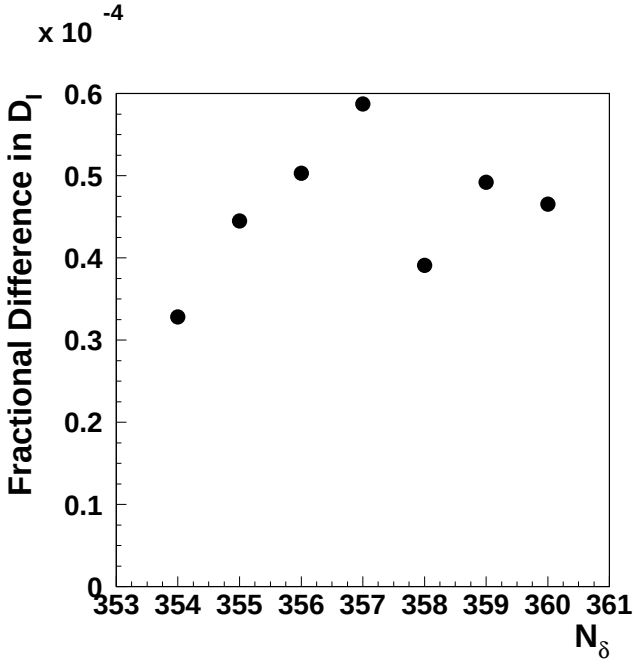
(a)

(b)

Figure 10.7. Effective detector response for the different source models—(a): seven source model ($F_s = 0.28$); (b): dark matter halo model ($r_s = 10$ kpc). All figures are shown in a Hammer-Aitoff projection of equatorial coordinates (right ascension right to left). The highest density in each panel corresponds to the lightest regions, the lowest density to the darkest regions.



(a)



(b)

Figure 10.8. The variation of $D_{\rm I}$ with respect to N_δ—(a): $D_{\rm I}$ over a range of values of N_δ for two similar finite event sets that have similar values for $D_{\rm I}$. (the dots indicate the values $D_{\rm I}$ for the first set while the stars indicate the values $D_{\rm I}$ for the second set); (b): the fractional difference in $D_{\rm I}$ for the same two sets over the range of N_δ values.

difference between values of $D_{\rm I}$ can fluctuate substantially even over small intervals of N_δ. While these fluctuations are typically much smaller than the difference in $D_{\rm I}$ values between any two sets, we will take $D_{\rm I}$ to be $<D_{\rm I}(N_\delta)>$ for the interval: $N_\delta = [354, 360]$ in order to minimize the statistical fluctuations in $D_{\rm I}$ between individual sets. This range was chosen to optimize our computational ability.

To account for the fact that we no longer have a smooth distribution for the calculation of the values of $P_{\rm i}$, we refer back to equation 10.13 to see how to calculate $D_{\rm I}$ for a finite set of elements. This only requires us to determine a value for $N_{\rm Dist}$. We can set this value based upon what value we wish for $n_{\rm i}$ in combination with equation 10.12 (i.e. $<n_{\rm i}>= \frac{1}{(\Delta n)^2}$ where Δn is the fractional Gaussian fluctuation of a bin with $n_{\rm i} =<n_{\rm i}>$). Then,

$$N_{\rm Dist} = \frac{\frac{4}{\pi}(N_\delta)^2 <n_{\rm i}>}{N_{\rm Shower}}. \tag{10.23}$$

If we then combine equations 10.9, 10.13, and 10.23; we obtain:

$$P_{\rm i}(N_\delta) = \frac{n_{\rm i}}{<n_{\rm i}>}\frac{\pi^3}{4(N_\delta)^4\Delta\Omega_\delta}. \tag{10.24}$$

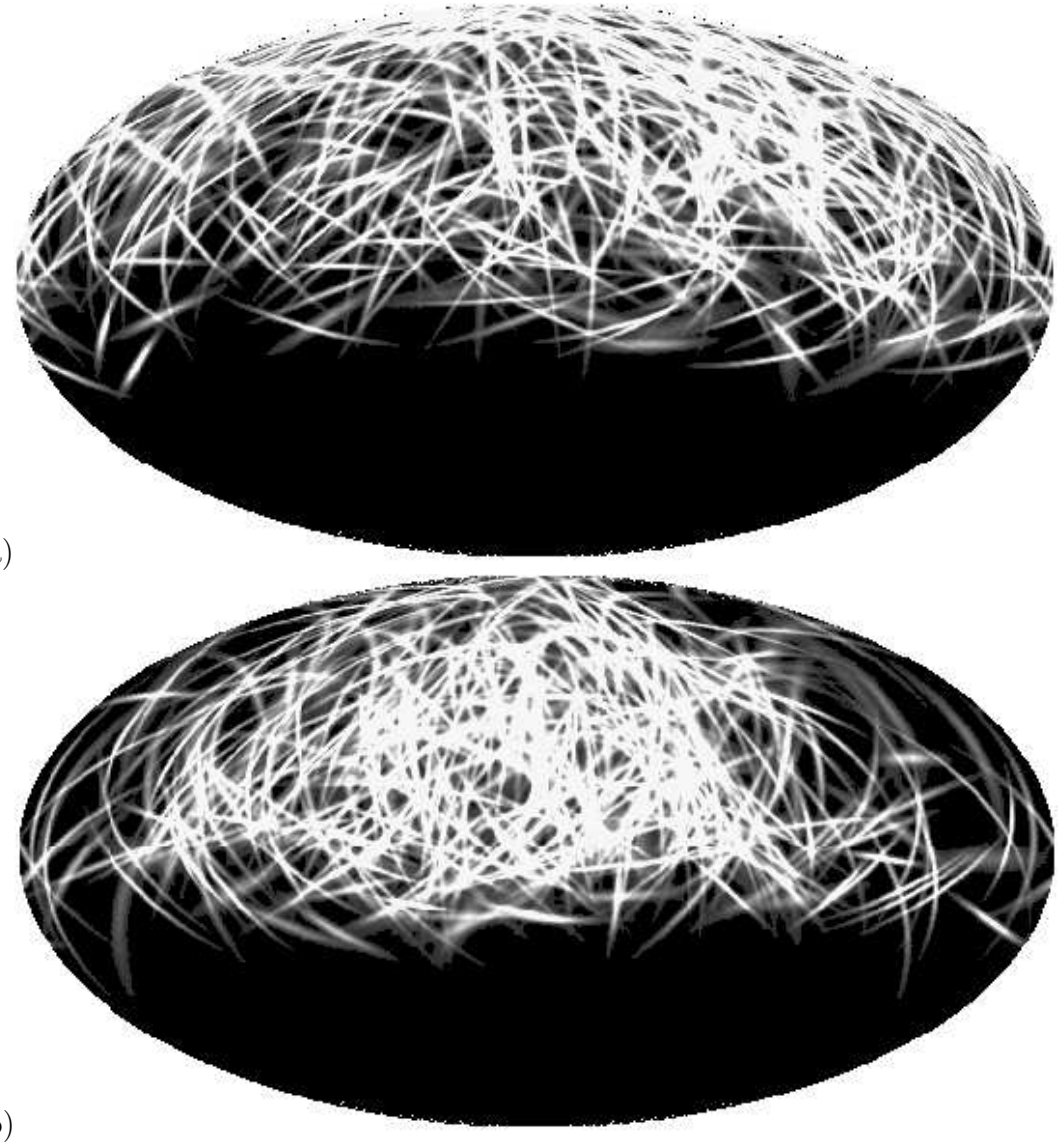
We can then calculate $D_{\rm I}$ from equation 10.14:

$$D_{\rm I} = \left\langle -\frac{1}{\log N_\delta}\sum_{i=1}^{N} P_{\rm i}(N_\delta)\log P_{\rm i}(N_\delta)\right\rangle,\ N_\delta = [354, 360]. \tag{10.25}$$

Thus, If we want $<n_{\rm i}>= 500$ and $N_{\rm Shower} = 500$, we find that $N_{\rm Dist} \simeq 1.65\times10^5$. If $N_{\rm Shower} = 2000$, we have $N_{\rm Dist} \simeq 4.1 \times 10^4$.

We now consider two cases: finite event sets with 500 events and finite event sets with 2000 events. These sets will have the angular resolution characteristics described in equations 10.5 and 10.6. The exposure will be modeled via the zenith angle and sidereal time distributions shown in Figure 10.3. Figures 10.9 and 10.10 contains examples of event sets with all four source models and $N_{\rm Shower} = 500$ and Figures 10.11 and 10.12 contains examples of events sets with all four previously described source models and $N_{\rm Shower} = 2000$.

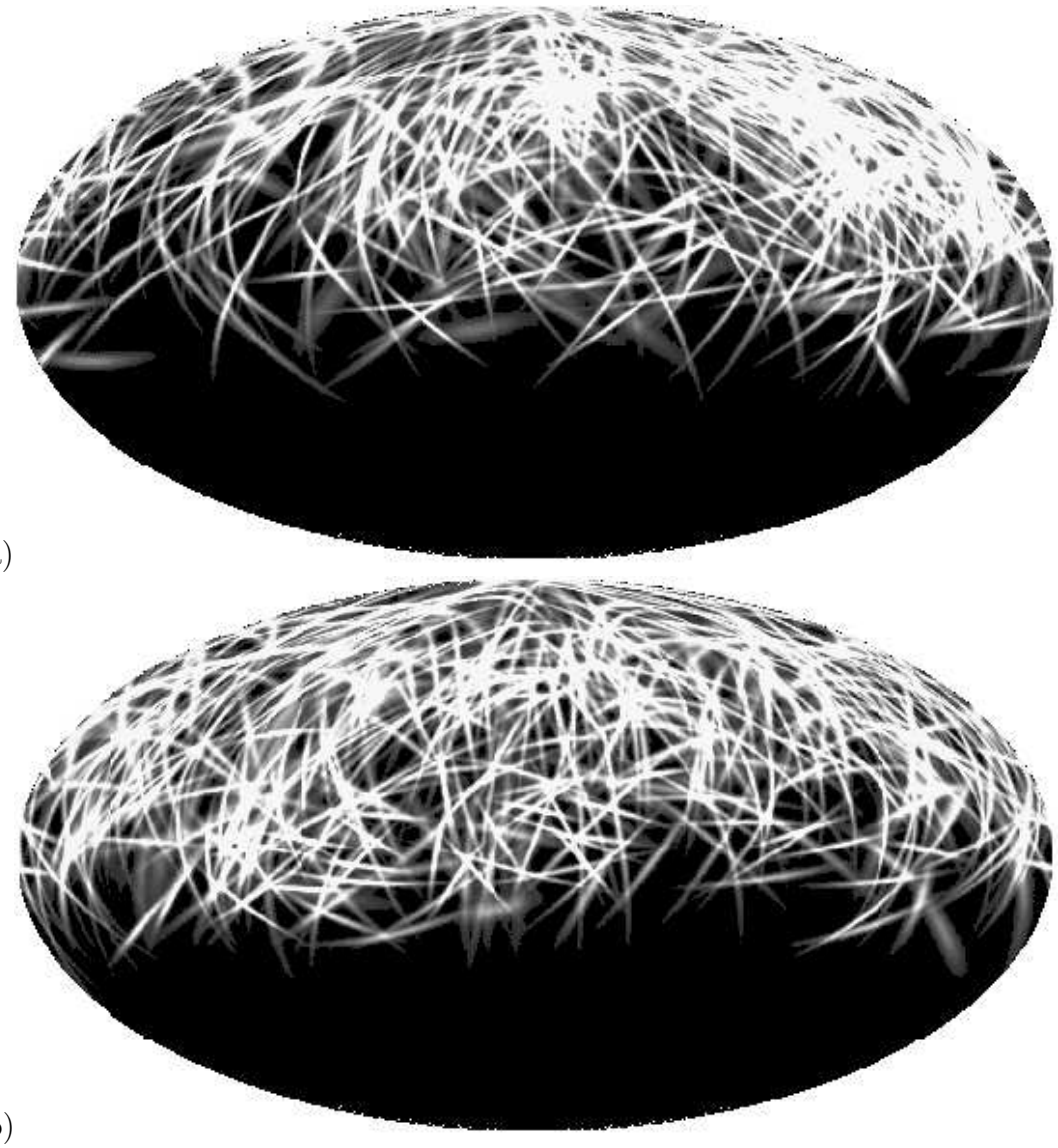
In Figures 10.9, 10.10, 10.11, and 10.12, one can see the that these distributions of arrival directions have a far greater degree of statistical fluctuation than the



(a)

(b)

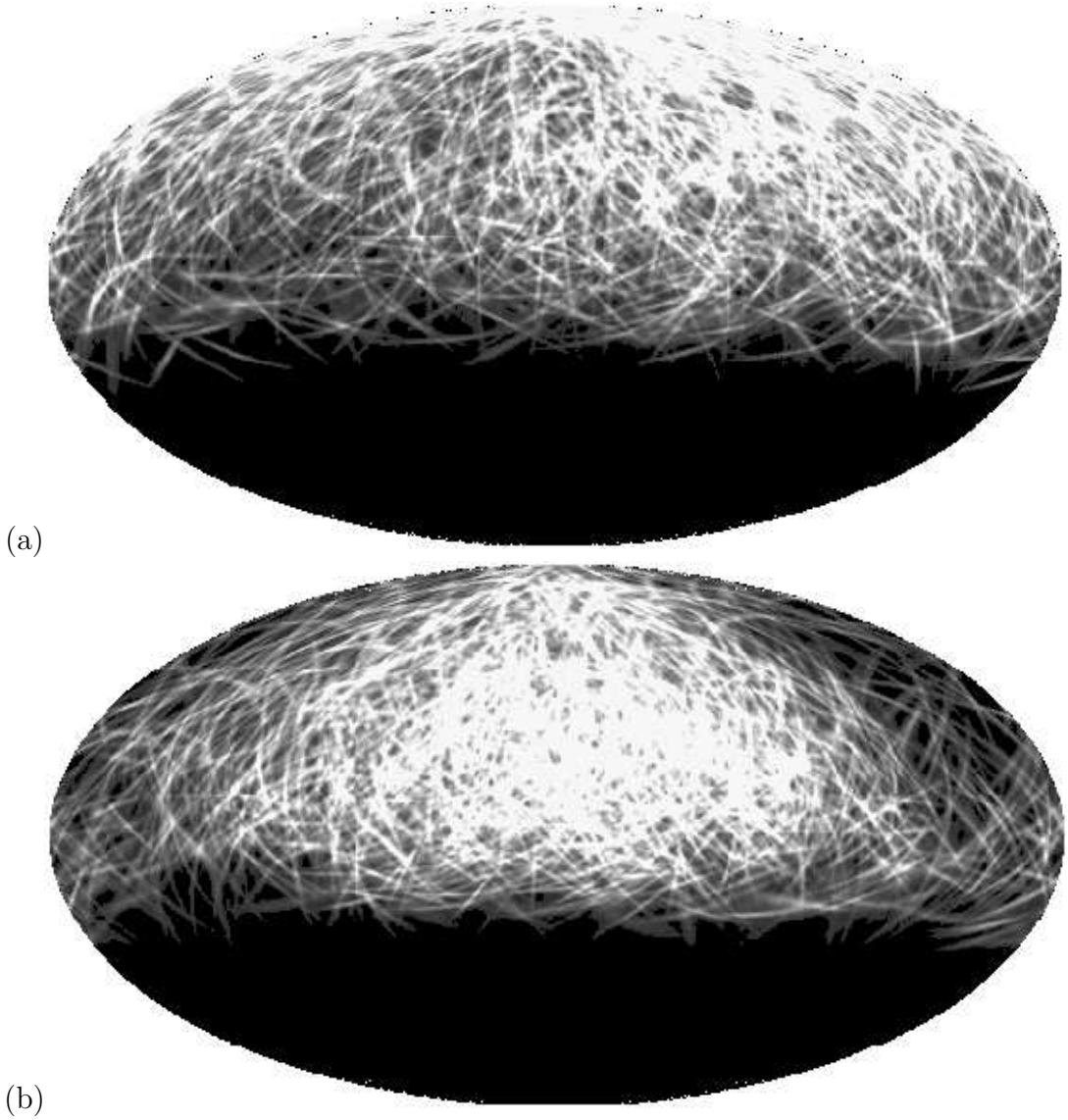
Figure 10.9. Simulated 500 event distributions for different source models — (a): Isotropic model; (b): dipole enhancement model ($\alpha = 1.0$). All figures are shown in a Hammer-Aitoff projection of equatorial coordinates (right ascension right to left). The highest density in each panel corresponds to the lightest regions, the lowest density to the darkest regions.



(a)

(b)

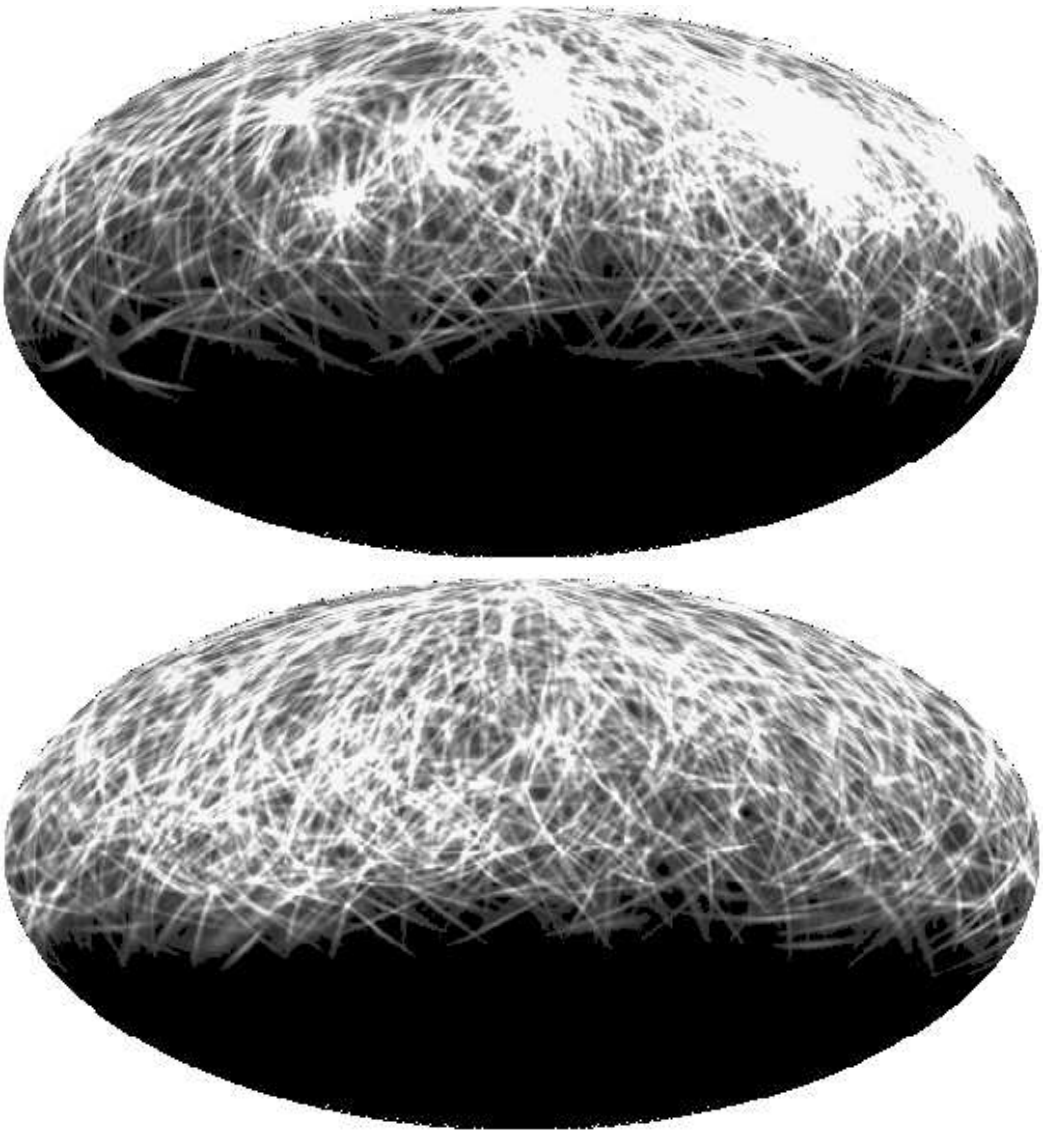
Figure 10.10. Simulated 500 event distributions for different source models — (a): Seven source model ($F_s = 0.28$); (b): dark matter halo model ($r_s = 10$ kpc). All figures are shown in a Hammer-Aitoff projection of equatorial coordinates (right ascension right to left). The highest density in each panel corresponds to the lightest regions, the lowest density to the darkest regions.



(a)

(b)

Figure 10.11. Simulated 2000 event distributions for different source models — (a): Isotropic model; (b): dipole enhancement model ($\alpha = 1.0$). All figures are shown in a Hammer-Aitoff projection of equatorial coordinates (right ascension right to left). The highest density in each panel corresponds to the lightest regions, the lowest density to the darkest regions.



(a)

(b)

Figure 10.12. Simulated 2000 event distributions for different source models — (a): seven source model ($F_s = 0.28$); (b): dark matter halo model ($r_s = 10$ kpc). All figures are shown in a Hammer-Aitoff projection of equatorial coordinates (right ascension right to left). The highest density in each panel corresponds to the lightest regions, the lowest density to the darkest regions.

smooth distributions shown in Figures 10.6 and 10.7. Because of the fluctuations in our simulated event samples, the value of $D_{\rm I}$ varies significantly (see Figures 10.13 and 10.14) from one simulated set to the next. In Figures 10.13 and 10.14, we examine the distribution of $D_{\rm I}$ values for $\sim$ 500 sets (of 500 and 2000 events respectively). We see that the distribution of 500 event samples have both a lower mean value and larger width than the 2000 event samples.

10.5 Application to Anisotropy Analysis

We now need to develop a scheme by which we can apply fractal dimensionality analysis to a real data set. In the case of a real data set, we will be dealing with only *one* value of $D_{\rm I}$. By itself $D_{\rm I}$ is insufficient to characterize the data set; $D_{\rm I}$ fluctuates a great deal due to variation in $N_{\rm Shower}$. However, a comparison between the value of $D_{\rm I}$ for a real event sample and a distribution of $D_{\rm I}$ values (with the *same* $N_{\rm Shower}$ and $N_{\rm Dist}$ values as the real data) for a series of simulated data sets of a given source model does provide a viable measurement of anisotropy.

We can demonstrate this by considering a *single* simulated event sample generated with an isotropic source model. We will suppose this sample to be our "real" data. We consider the isotropic simulated event samples shown in Figures 10.9a and 10.11a. We will once again stipulate that $<n_i>= 500$ which means that in the case of $N_{Shower} = 500$, we have: $N_{\rm Dist} \simeq 1.6 \times 10^5$ which leads to $D_{\rm I} = 1.89715$ and in the case of $N_{Shower} = 2000$, we have: $N_{\rm Dist} \simeq 4 \times 10^4$ which leads to $D_{\rm I} = 1.93920$.

In Figures 10.13 and 10.14, we demonstrated that for a fixed scaling parameter the distribution of $D_{\rm I}$ values for a large number of simulated sets fits well to a Gaussian curve. By establishing the relationship between the distribution of $D_{\rm I}$ values and the scaling parameter in each model, we can establish a 90% confidence interval on the scaling parameter for that model.

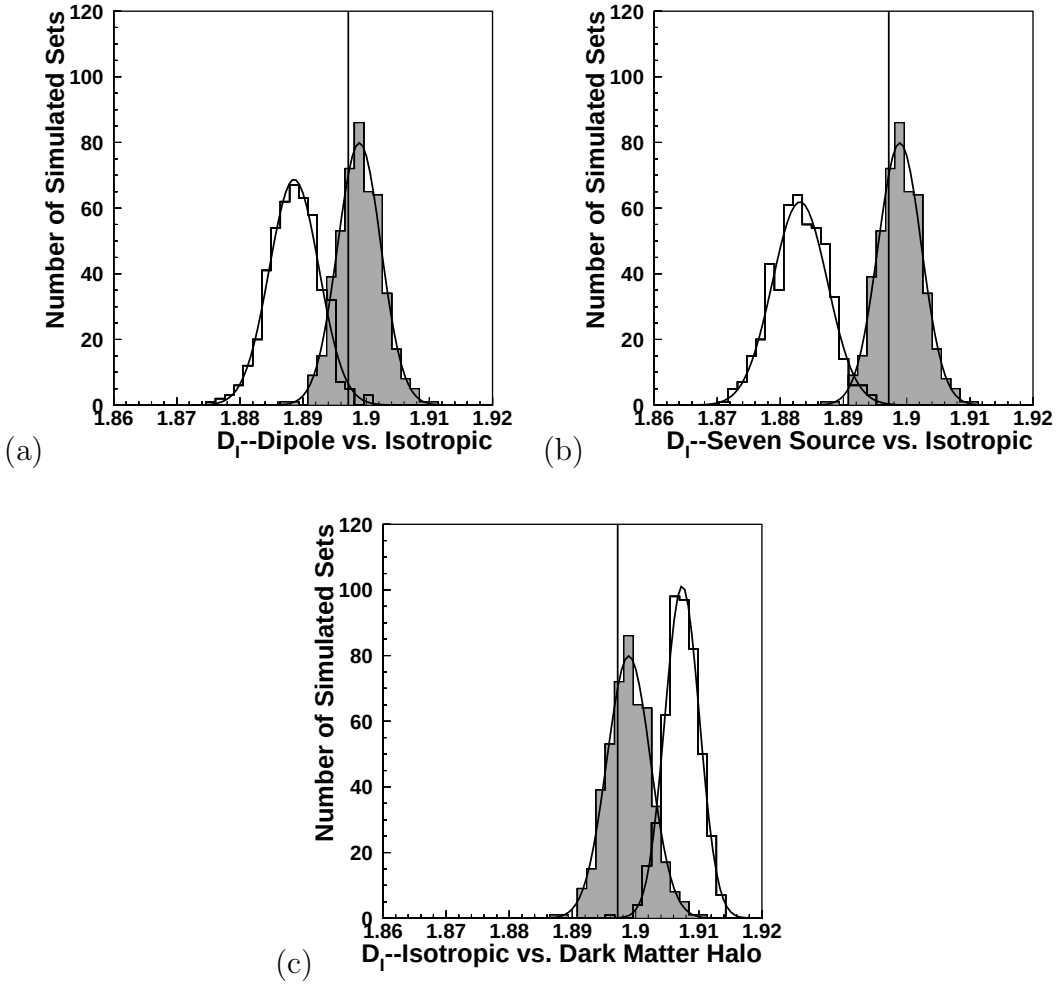


Figure 10.13. Comparison of distributions of $D_{\rm I}$ values for ~ 500 sets of 500 events between the isotopic source model (shaded) and the other three source models — (a): Dipole enhancement model vs. isotropic model ($\alpha = 1.0$); (b): Seven source model vs. isotropic model ($F_s = 0.28$); (c): Isotropic model vs. Dark matter halo source model ($r_s = 10$ kpc). The vertical line corresponds to the value of $D_{\rm I}$ for the "real" 500 event sample. In all cases the distributions of $D_{\rm I}$ values fit well to a Gaussian curve.

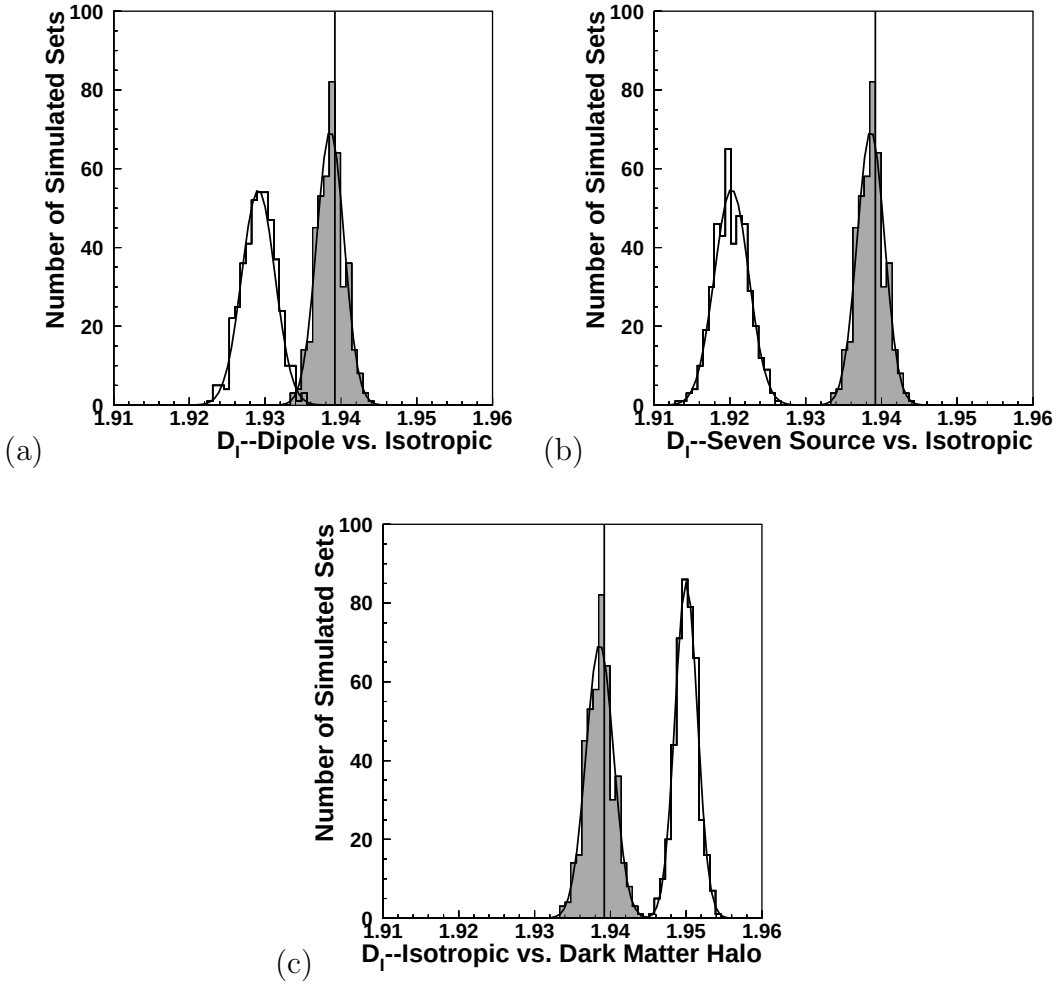


Figure 10.14. Comparison of distributions of $D_{\rm I}$ values for $\sim$ 500 sets of 2000 events between the isotopic source model (shaded) and the other three source models — (a): Dipole enhancement model vs. Isotropic model ($\alpha = 1.0$); (b): Seven source model vs. Isotropic model ($F_s = 0.28$); (c): Isotropic model vs. Dark matter halo source model ($r_s = 10$ kpc). The vertical line corresponds to the value of $D_{\rm I}$ for the "real" 2000 event sample. In all cases the distributions of $D_{\rm I}$ values fit well to a Gaussian curve.

10.5.1 Dipole Enhancement Source Model

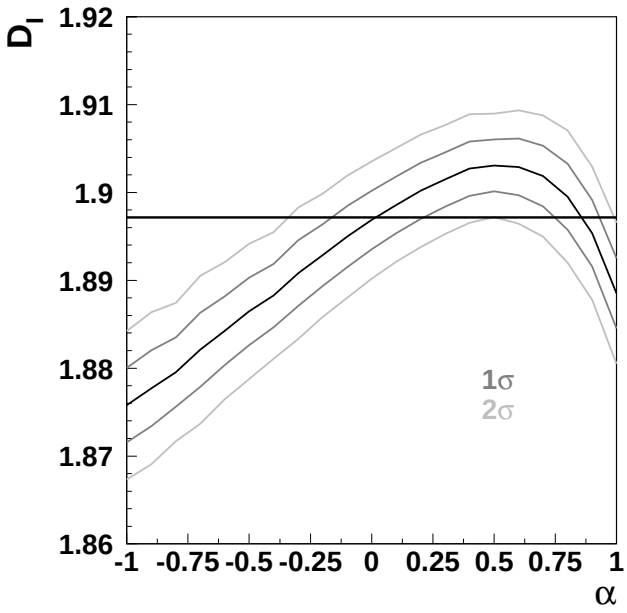
In the case of the dipole enhancement source model in equation 10.17, the scaling parameter is α. By varying α between -1 and 1, we develop a curve which will show the relationship between D_{I} and α. By considering the actual value of D_{I} for the "real" data set, we then establish a nominal value for α and a 90% confidence interval. The results for both $N_{Shower} = 500$ and $N_{Shower} = 2000$ are shown in Figures 10.15 and 10.16. In the case of our simulated isotropic set with $N_{Shower} = 500$, $\alpha = 0.02 \pm 0.21$ with a 90% confidence interval of $[-0.29, 0.36]$. In the case of our simulated isotropic set with $N_{Shower} = 2000$, $\alpha = 0.075 \pm 0.085$ with a 90% confidence interval of $[-0.065, 0.24]$. We notice that in Figures 10.15 and 10.16, there are potentially two suitable intervals of α that possess similar values of D_{I}. This emphasizes the necessity of making density plots to assure by visual inspection that the appropriate interval is chosen.

10.5.2 Seven Source Model

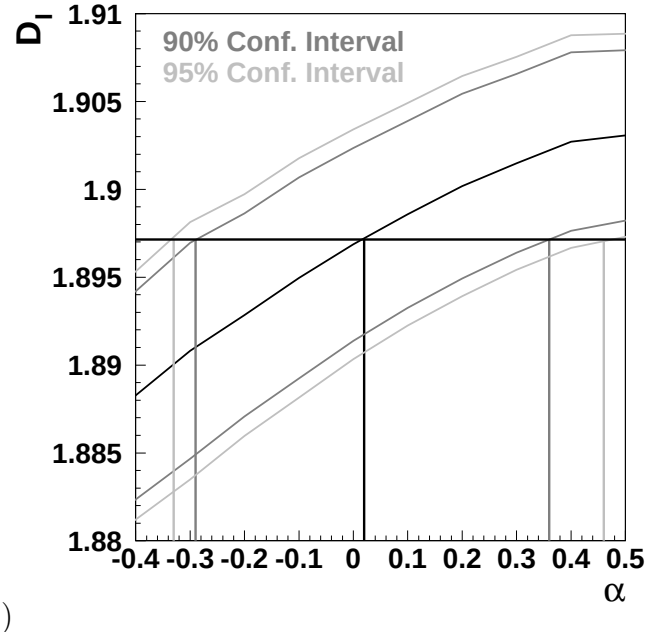
In the case of the seven source model, the scaling parameter is F_s, the fraction of the total event sample that is produced by the discrete sources. By varying F_s between 0 and 0.40, we develop a curve which shows the relationship between D_{I} and F_s. By considering the actual value of D_{I} for the "real" data set, we then establish a nominal value for F_s and a 90% confidence upper limit for F_s. The results for both $N_{Shower} = 500$ and $N_{Shower} = 2000$ are shown in Figures 10.17 and 10.18. In the case of our simulated isotropic set with $N_{Shower} = 500$, we have a 90% confidence upper limit of $F_s = 0.16$. In the case of our simulated isotropic set with $N_{Shower} = 2000$, we have a 90% confidence upper limit of $F_s = 0.04$. This tells us that our "real" data can have, at most, 16% (for 500 events) or 4% (for 2000 events) of these events coming from the seven sources.

10.5.3 Dark Matter Halo Source Model

In the case of the dark matter halo source model in equation 10.19, the variable parameter is r_{s}, the critical radius in the NFW profile [70]. By varying r_{s} between

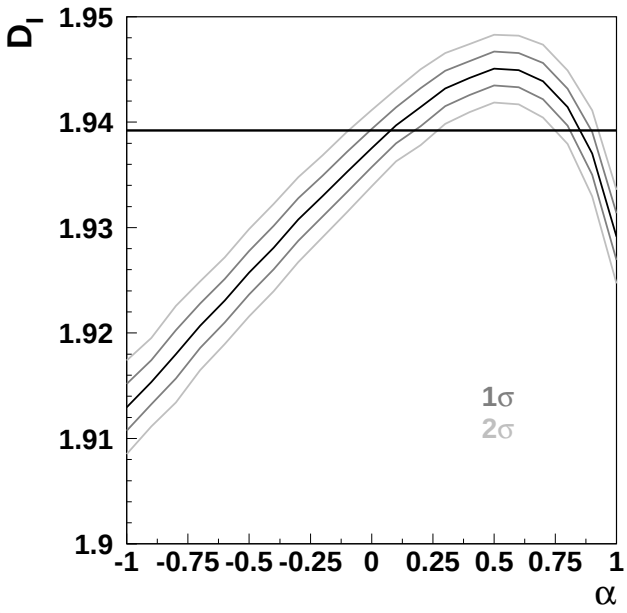


(a)

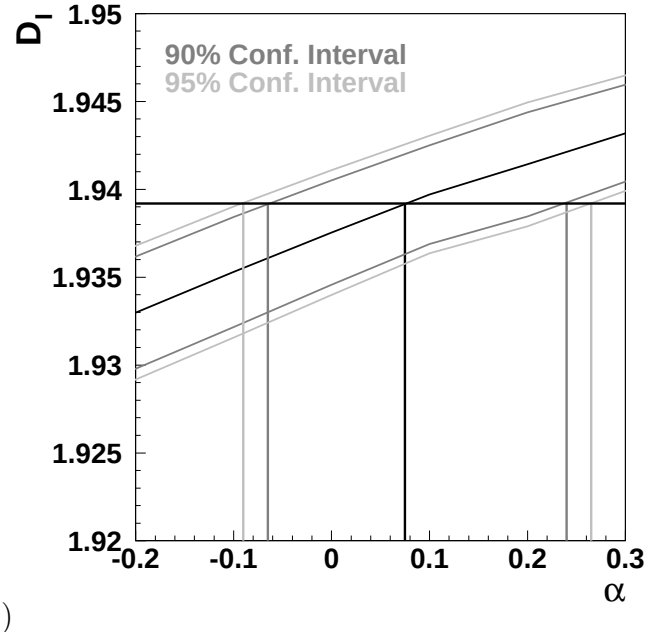


(b)

Figure 10.15. Study of the dependence of $D_{\rm I}$ upon α for a dipole enhancement source model—(a): $N_{Shower} = 500$ $\alpha = .02 \pm .21$; (b): $N_{Shower} = 500$ (zoomed); 90% confidence interval: $[-0.29, 0.36]$; In each case, the solid solid horizontal line indicates the value of $D_{\rm I}$ for the simulated isotropic data set. The vertical lines in (b) and (d) indicate the projection of the nominal value and 90% and 95% confidence intervals of α on the x-axis.

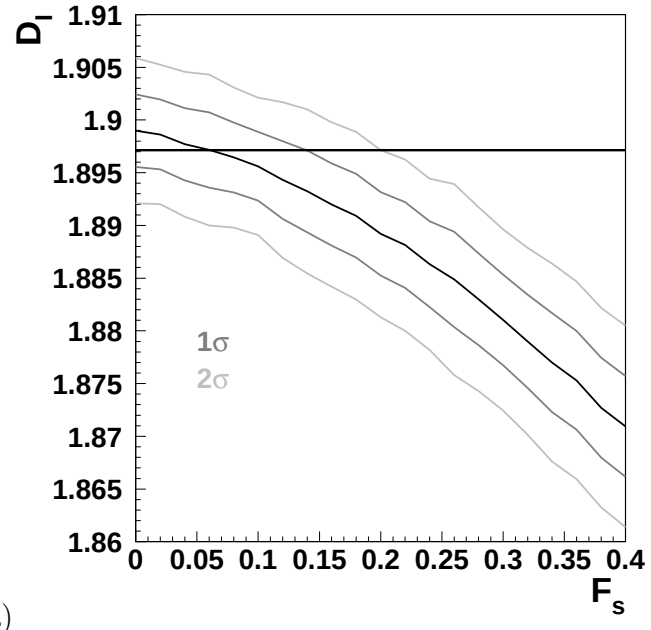


(a)

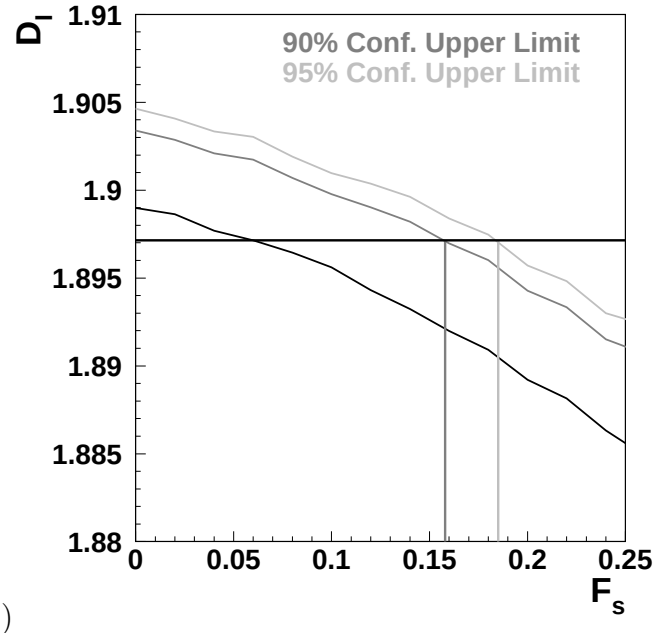


(b)

Figure 10.16. Study of the dependence of $D_{\rm I}$ upon α for a dipole enhancement source model—(a): $N_{Shower} = 2000$; $\alpha = .075 \pm .085$; (b): $N_{Shower} = 2000$ (zoomed); 90% confidence interval: $[-0.065, 0.24]$. In each case, the solid solid horizontal line indicates the value of $D_{\rm I}$ for the simulated isotropic data set. The vertical lines in (b) and (d) indicate the projection of the nominal value and 90% and 95% confidence intervals of α on the x-axis.

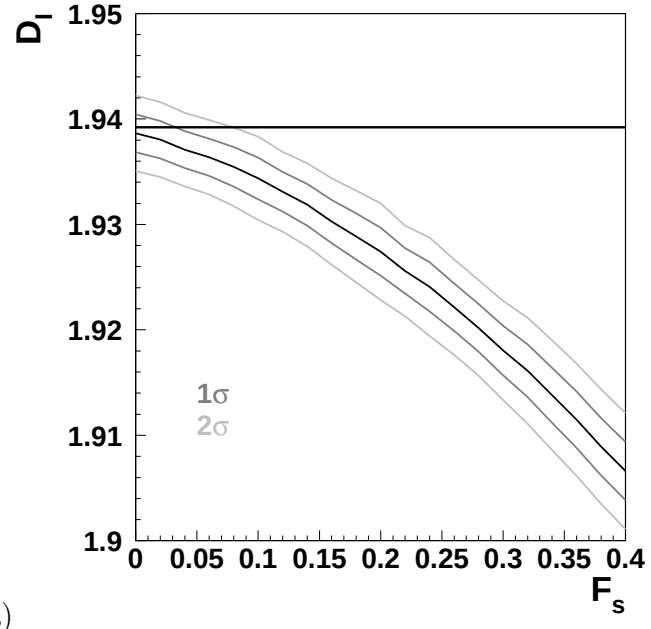


(a)

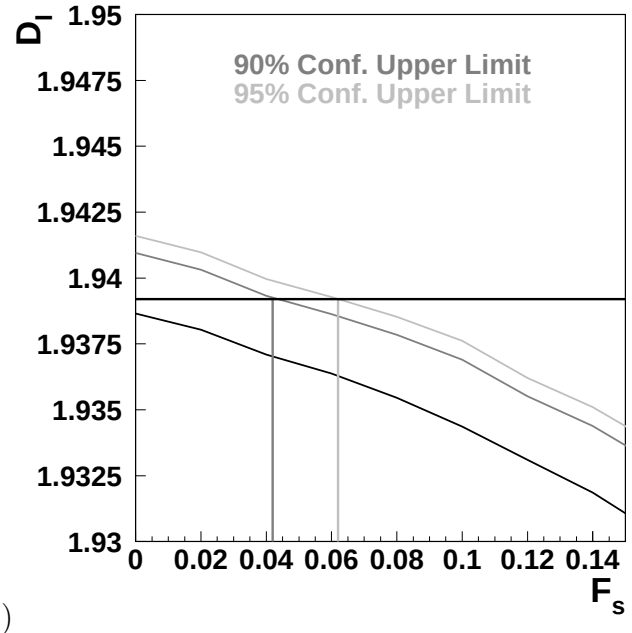


(b)

Figure 10.17. Study of the dependence of $D_{\rm I}$ upon F_s for the seven source model—(a): $N_{Shower} = 500$; (b): $N_{Shower} = 500$ (zoomed); 90% confidence upper limit=0.16. In each case, the solid solid horizontal line indicates the value of $D_{\rm I}$ for the simulated isotropic data set. The vertical lines in (b) and (d) indicate the projection of 90% and 95% confidence upper limits of $r_{\rm s}$ on the x-axis.



(a)



(b)

Figure 10.18. Study of the dependence of $D_{\rm I}$ upon F_s for the seven source model—(a): $N_{Shower} = 2000$; (b): $N_{Shower} = 2000$ (zoomed); 90% confidence upper limite=0.04;. In each case, the solid solid horizontal line indicates the value of $D_{\rm I}$ for the simulated isotropic data set. The vertical lines in (b) and (d) indicate the projection of 90% and 95% confidence upper limits of $r_{\rm s}$ on the x-axis.

5.0 kpc and 15.0 kpc we develop the curve which will demonstrate the dependence of $D_{\rm I}$ upon $r_{\rm s}$. We can then show that the dark matter halo source model can be rejected with at level $\geq 3.6\sigma$ for $N_{Shower} = 500$ and at a level of $\geq 7.0\sigma$ for $N_{Shower} = 2000$ for the full range of hypothesized values for $r_{\rm s}$. The results are shown in Figure 10.19.

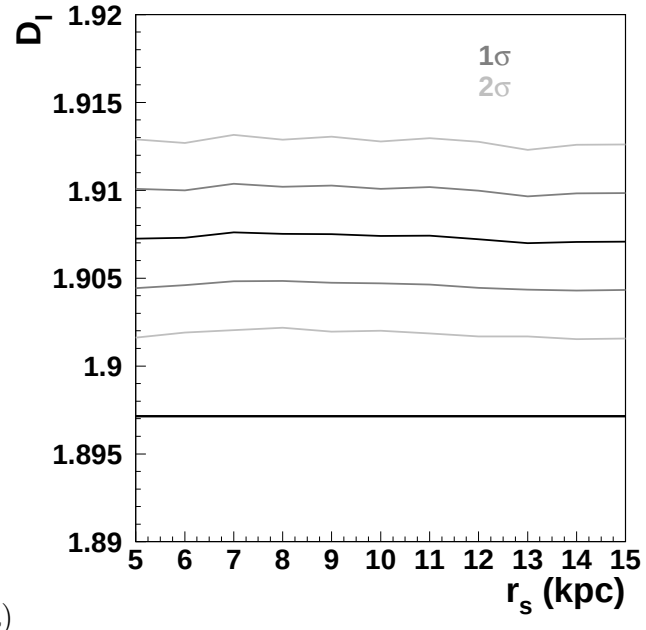
10.6 Discussion

Fractal dimensionality has several advantages over conventional anisotropy techniques. First of all, it naturally accommodates angular resolution. This is extremely important when considering event sets with asymmetric errors, when analysing event sets with variable values for the angular resolution (e.g., dependent on energy or geometry), or when combining multiple data sets from different detectors for a single analysis.

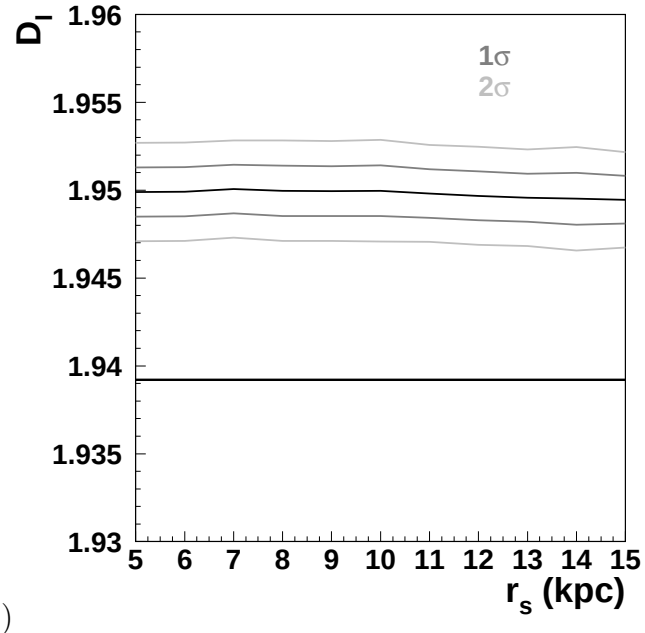
Another advantage that fractal dimensionality possesses is the ability to accommodate any aperture. Because this method makes a relative comparison between a sample and simulations using the same aperture, the physical aperture is simply folded into the analysis. This once again allows the combination of data from multiple detectors with *very* different apertures. It also allows the analysis of extremely complicated apertures without the need to include normalizing factors that needlessly complicate the predictability of the Poisson fluctuations in the data sample.

Perhaps the most striking feature of the fractal dimensionality method is that it only requires a *single* measurement of one's data. While fractal dimensionality will not always provide better statistical significance than a direct measurement for a specific anisotropy, the fact that one considers only a single measurement of the data, for any number of potential anisotropic models, provides one the means to simultaneously reject or accept all of those models without the ensuing statistical penalty.

A possible way of increasing the sensitivity of the fractal dimensionality method is by considering the general case of case of D_q. By varying the value of q to



(a)



(b)

Figure 10.19. Comparison of the distribution of $D_{\rm I}$ values for a dark matter halo source model with a full range of hypothesized values for $r_{\rm s}$—(a): $N_{Shower} = 500$, the dark matter halo model can be rejected at a level $\geq 3.6\sigma$; (b): $N_{Shower} = 2000$, the dark matter halo model can be rejected at a level $\geq 7.0\sigma$. In each case, the solid horizontal line indicates the value of $D_{\rm I}$ for the simulated isotropic data set.

something other than 1, it might be possible to increase the sensitivity of this method to various anisotropies.

The fractal dimensionality method does have some drawbacks. First and foremost is the potential for multiple solutions as was demonstrated in the dipole source model above. The method cannot be applied blindly. It requires a careful inspection of both the data sample and the simulated samples in order to resolve possible ambiguities. Another drawback is the amount of computation required to calculate D_{I} for a large number of simulated data sets. Producing just the plots in Figure 10.16a consumed the equivalent of $\sim$ 1000 CPU hours on 1GHz Athlon machine. Another limitation is the potential for different source models to effectively cancel each other out and yield a potentially deceiving value of D_{I} that resembles that of an isotropic source. One solution for this is to separately consider the value of D_{I} for different celestial regions in the data. Of course, this will incur a statistical penalty.

Two particularly useful roles for fractal dimensionality analysis are (a): as a first test to ascertain if a sample possesses the same heterogeneity as the expected isotropic background and (b): as a wholly independent observable that can be used to confirm the results of a direct measurement.

CHAPTER 11

APPLICATION OF FRACTAL DIMENSIONALITY TO THE GLOBAL DIPOLE SEARCH

The information dimension, $D_{\rm I}$ [66, 67], is a measure of the overall heterogeneity of a data sample. The smaller the value of $D_{\rm I}$, the more heterogeneous the sample is. From Chapter 10, the basic formula for calculating $D_{\rm I}$ is:

$$D_{\rm I} = \left\langle - \frac{1}{\log N_\delta} \sum_{i=1}^{N} P_{\rm i}(N_\delta) \log P_{\rm i}(N_\delta) \right\rangle, \; N_\delta = [354, 360], \tag{11.1}$$

where N_δ is the total number of declinational bins (with a range of values between 354 to 360) and:

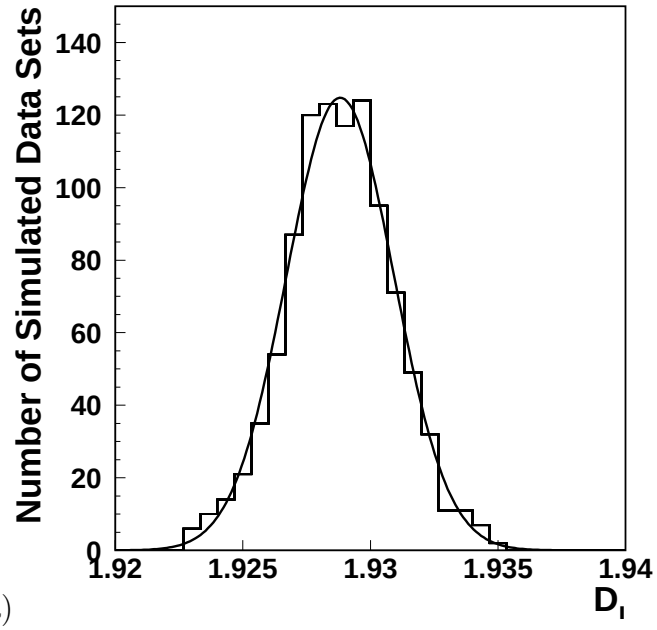
$$P_{\rm i}(N_\delta) = \frac{n_{\rm i}}{<n_{\rm i}>} \frac{\pi^3}{4(N_\delta)^4 \Delta\Omega_\delta}, \tag{11.2}$$

with $n_{\rm i}$ being the number of counts in a particular latitudinal bin, $<n_{\rm i}>$ being the average bin count over the entire sample and $\Delta\Omega_\delta$ being the area of that particular latitudinal bin. A detailed description of this method was given in the previous chapter. In this chapter, we report the results of applying this novel technique.

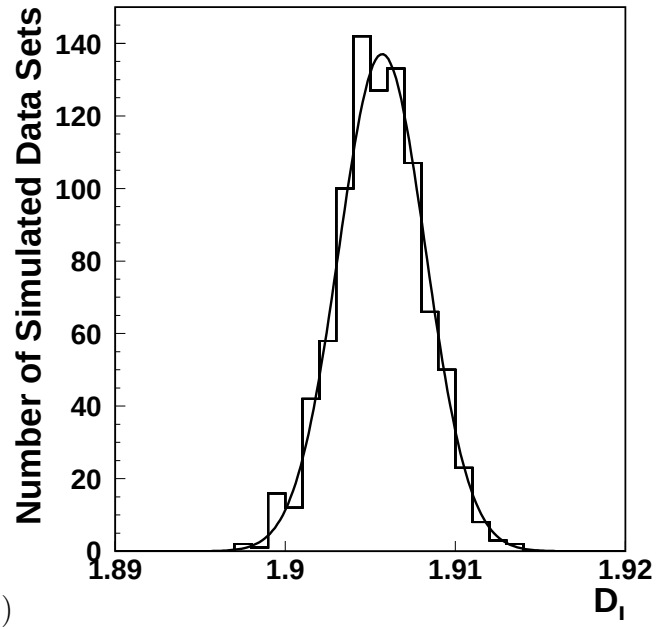
While the measurement of $D_{\rm I}$ is not necessarily the most sensitive tool available, it allows one to rule out any number of potential anisotropic source models with a single measurement. The general scheme that we followed is similar to what we used in the case of the dipole analysis previously reported in Chapter 8.

1. We calculated the value of $D_{\rm I}$ for the real data sample.

2. We created a total of 20,000 simulated data samples, 1000 each for 0.1 increments of α from -1.0 to 1.0. In Figure 11.1 we can see that distribution of $D_{\rm I}$ values for each α-value is Gaussian.



(a)



(b)

Figure 11.1. The distribution of $D_{\rm I}$ values for simulated data sets with a single α-value—(a) the galactic dipole source model with $\alpha = 1.0$; (b) the galactic dipole source model with $\alpha = -1.0$.

3. We constructed a curve consisting of the mean and standard deviation of D_{I} for each value of α.

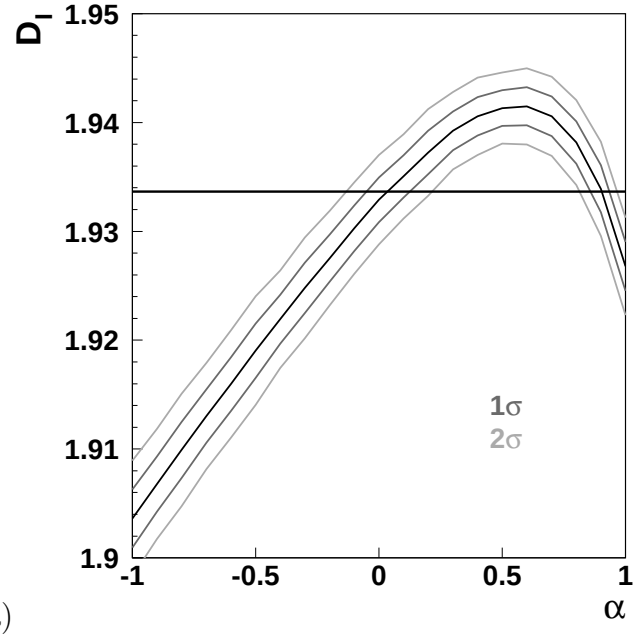
4. We then ascertained the preferred value of α and the 90% confidence interval for each dipole source model by referring the intersections of the 90% confidence interval curves with the actual value of D_{I} for the real data.

The results for all three dipole source models are shown in Figures 11.2, 11.3, and 11.4. The determination of α for both methods are compared in Table 11.1. The 90% confidence intervals for the determination α via the use of D_{I} are substantially larger. This is to be expected because the value of D_{I} is a single number that contains no *a priori* preference for a specific source model. Furthermore, in two cases there is a second solution to α that is excluded by considering the results of the $<\cos\theta>$ method. The important observation is that the results of the two methods are consistent. One advantage of the D_{I} method is that we can state all three 90% confidence intervals jointly, since they are all considering only a single measurement on the real data. In the case of the $<\cos\theta>$ method, we would have to consider a broader confidence interval for each individual model in order to have a simultaneous 90% confidence level for all three models.

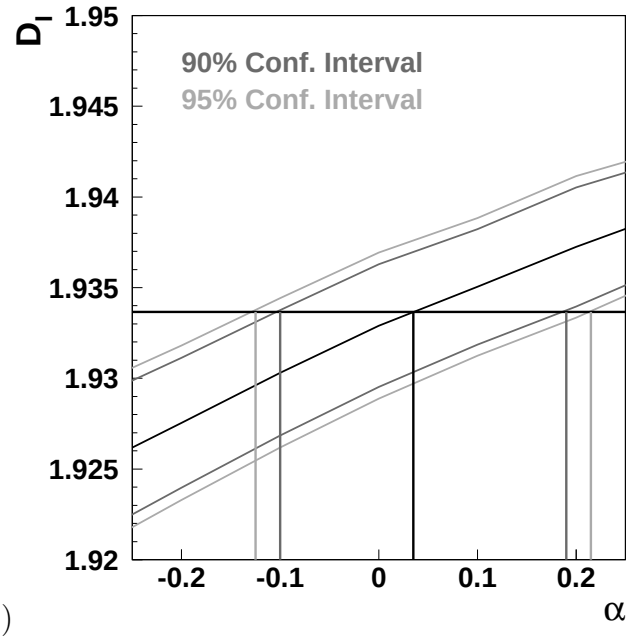
A possible future extension of this work is to employ a multi-parameter fit to improve the statistical precision. The information dimension, D_{I}, and $<\cos\theta>$ can be used simultaneously to narrow the confidence intervals for the dipole function

Table 11.1. Comparison of the estimation of α via direct fit, the value of $<\cos\theta>$ for the dipole function, and the value of D_{I}.

	1	**2**	**3**
SOURCE MODEL	α determined without considering angular resolution	α determined by the value of $<\cos\theta>$	α determined by the value of D_{I}
Galactic	-0.010 ± 0.055	0.005 ± 0.055	0.035 ± 0.09
Centaurus A	-0.035 ± 0.060	-0.005 ± 0.065	0.040 ± 0.095
M87	-0.005 ± 0.045	-0.010 ± 0.045	0.020 ± 0.100

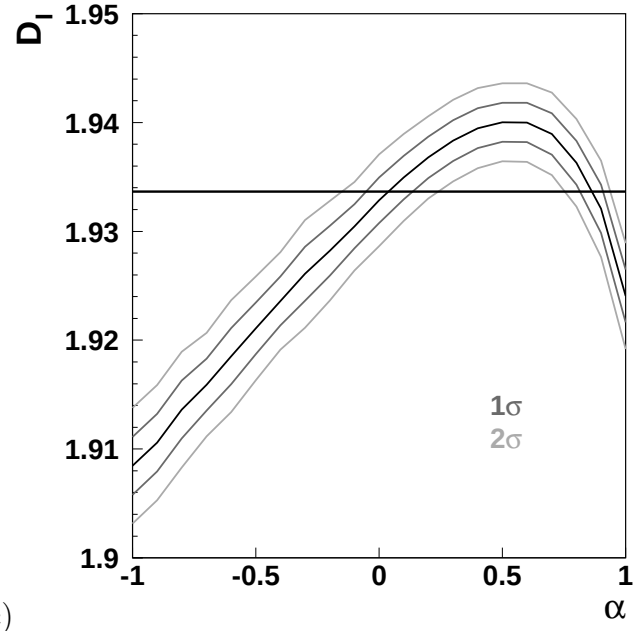


(a)

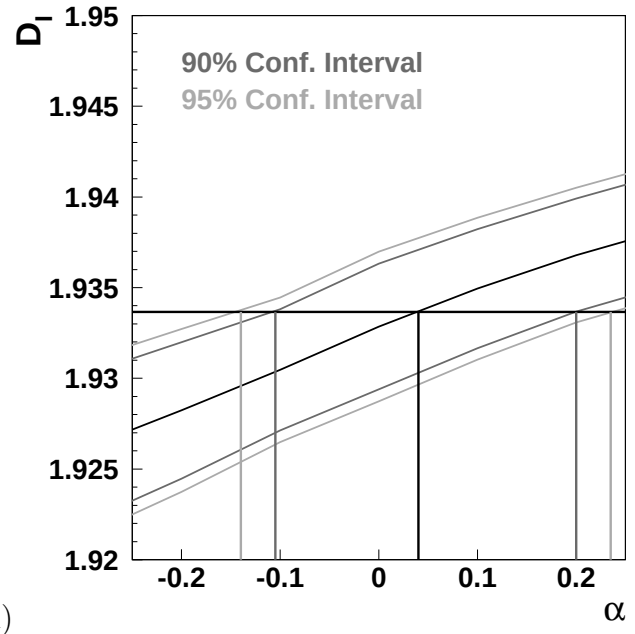


(b)

Figure 11.2. Estimations of the value of α for a dipole source models. The curves demonstrate the dependence of $D_{\rm I}$ upon α. The horizontal lines represent the value of $D_{\rm I}$ for the real data—(a) the galactic dipole source model for $\alpha = [-1.0, 1.0]$; (b) the critical region for the galactic dipole model: $\alpha = 0.035 \pm 0.090$ with a 90% confidence interval of: $[-0.100, 0.190]$.

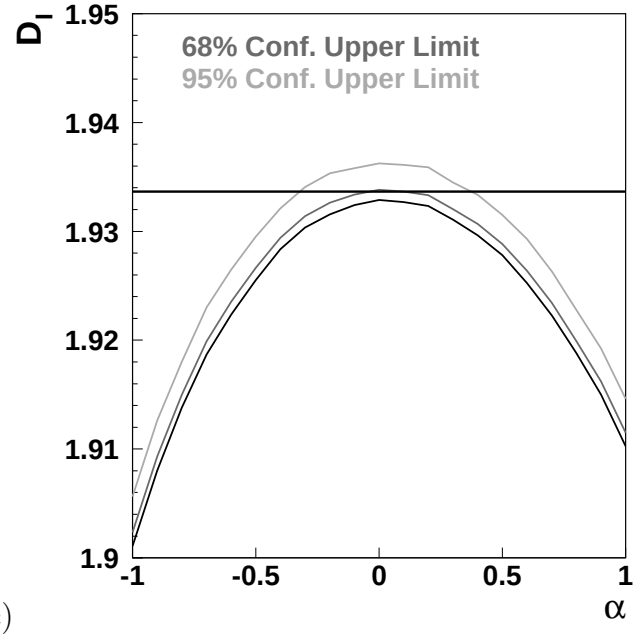


(c)

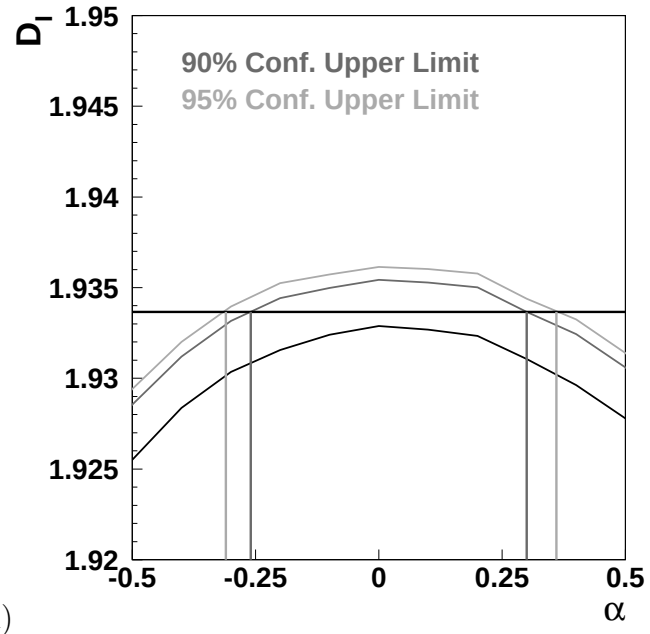


(d)

Figure 11.3. Estimations of the value of α for a dipole source models. The curves demonstrate the dependence of $D_{\rm I}$ upon α. The horizontal lines represent the value of $D_{\rm I}$ for the real data—(a) the Centaurus A dipole source model for $\alpha = [-1.0, 1.0]$; (b) the critical region for the Centaurus A dipole model: $\alpha = 0.040 \pm 0.095$ with a 90% confidence interval of: $[-0.105, 0.200]$.



(c)



(d)

Figure 11.4. Estimations of the value of α for a dipole source models. The curves demonstrate the dependence of $D_{\rm I}$ upon α. The horizontal lines represent the value of $D_{\rm I}$ for the real data—(a) the M87 dipole source model for $\alpha = [-1.0, 1.0]$; (b) the critical region for the M87 dipole model: $\alpha = 0.020 \pm 0.10$ with a 90% confidence interval of: $[-0.26, 0.30]$.

because the two quantities are not *a priori* correlated. With 50% more data already collected and another 50% more to be collected, we anticipate to perform such a study after the conclusion of HiRes observation.

APPENDIX

AN EXAMINATION OF THE SIGNIFICANCE OF THE REPORTED AGASA CLUSTERING PHENOMENON

In 1999 [47] and again in 2001 [25], the AGASA collaboration reported observing what eventually became seven clusters (six "doublets" and one "triplet") in its event set with estimated energies above $\sim 3.8 \times 10^{19}$ eV. Several different attempts have been made to ascertain the significance of these clusters with values chance probabilities as low as 4×10^{-6} [49] or as high as 0.19 [50]. However, the analyses that have been presented by the AGASA collaboration and others did not account for the angular sensitivity of the AGASA detector.

In the study reported in this appendix, we use simulated event sets in order to determine AGASA's intrinsic sensitivity to the observed clustering phenomenon. By doing so, we show that the actual sensitivity of the AGASA array to the observed phenomenon is much lower than the apparent significance originally reported.

A.1 Determining the Significance of the Observed AGASA Clustering Phenomenon

In order to determine the significance of the measured result, we compare the $<\cos\theta>_{[0^\circ,10^\circ]}$ value from the reported AGASA data to a distribution of values for many simulated AGASA-like event sets. In the simulations, we assume the events to be isotropically distributed in right ascension and that they have the declinational distribution presented in Uchiori et al. [55]. Following the lead of Evans et al. [56], we fit a normalized polynomial to this distribution:

$$N(\delta) = 0.323616 + 0.0361515\delta - 5.04019 \times 10^{-4}\delta^2 + 5.539141 \times 10^{-7}\delta^3; \quad \text{(A.1)}$$

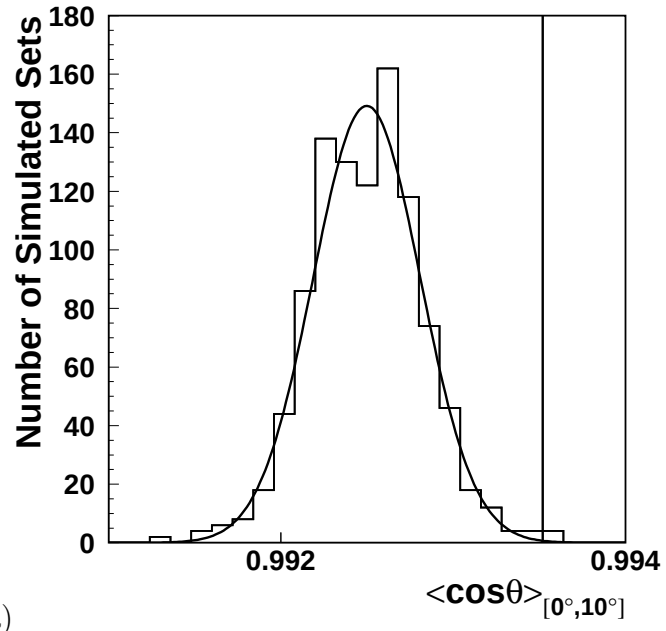
where $N(\delta)$ holds for $\delta = [-8^\circ, 87.5^\circ]$ and the maximum value of $N(\delta)$ is 1.

By simulating large numbers of sets, each with the same number of events as the real data set, we can then establish what the expected a distribution of $<\cos\theta>_{[0^\circ,10^\circ]}$ values is for isotropic event samples. In Figure A.1a we show the distribution of $<\cos\theta>_{[0^\circ,10^\circ]}$ values that we obtained for 1000 simulated isotropic data sets each of which contains 59 simulated events with energies (and thus angular resolutions) randomly sampled from the known values of the AGASA data set. The distribution has a mean value of 0.99250 and a standard deviation of 0.00033. This allows us to state the significance of the clustering phenomenon in the actual AGASA data set at 3.1σ or a chance probability of $\sim 10^{-3}$.

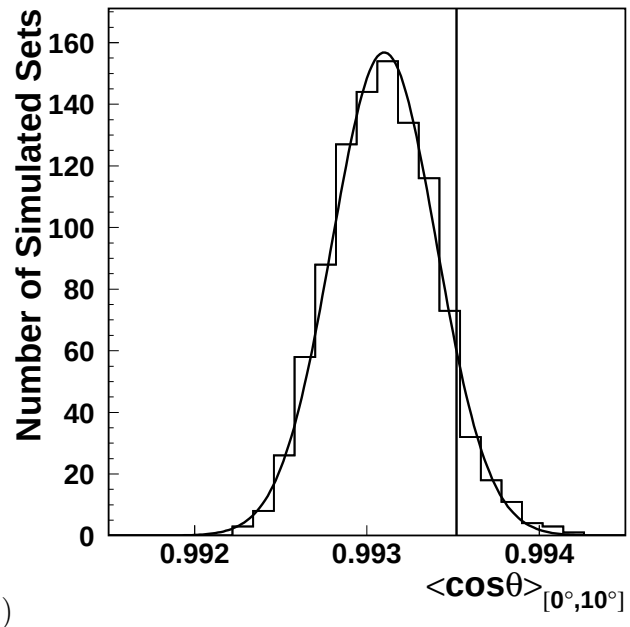
A.2 Establishing the Sensitivity of AGASA to Clustering Phenomena

We now determine the sensitivity of the AGASA detector to the observed clustering phenomenon. This can be done by calculating the value of $<\cos\theta>_{[0^\circ,10^\circ]}$ for sets that contain artificially inserted clusters whose angular separation is defined by the intrinsic angular resolution of the AGASA detector. In particular, the AGASA collaboration states that they find seven clusters consisting of six doublets and one triplet in a total set of 59 events. Thus, we calculate the value of $<\cos\theta>_{[0^\circ,10^\circ]}$ for simulated sets that contain 59 events with six doublets and one triplet. Figure A.1b shows the distribution of $<\cos\theta>_{[0^\circ,10^\circ]}$ values with a mean value of 0.99310 and a standard deviation of 0.00030.

What is remarkable about this distribution is that only 8% of the time does its value exceed that of the actual data sample. This suggests that the AGASA clustering signal is enhanced by a chance upward fluctuation due to nominal arrival directions which are actually closer together than what one would, on average, expect given the angular resolution.



(a)



(b)

Figure A.1. Distributions of $<\cos\theta>_{[0^\circ,10^\circ]}$ values for 1000 simulated sets—(a): isotropic simulated sets; (b): simulated sets with six "doublets" and one "triplet" artificially inserted. In both cases, the vertical line corresponds to the value of $<\cos\theta>_{[0^\circ,10^\circ]}$ for the reported AGASA data.

Next, we do a one-to-one comparison between the values in the distributions shown in Figures A.1a and A.1b. This done by repeated random sampling of the $<\cos\theta>_{[0^\circ,10^\circ]}$ value from each of the two distributions in Figure A.1. The scatter plot of the one-to-one comparisons of the values of $<\cos\theta>_{[0^\circ,10^\circ]}$ for the simulated sets with and without clustering is shown in Figure A.2. We find that the clustering signal of the isotropic simulation *exceeded* the signal of the set with artificial clustering $\sim 9\%$ of the time.

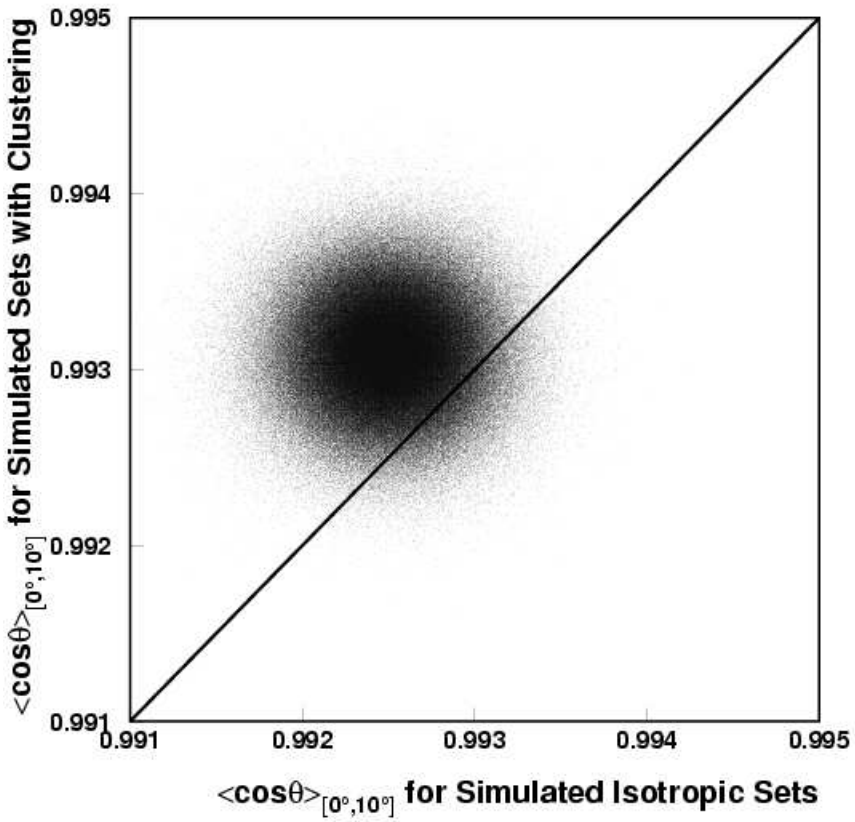


Figure A.2. A one to one comparison of the value of $<\cos\theta>_{[0^\circ,10^\circ]}$ for simulated sets with and without clustering at the level reported by AGASA

A.3 Conclusion

We conclude that while AGASA did indeed observe a clustering signal with a chance probability of $\sim 10^{-3}$, this is likely primarily due to a random a fluctuation in the nominal arrival directions of the observed events. The AGASA detector simply does not have the prerequisite angular resolution to observe clustering at the rate that is claimed with a significance better than a chance probability of $\sim 10^{-1}$. A much better determination of the presence of clustering can be expected with superior angular resolution such as that provided by HiRes stereo data [53] and eventually by Auger hybrid data [71].

REFERENCES

[1] J. Linsley, Phys. Rev. Lett., **10** (1963), p. 146.

[2] R. U. Abbasi et al., Phys. Rev. Lett., **92** (2004), p. 151101, [arXiv:astro-ph/0208243].

[3] ——, [arXiv:astro-ph/0208301].

[4] ——, Astropart. Phys., **21** (2004), p. 111, [arXiv:astro-ph/0309457].

[5] ——, Astropart. Phys., **22** (2004), pp. 139–149, [arXiv:astro-ph/0404366].

[6] B. T. Stokes, C. C. H. Jui, and J. N. Matthews, Astropart. Phys., **21** (2004), p. 95, [arXiv:astro-ph/0307491].

[7] J. Linsley, L. Scarsi, and R. B., J.Phys. Soc. Japan (supp. A-III), **17** (1962), p. 91.

[8] R. U. Abbasi et al., [arXiv:astro-ph/0407622].

[9] K. Greisen, Phys. Rev. Lett., **16** (1966), p. 748.

[10] G. T. Zatsepin and V. A. Kuzmin, Sov. Phys. JETP Lett., **4** (1966), p. 78.

[11] X. Bertou, M. Boratav, and A. Letessier-Selvon, Int. J. Mod. Phys., **A15** (2000), pp. 2181–2224, [arXiv:astro-ph/0001516].

[12] D. E. Groom et al., Eur. Phys. J., **C15** (2000), pp. 1–878.

[13] B. N. Afanasiev et al., in 24th ICRC, vol. 2, Rome, 1995, p. 756.

[14] M. Ave, J. Knapp, J. Lloyd-Evans, M. Marchesini, and A. A. Watson, Astropart. Phys., **19** (2003), pp. 47–60, [arXiv:astro-ph/0112253].

[15] D. J. Bird et al., Astrophys. J., **424** (1994), pp. 491–502.

[16] S. Yoshida et al., Astropart. Phys., **3** (1995), pp. 105–124.

[17] T. Abu-Zayyad et al., Astrophys. J., **557** (2001), pp. 686–699, [arXiv:astro-ph/0010652].

[18] D. J. Bird et al., Phys. Rev. Lett., **71** (1993), pp. 3401–3404.

[19] D. Heck, G. Schatz, T. Thouw, J. Knapp, and J. N. Capdevielle, *CORSIKA: A Monte Carlo code to simulate extensive air showers*, 1998. FZKA-6019.

[20] V. Berezinsky, M. Kachelriess, and A. Vilenkin, Phys. Rev. Lett., **79** (1997), pp. 4302–4305, [arXiv:astro-ph/9708217].

[21] P. Bhattacharjee, C. T. Hill, and D. N. Schramm, Phys. Rev. Lett., **69** (1992), pp. 567–570.

[22] R. Schodel et al., Nature, **419** (2002), pp. 694–696.

[23] M. Lemoine, G. Sigl, and P. Biermann, [arXiv:astro-ph/9903124].

[24] Auger Collaboration, *The Pierre Auger Project Design Report*, 1996. FERMILAB-PUB-96-024.

[25] N. Hayashida et al., [arXiv:astro-ph/0008102].

[26] S. R. Coleman and S. L. Glashow, Phys. Rev. D, **59** (1999), p. 116008, [arXiv:hep-ph/9812418].

[27] G. R. Farrar and T. Piran, Phys. Rev. Lett., **84** (2000), p. 3527, [arXiv:astro-ph/9906431].

[28] T. K. Gaisser and A. M. Hillas, in 15th ICRC, vol. 8, Plovdiv, 1977, p. 353.

[29] T. K. Gaisser, *Cosmic Rays and Particle Physics*, Cambridge University Press, 1990, ch. 15, p. 226.

[30] A. N. Bunner, *Cosmic Ray Detection by Atmospheric Fluorescence*, PhD thesis, Cornell University, 1967.

[31] P. Huntemeyer, in AIP Conf., vol. 698, 2004, pp. 341–344.

[32] J. Linsley, in 18th ICRC, vol. 12, Bangalore, 1983, p. 135.

[33] T. AbuZayyad, *The Energy Spectrum of Ultra High Energy Cosmic Rays*, PhD thesis, University of Utah, 2000.

[34] R. M. Baltrusaitis et al., Nuclear Instruments and Methods in Physics Research, **A240** (1985), pp. 410–428.

[35] P. Sokolsky, *Introduction to Ultrahigh Energy Cosmic Ray Physics*, Addison-Wesley Publishing Company, Inc., 1989, ch. 5, p. 51.

[36] F. Kakimoto et al., NIM A, **372** (1996).

[37] P. Sokolsky, *Introduction to Ultrahigh Energy Cosmic Ray Physics*, Addison-Wesley Publishing Company, Inc., 1989, ch. 15, p. 201.

[38] M. J. Kidd, *Properties of Extensive Air Showers Around* 10^{17} *eV*, PhD thesis, University of Illinois at Urbana-Champaign, 1997.

[39] P. L. Biermann, E. J. Ahn, G. A. Medina-Tanco, and T. Stanev, Nucl. Phys. Proc. Suppl., **87** (2000), pp. 417–419, [arXiv:astro-ph/0008063].

[40] G. R. Farrar and T. Piran, [arXiv:astro-ph/0010370].

[41] L. A. Anchordoqui, H. Goldberg, and T. J. Weiler, Phys. Rev. Lett., **87** (2001), p. 081101, [arXiv:astro-ph/0103043].

[42] N. Hayashida et al., in 26th ICRC, vol. 3, Salt Lake City, 1999, pp. 256–259, [arXiv:astro-ph/9906056].

[43] D. J. Bird et al., Ap. J., **511** (1999), p. 738, [arXiv:astro-ph/9806096].

[44] J. A. Bellido, R. W. Clay, B. R. Dawson, and M. Johnston-Hollitt, Astropart. Phys., **15** (2001), pp. 167–175, [arXiv:astro-ph/0009039].

[45] P. Sommers, Astropart. Phys., **14** (2001), pp. 271–286, [arXiv:astro-ph/0004016].

[46] L. R. Wiencke et al., in 27th ICRC, vol. 1, Hamburg, 2001, p. 635.

[47] M. Takeda et al., Astrophys. J., **522** (1999), pp. 225–237, [arXiv:astro-ph/9902239].

[48] M. Takeda et al., in 27th ICRC, vol. 1, Hamburg, 2001, p. 337.

[49] P. G. Tinyakov and I. I. Tkachev, JETP Lett., **74** (2001), pp. 1–5, [arXiv:astro-ph/0102101].

[50] C. B. Finley and S. Westerhoff, Astropart. Phys., **21** (2004), pp. 359–367, [arXiv:astro-ph/0309159].

[51] J. Bellido et al., in 27th ICRC, vol. 1, Hamburg, 2001, p. 364.

[52] J. Bellido et al., in 28th ICRC, vol. 1, Tsukuba, 2003, p. 425.

[53] R. U. Abbasi et al., Astrophys. J., **610** (2004), p. L73, [arXiv:astro-ph/0404137].

[54] M. Takeda et al., Phys. Rev. Lett., **81** (1998), pp. 1163–1166, [arXiv:astro-ph/9807193].

[55] Y. Uchihori et al., Astropart. Phys., **13** (2000), pp. 151–160, [arXiv:astro-ph/9908193].

[56] N. W. Evans, F. Ferrer, and S. Sarkar, Astropart. Phys., **17** (2002), pp. 319–340, [arXiv:astro-ph/0103085].

[57] D. Bird et al., in 23rd ICRC, vol. 2, Calgary, 1993, p. 51.

[58] ——, in 23rd ICRC, vol. 2, Calgary, 1993, p. 30.

[59] T. Stanev, P. L. Biermann, J. Lloyd-Evans, J. P. Rachen, and A. Watson, Phys. Rev. Lett., **75** (1995), pp. 3056–3059, [arXiv:astro-ph/9505093].

[60] M. A. Lawrence, R. J. O. Reid, and A. A. Watson, J. Phys. G, **17** (1991), pp. 733–757.

[61] T. A. Erogov, *Tokyo Workshop on Techniques for the Study of Extremely High Energy Cosmic Rays*, ICRR, Tokyo, 1993, p. 35.

[62] S. Yoshida et al., Astropart. Phys., **3** (1995), pp. 105–124.

[63] J. Linsley, in 8th ICRC, vol. 4, Jaipur, 1963, p. 77.

[64] A. N. Kolmogorov, Dokl. Akad. Nauk. SSSR, **124** (1959), p. 754.

[65] ——, Math. Rev., **21** (1959), p. 2035.

[66] J. Balatoni and A. Reyni, Publ. Math. Inst. Hung. Acad. Sci., **1** (1959), p. 9.

[67] A. H. Nayfeh and B. Balachandran, *Applied Nonlinear Dynamics: Analytical, Computational, and Experimental Methods*, Wiley, 1995, pp. 545–547.

[68] P. Grassberger, Phys. Rev. A, **97** (1983), p. 227.

[69] H. G. E. Hentschel and I. Procaccia, Physica D, **8** (1983), p. 435.

[70] J. F. Navarro, C. S. Frenk, and S. D. M. White, Astrophys. J., **462** (1996), pp. 563–575, [arXiv:astro-ph/9508025].

[71] L. A. Anchordoqui et al. (2004), [arXiv:astro-ph/0409470].