DISSERTATION

THE COSMIC RAY ENERGY SPECTRUM FROM 1-10 EXA ELECTRON VOLTS

MEASURED BY THE PIERRE AUGER OBSERVATORY

Submitted by

Robert Knapik

Department of Physics

In partial fulfillment of the requirements

For the Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2009

UMI Number: 3385171

UMI®

Dissertation Publishing

ProQuest®

COLORADO STATE UNIVERSITY
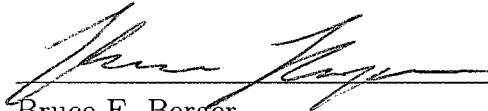
May 13, 2009

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED
UNDER OUR SUPERVISION BY ROBERT KNAPIK ENTITLED "THE
COSMIC RAY ENERGY SPECTRUM FROM 1-10 EXA ELECTRON VOLTS
MEASURED BY THE PIERRE AUGER OBSERVATORY" BE ACCEPTED
AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY.

Committee on Graduate work

_____
Robert J. Wilson

_____
Bruce E. Berger

_____
Thomas B. Borak

_____
Advisor, John L. Harton

_____
Department Chair, Dieter H. Hochheimer

ii

ABSTRACT OF DISSERTATION

THE COSMIC RAY ENERGY SPECTRUM FROM 1-10 EXA ELECTRON VOLTS

MEASURED BY THE PIERRE AUGER OBSERVATORY

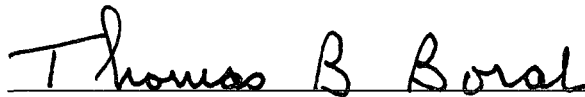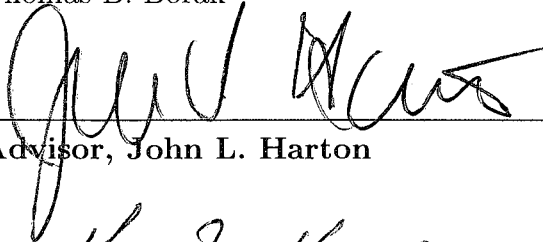The observed decrease in flux of cosmic rays as the energy increases can be described by power law with an almost constant spectral index for 12 decades of energy. Observing spectral index changes are used to constrain models for the sources of cosmic rays. The Pierre Auger Observatory was built to study the highest energy cosmic rays and combines two complementary techniques, a fluorescence detector and a surface detector. The surface detector is 100% efficient for energies above 3 EeV allowing for a flux measurement with low systematic uncertainties. This thesis describes the techniques developed to measure the flux of cosmic rays below 3 EeV while maintaining low uncertainties. The resulting energy spectrum confirms the previously measured change in spectral index observed by other experiments. Systematic differences in the measured energy spectra between experiments exist. Possible reasons for these differences and the astrophysical implications are discussed.

Robert Knapik
Department of Physics
Colorado State University
Fort Collins, CO 80523
Summer 2009

iii

Contents

Introduction

Cosmic rays are particles from space that are constantly bombarding our atmosphere. The origin of cosmic rays is a great mystery. It has long been the goal of cosmic ray experiments to identify candidate astrophysical sources. This has proved to be a very difficult task.

Unlike photons, which are neutral particles, cosmic rays are mostly charged particles. The magnetic fields of the universe bend the trajectories of cosmic rays so that the arrival direction does not generally point back to the source. Both the galactic and extragalactic magnetic fields are not precisely understood. This makes inferring the source location from the arrival direction of the cosmic ray very difficult. At high energies, the bending in magnetic fields is reduced so that large scale patterns in the arrival directions of cosmic rays could help narrow down possible candidate sources. For the highest energy comic rays, so called ultra high energy cosmic rays, the deflections might be small enough to detect a point source.

Cosmic rays are mainly atomic nuclei with small percentages of electrons at low energies. Protons, or hydrogen nuclei, dominate the low energy flux and this is naively expected since hydrogen is the most abundant element in the universe. However, for higher energies there are indications that the percentage of heavier nuclei, such as iron may be significant. Understanding how the primary composition, or the percentage of each particle type, changes with energy can also help rule out possible sources.

Cosmic rays have been detected with drastically differing energies and arrival rates. The energy spectrum of cosmic rays demonstrates how the flux of cosmic rays changes as a function of energy. The observed decrease in flux as the energy increases

can be described by power law with an almost constant spectral index for 12 decades of energy. A change in spectral index could be indicative of a change in source type or of propagation effects.

The arrival directions, primary composition and energy spectrum of cosmic rays are all important pieces of information that can help identity the sources of ultra high energy cosmic rays. This thesis work is dedicated to measuring the energy spectrum from 10^{18} and 10^{19} eV with the Pierre Auger Observatory. There have been two observed departures or spectral index changes above the energy of 10^{18} eV. One between 10^{18} and 10^{19} eV, and another between 10^{19} and 10^{20} eV. Previous experiments have not agreed on the exact energy of the spectral index change or on the magnitude of the change.

Detection of cosmic rays at these energies is difficult due to the low flux incident on Earth. The flux of cosmic rays at 10^{18} eV is only one cosmic ray per km^2 per year. For energies greater than 10^{20} eV, the flux is less than a few cosmic rays per km^2 per millennium. In addition to the low flux, detection techniques for these energies have large systematic uncertainties. The Pierre Auger Observatory was built to a scale 30 times larger than previous experiments to accommodate the low flux. Redundant techniques along with extensive monitoring provide for better understanding of systematic uncertainties. These factors have made the Auger Observatory the premiere cosmic ray experiment at ultra high energies. However, the Pierre Auger Observatory was designed for optimal detection of cosmic rays with energies above 10^{19} eV. To observe the lower energy spectral index change new techniques are needed.

After a few introductory chapters, this thesis describes the techniques I developed to measure the flux of cosmic rays in the non-optimized region from 10^{18} to 10^{19} eV, while maintaining low systematic uncertainties. Specifically, a chapter is devoted to how the energy of an event is measured and another focuses on measuring the detec-

tion efficiency of the observatory. A chapter has also been devoted to a hardware task that I was a leading member of. Its inclusion demonstrates the emphasis the observatory places on understanding systematic uncertainties. The final chapter contains the measured energy spectrum, a comparison with other experiments and a discussion on the possible astrophysical implications of my results. The measured flux in this thesis combined with work published in early 2008 from the Auger Observatory show both spectral index changes previously observed by other experiments but now with more statistics and arguably better understanding of systematic uncertainties. Work is still on going in the Pierre Auger Observatory to reduce systematic uncertainties, but this work already helps to constrain theories on the origins of ultra high energy cosmic rays.

Chapter 1

Ultra High Energy Cosmic Rays

The term cosmic ray dates back to 1912 when Victor Hess performed high altitude balloon experiments. He discovered that there was a source of radiation penetrating the atmosphere that was coming from outer space. In 1938 the physicist Pierre Auger noticed time coincidences of the signals detected from two particle detectors separated by distances of many meters. As the particle detectors were moved further apart the coincidences remained, only at a lesser rate. He theorized that the detectors were recording particles from the same phenomena. When subatomic particles interact with matter, a showering of particles can be observed. In the 1930's such showers had already been observed in cloud chambers by physicists on small scales. Auger realized that if a high energy particle interacted in air, then a similar shower would result. The phenomena turned out to be large showers of particles resulting from primary interactions high in the atmosphere. These showers were called extensive air showers (EAS). It had also been observed that the size of the shower was roughly proportional to the energy of the primary cosmic ray particle. A shower of such large scale implied that the energy of the primary cosmic ray would have been many orders of magnitude greater than any particle yet observed at that time.

For much of the first half of the twentieth century, the field of particle physics and cosmic ray physics were one in the same since there were not yet many terrestrial sources of energetic subatomic particles. With the advent of particle accelerators the two fields eventually separated. Particle accelerators have allowed for many precision measurements of the properties of subatomic particles. Air showers continued being detected from cosmic rays with larger energy than the particles being produced in particle accelerators. However, the nature of detection was far less precise than the controlled beams at terrestrial particle accelerators. Particle physicists went on to create and verify the standard model of particle physics. Cosmic ray physicists instead, started focusing on the search for the astrophysical sources and acceleration mechanisms of cosmic rays.

The detected intensity of cosmic rays at Earth, or the number of events incident per unit time, area and solid angle, changes with the energy of the cosmic ray. The differential change in intensity with is the flux of cosmic rays and can be approximated by a power law over a very large range of energies. The energy spectrum shows the cosmic ray flux versus energy. Flux is generally measured in units of $[km^{-2}yr^{-1}sr^{-1}eV^{-1}]$. Figure 1.1 shows the energy spectrum as measured by several experiments from (10^{14} -10^{20} eV). Due to the steepness of the flux, log-log scales are typically used. The center of mass energies of the interaction of a cosmic ray and an air molecule are also shown on the top axis of Figure 1.1. The energies of the largest human-made terrestrial accelerators are indicated. It is the hope of the particle physicist that a better understanding of cosmic rays at the highest energies will lead to information on hadron interactions.

At the highest energies the flux of cosmic rays is very low. The flux at 10^{18} eV is around one cosmic ray per football field per year, and at the highest energies, of around 10^{20} eV, the flux is below one cosmic ray per city block per millennium.
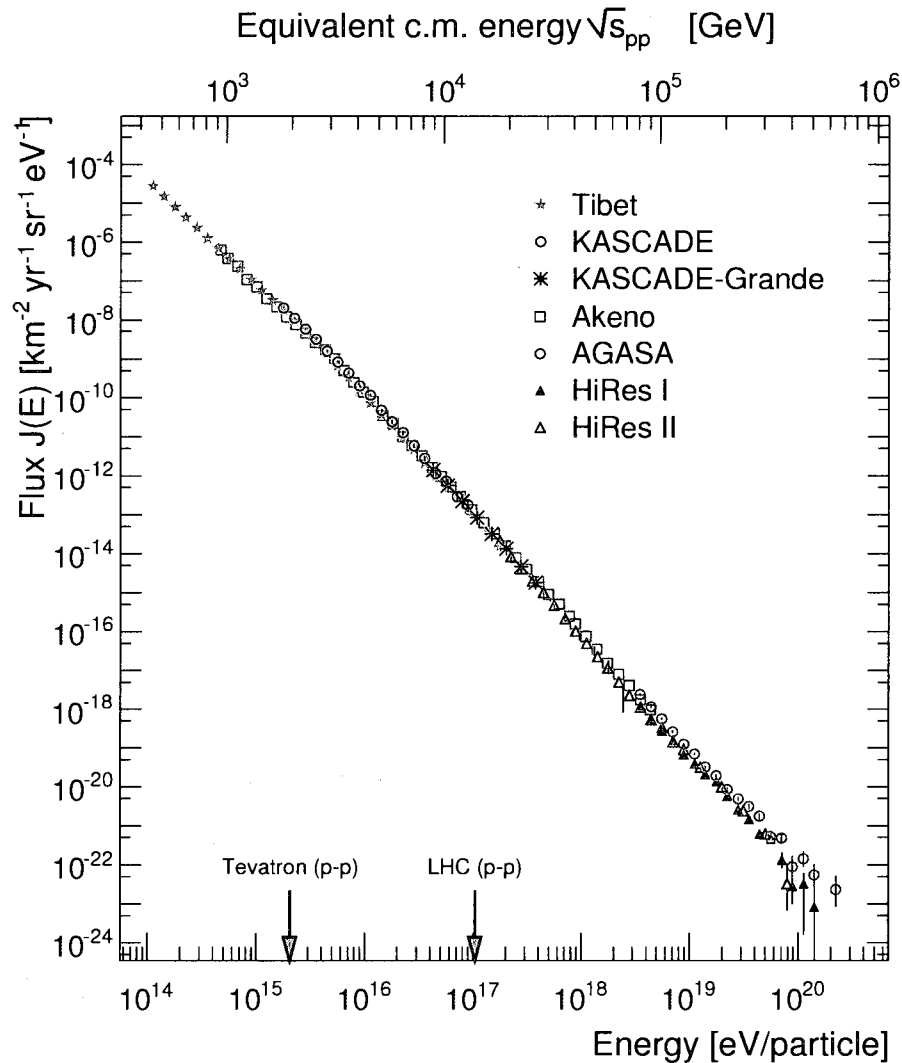
Data from other experiments collected by Ralf Engel as presented in [1]

Figure 1.1 The cosmic ray energy spectrum plots the flux of cosmic rays, J(E), versus energy, E. Log-log scales are typically used due to the steeply falling flux. Flux is measured in units of $[km^{-2}yr^{-1}sr^{-1}eV^{-1}]$. The lower y-axis shows the primary energy of the cosmic ray and the upper y-axis shows the equivalent center of mass energy of a proton-proton collision. The largest human-made particle colliers center of mass energies are indicated.

Chapter 2 will provide details on the detection methods needed over the entire range of energies shown in Figure 1.1 with a focus on the large aperture experiments needed for energies greater than 10^{17} eV. These experiments study what are called ultra high

energy cosmic rays (UHECR).

1.1   What is cosmic radiation?

Cosmic rays are mainly charged particles that have been accelerated to relativistic energies. Gamma rays, neutrinos and other neutral particles can also be considered cosmic rays, but charged particles, including protons and heavier nuclei, dominate the flux. At all energies the cosmic ray flux contains some percentage of heavier nuclei. This is expected because the matter density of the universe contains nuclei from the nuclear reactions in stars. The greatest cosmic abundances of nuclei are H, He, Si, C, N, O and Fe [2]. Hydrogen nuclei and protons are synonymous. The term primary composition is used to refer to this mix of different cosmic ray particles. Understanding the average composition at different energies is a major task of cosmic ray experiments.

The acceleration mechanisms for the entire range of the energy spectrum are not known. Also, the astrophysical objects responsible for the acceleration are only known at the lowest end of the energy spectrum. The nuclear processes in the Sun are responsible for the cosmic ray flux detected at Earth for energies around 10^9 eV. Galactic super nova remnants (SNR) are theorized to be the most likely sources of cosmic rays from $\approx 10^{12}$ to 10^{16} eV [3]. In the ultra high energy region, very powerful accelerators are needed to accelerate particles to such extreme energies. The sources are most likely extragalactic above 10^{19} eV, but between 10^{16} and 10^{19} eV, there could be contributions from both galactic and extragalactic sources. Observatories can detect certain average cosmic ray observables to try and understand the origins of cosmic rays.

The aim of this thesis work is to measure the energy spectrum in detail from

$10^{18} - 10^{19}$ eV, which is just one cosmic ray observable that will aid in the search for the origins of the highest energy cosmic rays. By combing the information from all observables together, many classes of astrophysical objects can ruled out and the true sources eventually identified.

1.2 Cosmic ray observables

To learn about cosmic rays, there are three main observables that one tries to measure. The first is the energy spectrum, which plots flux verses energy. The second is the arrival directions of the cosmic rays. The third is the mass composition, or the type of particle the cosmic rays are. The following subsections briefly describe these measurements.

1.2.1 Energy spectrum

If cosmic rays were only protons, were only accelerated by one mechanism and the sources were evenly distributed throughout the galaxy, the power law nature of the energy spectrum over such a large dynamic range would be expected. The slope of the power law would depend only on the strength of the sources and any propagation effects. In this situation the energy spectrum would be relatively featureless. The fact that the energy spectrum has so few features is remarkable. There are however, three main deviations, or spectral breaks in the power law behavior above 10^{12} eV. These deviations could be the result of a change in composition, a change in acceleration mechanism, a change in source type or distribution, interactions during propagation or some combination of these effects.

Observatories measure the cosmic ray flux incident on earth. This flux will be different from the injected flux at the sources. Interactions with magnetic fields, the

interstellar medium and cosmic background radiation (starlight, cosmic microwave background, etc.) will modify the observed flux on earth. Evidence from astronomy and other astroparticle experiments can help map out the magnetic fields and background radiation. These effects need be taken into account when analyzing the observed cosmic ray spectrum. The deviations from a strict power law could be from propagation effects alone. This could imply that only one type of astrophysical accelerator is responsible for the observed flux. However, if deviation is present that is not explained by propagation effects, then this could be evidence for a possible new source or different acceleration mechanism. This could be verified if the change in energy spectrum and a change in arrival direction distribution happen at the same energy. A change in mass composition could also manifest as a spectral feature. If the mass composition changes to heavier nuclei, the bending in magnetic fields will increase such that the arrival direction distribution could change as well.

1.2.2 Arrival directions

As charged particles propagate from their source they are subject to forces from the galactic magnetic field that modify their trajectories. The force is proportional the charge of the cosmic ray, so iron nuclei are subject to 26 times the force as protons. Below $\approx 10^{16}$ eV, the magnetic field forces effectively randomize the trajectories so that cosmic rays form an isotropic flux in the galaxy. The detected arrival direction of a cosmic ray has no information on the source location at these lower energies. The galactic magnetic fields also effectively contain these cosmic rays in our galaxy. Only above $\approx 10^{17}$ eV do cosmic rays have enough energy to efficiently escape from our galactic magnetic field [4]. The magnetic fields of other galaxies are thought to have similar strengths so that above this energy extragalactic sources might begin to contribute to the flux detected on earth. The contributions of extragalactic cosmic

rays and the deficit of galactic cosmic rays could cause a change in the observed energy spectrum.
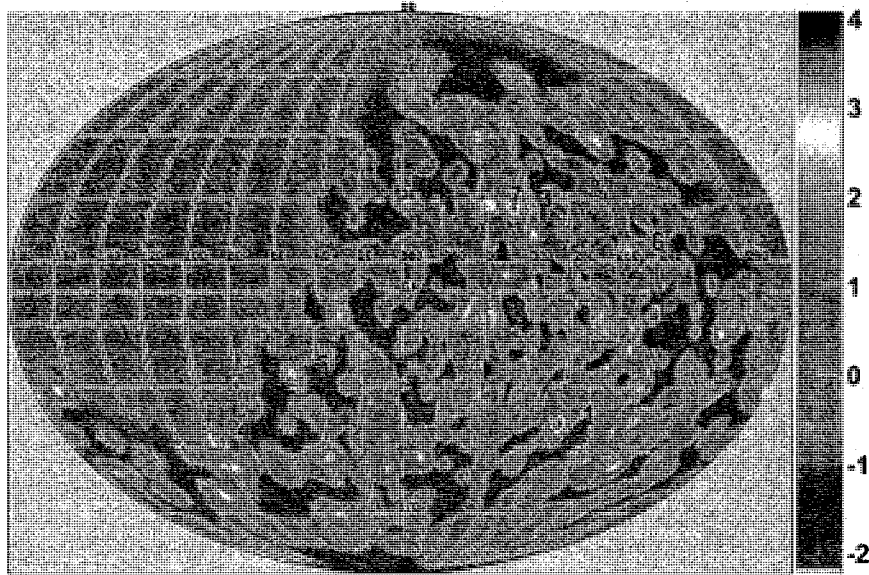
Figure 1.2 A sky map from the Pierre Auger Observatory for cosmic rays of energy around 10^{18} eV. The colors represent the excess or deficit of the number of cosmic rays detected in each region of the observable sky. A value of 4, represents a 4σ excess of events, and a value of -2 represents a 2σ deficit. The Observatory is located in the southern hemisphere so a portion of the universe is not observable (upper left hand side). The small excesses and deficits seen are compatible with an isotropic sky distribution given the current number of events detected.

At high enough energies, the deflections in the galactic magnetic fields may be small enough so that the cosmic rays could point back to their sources. A few candidate sources have been proposed by previous experiments but none have been verified with the high statistics available from the Auger experiment [5]. At this time, only an isotropic flux has been detected for energies around $10^{18} - 10^{19}$ eV. In fact only for the highest energies, greater than $10^{19.5}$eV, have there been any data that shows an anisotropic arrival direction of cosmic rays [6].

1.2.3 Mass composition

At lower energies (below 10^{12} eV) the individual particle types can be distinguished and plotted separately. However, at higher energies distinguishing between particle types becomes more difficult and typically only the all-particle spectrum can be made. Particles of different mass will have different observable effects but there is significant overlap in these effects between different nuclei. Only on a statistical basis can the fraction of the flux each nucleus represents can be deduced from the all particle flux. This is a difficult task that requires vast statistics and detailed simulations of air shower properties.

The detected mass composition at Earth will not be the same as composition at the source. Propagation in the interstellar and intergalactic medium will cause nuclear spalation of atomic nuclei. This process is shown in Figure 1.3 for Be^7 interacting with background starlight. Nuclear spalation causes a different observed flux on earth from the flux at the source. If other sources are of information are used then it could be possible to learn about the source composition. The difference between the source and observed flux could then give information about the density of the background starlight or magnetic fields strengths.

These subsections have tried to illustrate the challenges associated with learning about the sources of cosmic rays from the measured properties detectable on earth. Measurements of the energy spectrum, mass composition and arrival direction information all need to be interpreted together. Only with these observables along with information, both experimental and theoretical, on magnetic fields and photon backgrounds from other experiments can we learn about the sources of cosmic rays. While this may seem like an impossible task, much progress has been made in interpreting the few major deviations from a strict power law in the cosmic ray energy spectrum.

$$Be^7 + \gamma_{IR} \rightarrow Li^6 + p$$
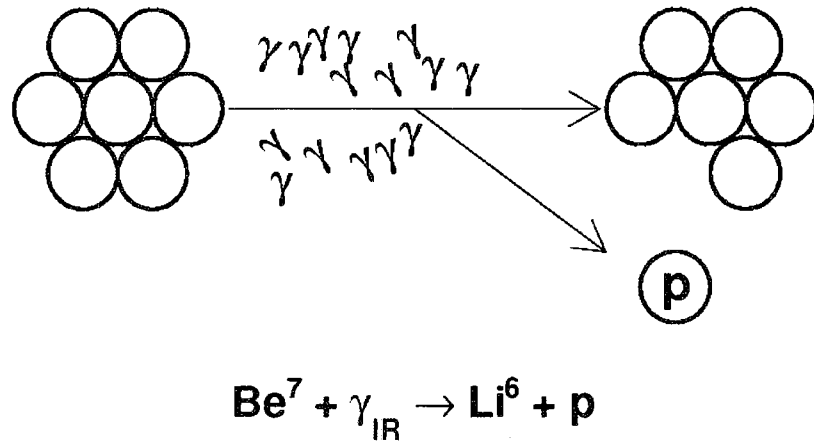
Figure 1.3 As atomic nuclei travel in the interstellar medium, they will interact with background radiation that causes the nuclei to break down. This process is schematically shown.

1.3 Spectral features in the energy spectrum

The spectral breaks in the energy spectrum can be seen more easily if the spectrum is plotted as $E^\gamma J(E)$, where γ is some constant spectral index. When this is done, the small changes in spectral index become apparent. This is common practice in the field, however, this procedure also magnifies any systematic differences in energy assignment between experiments. For energies above 10^{17} eV, typical values for the uncertainty on the energy measurement are 20-30%. A discussion of the difficulties of comparing data from different experiments with such large systematic uncertainties is left for the final chapter where the results from this thesis are presented. For now the systematic differences between experiments are ignored and only the general features are discussed.

In Figure 1.4, the data from several experiments has been scaled by E^γ, where $\gamma = 2.7$ is close to the general spectral index observed over the entire energy range. The first spectral break can be seen in occurring between 10^{15} and 10^{16} eV. This break
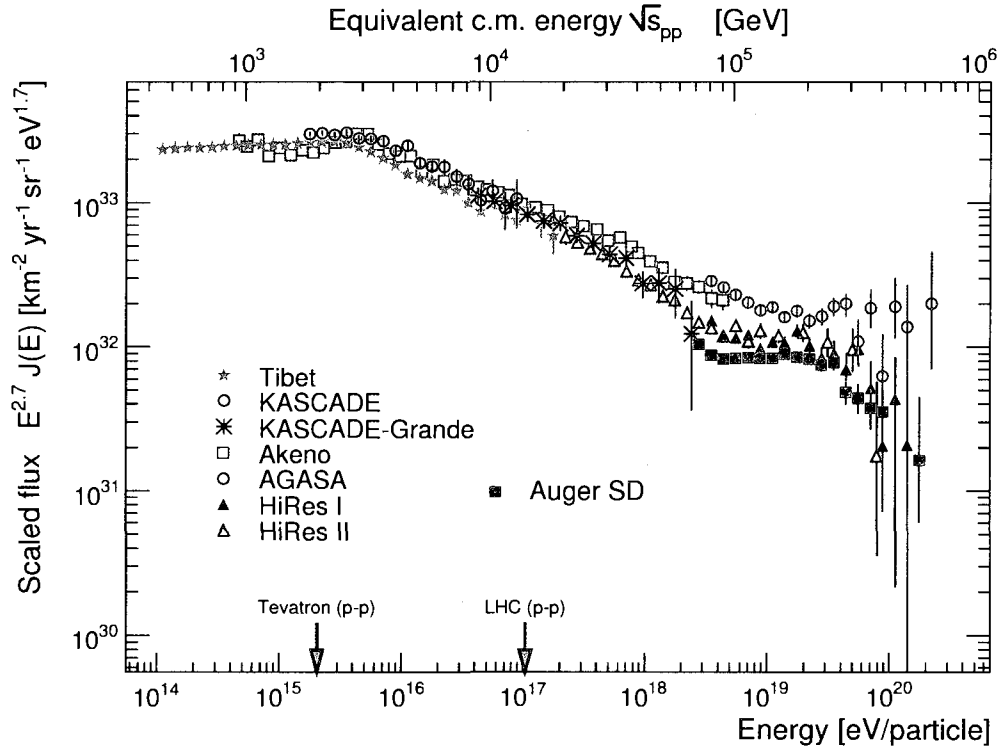
Figure 1.4 The cosmic ray energy flux versus energy, E, multiplied by $E^{2.7}$. In this representation the deviations from the overall power law can be identified clearly. The breaks between $10^{15} - 10^{16}$ eV and between $10^{18} - 10^{19}$ eV are called the knee and ankle respectively.

in the spectrum has been call the *knee* due to the downward bend that is similar to a leg having a bend at the knee.

1.3.1 The knee

It is strongly believed [4] that there must be different strength accelerators responsible for the observed cosmic ray flux. Each accelerator will have a cutoff at its specific maximum acceleration energy. The simplest explanation for a change in spectral index is that one source has turned off and another source has turned on. Each source has a different spectral index, flux normalization and maximum acceleration energy. The

break point is then the energy at which the flux contributions from each source are equal. Below the break point one source dominates, and the other source dominates above. By looking at the energy spectrum alone, the knee could be explained by assuming a two source scenario.

As emphasized in the previous section, all cosmic ray observables need to be interpreted together along with other sources of information about the universe to make sense of a spectral index change. Arrival direction information shows nothing but an isotropic sky for energies from 10^{14} to 10^{16} eV, so little can be learned from the specific directions of each cosmic ray. However, an isotropic sky can put limits on the distribution of sources. The current known bounds on the strength of galactic magnetic fields puts a constraint on the magnitude of the density of sources in the galaxy. Along with arrival direction information the most valuable insight at these energies comes from mass composition studies.

The KASCADE [7] experiment has made attempts to isolate individual nuclei and look at their spectra in the knee region. In Figure 1.5, the scaled differential flux for different elemental groups are displayed. A drop off, or a separate knee, for each nucleus occurs at different energies. The energy for each knee increases roughly proportional to the charge of the elemental group. The Iron group (Fe) shows no knee-like feature and the flux may stay flat beyond the limits of the KASCDE detector (beyond 10^{17} eV, where 1 GeV = 10^9 eV).

In a simplified picture of a cosmic ray accelerator there will be a maximum energy, E_{max}^p to which a proton, with charge $Z = 1$, can be accelerated. The maximum energy that each other type of nucleus can achieve is proportional to the charge, Z, of the particle. For Iron, with $Z = 26$, $E_{max}^{Fe} = 26 \times E_{max}^p$. This implies that given an individual accelerator, the spectrum will continue on past E_{max}^p due to the presence of heavier nuclei. This fact can lead to a larger fraction of heavy nuclei, as compared
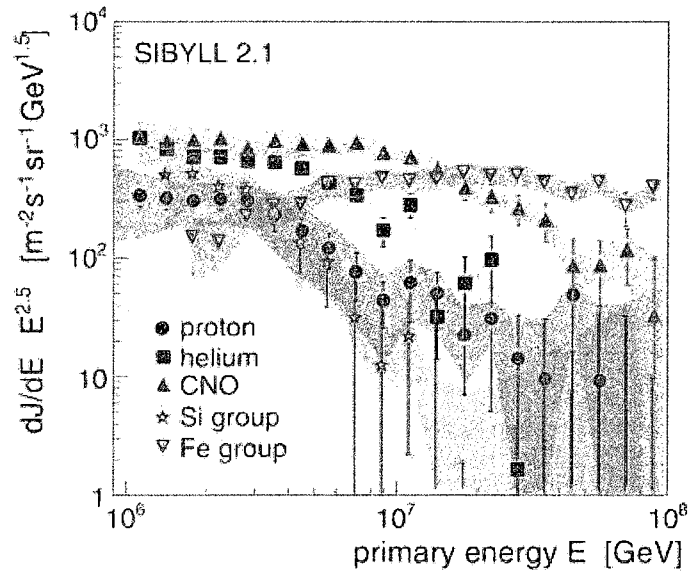
Figure 1.5 The cosmic ray energy spectrum broken up by primary particle type. The fall off in flux is seen at a different location for each particle. These results are from the KASCADE collaboration. 1 GeV $= 10^9$ eV.

to protons, at higher energies. This appears to be the situation at the knee.

Candidate astrophysical sources for cosmic rays are galactic super nova remnants (SNR) [4]. SNR being the site of acceleration agrees with the data from the KAS-CADE experiment reasonably well up to $\approx 10^{17}$ eV. For SNR, E^p_{max} is theorized to be a few times 10^{15} eV, so that E^{Fe}_{max} would then be $\approx 10^{17}$ eV. This implies that above $\approx 10^{17}$ eV the acceleration process theorized for SNR is not strong enough to contribute significantly to the observed cosmic ray flux. A new accelerator candidate is needed. However, to explain cosmic rays up to $\approx 10^{16}$ eV, SNR fit certain criteria necessary to explain all the cosmic ray observables.

1. X-ray and gamma ray observations of near by SNR show signatures that could be constant with π^0 decays. The presence of mesons is a necessary condition for most cosmic ray acceleration theories.

2. SNR have approximately the right density, maximum acceleration energy and particle acceleration efficiency to account for the magnitude of flux below 10^{16} eV.

3. Confinement of cosmic rays is very efficient below 10^{16} eV based on the strength of galactic magnetic fields [4]. Cosmic rays accelerated in the galaxy will remain in the galaxy.

4. The source spectrum from a SNR is theorized to have a power law dependence on the energy, E, such that source flux, $J_{source} \propto E^{2.0}$ for most likely acceleration mechanisms. The observed flux at earth will be modified from the source flux. The magnetic field strength and SNR density are not know very accurately but their estimated values are the correct magnitude to account for the change from a source index of 2.0 to an observed index at Earth of ≈ 2.7 [4].

There is still much theoretical and experimental work needed to confirm that SNR are the definitive sources of cosmic rays up to $\approx 10^{17}$ eV, but the evidence continues to support this hypothesis. The Pierre Auger Observatory does not have sensitivity to cosmic rays at energies around the knee but other next generation experiments are being built to understand the origins of cosmic rays at these energies.

1.3.2 The ankle

The next spectral change in the energy spectrum occurs between 10^{18} and 10^{19} eV. The flux begins to recover to a spectral index, $\gamma \approx 2.7$ (near horizontal on the scaled flux plot in Figure 1.4). This feature is called the ankle since the flux bends upwards just as the ankle does on a leg. The cause of this feature is currently uncertain, but there are two classical scenarios proposed to explain it. Before explaining the

scenarios, there are three key factors that need to be considered when trying to explain the ankle.

1. The sources of the cosmic ray flux must be make a transition from galactic-dominant sources to extragalactic -dominate sources at some energy. Cosmic rays will have sufficiently high energy to escape confinement by magnetic fields in other galaxies at high energies. It is believed that no galactic sources are strong enough to accelerate cosmic rays to the very highest energies.

2. As extragalactic cosmic rays propagate from their sources they will interact with the cosmic microwave background (CMB). Extragalactic protons of energy greater than 10^{18} eV (roughly the energy threshold for the following interaction), that travel large distances, will loose energy due to pair production.

$$p + \gamma_{CMB} \rightarrow p + e^+ + e^- \tag{1.1}$$

Galactic cosmic rays would not travel distances far enough to lose a significant fraction of their energy from this process.

3. The fraction of heavy nuclei in the flux before and after the ankle may be significantly different.

Both theoretical scenarios take these three factors into consideration but in different ways. The first scenario is just the simplistic view of a two-source model; below the ankle one type of source dominates and another type of source dominates above. Unlike the knee, there is a good reason a priori to think the two-source scenario is plausible. No known galactic objects are theorized to be strong enough to accelerate cosmic rays to energies above 10^{20} eV. An extragalactic accelerator is needed for the highest energies. The bounds on the strength of the galactic magnetic field rule out galactic confinement of cosmic rays above $\approx 10^{17}$ eV. With these a priori insights the

two-source scenario is strengthened by interpreting the ankle as the transition from a galactic dominated flux to an extragalactic flux.

To identify the actual sources of the cosmic rays in this scenario more work is needed. Since galactic SNR only have a theoretical maximum acceleration energy of $\approx 10^{17}$ eV, a different type of galactic source would be needed to explain the flux up to the energies near the ankle (between 10^{18} and 10^{19} eV). While a few candidate sources have been proposed, such as X-ray binary stars or rapidly rotating pulsars, no compelling evidence has confirmed these sites as cosmic ray accelerators. This new galactic source of cosmic rays could produce a flux that has a mixed composition of light and heavy nuclei or a pure proton source. Future observations of galactic objects with X-ray and γ-ray telescopes will help constrain the possible source candidates and the expected mass composition from the sources.

The second scenario is that the ankle is caused by propagation effects of extragalactic protons. In this scenario the transition to an extragalactic flux occurs at a lower energy, $\approx 10^{17}$ eV, consistent with SNR being the only source of galactic cosmic rays. Unlike the first scenario, there are not as many uncertain free parameters in this theory. The only assumptions are that the extragalactic cosmic ray flux is composed of 100% protons traveling from uniform sources throughout the universe. As protons travel cosmological distances, energy loses from pair production will occur as long as they have energy above the threshold energy for the interaction (a few times 10^{18} eV). The observed flatting of the flux is then just a manifestation of higher energy cosmic rays losing energy which causes a build up of cosmic rays at a lower energy due to the pair production interaction with the CMB.

Both of these scenarios have a certain elegance in their interpretation. The simplistic idea of a transition from a galactic to an extragalactic flux is compelling. As well as the idea that the ankle is simply propagation effects of protons traveling cos-

mological distances. However, unlike the general agreement between the experimental data, the SNR accelerator theory to explain the knee, the experimental evidence at this time is not strong enough to confirm or deny either scenario about the ankle.

Each ankle scenario has some free parameters that can be modified to fit the experimental data to some degree. But there are some signatures that are crucial to each theory that can be ruled out with more accurate measurements from comic ray observatories. The focus of this thesis is to do just that. This thesis measures the energy spectrum from 10^{18} to 10^{19} eV. More details on the interpretation of the ankle will be discussed in Chapter 7 where the results from this thesis are presented. This work and other preliminary work on mass composition and arrival direction studies are starting to distinguish between these two scenarios.

1.3.3 The end of the cosmic ray energy spectrum

The third major spectral feature happens between 10^{19} and 10^{20} eV where the flux begins to steepen once more. Above a few times 10^{20} eV no cosmic rays have been detected so this steepening is commonly referred to as a suppression or cut-off in the flux. This feature could be interpreted simply as the end of one accelerator and instead of a different accelerator turning on, the cosmic ray flux just ends. This interpretation appears very plausible, since even the most powerful astrophysical accelerators must have some finite limit. However, there is another energy loss threshold at around 6×10^{19} eV that needs to be considered.

Along with pair production losses from interactions with the CMB, protons will loose energy to pion production above around 6×10^{19} eV. This energy loss is commonly called the Greisen-Zatsepin-Kuzmin cutoff or the GZK cutoff. These three physicists predicted this effect in the 1960s, shortly after the CMB was discovered and years before any cosmic rays with these energies had been observed [8] [9]. As

protons travel greater and greater distances they will continue to loose energy until they drop below the pion production threshold.

The suppression of the flux can be interpreted as a propagation effect under two simple assumptions. Firstly, the sources of the highest energy cosmic rays have a spacial distribution the same as the matter density of the universe. Secondly, the power law governing the cosmic ray flux from 3×10^{18} and 3×10^{19} eV continues with the same spectral index of ≈ 2.7 up to at least 10^{21} eV. No assumption has to be made about mass composition because, coincidently, at energies above a few $\times 10^{19}$ eV, atomic nuclei will interact with the CMB photons and lose energy through spalation in the same way they do at lower energies with infrared and starlight background photons. Both a proton dominated and a mixed composition flux will have a similar suppression feature under the above assumptions (Actually, in the mixed composition case, the fraction of Iron nuclei must be relatively high [10]).

This implies that cosmic rays with energies near 10^{20} eV can only reach earth if they travel relatively short distances (from a cosmological point of view). At 10^{20} eV extragalactic protons can only travel distances of around 100 Mpc before losing significant energy from pion production (pc is a parsec, 1 pc $=3.26$ light years, 1 Mpc $=10^{6}$ pc. The disc of the Milky Way galaxy is roughly 30,000 pc in diameter and the Andromeda galaxy is around 0.8 Mpc away). At large distance scales (500 Mpc or more), the matter density in the universe is uniform. However at smaller scales, (100 Mpc) the matter density is lumpy such that if the above assumptions are true, the arrival directions of cosmic rays will start to look anisotropic above 6×10^{19} eV.

More statistics at the highest energies, combined with further results on anisotropy and composition are needed to definitively claim that the suppression of the cosmic ray flux is the result of the GZK effect. The latest results at these energies are discussed in the final chapter.

1.4 Measuring the energy spectrum with the Pierre Auger Observatory

For cosmic rays with energy greater than 10^{17} eV the experimental evidence is plagued by difficulties of the low flux and large systematic uncertainties. As explained in the previous section, no firm conclusions on the origins of cosmic rays have been established at these energies. Properly treating systematic uncertainties from the various experiments is a difficult task that the cosmic ray community has yet to formalize leading to, in some cases, contradictory results.

The most recent experiments were built to combat these difficulties. KASCADE-Grande is currently taking data that will extend the energy spectrum measurements of KASCADE up to 10^{18} eV. The Pierre Auger Observatory (or simply Auger) was built to study cosmic rays with energy above 10^{19} eV, but the work that is presented later in this thesis extends measurements of the energy spectrum to just below 10^{18} eV. Another experiment called TA is currently being built in Utah that will be able to study cosmic rays above 10^{17} eV and provide a good overlap with KASCADE-Grande and Auger [11]. Auger is also planning a low energy enhancement to be built in the next few years. This enhancement, combined with the techniques of this thesis will extend the Auger capabilities down to 10^{17} eV as well.

The Pierre Auger Observatory was designed to measure the properties of the highest energy cosmic rays with greater accuracy and more statistical power than previous measurements. The observatory is over 30 times larger than previous experiments and uses more advanced hardware. Complementary cosmic ray detection techniques have been combined for the first time at a single location. The large size of the collaboration (over 60 institutions and 400 people) has allowed for more emphasis to be placed on monitoring and reducing systematic uncertainties than previous experiments with

only a few dozen members.

While designed for energies greater than 10^{19} eV, the Auger Observatory detects cosmic rays down to $\approx 10^{17}$ eV. The first energy spectrum conservatively only used data where the detection efficiency is 100%. This result was recently published [12] and shows the energy spectrum from $10^{18.45}$ eV and above. By using data where the detection efficiency was 100%, the systematic uncertainty on the flux calculation was minimized. As a result of this lower energy bound, the spectral feature of the ankle is not clearly seen in this result (the red solid square points in Figure 1.4). The work reported in this thesis introduces techniques to measure the flux and energy of the cosmic rays detected below the energy threshold of $10^{18.45}$ eV while maintaining low systematic uncertainties. Extending the spectrum down to energies below 10^{18} eV allows for the complete observation both the ankle and the suppression of the flux.

The ability to observe multiple spectral features with the same observatory is very important. Due to the difficulties in measuring the low flux of cosmic rays above 10^{17} eV, experiments typically have large systematic uncertainties (20-30%) on the energy and sometimes equally large uncertainties on the flux measurement. This can be seen in the larger spread between the data points from different experiments in Figure 1.4. The relative data uncertainties from data point to data point inside each experiment are typically much smaller (a few percent). If multiple spectral features are seen by one experiment, then the relative energy difference between the features is known more accurately than the absolute energies of the features. The relative energy difference between the ankle and the suppression can be used to rule out certain possible source scenarios.

This chapter has served as an introduction to ultra high energy cosmic rays and has provided motivation for studying the spectral features in the energy spectrum. The next chapter gives an introduction to how cosmic rays are detected in general.

Chapter 3 discusses how the properties of the cosmic rays, such as energy and arrival direction, are reconstructed at the Pierre Auger Observatory. The Auger Observatory is putting much effort in minimizing systematic uncertainties. Chapter 4 outlines the details of one such effort of which I have been a leading member. Chapters 5 and 6 introduce the techniques I developed to use the Auger data at energies lower than reported in [12].

The final chapter, shows the resulting energy spectrum from $10^{18} - 10^{19}$ eV and discusses the associated uncertainties. Two past experiments had made measurements of the cosmic ray energy spectrum from 10^{18}-10^{20} eV. Both these experiments are finished taking data and their results had contradictory claims about the spectral features in the energy spectrum. With a little over four years of data talking, the Auger Observatory has already detected more cosmic rays above 10^{18} eV than both these experiments combined. The difficulties of comparison with other experiments and possible astrophysical implications of the measured energy spectrum of cosmic rays from 10^{18}-10^{20} eV are also discussed in the final chapter.

Chapter 2

Cosmic Ray Detection

2.1 Introduction

Many detection techniques of subatomic particles have been invented during the rather short history of particle physics. All the techniques involve detecting macroscopic effects induced in matter by interactions with the subatomic particles. A calorimeter is one such detection instrument. The word calorimeter comes originally from devices that were used to measure the heat released in a chemical reaction (calor means heat in Latin). However, in particle physics calorimeters are used to measure the energies of a subatomic particles. When a particle enters a calorimeter it will interact with an atom inside the detector. This interaction initiates a shower of secondary particles that share the initial energy of the primary particle. The secondary particles in the shower will then either interact again or will be absorbed by the material in the calorimeter. With appropriate detectors monitoring the calorimeter, the number of secondary particles as well as how much the shower spreads laterally can be deduced. Both the number of the particles in the shower and the lateral spread, are related to the energy of the primary particle. More detailed analysis of the shower

can extract other information such as the type of primary particle and details of the primary interaction.

Typically, calorimeters are composed of a dense material that provides many targets for the incident particle to interact with. Interlaced with the dense material is some sort of gas or wire chamber that can detect certain properties of the secondary particles. Liquid and plastic materials are also sometimes used. Depending on the type of materials in the calorimeter, different information is easier to extract from the shower of secondary particles. Large particle accelerator experiments may have more than one type of calorimeter in use because different material calorimeters work better with different types of particles. There is much active research in building more sensitive and different types of calorimeters for use in particle physics experiments.

Cosmic rays can be directly detected by placing calorimeters in high altitude balloons or on satellites. These devices can observe the cosmic ray before it interacts with a molecule in the atmosphere. This is only practical for low energy cosmic rays because the flux of cosmic rays drops off very steeply with energy. Above $\approx 10^{14}$ eV observatories with large acceptances are needed. Building larger and larger balloon or satellite observatories becomes cost prohibitive. Fortunately, indirect measurements have been developed for cosmic rays above 10^{14} eV. When a cosmic ray with these energies interacts with the atmosphere, the predicted large shower of secondary particles can be detected in the atmosphere. The atmosphere acts as a calorimeter and is observed with ground based instruments. To monitor large volumes of atmosphere, ground based observatories can be built relatively cheaply. In Figure 2.1, an illustration of a shower in the atmosphere from a cosmic ray is shown.
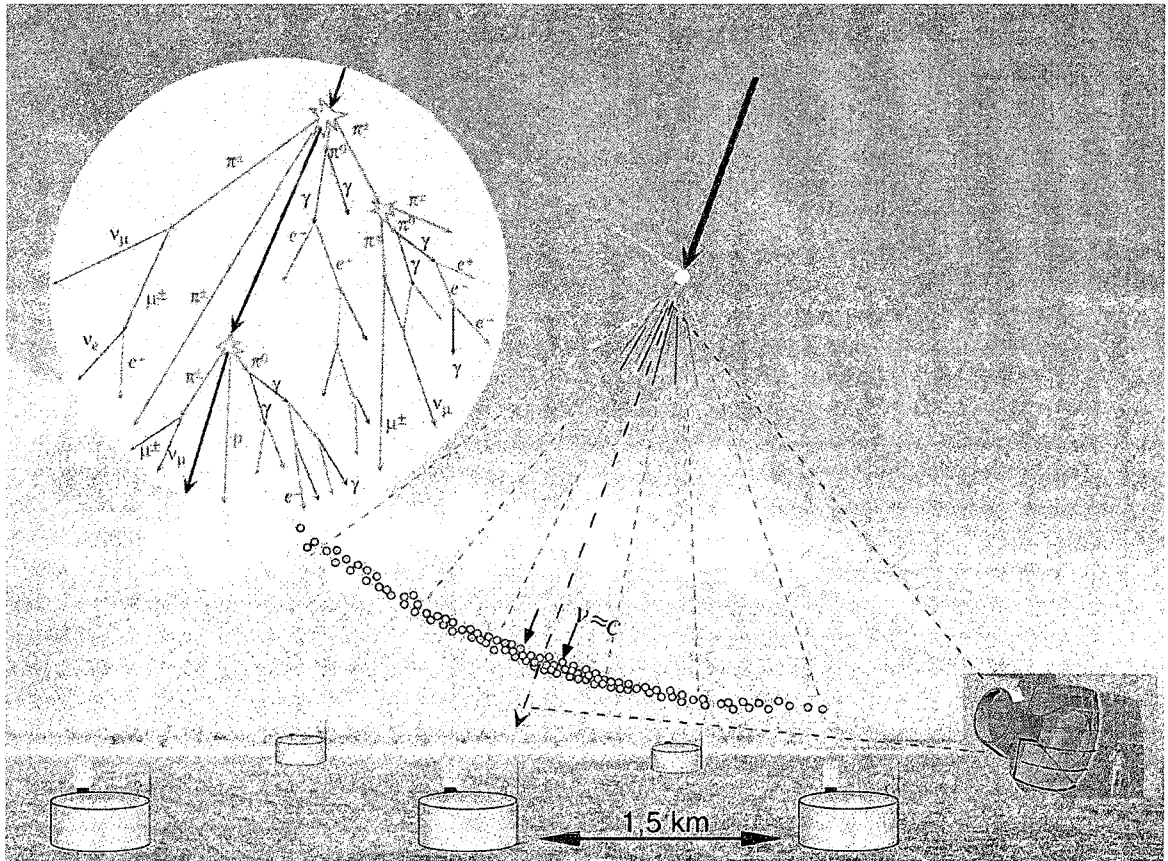
Figure 2.1 The atmosphere acts as a calorimeter for cosmic rays. When a cosmic ray enters the atmosphere, it will interact and start a cascade, or shower of secondary particles. For cosmic rays above 10^{14} eV the shower of secondary particles in the atmosphere can be detected from ground based observatories.

2.2 Extensive Air Showers

Different types of cosmic rays, including photons, nucleons, nuclei or neutrinos, will induce air showers with different characteristics. The showers from nucleons and different nuclei have similar enough properties that make distinguishing the composition of the primary particle on an event by event basis impossible. However, determination on a statistical level is possible and is a major challenge of current cosmic ray detectors. Photon-induced showers have more distinct features. At the energies near the ankle in the cosmic ray energy spectrum ($\approx 10^{18.5}$ eV), the limit on the fraction of photon cosmic rays is less than 1% [13].

Neutrinos at these energies can also be distinguished from hadron showers. Neutrinos have a very low cross section with matter so that they can travel through large volumes of matter (earth and/or atmosphere) before initiating a shower. A neutrino could travel through a portion of the earth such that it will initiate an air shower that will appear to be traveling up from the ground rather than down from the atmosphere. This makes identifying neutrinos possible since no other known particle could have an upward going shower. Nuclei and photon induced air showers are observed by looking for showers coming downwards in the atmosphere only. However, no neutrinos have yet to be found at the highest energies. The Auger collaboration has recently released a limit on the tau-neutrino flux [14].

A primary cosmic ray initiates a shower of secondary particles based on the initial interaction of the cosmic ray with a molecule in the atmosphere. This initial reaction occurs after the primary particle travels through enough matter to interact based on the relevant cross section. At high energies the interactions with the atmosphere all take place inside of the troposphere, which extends from ground level to 10-20 km above the surface of the earth. This region of the atmosphere is where all the weather

we experience on earth comes from, but it still can be adequately described based on average quantities such as temperature and pressure. The approximation of an exponential dependence of atmospheric density on altitude is valid in the troposphere [15].

The first interaction divides the energy of the primary cosmic ray into the secondary particles that continue to travel in the atmosphere. After a sufficient distance these secondary particles will interact with more air molecules, further dividing the initial energy into more secondary particles. The initial energy eventually gets divided among the secondary particles so that the average energy per particle is below a critical energy. Below this critical energy the secondary particles begin to be absorbed by the atmosphere, rather than creating more particles from interactions. The critical energy depends on the secondary particle type.

Since the atmospheric density changes with altitude the measure of slant depth is more appropriate than altitude. Slant depth is a measure of the amount of matter traversed and is measured in units of grams per square centimeter, $[\frac{g}{cm^2}]$ (details on how slant depth is calculated are given in the next chapter). The symbol commonly used for slant depth is X, where X_1 denotes the depth of first interaction and X_{max} the depth of shower maximal development (meaning the maximum number of particles in existence). Before X_{max}, most of the particles are above the critical energy so that absorption is low. After X_{max}, new secondary particles are still being created but, absorption starts to take over and the number of secondary particles in the shower declines.

At energies around 10^{18} eV over 10^9 particles have been created by the time the shower reaches detection level. Even though the entire shower can be contained in a finite volume of atmosphere accurate detection of all 10^9 particles is generally impossible. Cosmic ray detectors are generally designed to measure only the most prominent signatures, or the average characteristics of an extensive air shower.

The typical number of particles of each type as a function of slant depth are shown in Figure 2.2. This figure is based on a simulated shower with energy 10^{20} eV. Since such large numbers of particles are involved, it is useful to think of the shower as consisting of three components; the hadronic, the muonic and the electromagnetic. The hadronic component consists of baryons and mesons. The muonic component consists of positive and negative muons. The electromagnetic component consists of electrons, positions and photons. Each component has its own development characteristics and detectable signatures that can be observed with different types of detectors. In general these characteristics and signatures depend on the both the energy and type of cosmic ray primary. It should be noted that neutrinos are also created in the air shower but the fraction of energy transfered into neutrinos is small and is generally impossible to detect. The neutrino component will be ignored in the following discussion.
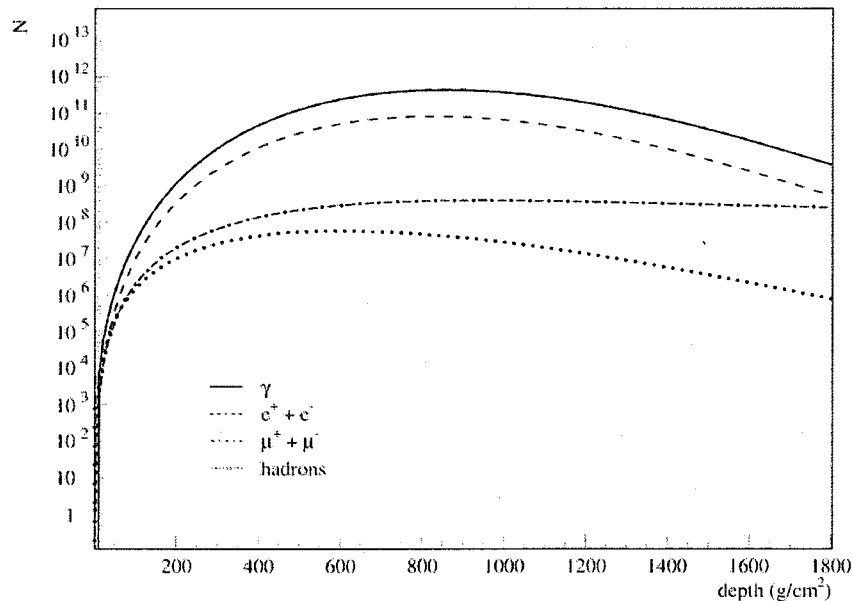
figure from [16]

Figure 2.2 The number of particles of each type in an air shower of energy 10^{20} eV. This is based on a simulated shower.

2.2.1   Hadronic components of an EAS

The first interaction of the primary proton and an air molecule will create many mesons and baryons. These particles will share the energy of the primary cosmic ray and have momenta approximately parallel so they will continue to travel down in the atmosphere. The new particles will either decay or interact with another air molecule. To a first approximation only pions and to a lesser extent kaons are created in significant numbers [17]. The kaons that decay before interacting will create muons, neutrinos and more pions. Positive, negative and neutral pions will be created in equal numbers [17]. The short lived neutral pions (lifetime of 8.4×10^{-17} s compared to 2.6×10^{-8} s for changed pions) will decay into two photons before interacting while the charged pions that do not interact again will decay into muons and neutrinos. All subsequent interactions and decays act as sources for the other components of the air shower. After only a few interaction lengths (50-80 $\frac{g}{cm^2}$) the fraction of hadrons becomes small. At shower maximum, the hadronic part of the air shower represents only a small fraction of the total number of particles (see Figure 2.2) and less than 1% of the total energy remaining in the shower.

The hadrons can be thought to reside only along the axis of the shower. At ground level there will be some lateral spread, maybe 100-200 m. This is small compared to either the muonic or the electromagnetic components which extend out for several kilometers [2]. The hadronic core of the shower is surround by the muonic and electromagnetic components, and by the time it reaches ground level, separating the three components close to the shower core would require a complex detector with multiple calorimeters.

The number hadrons at any stage of the shower could be used to learn about the primary energy, mass composition or the first interaction that initiated the shower [17]. However, the hadrons that survive at detection level are hard to separate from

the other components. Currently, for energies above 10^{18} eV, there are no cosmic ray detectors capable of probing the hadronic component with enough sensitivity to use the information from the shower core.

2.2.2 Muonic components of an EAS

As mentioned above muons are created from the decay products of the various types of mesons found in the hadronic component of the shower. In the early stages of shower development π^+, π^- and π^0 particles are created in roughly equal numbers and approximately share the energy equally as well [17]. The pions will either have a greater probability to decay or interact depending on their energy. The lifetime of a charged pion (2.6×10^{-8} s) is sufficiently long to allow high energy pions to interact before decaying. Once the average pion energy drops below a critical energy, the decay process dominates and the creation of many muons begins. Regardless of the π^0 energy, decay dominates over interaction and two photons are created that feed the electromagnetic component.

Initially, when the average pion energy is above E_{crit}, one third of hadrons get converted into electromagnetic energy and two thirds remain hadronic. When the average energy is below the pion critical energy, a substantial fraction of the total energy is already in the electromagnetic component from all the π^0 decays. At this stage of shower development approximately 10% of the total energy is remaining in the hadronic component. As the charged pions begin to decay the muonic component rapidly is fed with this remaining energy. Since pion decay dominates over interaction, the hadronic component creates fewer and fewer pions. Charged pions continue to feed the muon component throughout the shower development, but by far and away the majority of the muons are created when the average particle energy is just below E_{crit} which is 115 GeV [2].

The muonic component mainly loses energy from decay and only weakly by ionization. Most of the muons created have high enough energy to traverse many tens of kilometers before decaying (the lifetime of a muon is 2.2×10^{-6} s). All but the lowest energy muons reach ground level before they decay. The total number of muons in a shower is a quantity that is proportional to the primary cosmic ray energy.

2.2.3 Electromagnetic components of an EAS

The electromagnetic component consists of electrons, positrons and photons. Almost all the initial energy eventually ends up in this component. As the shower develops, there is a continual influx of energy from the hadron component through π^0 decays. The number of electromagnetic particles develops fast. Once energy in the shower is transfered to the electromagnetic component it only leaves through absorption in the atmosphere. No energy is transfered to other components. Each high energy photon from a π^0 decay will undergo pair production creating an electron and a positron. Pair production will occur as long as the photon has more energy then the combined mass of the electron and positron, (≈ 1.02 MeV). The electron and positron will undergo interactions that only create electromagnetic particles. The end result is that from each photon, an individual electromagnetic shower of particles is created.

The shower does not continue indefinitely because below a critical energy, ionization energy losses become more important than new particle production processes. In this way the atmosphere acts as a calorimeter, absorbing the energy of the electromagnetic shower. The general behavior of the electromagnetic component is to quickly rise to a maximum and then almost as quickly decrease. The electromagnetic shower maximum, X_{max}, is a well defined observable that contains valuable information about the shower.

2.3 Types of Detectors

2.3.1 Fluorescence Detectors

The electromagnetic component of the EAS excites the nitrogen in the atmosphere. After excitation the nitrogen will quickly fall back to its ground state and a UV photon is emitted. This fluorescence emission is isotropic and is concentrated around the shower axis where the density of particles is highest. These photons can be detected by sensitive telescopes on clear moonless nights. The fluorescence yield of nitrogen is composed of several distinct lines corresponding to specific orbital excitation states. The most dominant lines are in the region from 280-440 nm [18]. Accurate knowledge of the fluorescence yield is needed to convert the number of detected photons to number of charged particles in the EAS. The number of particles can then be used to find the amount of energy deposited in the atmosphere. This calorimetric measure of the energy deposited by the electromagnetic component is completely independent of any unknown hadronic interactions at high energy. At typical detection depths, roughly 90-95% of total energy in the shower has been transfered to the electromagnetic component. To obtain the full energy of the shower, a correction for the energy still residing in the other components is needed.

2.3.2 Surface Detectors

A surface detector array samples the lateral spread of an EAS at one point in its development, namely, when the shower intersects the ground. Due to the lateral spread of the various components, an EAS appears to move through the atmosphere like a thin disc. The thickness of this disc arises from the slight differences in speeds and from the differences in track lengths traveled for the various secondary particles. The thickness is small (hundreds of meters) compared to the lateral spread (tens of

kilometers) at ground level so the shower front can be thought to have no thickness. The timing of the shower front passing ground level can be used to find the arrival direction of the primary cosmic ray and the shape of the shower front as well.

The direct observables in a SD are the arrival direction and the lateral density of particles. The lateral density can be related to the primary energy. Depending on the type of surface detector, both the electromagnetic and muonic lateral density can be measured. These components have differing development characteristics. The number of electromagnetic particles rises quickly before shower maximum and drops just as quickly afterwards. Whereas after shower maximum, the muonic component has roughly constant number of particles (see Figure 2.3). Having sensitivity to both these components can yield information on the primary particle type (i.e. proton, iron nucleus, photon, etc.).
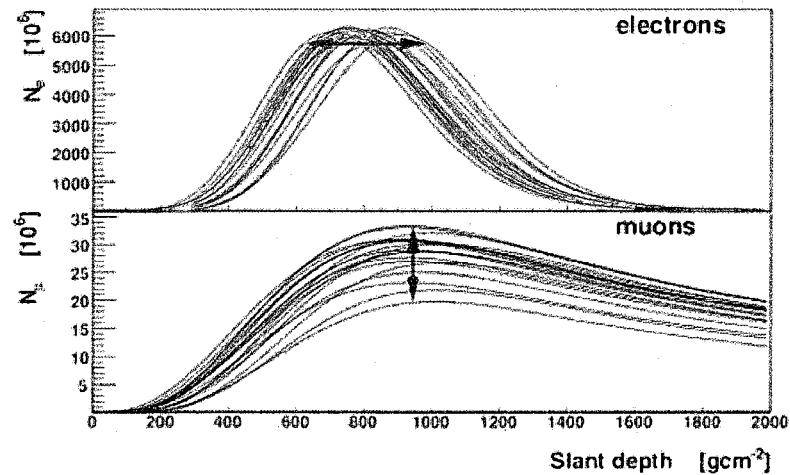
figure from [19]

Figure 2.3 The number of particles in the muonic and electromagnetic components as a function of shower depth. Each of the 15 lines are for a simulated showers with energy of 10^{19} eV.

2.3.3 Other types of detectors

There are other types of cosmic rays detectors besides surface and fluorescence detectors. The Cerenkov light that is produced in the forward cone along the shower axis can also be detected by sensitive telescopes. The Cerenkov light produced is a continuous spectrum of UV and visible light that is much broader than the fluorescence light produced in the air shower (more details on the specific spectrum of fluorescence light will be discussed in the next chapter). Gamma ray astronomy makes use of Cerenkov telescopes in the TeV range. Radio emission, from the numerous electron positron pairs in an air shower, is also a signature of cosmic ray air showers. Radio detection is very difficult and has not proved to be practical. However, due to technological advancements in electronics, radio detection has see a resurgence in recent years and prototype radio detectors are being deployed at various observatories, including the Pierre Auger Observatory [20].

This chapter has provided a brief introduction to the detection of cosmic ray air showers. The main points that should be taken away from this chapter are that the general properties of primary cosmic rays can be deduced from the extensive air shower that is produced in the atmosphere. The next chapter will discuss the details of how the Pierre Auger Observatory reconstructs the air showers of cosmic rays with energies greater than 10^{18} eV.

Chapter 3

The Pierre Auger Observatory

3.1 Introduction

There are a few key items to remember from the phenomenology introduced in the previous chapter. An air shower moves through the atmosphere as a thin disc of particles traveling at nearly the speed of light. The density of particles is cylindrically symmetric around the shower axis. As the shower propagates in the atmosphere from first interaction, X_1, through shower maximum, X_{max} and finally to detection level at ground, X_{ground}, the lateral distribution of particles is changing (see Figure 3.1).

As an extensive air shower (EAS) moves through the atmosphere the isotropically emitted fluorescence light allows the longitudinal development to be observed by a fluorescence detector. Along with the fluorescence light, Cerenkov light is emitted that has to be taken into account in reconstruction of the longitudinal profile. Almost 100% of the fluorescence light is emitted from the shower axis because the density of particles drops steeply away from the shower axis. The lateral density of the particles is observed when the shower front reaches the ground where it can be detected by a surface detector. The density of particles in this disc is described by the lateral
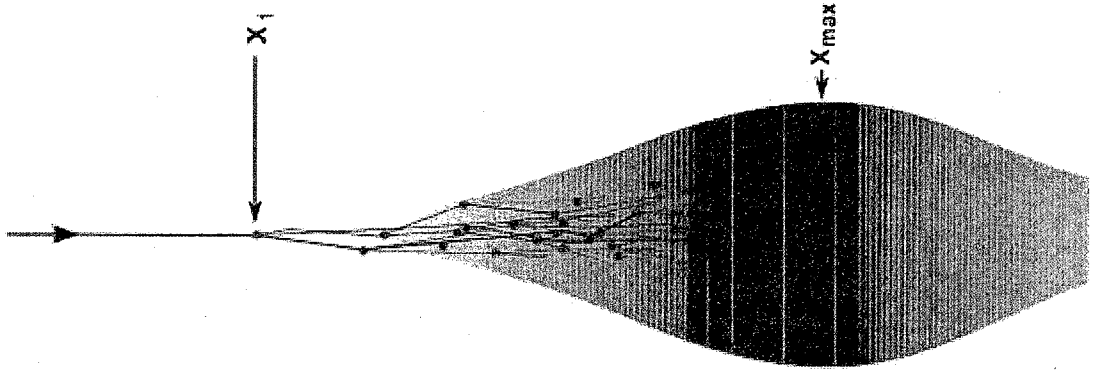
Figure 3.1 Schematic view of air shower development.

distribution function (LDF).

The Pierre Auger Observatory combines two air shower detection methods at one site, a fluorescence detector (FD) and a surface detector (SD). Both a FD and a SD can provide a measurement of the shower core location, arrival direction and energy. The Auger SD consists of approximately 1600 water Cerenkov detectors (WCDs) spread over 3000 km^2 over-viewed by four FD sites. By combining both a SD and a FD at one location the observatory is able to look at the subset of showers that are detected by both techniques simultaneously (see Figure 3.2). These events are called hybrid events and they can be scrutinized in great detail. The two detectors observe complementary aspects of the air shower. The SD measures the density of particles at one depth of shower development and is sensitive to the muonic and electromagnetic components of the shower. The FD is sensitive to the electromagnetic component only. The FD views the entire longitudinal profile of the shower and makes a calorimetric energy measurement. The hybrid data are valuable for cross checking systematic uncertainties in both detectors and are very important to the final analysis (Chapters 5 and 6). The main reconstruction details of each detector are introduced in this

chapter to provide the needed background for the later analysis work.
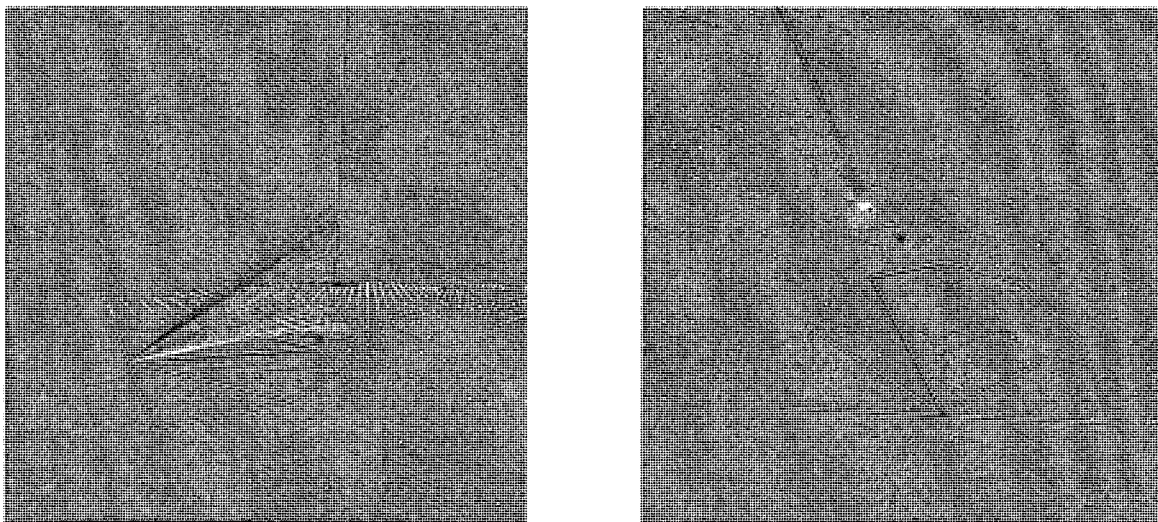
Figure 3.2 Side view (left) and birds eye view (right) of a near vertical hybrid event of energy 10^{19} eV. The white dots represent individual WCD locations and the magenta wedges show the field of view of the six telescopes in one of the FD locations. The WCDs are colored according to the timing of their recorded signal and the size indicates the magnitude of the signal. The red line shows the reconstructed shower axis from the FD. The amplitude and the timing of the fluorescence light induced by the air shower is indicated by the size and color of the circles on the shower axis.

The Auger Observatory is located at an altitude such that all air showers with energy greater the 10^{17} eV are detected after shower maximum. To relate the altitude with slant depth, the zenith angle of the air shower is needed. As the zenith angle goes from zero degrees (vertical shower) to 90 degrees (horizontal shower), the amount of atmospheric matter traversed increases by the inverse of the cosine of the zenith angle, θ. Thus the slant depth at different angles at ground altitude, X^{θ}_{ground} increases for larger angles.

$$X^{\theta}_{ground} = \frac{X^{0}_{ground}}{cos(\theta)} \qquad (3.1)$$

The slant depth for vertical showers, X^{0}_{ground} is 870 $\frac{g}{cm^2}$. For showers with zenith angle of 60 degrees the slant depth grows by a factor of two, $X^{60}_{ground} = 1740 \frac{g}{cm^2}$.

3.2 Surface detector reconstruction

The SD is the workhorse of air shower detection because it operates twenty-four hours a day 365 days a year with little dead time. This is in contrast to a FD that only can operate on clear moonless nights ($\approx 10\%$ duty time). The Auger SD is a sparse array of 1600 water Cerenkov detectors placed in a 3000 km^2 triangular grid with 1.5 km spacing. Each WCD is a cylindrical water tank 1.55 m tall with a 10 m^2 base (see Figure 3.3). The tank has interior lining made from diffusively reflective material and filled with 12 metric tons of purified water. Three 9 inch Photonis XP1805 photomultiplier tubes (PMTs) collect the Cerenkov light due to air shower secondary particles traveling faster than light in the purified water. Signals are digitized at 40 MHz by 10 bit Flash Analog-to-Digital Converter (FADC) with a dynamic range extending from a few to about 10^5 photoelectrons. Each WCD has a wireless communications link to the central campus as well as global positioning system (GPS) unit. The electronics need 10 Watts of power which is provided by two 12 Volt batteries that are recharged by solar panels (see Figure 3.3). The reconstruction methods for the arrival direction, shower core location and energy are described in the following subsections.

3.2.1 The signal in a WCD

Each surface detector event will trigger multiple WCDs providing both signal amplitude and timing information. The signal detected in each WCD depends on the density of particles in the shower front at that location. The signals close to the shower core will be large and are predominately due to electromagnetic particles (positions, electrons and photons). At large core distances (greater than 1 km) muons become the dominate source of the signal. When an electron or positron from the shower enters the WCD a burst of Cerenkov light will be produced. The particle will
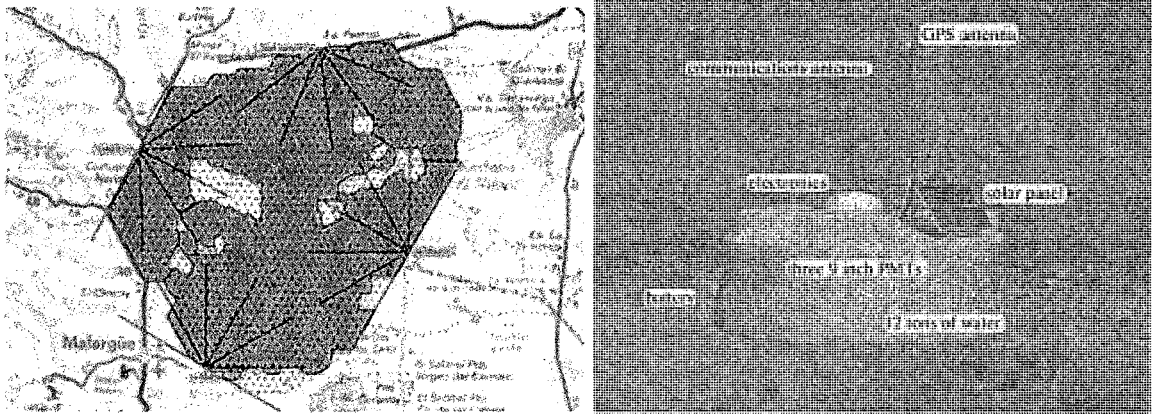
Figure 3.3 The surface detector is comprised of 1600 individual WCDs arranged on a hexagonal grid with surface area of 3000 km^2 (left). Each WCD (right) is a self contained unit with 3 PMTs, electronics, communications and solar power. There are four fluorescence detector sites overlooking the SD.

also be losing energy due to ionization. The particle will only traverse a few tens of centimeters in the water before Cerenkov light will stop being produced. When a photon enters a WCD, an electromagnetic shower will start that also produces a burst of Cerenkov light before quickly losing all its energy. Due to the high density of electromagnetic particles the Cerenkov light is generally detected in bunches that broaden out the signal in time to a few hundreds of nanoseconds.

Muons from air shower detected by the SD have a mean energy of around 1 GeV. Unlike the electromagnetic particles, these muons will pass all the way through the WCD producing a constant track of Cerenkov light. The signal from a muon will be comparatively sharp in time, lasting 50-100 ns. The amount of detected light from a muon in a WCD is roughly independent of muon energy. The main factor is the track length of the muon in the water. If a muon is traveling vertically then the track length is 1.5 m. If the muon is moving at a different angle with respect to the WCD then the track length can vary from 0 to 2 m (the low track lengths result from muons clipping the corners of a WCD). Figure 3.4 shows the FADC traces of three different
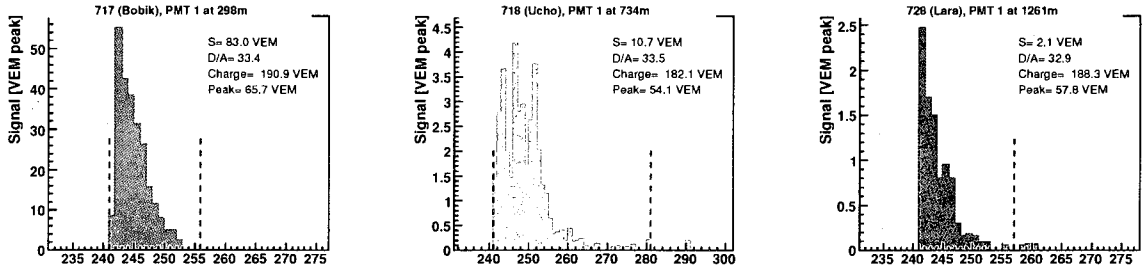
Figure 3.4 This signal in a WCD 300 m from the shower core (left), at 750 m (center) and at 1250 m (right) from a 10^{18} eV shower. Each WCD is given an unique name and identification number (i.e. the station on the left is named Bobik and numbered 717).

WCDs at different distances from the same event.

The signal produced from a vertical muon passing through a WCD is used as a reference value and all other signals can be compared to this. Thus, all signals can be converted into what is called a vertical equivalent muon (VEM) unit. The VEM is the unit used by Auger for all the signals in the SD stations. Small differences in the sensitivity of each WCD exist due to differing PMT gains, electronic amplification gain, water quality, optical couplings, etc. Seasonal effects, like temperature and pressure also can change the response of each WCD. To account for these differences, the response of each WCD to the constant background of muons is monitored continuously [21]. The background muons are from the relatively high and constant flux of lower energy showers incident on the atmosphere. This monitoring provides a relative calibration for each WCD. The electronic gains for each individual PMT in each WCD are adjusted to compensate for any differences in response so that the signals recorded from all three PMTs are equally responsive to low and high signals.

The signal is recorded by a FADC with 25 ns time bins. The total signal is used in the shower reconstruction, by integrating the signal from the FADC trace. The start time is defined by the time bin where the integrated signal is ten percent of the
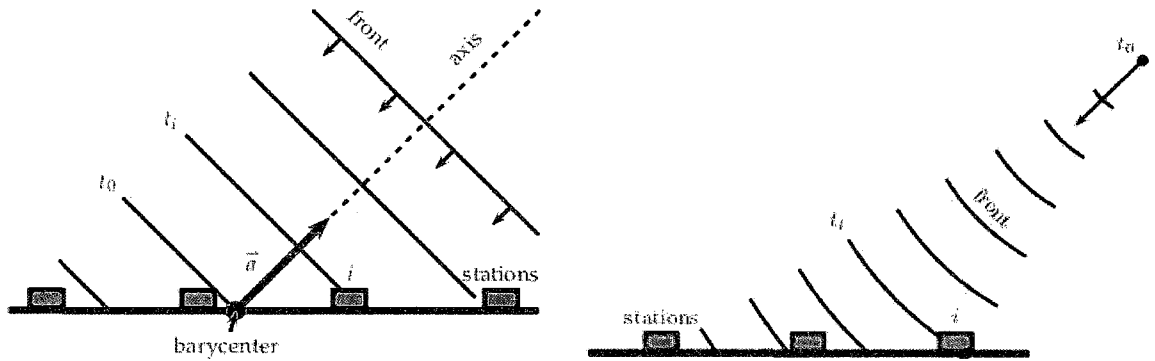
Figure 3.5 The shower plane is defined by the timing from the individual WCD stations. A flat plane fit is a first approximation but a curved shower front is more representative of the true shape.

total signal, t_{10}. The start time and integrated signal are the two parameters used for the reconstruction of the geometry and energy of the air shower.

3.2.2   The shower front

The shower front can be approximated as a plane perpendicular to the shower axis. The geometry of the shower plane with respect to the local ground coordinates can be found by using the timing of the signals detected in the WCDs. The barycenter of the shower front is a first estimate of the impact location of the shower axis and the ground. This is called the shower core location. The barycenter is calculated by the weighted average of the triggered stations in the event. After the first approximation of a plane shower front, a shower core location is found. Based on this shower core, another fit is done with a more realistic curved shower front. From this procedure the azimuth, ϕ, and zenith θ, angles are found.

3.2.3   The Lateral Distribution Function

The lateral distribution of shower particles changes with slant depth as the shower develops in the atmosphere. At a given stage in shower development the lateral distribution function (LDF) can be described as a one dimensional function of distance. The density is high near the center of the disc and falls off steeply with distance. For experimental purposes the actual density of particles is rarely used in large surface detectors. Instead, the expected signal in the detector as a function of distance is used. The LDF used by Auger is a modified NKG [2] function (named after the three physicists who developed it Nishimura, Kamata and Greisen).

$$S(r) = S_{1000}\left(\frac{r}{1000}\right)^{\beta}\left(\frac{r+700}{1700}\right)^{\beta} \tag{3.2}$$

Equation 3.2 describes the expected signal, S(r), in a WCD at a distance, r, from the axis of the shower. The parameters, β and S_{1000} are dependent on the energy (E), composition (A) and zenith angle (θ) of the shower. Recalling from chapter 2, the composition refers to the primary particle type and as discussed above, zenith angle relates to slant depth of shower detection. The Auger LDF has been tailored to the specific types of signals detected by the Auger WCDs. The parameter S_{1000} can be though of as the signal a WCD would have recorded at 1000 m from the shower axis.

$$S_{1000} = S_{1000}(E, A, \theta) \tag{3.3}$$

$$\beta = \beta(E, A, \theta) \tag{3.4}$$

The signals and locations of the WCDs are used to fit the above LDF for the core location, (x,y), S_{1000} and β. The fit is done iteratively first with only the shower core, and S_{1000} left to float. Then if there are more than 5 WCDs in the event, a further

fit is done with β also allowed to float. The resulting LDF parameters, S_{1000}, β, and core location coordinates (x,y) together with the angles from the shower front fit, ϕ (azimuth) and θ (zenith), completely describe the event. The arrival direction in the sky and the lateral density of particles in the air shower are now determined.

3.2.4 The shower size S_{1000}

The integral of the LDF is proportional to the total number of particles in the air shower at that specific stage of development. This integral is only proportional because rather than counting the total number of particles, the total signal from all particles is being estimated. As noted above there are different weights assigned to different types of particles depending on particle type and track length. This is not a problem because the integral of the total signal still is proportional to the energy of the shower. However, a problem does arise because the true functional form of the LDF is not known very well.

The functional form in equation 3.2 is only an approximation and other functions can describe the Auger data equally well. While other functional forms do exist, they all end up having a slope parameter similar to β. Larger uncertainties arise from lack of knowledge of the slope parameter β in each functional form than from the uncertainties associated with different functional forms. The large intrinsic fluctuations in the shower development lead to large fluctuations in the total number of particles detected by the SD at ground level. Hillas [22] has made the argument that the signal at a far distance from the shower core suffers from much smaller intrinsic fluctuations. So rather than calculating the integral of the LDF, only the density at one specific location is used to characterize the size of each shower. This density can then be calibrated to the primary energy of the cosmic ray.

The particular distance from the core that minimizes the uncertainties due to lack

of knowledge of the slope parameter β depends on the layout of the surface detector being used. It has been demonstrated that for the Auger SD the distance of 1000 m has this effect [23]. For this reason the signal at 1000 m, S_{1000}, is used to estimate the lateral spread of the shower at ground level. With aid from the hybrid events or simulations, S_{1000} can be related to shower energy. The procedure using hybrid events for this energy calibration will be shown explicitly in chapter 5.

3.3   FD and hybrid reconstruction

As the shower propagates through the atmosphere, the charged electromagnetic particles excite the molecules of nitrogen. After excitation, the nitrogen will quickly fall back ($\approx$ 40 ns) to its ground state and emit UV photons. This fluorescence emission is isotropic and is concentrated around the shower axis where the highest density of particles is. The absolute intensity of fluorescence light is large enough to be detected above the night sky background by sensitive telescopes, but only on clear moonless nights. When the moon is more than half full, the background light drowns out the fluorescence light from air showers. The fluorescence yield of nitrogen is composed of several distinct lines corresponding to specific orbital excitation states. The dominant lines are seen in Figure 3.6. Accurate knowledge of the fluorescence yield is needed to convert the number of detected photons to number of charged particles in the EAS. The number of particles can then be used to find the amount of energy deposited in the atmosphere by the electromagnetic component of the air shower. The electromagnetic energy is roughly 90-95% of the total energy so a correction is only needed for the remaining 5-10%. The correction is different depending on the primary composition assumed and for different simulations used. This correction is very model dependent but the differences in the models introduces a relatively small uncertainty
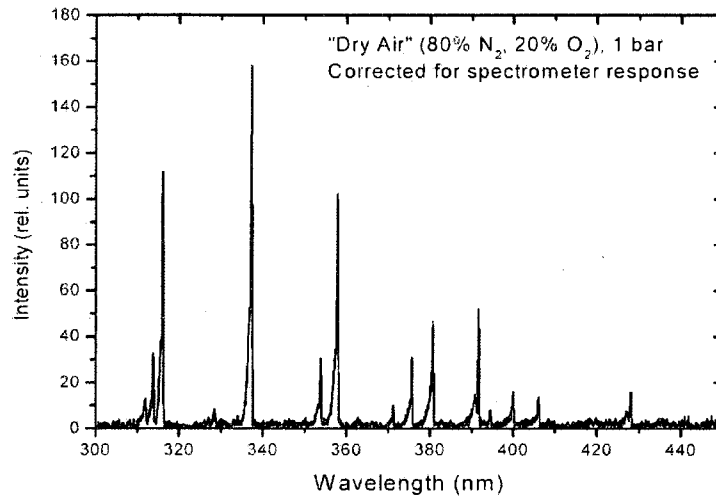
Figure 3.6 The fluorescence yield of nitrogen has several distinct lines.

on the total energy of the shower. There is an estimated 4% uncertainty assigned for this correction. The total systematic uncertainty on the FD energy measurement is 22% (more details will be given in Chapter 7).

Each of the four FD sites contains six fluorescence telescopes each with a field of view of a 30 degrees in azimuth and 28.6 degrees in elevation (see figure 3.7). This is a total of 180 degrees in azimuth and 28.6 degrees in elevation for each site. There is a UV filter ($\approx$ 300-410 nm) on each 1.10 m radius aperture to eliminate background light. A 3.5 m $\times$ 3.5 m spherical mirror collects and focuses the light onto a camera. Schmidt optics, are used to reduce spherical aberrations while maximizing the spot size on the camera. The camera contains 440 pixels with each pixel corresponding to a Photonis XP3062 PMT. The camera lies on the focal surface of the spherical mirror of 1.743 m radius. The pixels are approximately hexagonal with a side to side distance of 45.6 mm, corresponding to an angular size of 1.5 degrees [24]. More details on the FD electronics and optics are discussed in chapter 4.

Previous experiments with fluorescence detectors did not have a surface detector so

the reconstruction of all shower parameters was done with only FD information. Since Auger is a hybrid detector almost every FD event also has information from the SD. The timing from just one WCD in the surface detector helps in the FD reconstruction tremendously. For this reason, the Auger collaboration only reconstructs FD events that also have information from at least one WCD from the surface detector. In this way all FD events reconstructed are hybrid events. To properly reconstruct an SD event at least three WCDs need to trigger. If no SD reconstruction is possible but there is still at least one WCD with information the event is still considered a hybrid event.

The first step in reconstruction of a hybrid event is to reconstruct the shower axis. The timing information from the FD pixels and the SD station are used. Then the detected FD signal is converted from a light flux at the aperture to a light flux at the shower axis by propagating the light back through the atmosphere. Accurate knowledge of the atmospheric clarity and the calibration of the detector components are needed to relate the signal detected to the light emitted from the shower axis. The light produced along the shower axis then can be linked to the amount of energy the shower deposited in the atmosphere leading to the calorimetric energy measurement. The following sections highlight the relevant details.

3.3.1 The shower axis

The fluorescence light emitted from the shower axis can be thought of as a point source of light moving at the speed of light along the shower axis. The track of light detected in the camera of the FD defines what is called the shower detector plane (SDP). The SDP is easy to obtain. The track the shower makes in the cameras (see Figure 3.8) defines the SDP to within 0.1 degrees of accuracy (as long as there are at least 10 triggered pixels in the reconstructed event).
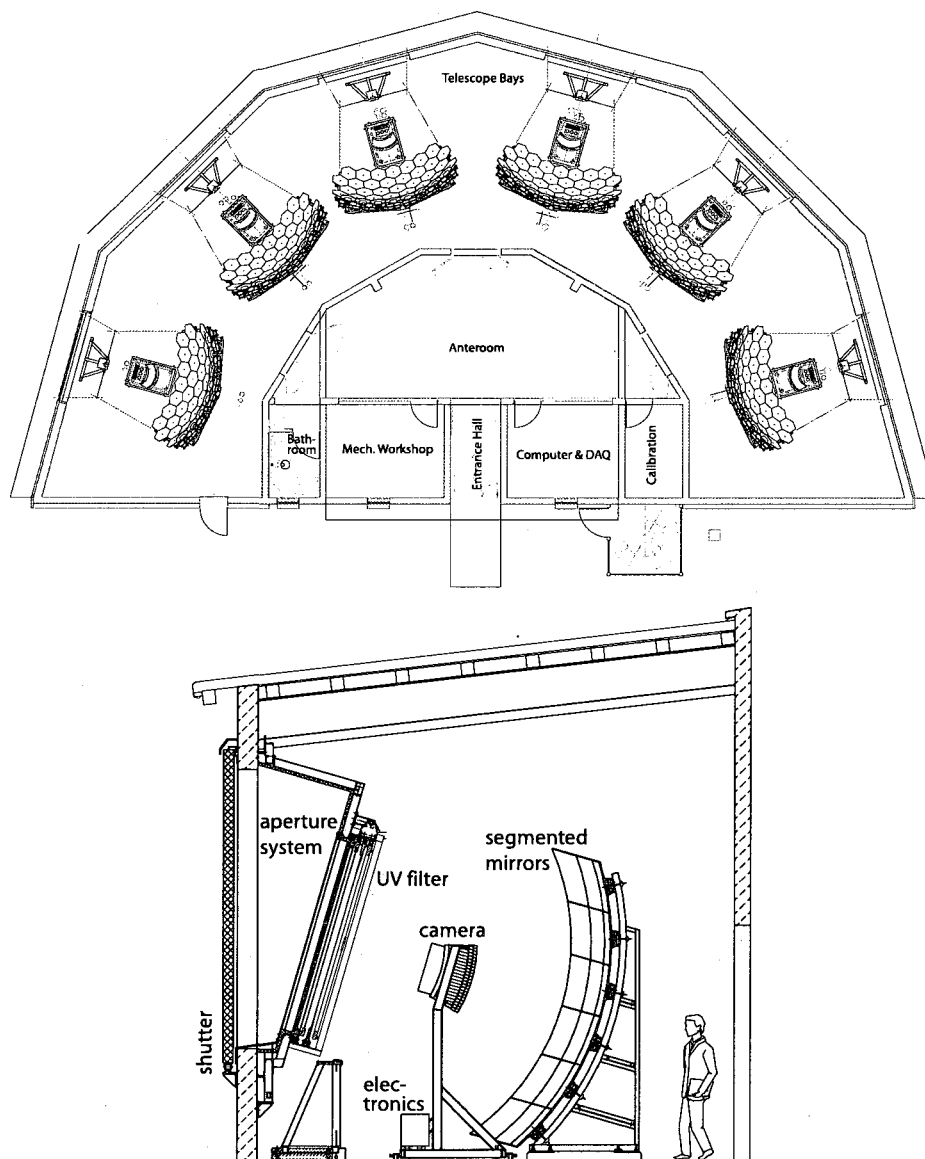
figures from [24]

Figure 3.7 Each of the four FD sites has six telescopes.
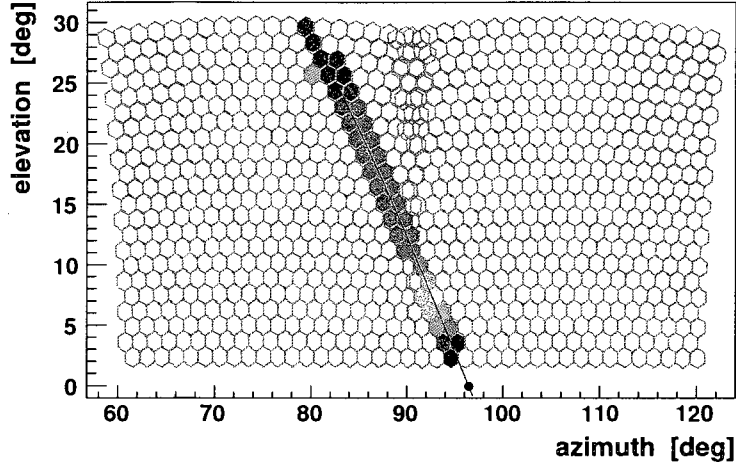
Figure 3.8 The track made in the camera from a 10^{18} eV shower. The triggered pixels are shown in colors corresponding to the time each pixel detected light from the shower.

To find where the axis of the shower is inside of the SDP the timing of the detected pulses in each pixel are used. In a hybrid detected shower the SD will also record a signal in one or more stations. The timing of the WCD with the largest signal along with the timing in the triggered FD pixels is used to determine the shower geometry inside the SDP.

Each triggered pixel in the camera has a specific field of view. The angle the ith pixel in the event makes with the ground is defined as χ_i. The time the ith pixel detects the pulse is called t_i. The SD station with the largest signal is also used. The shortest perpendicular line from the shower axis to the FD is called R_p. The angle, χ_0, is the angle the shower axis makes with the ground in the SDP. R_p and χ_0 completely define the shower geometry in the SDP (see figure 3.9). The equation that relates these parameters is,

$$t_i = T_0 + \frac{R_p}{c}\frac{\tan(\chi_0 - \chi_i)}{2} \qquad i = SD, 1, 2, \dots \qquad (3.5)$$
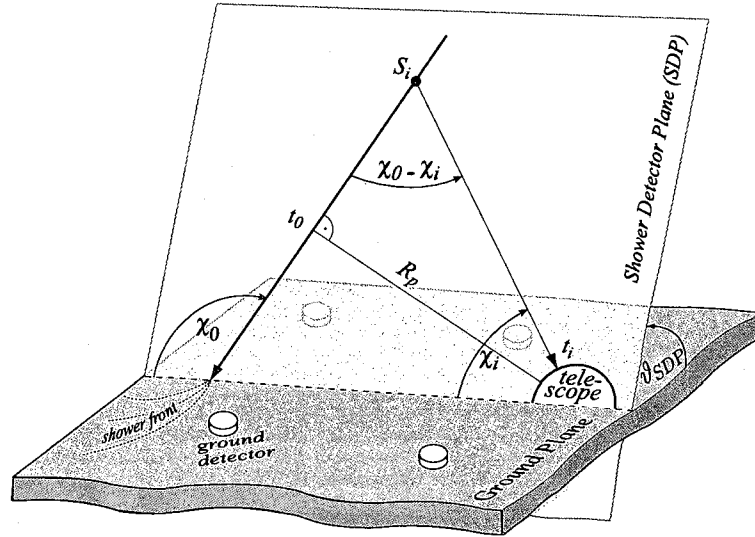
figure reproduced from [24] and modified for hybrid reconstruction

Figure 3.9 The basic geometry of an event detected by an FD

This equation is nonlinear and is solved by a minimization routine to get R_p and χ_0 . Before minimization, T_0, the time the shower impacts the ground is calculated from the WCD timing information. The timing of the signal in the WCD is called t_{SD}, and the angle χ_{SD}, is found by projecting the station into the SDP along a line pointing vertically from the station. By using the SD to get T_0, the fit is constrained and allows for an accurate geometrical reconstruction. If the information from the WCD is not used then the shower axis can still be reconstructed but has a larger uncertainty. Figure 3.10 demonstrates this.

The arrival direction of the primary cosmic ray is first determined by converting the SDP angle, R_p and χ_0 angle to local variables, θ and ϕ. The zenith angle, θ is measured from the vertical. The azimuth angle ϕ is measured counterclockwise from above with zero corresponding to due east. The angular accuracy is better compared with other fluorescence detectors that cannot operate in hybrid mode [25]. This geometrical accuracy is very important for determining the slant depth of shower maximum, X_{max}.
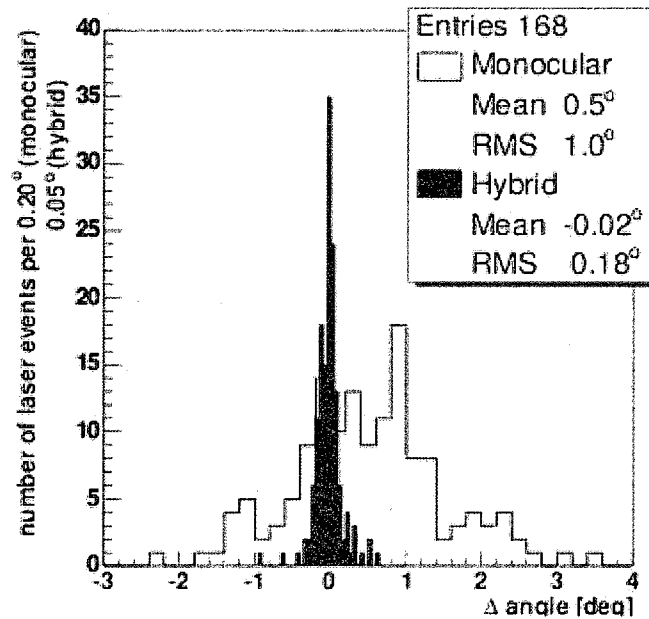
figure reproduced from [25]

Figure 3.10 Laser tracks with know geometries that also simultaneously trigger the SD were reconstructed using the minimization routine in equation 3.5. First without SD information, monocular reconstruction. Then with SD information, hybrid reconstruction. The angular accuracy the hybrid reconstruction has a significantly narrower width.

3.3.2 Detector calibration and the atmosphere

Each pixel along the track of the shower records a signal that is proportional to the incident light on that pixel. To relate that signal (given in ADC counts) to a light flux at the aperture, a calibration of all the detector components is needed. The calibration of the FD telescopes is given in units of photons at the aperture per ADC count, $\frac{\gamma_{app}}{ADC}$. The details of the calibration will be discussed at great length chapter 4.

The propagation of light through pure air has been very well studied and is governed by Rayleigh scattering. While the location of Auger was chosen because of the clear skies that the pampas of Argentina have, most nights still have some clouds and aerosols (dust, water vapor, etc.). In order to classify the clarity of the atmosphere several monitoring techniques are used. There are many different devices used to monitor the atmosphere and details can be found in [26]. The use of these devices provides an accurate characterization of the local conditions specific to the Auger site. When the atmosphere is clear there is a low systematic uncertainty on propagating the light from the FD aperture to the shower axis. The monitoring equipment allows for quality cuts on the data to insure low systematic uncertainties.

3.3.3 The longitudinal profile

As the electrons and positrons in the shower move through the atmosphere, both fluorescence and Cerenkov light is produced. The fluorescence light is emitted isotropically while the Cerenkov light is emitted in a narrow forward cone along the shower axis (see Figure 3.11). Depending on the shower geometry in relation to the FD building, the amount of Cerenkov light detected can vary drastically (however, the production of Cerenkov light will not). If the forward cone of a shower is pointed
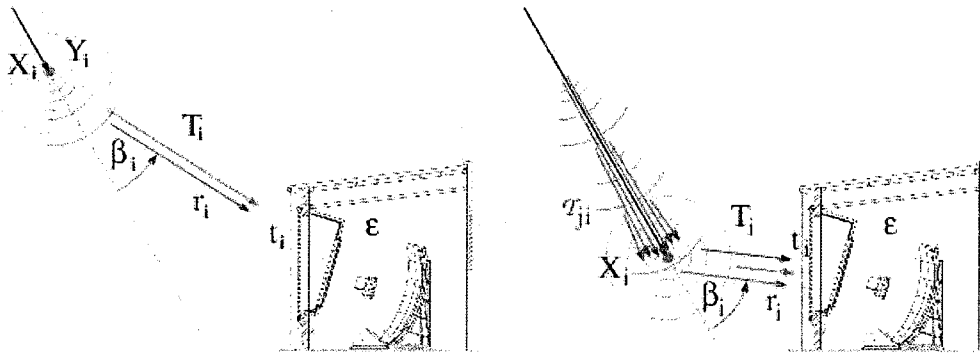
Figure 3.11 A schematic view of the light production from the shower axis. Direct fluorescence light in green and Cerenkov light in red are detected (left). Scattered light (right) also is detected and the amount of scattering depends on the clarity of the air. Monitoring of the atmosphere is needed to make corrections for any aerosols or clouds located between the shower axis and the FD.

away from of the FD field of view then zero Cerenkov light will be detected and all the light detected can be attributed to fluorescence light. Conversely, if the shower axis is aligned with the field of view of the FD then the signal will be dominated by Cerenkov light. The ratio of Cerenkov to fluorescence light ranges from almost zero to almost 100%. Typically a quality cut that limits the fraction of Cerenkov light to less than 50% is used.

A novel reconstruction algorithm which treats both the fluorescence and Cerenkov light as signal has been developed by Auger collaborators [27]. This is in contrast to other techniques that use iterative processes to try to subtract the Cerenkov light contribution from the total light detected.

The longitudinal profile is a plot of the amount of energy deposited per slant depth, $\frac{dE}{dX}$, verses slant depth, X. A reconstructed profile is shown in Figure 3.12. The amount of fluorescence light produced depends on the amount of energy deposited and on the fluorescence yield, $Y_F(X)$ of the nitrogen, which also has a dependence
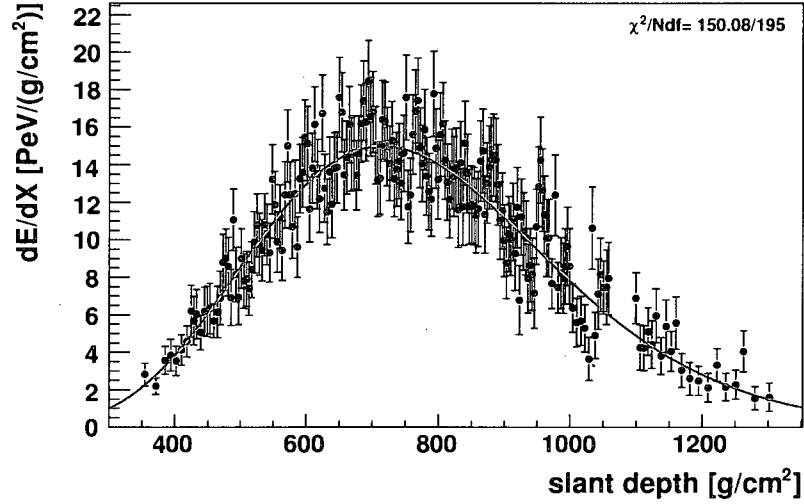
Figure 3.12 The profile plot. The integral is a measurement of the total energy deposited in the atmosphere by the EAS. The peak is the depth of shower maximum, X_{max}. Both the fluorescence and Cerenkov light were treated as signal to convert the light detected at the FD aperture (figure 3.11) to $\frac{dE}{dX}$.

on slant depth. Experiments have previously measured the yield, $Y_F(X)$ [28], but new measurements are being made by contemporary experiments to improve on the understanding of uncertainties on this quantity [29]. The Cerenkov light production also depends on the energy deposited and has an analogous yield factor, $Y_C(X)$ which can be obtained analytically [27]. Along with these yield factors the light distribution patterns for fluorescence (isotropic) and Cerenkov (forward peaked) light are needed.

To take into account all these factors the minimization of a complicated approximate functional form is needed. The choice of function used is a Giasser-Hillas function [2] which has historically been used in event reconstruction. The new reconstruction technique used is well described by a modified version of this function.

$$f_{GH} = \frac{dE}{dX}(X) \quad = w_{max} \left(\frac{X - X_0}{X_{max} - X_0} \right)^{(X_{max}-X_0)/\lambda} e^{(X_{max}-X)/\lambda} \qquad (3.6)$$

This function has four free parameters: X_{max}, the maximum energy deposit w_{max} and two shape parameters X_0 and λ. The integral of this equation is the total energy deposited in the atmosphere. Since this procedure has only accounted for the electromagnetic component of the air shower a correction is needed to obtain the total energy of the primary cosmic ray. This correction is dependent on the composition and the high energy hadronic interaction model assumed. The correction introduces small a systematic uncertainty of 4%. As mentioned above this is small compared to the overall 22% uncertainty on the energy (chapter 7 will discuss all the systematic uncertainties in more detail).

Chapter 4

The Calibration of the

Fluorescence Detectors

As explained in chapter 3, there are six telescopes located in each of the four FD buildings. Each of these 24 telescopes has a 440 pixel camera giving a total of 10560 pixels. Each pixel has an individual PMT (photomultiplier tube) and associated electronics which measures an ADC response proportional to the incident number of photons on the telescope aperture. Accurate knowledge of this conversion from ADC counts to photons is needed to reconstruct the energy of extensive air showers. The calibration constant, cc, is defined as the number of photons incident at the aperture of the telescope per ADC count response from the data acquisition system.

$$cc = \frac{photons}{ADC} \tag{4.1}$$

The calibration effort is a multifaceted operation involving many people from various institutions in the Auger collaboration. There are three main aspects of the calibration; the absolute, the relative and the multi-wavelength calibration. Redundant techniques are used, when possible, to reduce the impact of systematic uncertainties

in each operation.

This chapter reports on the calibration techniques that have been used in published physics results. The multi-wavelength and relative calibration results are all linked to the absolute calibration results from June 2005. The calibration is an evolving ongoing effort and one is constantly trying to reduce systematic uncertainties. Where appropriate, more recent results from lab work have been included on the systematic uncertainties from the June 2005 calibration.

4.1 Introduction

The main components of the telescope are the UV-filter at the aperture [30], the spherical mirror and the PMT camera. The UV-filter is designed to only accept wavelengths of light from 280-425 nm. The filter is needed to reduce the background light from the night sky to allow the detection of the faint fluorescence light from an EAS. The fluorescence light emitted from the nitrogen in the atmosphere is composed of several distinct lines with the most dominant lines at around 320, 337, 355, 380 and 390 nm (see Figure 4.1).

Along with the major components of the telescope, there are many other minor components that are needed to optimize the collection efficiency. The efficiencies of the major components, along with some of the minor components were individually measured [31]. These measurements, along with some manufacture specifications, combined to form a piecewise calibration of the entire telescope (see Figure 4.1). This piecewise calibration gives a conversion from ADC counts to photons for the wavelengths from 280-425 nm.

While the attempt has been made to try to take into account all the individual components in the piecewise calibration, the combined effects of all the components
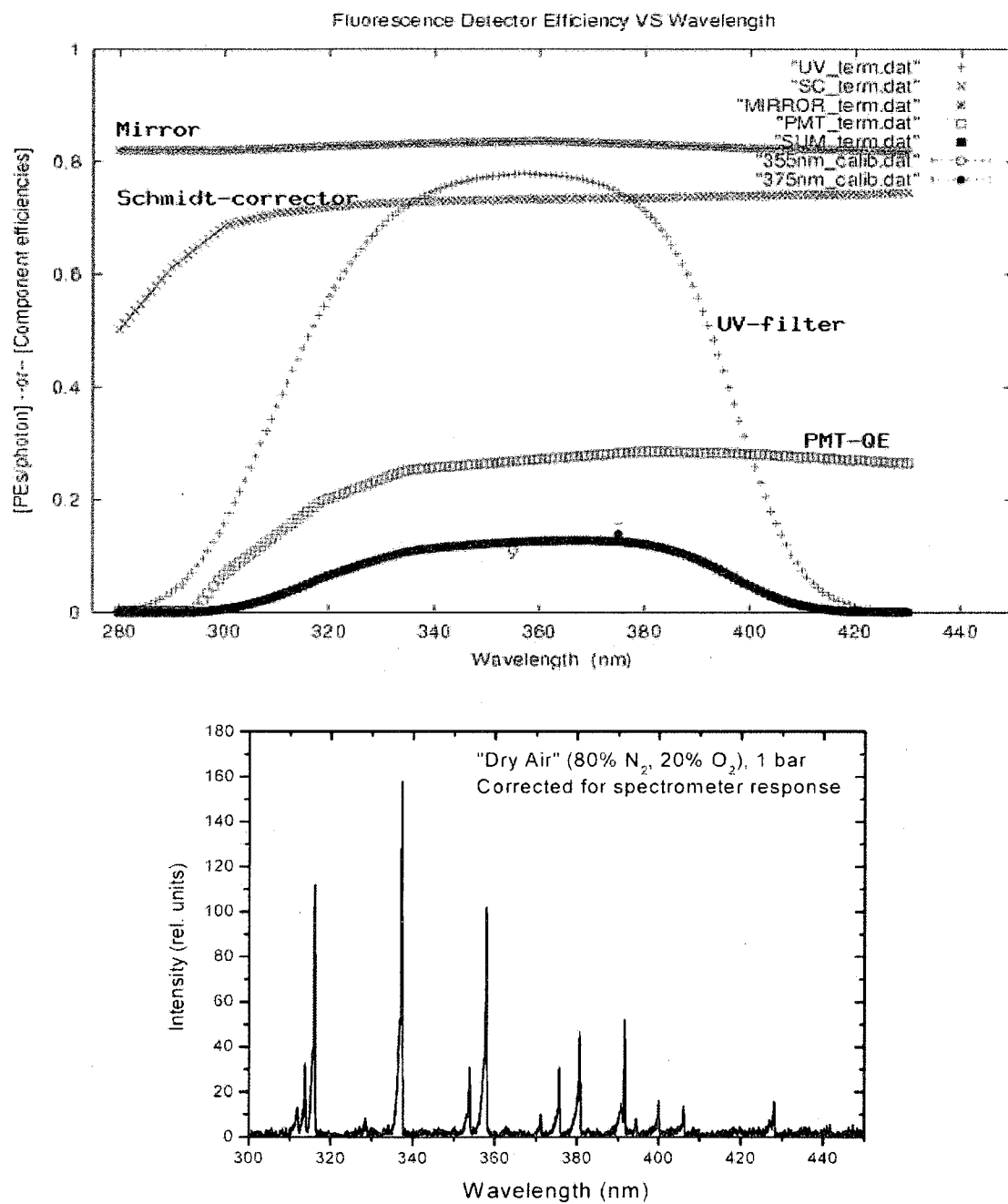
Figure 4.1 The piecewise calibration (top) and the nitrogen fluorescence spectrum (bottom).
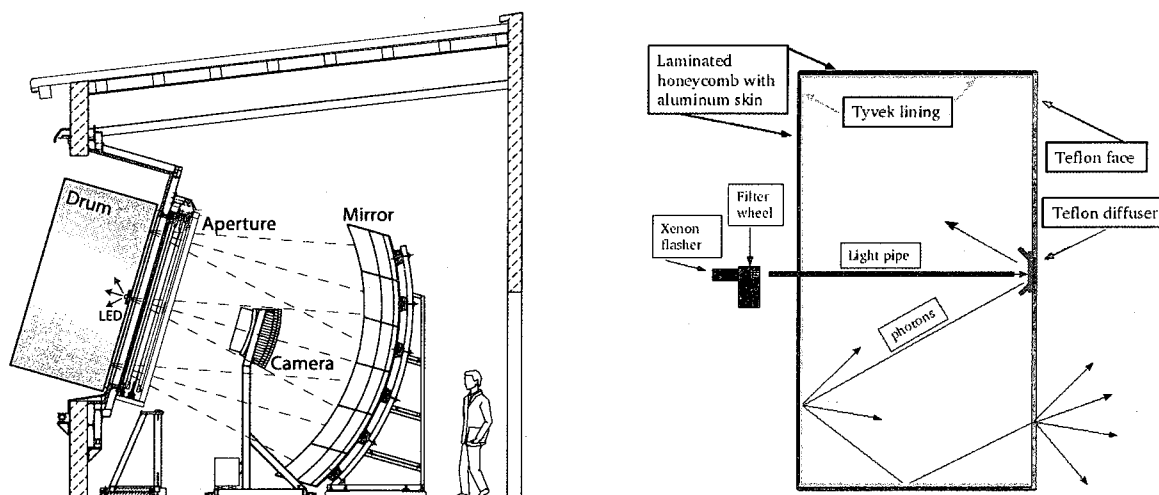
Figure 4.2 Schematic drawing of the calibration drum mounted in the aperture of a FD telescope. The photons from the drum pass through the UV-filter, then get focused by the mirror on to the individual PMTs of the camera.

once assembled could change the overall efficiency. A sophisticated ray-tracing program would be needed to combine all the effects and estimate the overall efficiency. While possible, this is not done because it would be hard to convince ourselves that all effects were properly being accounted for, and thus be difficult to estimate uncertainties. The combined effect of all the components including the aperture projection, optical filter transmittance, reflection at specific optical surfaces, mirror reflectivity, PMT light collection efficiency, electronic gain, ADC conversion etc. as they are used can only really be determined once the telescope has been assembled. To take into account these effects an end-to-end technique has been developed.

4.2 End-to-end calibration

A portable light source that mounts in the aperture of each FD telescope was designed to uniformly illuminate all 440 pixels in a single camera simultaneously. Due to the appearance of this portable light source it is referred to as the "drum". The drum
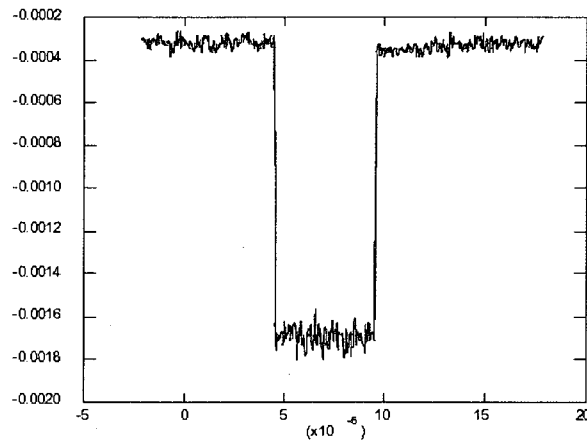
Figure 4.3 The LED driver can be switched from pulsed to continuous operations. The resulting intensities of the LED are compared in these two modes of operation. The pulse is square to within 1% but the relative amplitude of the flat top of the pulse has a 5% uncertainty when compared to the amplitude when the LED is run in continuous mode.

light source is a cylinder 1.4 m deep with a 2.5 m diameter shown schematically in figure 4.2. A light pipe runs from the front face to the back of the drum along the center axis. The drum can be used with UV LEDs for the absolute measurement or with a xenon flasher for measurements at different wavelengths. The broad spectrum of a xenon flasher necessitates the use of notch filters at wavelengths in the 280-425 nm region of interest.

A constant current LED driver is used to power the LEDs that illuminate the drum. The LED driver creates a square pulse of current that in turn creates a square output of light. The rise time of the pulse is less than 100 ns and the top of the pulse is flat at the 1% level (see Figure 4.3) . The LEDs can be pulsed on and off with a 1, 3, 5 or 10 μs long pulse as well as be powered in continuous mode, providing a constant light output. The LEDs are mounted inside the light pipe running through the center of the drum.

Where the light pipe meets the front face of the drum a Teflon© diffuser is mounted

in a reflecting cup that directs the light to the walls and back surfaces of the drum which are lined with the diffusively reflective material Tyvek©. The front face of the drum is made from Teflon© which is a diffusively transmitting material. All the photons make multiple bounces off the diffusive surfaces making the face of the drum illuminate uniformly in all directions. A surface with this property is know as a Lambertian surface. A Lambertian surface viewed at an angle θ appears equally bright as viewing the surface perpendicularly (see figure 4.4). However the integrated intensity when viewed at an angle θ is reduced by the cosine of the angle. This is because the cross section of viewable surface area has been effectively reduced by $cos(\theta)$.

When the drum is mounted, it completely covers the aperture, blocking out all external light. Once mounted, the LEDs are pulsed and an ADC readout from the FD is recorded. Each of the 440 pixels in a camera subtends a solid angle of 1.5 square degrees. These pixels view incoming photons from the entire drum surface. Each pixel views the drum at a different angle, varying from 0 to 22 degrees. If the drum is a Lambertian surface then calculating the number of photons each pixel sees from the drum changes by only the cosine of the viewing angle. The calibration from ADC counts to photons for all 440 pixels in each camera then only relies on the knowledge of the drum intensity.

The drum light source thus needs to be characterized in two ways. The absolute intensity is needed we must verify that the drum is Lambertian out to at least 22 degrees. A large lab was designed to make these measurements. The lab consists of a large dark room, 4 x 18 x 4 m, and a 150 x 50 x 30 cm dark box with enclosed optical equipment.
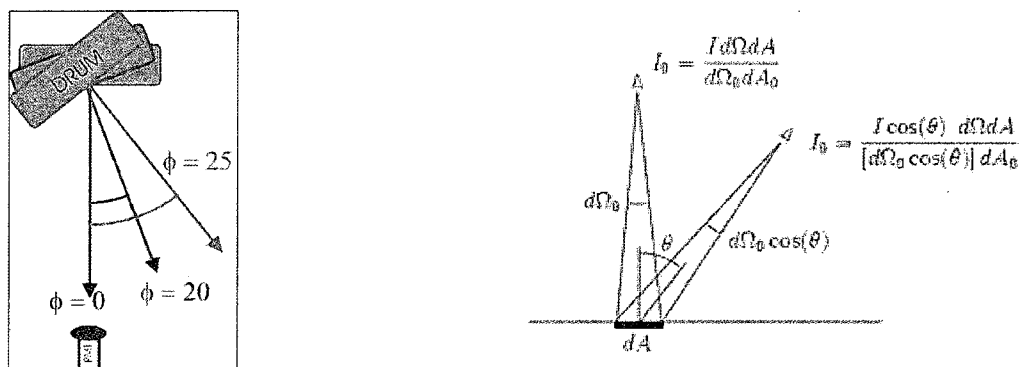
Figure 4.4 The rotation of the calibration drum (right) and the details of a Lambertian surface (left). The intensity remains the same as a Lambertian surface is rotated, but the number of detected photons decreases by $\cos\theta$.

4.2.1 Drum Uniformity

The drum uniformity at a 0 degree viewing angle is verified with a CCD image. A CCD camera is placed 14 m away, which is far enough away such that all incoming photons are incident approximately parallel on the CCD camera. All photons are within ≈ 5 degrees of parallel so any angle of incidents effects are less than 1% corrections. The image in Figure 4.5 (left) was analyzed with software by dividing the image into concentric rings. The intensity of the pixels in each region form a Gaussian histogram (middle). The means of these histograms are separated by only $\pm$ 2%.

The four concentric regions in figure 4.5 do not cover the entire surface of the drum. The very center region is not used in this comparison because it is known to be less intense. This is due to the reflecting cup blocking most of the photons. This area can be neglected when looking at the total surface intensity because it represents a small fraction of the total surface area (less then 1%). The outermost region is also not compared because it rests outside the aperture of the FD telescopes when mounted. This ring is less intense due to unavoidable edge effects and this is why the drum was built to be larger than the 2.2 m aperture of the fluorescence
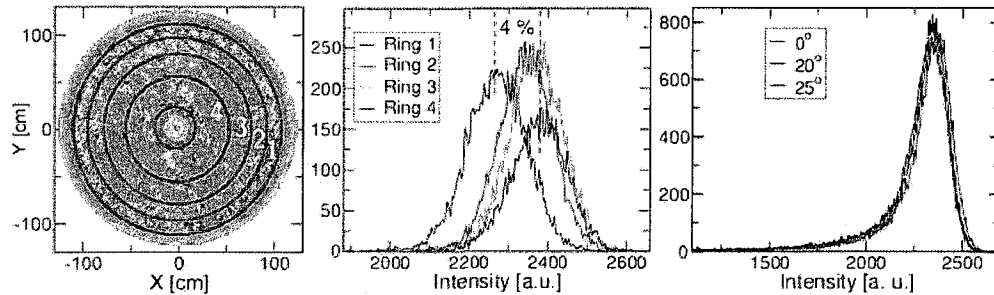
Figure 4.5 From left to right:Color-enhanced CCD image at 0 degree drum angle, showing the defined rings for relative intensity analysis. Production deformations in the Teflon© material can be seen; A plot of the observed pixel intensities in the defined regions of the drum, shown in the previous figure; The results of angular relative intensity measurements at 0, 20, and 25 degrees.

telescopes. For all the intensity measurements a mask is placed on the drum so that the outer ring is blocked off. Only the center area (2.2 m diameter) of the drum surface contributes to the measurements.

CCD images were taken of the drum rotated at angles from 0 to 25 degrees. Figure 4.5 (right) shows that the mean and the distribution of intensities do not change as the drum is rotated. The integral of each of these histograms is reduced by the cosine of the angle as expected for a Lambertian surface.

Perfect drum uniformity is desirable but not strictly needed. The nonuniformities in illumination can be measured and taken into account by a ray-tracing program. The advantage of near perfect uniformity is that reliance on the ray-tracing program is small. The non-uniformities that do exist on the surface of the drum have been mapped from the CCD images.

The dominant source for corrections to uniformity comes from the camera shadow effect. When the drum is in the aperture the PMT camera shadows some of the incident light. The camera shadows a different portion of the aperture depending on

the viewing angle of the individual pixel. This shadowing of light can be thought of as just another efficiency factor of the telescope. The end-to-end technique will take this effect into account along with all of the other effects from the components of the telescope. However the small non-uniformities that do exist on the surface of the drum cause this shadowing effect to depend on what part of the surface drum is shadowed by the camera. Corrections are made for each pixel viewing angle that are each less than 1.5% [33].

4.2.2 Absolute Drum Intensity

The absolute calibration of the drum light source is based on a set of Si photodiodes, calibrated at NIST [34]. The photodiode response is read out as a current by a pA sensitive Kiethley brand electrometer. Even with pA sensitivity, the small surface area and low response of the NIST calibrated photodiode precludes detection of the small photon flux from the drum surface directly. Additionally the calibration of the photodiode is only valid when measuring continuous beams of light. Since the drum light source needs to be powered in pulsed mode to more closely represent an air shower signal this poses another challenge. A technique is used to accurately scale the low level intensity of light from each drum pulse to a continuous beam of light 2-3 orders of magnitude more intense. The intense beam can be measured with the NIST calibrated photodiode to obtain the absolute number of photons. The scaling factor is then applied to obtain the absolute drum intensity.

The scale factor mentioned above depends on relating a pulsed intensity to a continuous intensity. The LED driver that powers the drum can be switched from pulsed to continuous operation. The squareness of the pulsed current allows the light outputs in pulsed and continuous mode to be compared easily (see Figure 4.3). The width of the pulse is needed to calculate the duty cycle between pulsed and
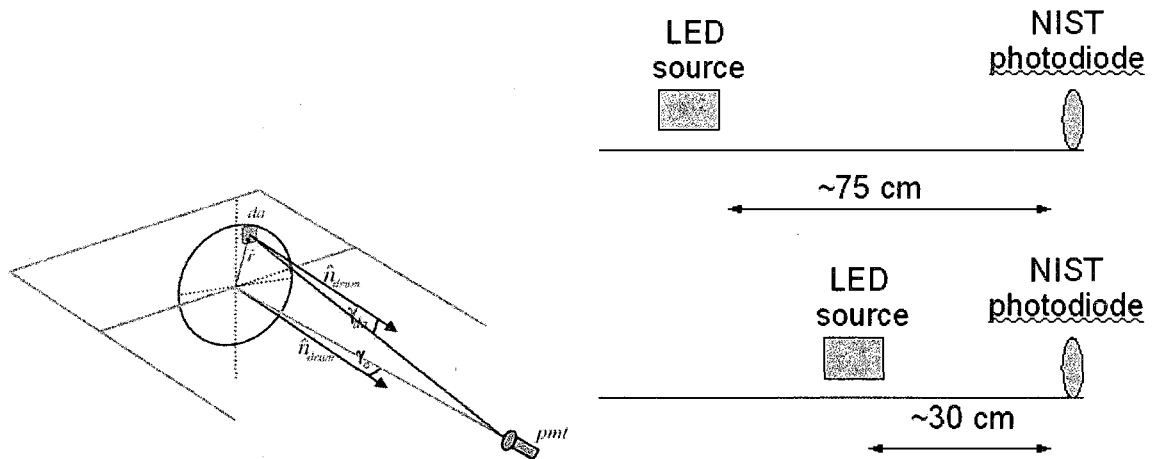
Figure 4.6 The lab PMT setup in the dark hall (left) and a in the dark box (right). In the dark hall, the PMT is located at a distance 14 m. The light the PMT measures is effectively the same as measuring the flux normal to a point source at a distance of 14 m. In the dark box, a diffuse led light source is produces a flux of equal intensity on the lab PMT. The PMT is replaced by a NIST calibrated photodiode to measure the flux, which is proportional to the drum intensity.

continuous operation. The relative amplitude of the top of square pulse compared with the continuous light output is also needed. These two values need to be well known to correctly scale the number of photons the LED emits between pulsed and continuous operation. Currently the uncertainty on this scale factor is 5%. This is the largest uncertainty in the absolute measurement and due to the specifics of the LED driver used is hard to measure more precisely. An upgraded LED driver is in development to address this uncertainty but is not discussed.

The drum is placed in the large dark room roughly 14 m away from a reference PMT (see Figure 4.6) to measure the surface intensity. The LED driver sends 5 μs pulses to the drum and the intensity is recorded by the lab data acquisition system. A histogram of the integrated signal in ADC counts is used to extract the intensity of the drum. The histograms are Gaussian (see Figure 4.7) and are a relative measure of
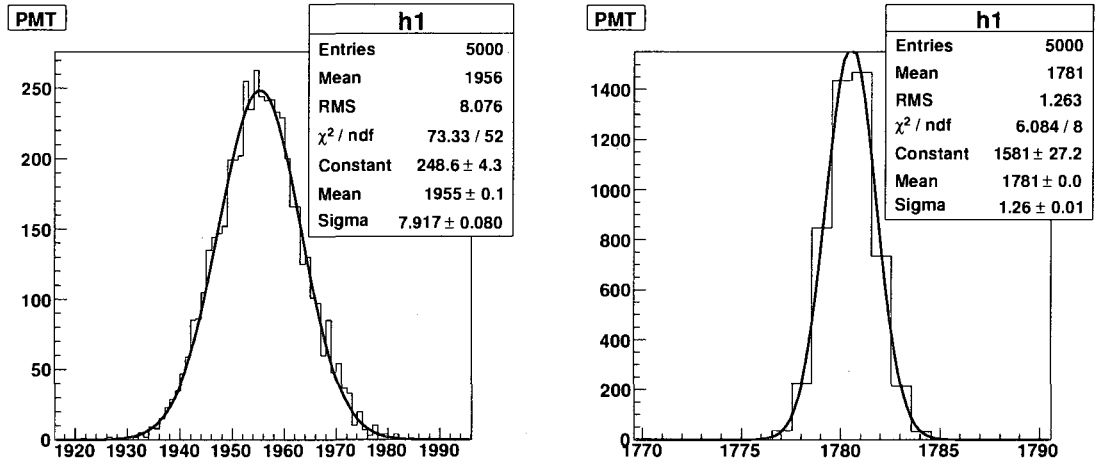
Figure 4.7 The histogram of the foreground (left) and the background (right) as measured by PMT_{lab}. The difference of these histograms centroids is the relative measure of the drum intensity that is later matched by the light source on the optical bench.

the drum intensity. The number of photons detected by the PMT per pulse, N_{PMT}, is proportional to the PMT histogram centroid, C_{PMT}.

$$N_{PMT} \propto C_{PMT} \tag{4.2}$$

At the distance of 14 m, each area element, dA, on the drum surface contributes to the intensity measurement. All the photons incident on the PMT are within 5 degrees of parallel so effects due to angle of incident are small. The light the PMT measures is effectively the same as measuring the flux normal to a point source at a distance of 14 m. This implies that drum can be approximated as a point source. The flux of photons detected at the PMT, Φ_{PMT}, is then just the number of photons, N_{PMT} divided by the area of the PMT, A_{PMT}.

$$\Phi_{PMT} = \frac{N_{PMT}}{A_{PMT}} \tag{4.3}$$

Only the distance is then needed to relate this detected flux to the number of photons

emitted normal, or from 0 degrees, off the drum surface, N^0_{drum}.

$$N^0_{drum} = \Phi_{PMT} * d^2 \qquad (4.4)$$

Since we demonstrated that the drum emits photons uniformly, the average drum intensity emitted at 0 degrees off the surface the drum, I^0_{drum} (measured in units of $[\frac{photons}{cm^2}]$) is just N^0_{drum} divided by the surface area of the drum, A_{drum}. Putting this all together we find that I^0_{drum} is proportional to N_{PMT} and the square of the distance to the drum.

$$I^0_{drum} = \Phi_{PMT} * d^2 \times \frac{1}{A_{drum}} \qquad (4.5)$$

A couple of geometric measurements, along with a conversion factor, k_{PMT} that relates the number of ADC counts in the PMT histogram to number of photons detected are all that is needed to calculate the drum intensity.

$$I^0_{drum} = \frac{N_{PMT}}{A_{PMT}} * d^2 \times \frac{1}{A_{drum}} = \frac{k_{PMT} \times C_{PMT}}{A_{PMT}} \times d^2 \times \frac{1}{A_{drum}} \qquad (4.6)$$

Rather than finding k_{PMT}, the PMT is used only as a transfer mechanism to find the flux of photons at 14 m. On an optical bench the same PMT is placed ≈ 1 m in front of a LED light source. This light source uses the same LED driver that powers the drum and LEDs with the same wavelength distribution. The LEDs are bright enough to read a signal with the photodiode but will saturate the PMT at these close distances. A neutral density filter is used to reduce the intensity for the PMT measurement. An accurate measurement the intensity reduction is needed and the details of this measurement will be discussed below.

The detected light intensity can be further adjusted by changing the distance between the source and the PMT. Histogram centroids are recorded in a range that covers the drum intensity recorded in the dark room at 14 m. At each distance from the PMT, the neutral density filter (NDF) is removed and the LED is switched from

pulsed to continuous mode. The calibrated photodiode replaces the PMT and the current, i_{pd}, is read out. By repeating this procedure at several distances, a relationship between photodiode response and PMT response is found. This relationship is found to be very linear (see Figure 4.8).

χ^2 / ndf	0.2408 / 4
p0	-0.001809 ± 0.003371
p1	0.0005518 ± 1.802e-05
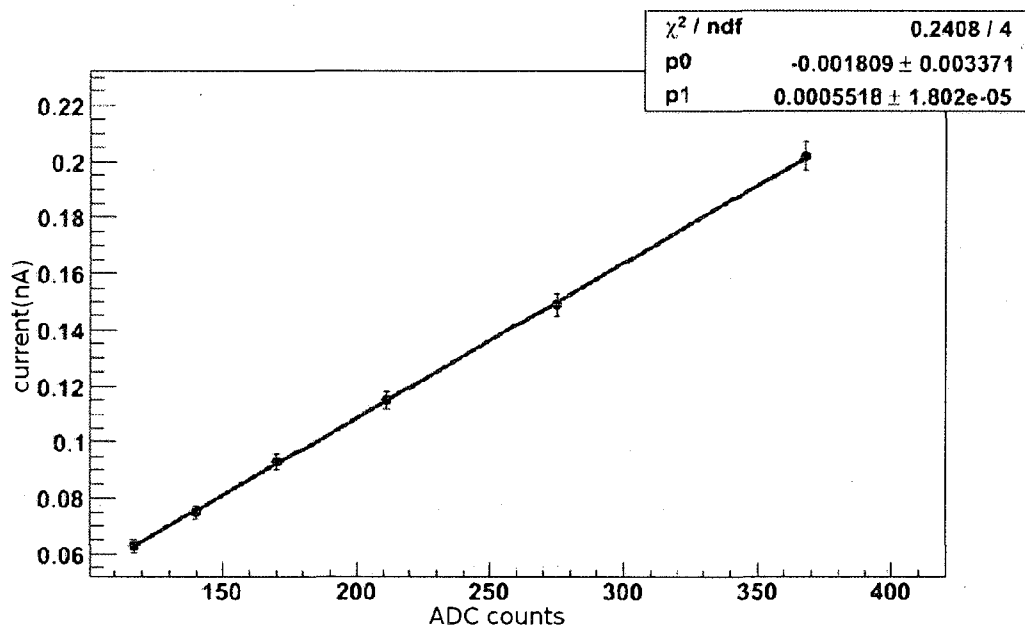
Figure 4.8 The DC current in the NIST photodiode versus the integrated response in the PMT in AC mode. The errors in the NIST current are 2% and the errors in the PMT integrated response are 1 ADC count.

Using the calibration of the photodiode, k_{NIST} in Amps per Watt of incident light, the number of photons detected per second, N_{pd}, is found from the current readings.

$$N_{pd} = k_{NIST} \times \frac{\lambda}{hc} i_{pd} \tag{4.7}$$

This can be turned into a flux of photons incident on the photodiode by dividing by the area, A_{pd} (the active area of the photodiode is determined by a mask with precisely known area given by NIST).

$$\Phi_{pd} = \frac{N_{pd}}{A_{pd}} = k_{NIST} \times \frac{\lambda}{hc} \frac{i_{pd}}{A_{pd}} \tag{4.8}$$

The flux incident on the PMT can now be found by reducing Φ_{pd} by the NDF attenuation factor and the duty cycle of the pulsed light, which is just the ratio of the pulse width to one second, w, of the pulse (here both w and NDF are dimensionless).

$$\Phi_{PMT} = \frac{\Phi_{pd}\, w}{NDF} \qquad (4.9)$$

The average drum surface intensity, I^0_{drum}, from equation 4.6 can now be found,

$$I^0_{drum} = \Phi_{PMT} * d^2 \times \frac{1}{A_{drum}} = \frac{\Phi_{pd}\, w}{NDF} \times d^2 \times \frac{1}{A_{drum}} \qquad (4.10)$$

This work on the optical bench is done immediately after the dark room measurement so that any environmental effects on the PMT will be negligible. The NDF attenuation is measured by comparing the ratio of the light detected with the filter in to the light detected with the filter out. These measurements are made in continuous mode using a photodiode. Due to the low light needed to match the drum intensity the light source is very dim when the filter is in. This measurement is not much higher than the background reading on the electrometer so the background subtraction is very critical. During the June 2005 calibration work, the cables and connectors were replaced to lower the background noise which allow a more accurate NDF measurement.

A series of measurements were done with different setups. Ideally the NDF would be measured in the exact geometrical setup used in the dark box measurements, but the light output would be too close to the background to obtain a reliable reading. Measuring the NDF in slightly different setups allows for a higher current reading but introduces systematic errors. The values obtained in the various setups ranged from 249 to 258 with an average of 252. The standard deviation in the series of measurements gives an estimate of the systematic error and is found to be 2%. Another source of systematic uncertainty is due to the geometrical differences in the filter wheel when the filter is in and when it is out. Based on recent measurements this effect is estimated to be less than 2.5%. The LED warms up during this measurement

and can change intensity. The heat up of the LEDs plateaus at a certain temperature after 5-10 minutes. The LEDs were allowed to warm up to a stable temperature so any intensity change due to temperature would be less than 1%. The combined uncertainty is estimated to be 3% on the NDF attenuation measurement.

The calibration of the photodiode, given in amps per watt incident, has a three sigma uncertainty of 1.5% [34]. However, the measurement of the current in the experimental setup suffers from a larger uncertainty due to temperature effects of the LED driver and the LEDs used. When LEDs are run in continuous, or DC, mode the junction temperature rises causing the light output of the LED to change. The transistors in the LED driver also heat up causing the driving current to be modified. The effects partially compensate but the LED temperature is the larger of the two effects.

The LED junction temperature is not easily measurable but when the junction temperature changes, the temperature of the entire encasing of the LED also changes. The encasing temperature is easily measured and the light output of the LEDs has been measured as a function of this temperature. There is of course a time lag between the junction heating, which takes only microseconds, and the encasing temperature, which takes on the order of milliseconds to change. Figure 4.9 shows the temperature dependence of one of the LEDs used in the calibration procedure. When the LED is switched on in DC mode, the temperature rises before a stable reading on the electrometer can be read out. Readings are made before the temperature changes significantly so that the light output in pulsed and continuous operations can be compared equitably. It takes around three seconds to get at stable reading on the electrometer so uncertainty on this reading thus depends on how much the light out put changes in this time. The steeper regions of the curve in Figure 4.9 change in a shorter time (1-2% per second) than the more flat regions (0.5% per minute). The

ambient temperature in the lab establishes where on the curve we start and lab is not temperature controlled. Different LEDs have different temperature curves but in general the uncertainty on the current reading can be kept below 2%.
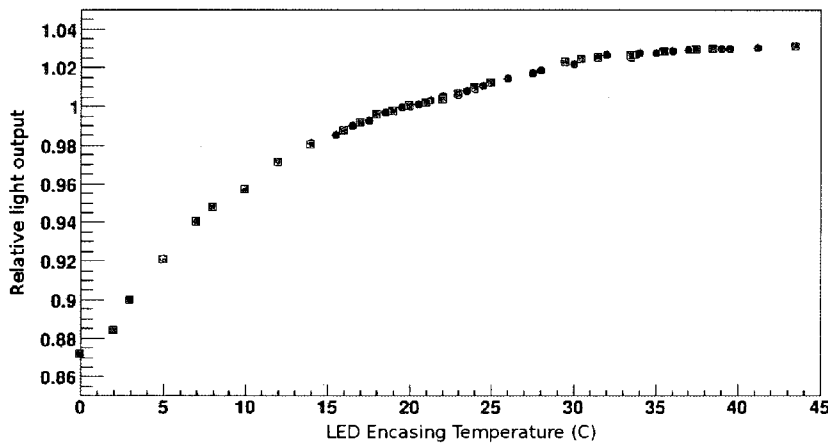
Figure 4.9 Temperature studies of the LED are important in understanding the systematic uncertainties between pulsed and continuous modes of operation of the LED driver.

Converting the current in the photodiode to number of photons incident is done by using the calibration of the photodiode, k_{NIST}, to get the power incident. Then the wavelength of the LED, λ; and the area of the photodiode, A_{pd} are used to convert to number of photons incident.

4.3 Calibration constants

The drum is transported in a specially designed trailer to the each of the FD buildings. At the FD the drum is placed into each aperture and pulsed 400 times, a typical response is shown in Figure 4.10. The recorded response is integrated and an average ADC count is calculated. There are small corrections that need to be made for individual pixel size, camera shadow effects (section 4.2.1) and reflections off the UV

filter. The drum will have a slightly different intensity at the FD as compared to the measured lab intensity, due to the ambient temperature difference between the two locations. The temperatures in the lab and at the FD are logged continuously and corrections are applied for differences. The results presented below were from a calibration trip during which the temperature variation were not extreme. No large corrections were needed. Systematic uncertainties are included for all these corrections and effects.
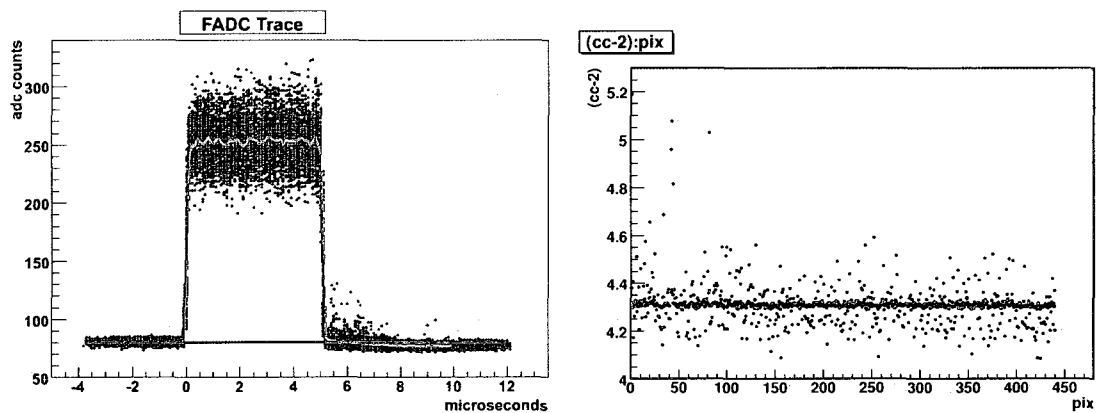
Figure 4.10 A typical response of a camera to 400 pulses from the drum (left). The black dots are the individual ADC counts for each time bin, the red line is the average signal and the blue is the average pedestal. The integral of all the pulses is used to calculate the ADC counts for the calibration constant for each pixel. The constants for one camera as shown (right) before and after flat-fielding. The black points are the calibration constants before the adjustments to the electronic gains. The red points are after the adjustments.

The drum is not a monochromatic light source. The wavelength distribution of the drum convolved with the FD wavelength acceptance (see next section) needs to be taken into account. For the 375 nm LEDs a 3.5% correction is needed. In order to reduce some of the uncertainty, some new techniques with high powered LEDs have been tested. These LEDs have larger distribution widths and different central

wavelengths, but can have as much as 10 times more light output. The extra light output is useful for the NDF measurement and to be far above background for the photodiode measurements. However, the larger wavelength distribution combined with the FD wavelength acceptance needs a large, 16-20%, correction for the type of 380 nm LED we use. The uncertainty on the correction also is increased, so currently only the 375 nm LEDs are used to generate the calibration constants.

The calibration constants are defined as the number of photons each pixel sees off the drum surface per ADC count,

$$cc = \frac{N_{pixel}}{ADC} \tag{4.11}$$

Relating I_{drum}^0 to the number of photons that will be visible by each pixel in the FD cameras N_{pixel} is the next step. The solid angle, $d\Omega_{pixel}$, subtended by each pixel, the area of the aperture, A_{app} combine with the drum intensity to get N_{pixel}.

$$N_{pixel} = I_{drum}^0 \times d\Omega_{pixel} \times A_{app} \tag{4.12}$$

During calibration in the lab, the area of the drum is masked to be the same size as the aperture area. So the area of the drum cancels with the area of the aperture. Inserting the remaining factors that make up I_{drum}, the number of photons each pixel sees per pulse is,

$$N_{pixel} = k_{NIST} \times \frac{\lambda}{hc} \frac{i_{pd}}{A_{pd}} \times d^2 \frac{w}{NDF} \times d\Omega_{pixel} \tag{4.13}$$

It should be noticed that no information from the lab PMT remains in the final equation.

Based on the calibration constants the electronic gains from each pixel are adjusted so that the response is uniform across the camera. Figure 4.10 shows a plot of the calibration constants verses pixel id number for one camera before and after these adjustments. The camera is adjusted to have flat response. Each pixel views the

aperture at an angle. Given the approximate Lambertian surface of the drum the number of photons in each pixel field of view can be calculated given the absolute drum intensity I_{drum}. These $\cos(\theta)$ adjustments are not made in the calibration constants for historical reasons. The $\cos(\theta)$ adjustment eventually gets canceled out inside the reconstruction algorithms. If this effect were included the calibration constants would have a clearly identifiable geometrical pattern, instead of the relatively flat distribution with pixel number as seen in Figure 4.10 (right).

4.4   Multi-wavelength calibration

All the components of the fluorescence telescopes vary in their transmittance and or reflectance depending on the wavelength of light being used (see Figure 4.1). The initial piecewise measurement makes a first attempt at combining all the wavelength dependencies. The resulting curve is a good approximation to what the response should be for the entire detector. However, to take into account all the varying effects from the different components an end-to-end measurement is once again needed.

For multiple wavelength measurements, a xenon flasher is mounted at the end of the pipe in the back of the drum. The xenon flasher provides a pulse covering a broad UV spectrum, in a time period of a few hundred nanoseconds. Notch-filters are used (see figure 4.11) to provide output in bands of wavelengths of $\approx$20 nm FWHM. The filters are mounted in a filter wheel that attaches to the back of the drum. Notch-filters were chosen at five wavelengths inside the range of the FD UV filter located at the telescope aperture. The filter transmissions are centered at 320, 337, 355, 380 and 405 nm.

Histogram centroids, C^{λ}_{PMT} for each filter were made using the same reference PMT as in the absolute measurements. The relative quantum efficiency, QE, of the
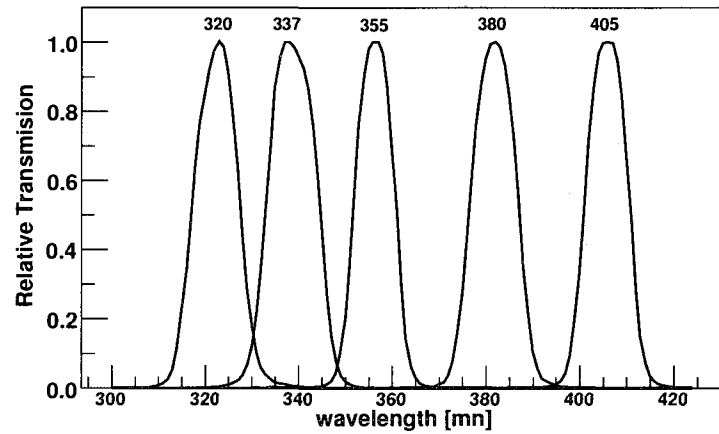
Figure 4.11 Relative transmission of the notch-filters used for multi-wavelength calibration. The nominal wavelength is indicated for each filter.
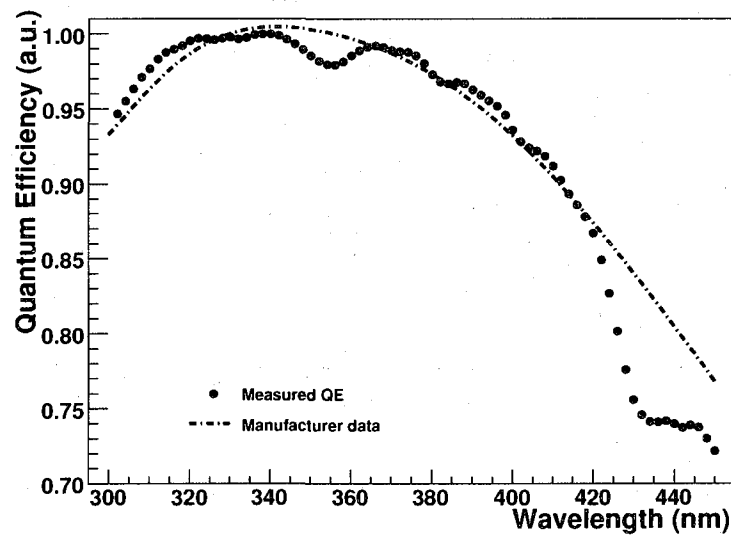
Figure 4.12 The measured quantum efficiency of the lab-PMT used in this work. The disagreement above 420 nm is beyond the detector input filter cutoff and therefore non relevant for this work.

reference PMT from 300-440 nm was also measured (see Figure 4.12). At each wavelength the recorded response from the reference PMT, combined with the quantum efficiency, yield a quantity that is proportional to the intensity emitted from the drum, I^λ_{drum}. The absolute intensity is not calculated and only the relative intensity between each of the filters is known.

$$I^\lambda_{drum} \propto C^\lambda_{PMT}/QE^\lambda \tag{4.14}$$

Once the relative drum intensity with each of the five filters is known the drum is taken to the fluorescence buildings. The drum is mounted in the aperture and the drum is pulsed roughly 400 times with each filter. Figure 4.13 shows the response of one camera to the drum powered by the xenon flasher with a specific wavelength filter. The histogram of the integrated pulses for a typical pixel is also shown.
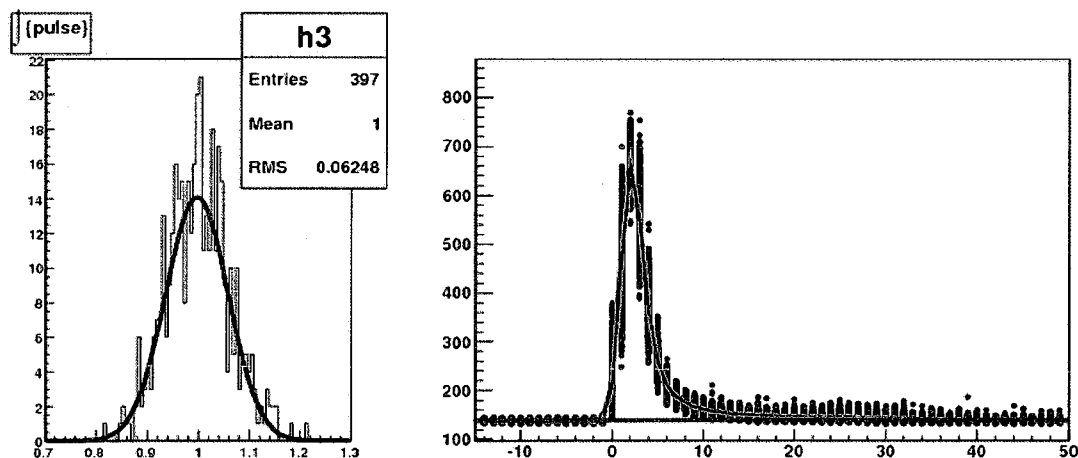
Figure 4.13 The traces from LL4 from the August 2006 trip for pixel 220 with the 337 filter. The distribution of the integrated signal has a RMS ≈ 6%. The black dots are the individual events, the red line is the average signal and the blue is the average pedestal.

Figure 4.14 shows the ratio of the 337 nm filter to the 380 nm filter for all the 440 pixels in four different cameras (the mean of 0.85 indicates that the FD at 337 nm is roughly 85% as efficient compared to 380 nm). Each color is a histogram of all

440 pixels in each camera and the gray histogram is the combined histogram of all four cameras. The individual distributions all have very narrow widths ($\approx$ 1%). The different cameras in each FD building have very similar responses. The RMS of the combined distribution is only 3%.

The systematic uncertainty for this measurement is small since most effects cancel in the normalized ratios used in each step. The estimated uncertainty for each filter measurement is on the order of a few percent with contributions from the measurement of the PMT relative quantum efficiency, notch filter transmission and xenon lamp source spectrum. To generate the final multi-wavelength curve only the average response of all the pixels in all cameras is used. This is done because the differences measured between pixels in each camera and between cameras is of the same magnitude as the systematic uncertainty. This average response at the FD combines with the lab intensity measurement to form a curve of relative camera response shown in Figure 4.14 (right). Interpolation between the measured points is based on the response curve from the piecewise measurement described in section 4.1. Corrections made for the filter widths and the statistical uncertainties for the measured points are included in the interpolation. This curve is normalized to the absolute measurement made with LEDs at 375 nm and used in FD event reconstruction.

4.5   The relative calibration system

Before and after each night of data taking a relative calibration is performed. The system in each FD building uses a set of optical fibers that distribute light from a high power LED to diffusers located in the centers of all six mirrors that simultaneously illuminate all the pixels [35]. Night to night fluctuations during each data taking shift of 14 days are shown in Figure 4.15.
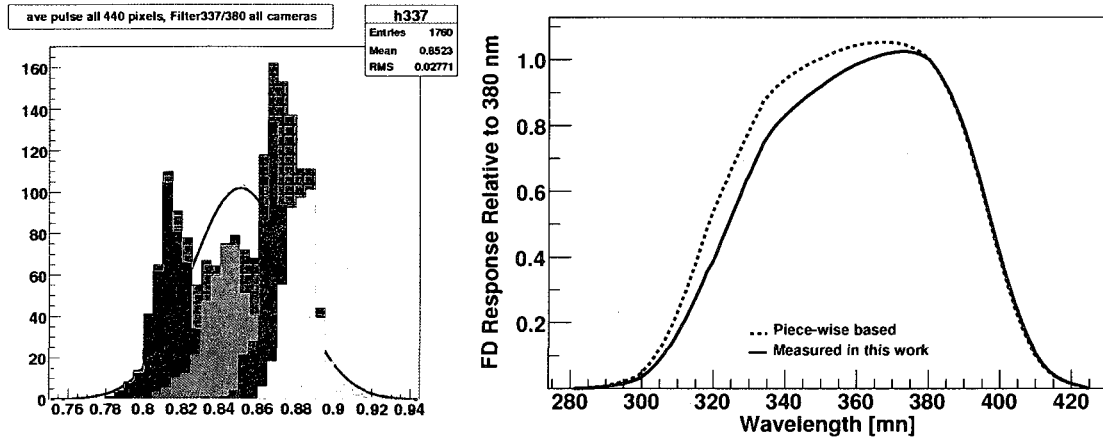
Figure 4.14 The figure on the left is the ratio of the FD response with the 337 nm filter and the 380 nm filter. The filter width correction have not been applied. Each color represents a different camera, the gray is the combined histogram of the four different cameras. Each camera has a narrow distributions. There is some variation from camera to camera. The average measured FD wavelength response for these four cameras is compared with the original piecewise curve made from manufacturer data.
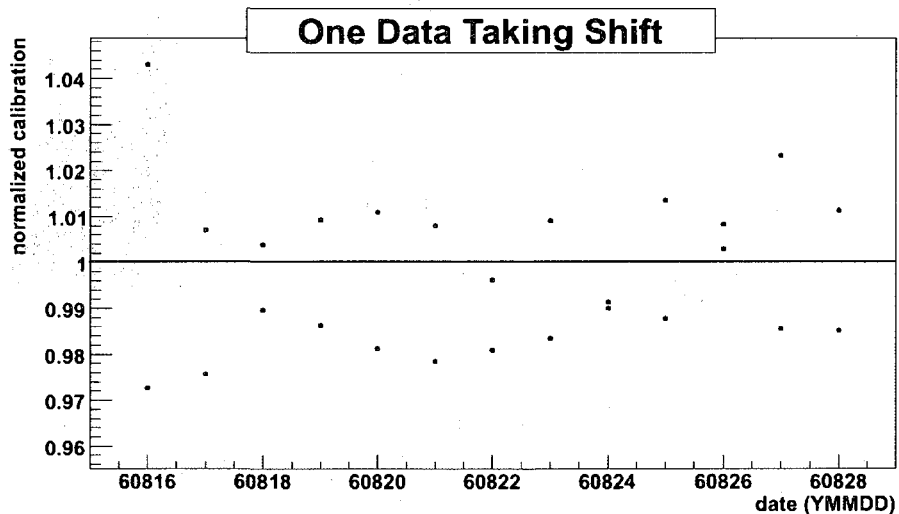
Figure 4.15 The general trend in the response of a camera during one data taking shift. Each point represents the average camera response. Except for the first night of the shift, the night to night fluctuations are small $\approx 2\%$.

Long term seasonal trends between data shifts have been identified and are shown in figure 4.16. Various absolute calibrations and laser cross checks (next section) have verified these long term trends. In order to reduce systematic affect from these trends the relative calibration is used to adjust to the absolute calibrations and track the relative drifts in camera response.
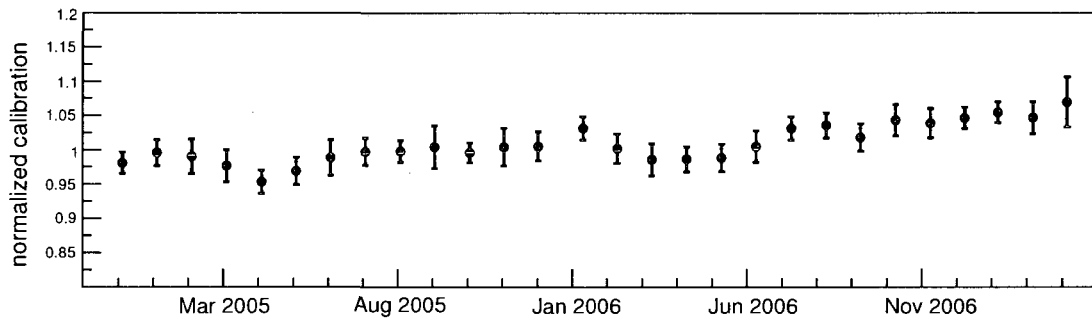
Figure 4.16 The change in the response of camera 4 Los Leones for almost a two year period. Each point represents the average camera response for one data taking shift (around 14 nights). The error bars indicate the one sigma night to night fluctuations. There is a general upward trend as well as up and down oscillations. The oscillations correlate with the seasons (January is summer and August is winter since the observatory is located in the southern hemisphere)

The number of ADC counts measured at the FD, the drum intensity calculated in the calibration lab and all the various corrections combine to give the calibration constants defined as photons per ADC count for each pixel in all the cameras at all the eyes. However, the calibration constants are valid for that night only due to the drift in response observed in the PMTs. In order to link the drum calibration to the relative calibration a reference relative calibration must be taken immediately after (or before) the drum is pulsed at the aperture. If the camera is allowed to warm up properly this reference run can be linked to the drum run with less than 1% uncertainty. This uncertainty is estimated from the combined effects of detector stability, systematics
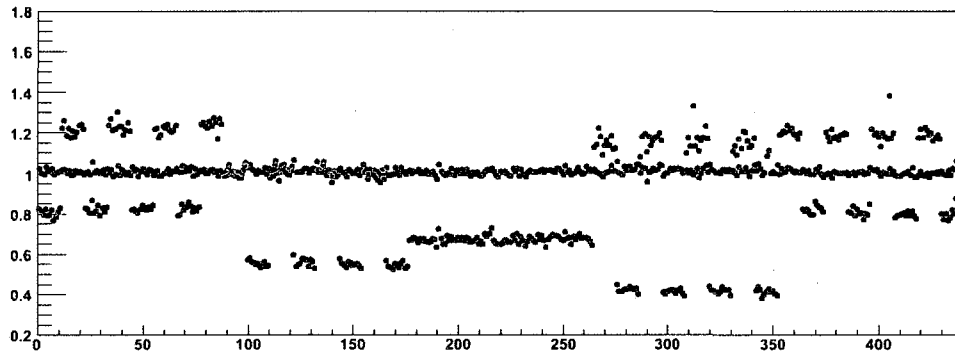
Figure 4.17 Several of the high voltage boards were replaced in this camera. The black points are the relative calibration constants before the change and the red are after the change took place. Hardware changes such as this motivate the start of a new epoch in the calibration database.

on the pulse integration, light source stability and operator error.

We have defined epochs of time representing stable periods of FD camera operation from December 2004 to the present. In each of these epochs a new reference relative calibration run was chosen to make an adjustment to the reference absolute calibration. The epochs span time periods from a few weeks to a few months. Hardware changes (Figure 4.17), seasonal variations (Figure 4.16) and any other variations motivate beginning of a new epoch. There are nightly fluctuations in response that vary at the 3% level in each epoch.

Since this relative calibration system is located inside the aperture, it is not an end-to-end measurement. There are possible effects that are not measured that may be significant. These effects could come from dust or dirt buildup on the UV filter or mirrors. Other relative calibration systems are used to measure these effects. Currently, the light sources in these systems are not stable enough to be operated every night consistently so there is not an extensive amount of night to night data available. By looking at this data long term changes can be investigated. Using

the limited data, no large change in response have been seen that the calibration system inside the aperture has not already taken into account. The possible effects that are not being measured on a night to night basis as measured by the inside aperture calibration system is estimated to introduce less then 2% uncertainty based on the [36].

During data taking shifts, the relative calibration can only be taken before and after data taking for practical reasons. For this reason there will be some uncertainties introduced even though the statistical error in the relative calibration is very small. Studies were done to evaluate these errors and they are estimated to be less than 2% when the FD is stable [36].

4.6   Crosscheck of the absolute calibration

A portable laser system has been created to provide a cross check of the drum calibration [37]. The laser is driven out into the field approximately 4 km in front of the FD buildings and fired vertically. The Rayleigh scattering cross section of the atmosphere is well known. This provides a flux of photons that can be predicted when the laser beam energy is measured by a calibrated energy probe. This technique includes all the components of the telescope giving an end-to-end calibration with systematic uncertainties that are completely independent from the drum calibration.

Using this method, only a track of pixels can be illuminated at one time (see figure 4.18). For other pixels in the camera to be calibrated the laser system has to be moved. Calibrating all the cameras in this manner is impractical so only a few pixels in certain cameras are cross checked in this way. The systematic uncertainties for this technique are less than 12%. The calibration of the energy probe ($\approx$ 10%) along with atmospheric corrections for non-perfect Rayleigh scattering ($\approx$ 5%) are
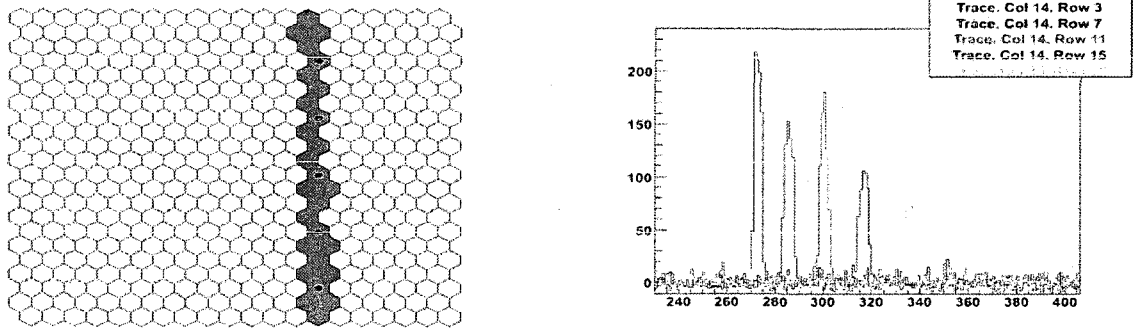
Figure 4.18 The roving laser illuminates only a track of pixels in one camera. The pulse shape in various pixels is shown on the right.

the main sources of uncertainty in this technique.

Figure 4.19 shows the consistency of the absolute measurements when the relative calibration is taken into account. The relative calibration system is used to track any changes in the response of the cameras due to hardware changes and seasonal trends from a reference absolute calibration.
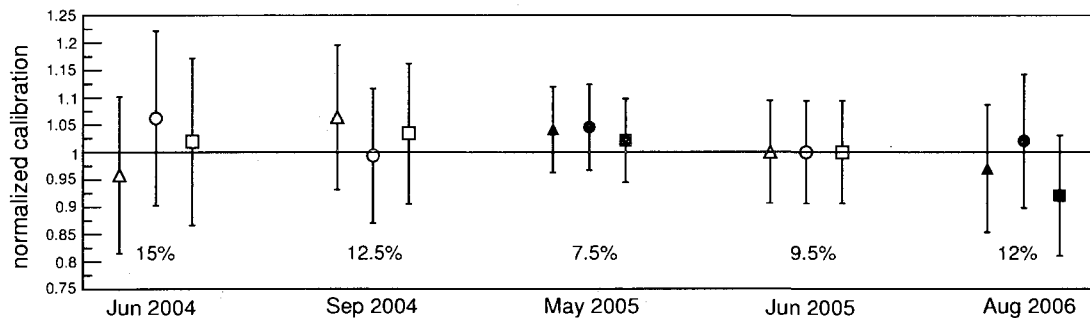
Figure 4.19 The absolute drum calibrations (hollow shapes) and the laser cross checks (solid shapes) for the past 3 years. The different shapes represent different cameras. The same three cameras are shown for each of the calibration trips. All points have been normalized to the June 2005 drum calibration. The systematic uncertainty for each trip is stated below each group of points. The uncertainties include both correlated and uncorrelated effects for each of the trips.

4.7 Uncertainties in the FD absolute calibration

The main sources of uncertainties in the FD calibration come from the variable intensity light source used to transfer the drum intensity to the NIST photodiode. Efforts to reduce the larger sources of uncertainties are ongoing. Table 1 summarizes the main uncertainties.

Absolute calibration:	
Drum intensity transfer to photodiode:	
pulse/continuous scale factor	5%
neutral density filter	3%
photodiode current readout	2%
NIST calibration [34]	1.5%
Temperature effects:	
Drum intensity differences between lab and FD	3%
PMT response differences between dark box and dark hall	2%
Reflections:	
at FD	1%
in lab	3%
Other:	
Wavelength distribution effects	3%
Drum non-uniformities	1%
ADC signal readouts	1%
Geometrical (alignments, areas, etc.)	1.5%
Relative calibration:	
reference measurement	1%
PMT response	3%
other effects (dust etc.)	2%
Total	**9.5%**

Table 4.1 Table of present uncertainties.

Chapter 5

Determining the energy of surface detector events

The first three chapters provided an introduction to cosmic rays, extensive air shower detection and the general reconstruction techniques of the Auger Observatory. Chapter 4 was an interlude into a discussion of a hardware task in which I was a leading member. The calibration of the fluorescence detector is important for not only the energy scale of the fluorescence detector, but for the surface detector as well. This chapter begins the analysis part of the thesis by discussing how the surface detector energy scale is determined. The next chapter calculates the efficiency of the surface detector. Finally, Chapter 7 takes the results from Chapters 5 and 6 to measure the energy spectrum of cosmic rays from $10^{18} - 10^{19}$ eV.

Recalling the information from previous chapters, a surface detector samples the density of particles in an air shower at a certain depth of shower development at several distances from the shower core. The density of particles at a distance far from the core, such as 1000 m, has been shown to relate to the energy of the primary cosmic ray. In the Auger data analysis, the density at 1000 m from the shower axis

is interpolated by fitting a function to the sampled densities. The interpolation uses information from the stations less that 1000 m and greater than 1000 m from the shower core. To relate the density measurement to the energy of the primary cosmic ray, the relationship between these two quantities is needed. Previous experiments have used Monte Carlo simulations to relate the measured shower density to primary energy. This strategy is difficult for the highest energy cosmic rays because the properties of the initial interactions in the air shower are beyond the limits that particle accelerators can probe. The specific parameters of high energy interactions used in the simulations must be extrapolated to the higher energies. The validity of these extrapolations is hard to quantify and thus large systematic uncertainties are introduced.

A fluorescence detector on the other hand, makes a calorimetric measurement of the energy. Accurate knowledge of the nitrogen fluorescence yield, atmospheric conditions and detector calibration are needed. All three factors can be measured directly. The atmospheric monitoring was touched on in Chapter 3 , the detector calibration was discussed in chapter 4, and a group inside the Auger collaboration is currently working to measure the nitrogen yield with better accuracy than previous experiments [29].

Both a surface detector (SD) and a fluorescence detector (FD) have advantages and weaknesses. A SD has a nearly 100% duty cycle and accumulates statistics quickly. A FD can only be operated on clear moonless nights which translates to a duty cycle of around 10%. The systematic uncertainties on the energy determination of a FD event are well understood, where as for a SD event, they are hard to quantify. The validity of the extrapolations used in the SD simulations can not be tested by accelerator data. The Pierre Auger Observatory has combined the two types of cosmic ray detectors at the same site. The Auger SD and FD are used together

to combine the advantages of both individual detectors. The well observed events that are reconstructed by both the SD and the FD, called golden hybrid events, are used for many systematic studies. The parameters of an air shower detectable by a SD can be compared to the parameters detectable by the FD to look for correlations. Specifically, this chapter discusses how the density of particles in an air shower at 1000 m from the shower core, S_{1000}, relates to the calorimetric energy measurement of the FD.

5.1 Golden hybrid events

The golden hybrid data set has an energy dependent acceptance. The acceptance of the FD grows with energy since higher energy showers produce more fluorescence light and thus can be seen at larger distances from the FD site compared with lower energy showers. The SD data on the other hand has uniform acceptance at high energies and only simple geometric cuts are needed at low energies. Details of the SD acceptance will be discussed in Chapter 6. The changing acceptance and limited field of view of the FD require strict quality cuts to insure an unbiased data set.

5.1.1 FD specific cuts

The quality cuts for the hybrid data used in this work were chosen to be similar to the cuts used in the Auger elongation rate analysis [38]. The elongation rate is a measure of the average depth of shower maximum, $< X_{max} >$, as a function of energy (see Figure 5.1 (bottom)). These cuts were chosen because they are unbiased in X_{max} which is an observable that correlates with primary composition. The golden hybrid data set is then representative of the average composition of cosmic rays. The major cuts are highlighted below.

- Cerenkov light less than 50%

- Distance from shower core to the WCD with largest signal less than 1000 m

- Almost the full profile of each event must be observed. Each shower enters the field of view of the FD at a slant depth X_{low}, and exits at X_{up} (which is normally ground level). The observed slant depth is then $X_{obs} = X_{up} - X_{low}$. The amount of slant depth needed to view shower profiles changes with energy. At lower energies a larger X_{obs} is needed because the shower develops higher in the atmosphere (see Figure 5.1 (top)).

- The statistical uncertainty from the profile fit for X_{max} less than 40 g/cm^2 ($\approx$5%) and for E_{Fd} less than 20%

- No large clouds, dust or fog are in the FD field of view

- A measurement of the Vertical Attenuation optical depth (VAOD) within an hour of the event (while the atmosphere is constantly monitored, there are technical glitches that cause some data to be unmonitored. These glitches were prevalent in the early working days (2004-2007) but are currently rare).

These cuts reduce the statistical sample of hybrid events by a factor of 20. However, they produce a sample of events that is well reconstructed, unbiased in X_{max}, and has low statistical uncertainties.

5.1.2 SD specific cuts

The cuts needed for the SD data are robust and are not very restrictive. As long as an event has a core location inside the array of water Cerenkov detectors (WCDs) and is of high enough energy it will be well reconstructed. The SD is fully efficient
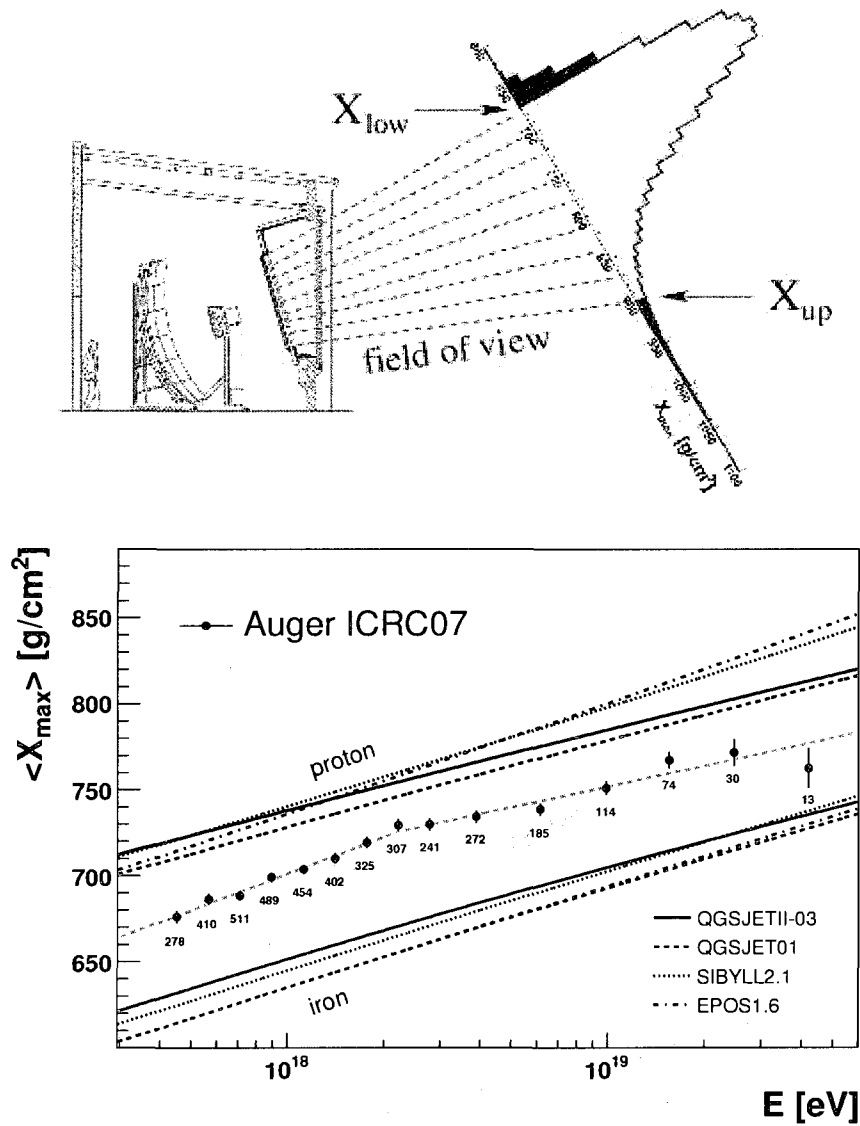
Bottom figure from [38]

Figure 5.1 In general the incomplete field of view of the fluorescence detector will have a bias on X_{max} distribution. This bias can be removed by only using the events when the majority of the profile is in the field of view (top). The field of view cut and the rest of the quality cuts were based on the elongation rate analysis (bottom). These cuts insure an unbiased X_{max} distribution.

for showers above $10^{18.5}$ eV. The efficiency drops for lower energies but the SD can still detect showers with energies down to $10^{17.5}$ eV. The minimum number of WCDs needed by the SD trigger is three. For the low energy region of interest in this thesis work ($10^{18} - 10^{19}$ eV) the vast majority of SD events have this minimum of three triggered WCDs. Chapter 6 will discuss the triggering and event selection criteria in greater detail.

With only three signals to fit a lateral distribution function (LDF), the statistical uncertainties from fitting are large. The average uncertainty on S_{1000} from the LDF fit for events with energy less than $10^{18.5}$ eV is 14% compared to higher energy events with is 6% or less statistical uncertainty. Besides the LDF fit uncertainly there are other uncertainties in the S_{1000} measurement which need to be considered when using S_{1000} as an energy estimator. Shower to shower fluctuations due to shower development are estimated to be 10% [39]. This fluctuation is due mainly to differences in depth the primary cosmic ray penetrates in the atmosphere before interacting for the first time. Another uncertainty is introduced by lack of knowledge of the true LDF (discussed earlier in section 3.2.3). This uncertainty is determined on an event by event basis and is typically less than 5%. These three uncertainties are combined in quadrature to obtain an average total uncertainty of 18% on the S_{1000} measurement for these low energy events.

5.1.3 The golden hybrid data set

The golden hybrid data set consists of the events that were well reconstructed with both the FD and the SD. The golden hybrid data is a subset of the SD data, but it is not a random subset. The average properties of these two sets do have differences. The angular distributions, in both zenith and azimuth angle, are very different. The azimuth differences will not contribute any biases because the SD reconstruction has a

negligible azimuth dependence. The zenith angle differences can be significant since measuring showers with different zenith angles with the SD implies measuring the lateral density at a different stage of shower development. Due to the quality cuts used the bulk of the well constructed golden hybrid data is between 25 and 50 degrees. For this reason and for SD efficiency reasons (see Chapter 6), only the data between 30 and 45 degrees is used in this analysis.

Besides zenith angle, there are other differences, such as temperature that could introduce a bias. All the hybrid data is taken at night and thus will have a average temperature lower than the average SD data temperature. The nights of winter are longer than summer so more hybrid data is taken in the winter, further biasing the temperature differences. The temperature effects on the SD data have been documented [40]. These effects only introduce a small bias in the SD energy determination because the extreme temperature variations only produce a maximum of 10% systematic difference on the S_{1000} value and the extremes represent only a small percentage of the total data. However, temperature can make a significant difference on the threshold signal detectable in each WCD and thus, will affect the trigger efficiency of the SD (more details on this will be discussed in chapter 6).

5.2 Energy calibration of the SD data

In Figure 5.2 the relationship between S_{1000} and the energy from the fluorescence detector, E_{FD}, is displayed for the golden hybrid events with zenith angles between 30 and 45 degrees that pass the quality cuts. The plot is on a log-log scale and is used to create what is called a calibration curve. The curve being a simple function that relates S_{1000} to energy. For each point in Figure 5.2 the 10% shower-to-shower uncertainty is not included on the S_{1000} value. Since this 10% uncertainty adds to

each event equally, it would not assign different weights to different points and thus not affect the fit. The other individual statistical uncertainties S_{1000} and the profile fitting uncertainties on E_{FD} are included on each point.

By the looks of the plot, a good choice for functional form between S_{1000} and E_{FD} would be linear.

$$log_{10}(S_{1000}) = m \; log_{10}(E_{Fd}) + b \tag{5.1}$$

With m, the slope and b, the intercept of the line. The relationship is purely empirical and by plotting the E_{FD} on the x-axis we treat the calorimetric energy measurement as the independent variable and S_{1000} as the dependent variable. This equation can then be inverted to obtain a conversion from S_{1000} to energy, which is called the surface detector energy, E_{SD},

$$log_{10}(E_{SD}) = \frac{log_{10}(S_{1000}) - b}{m} \tag{5.2}$$

This simple linear fit to the data in Figure 5.2 provides an energy calibration equation for the SD data. A linear relationship has phenomenological support and is intuitive if we recall what was discussed in Chapter 2.

As the shower develops, the initial energy of the primary cosmic ray is divided among the secondary particles produced. At deeper depths the number of secondary particles goes up and the average energy per particle goes down. These trends continues until that the average energy per particle is below a critical energy. Below this critical energy the secondary particles begin to be absorbed by the atmosphere, rather than creating more particles from interactions. The number of secondary particles thus grows rapidly, peaks at a maximum and then diminishes as the shower is absorbed in the atmosphere. Simple models of air showers predict that the total number of secondary particles at shower maximum grows proportionally with the energy of the primary particle [2].
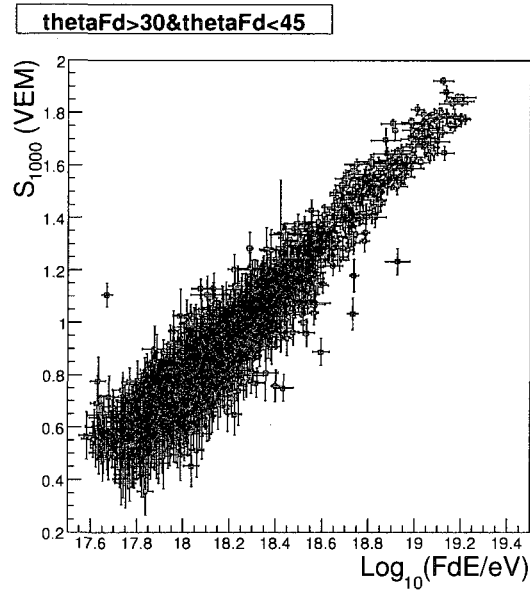
Figure 5.2 The relationship between S_{1000} and E_{FD} .

Despite the apparent linear relationship in Figure 5.2, we can not automatically assume proportionality because of model predictions. The shower simulations at these energies are not reliable and do not reproduce data very well [41]. Also, the SD does not measure directly the number of secondary particles at shower maximum. Only a small percentage of showers are actually detected exactly at shower maximum. In fact, all of the showers in Figure 5.2 were measured after shower maximum. Additionally the quantity S_{1000} is not the total number particles, but it is only related to the total number of particles at the specific depth of detection. The relationship between S_{1000} and total number of particles depends on the stage of shower development at which it is detected.

5.3 What is S_{1000} a measure of?

The SD samples the air shower at a specific depth of shower development, X_{ground}, which depends on the zenith angle. The maximum number of particles in the air shower occurs at X_{max}, which depends on the energy and composition of the primary particle. Therefore, the relationship between S_{1000} and total particles (thus the energy of the primary particle) will depend on the zenith angle and on the energy of the shower as explained below.

The slant depth at ground level from 30-45 degrees changes from ≈ 1000 to $1230\ g/cm^2$. These depths are well past the average shower maximum, $< X_{max} >$ of 700 -$740\ g/cm^2$ for the energy range of interest, $10^{18} - 10^{19}$ eV. After shower maximum the shower is losing electromagnetic particles to ionization rapidly. The average signal at 1000 m from the shower core should reflect this. For showers detected at larger depths (larger zenith angle), the signal at 1000 m from the shower core should be reduced. By making a calibration equation to all the data from 30-45 degrees, we are effectively saying that all the events should have the same S_{1000} value even though they are detected at different slant depths. This is inherently incorrect, but the 30-45 degree range seems to be small enough such that no drastic departures from a linear relationship Figure 5.2 are seen. Ideally, a two dimensional fit in both zenith angle and S_{1000} could be made to obtain a relation to energy. There are not enough golden hybrid events to do this yet but the use of smaller zenith angle bins is possible and the results are shown in section 5.4.

The measured quantity, S_{1000}, is an average characteristic of the shower that depends on both the electromagnetic and the muon content of the shower at X_{ground}. After X_{max}, the electromagnetic and the muon components of the shower are attenuating in the atmosphere, but at substantially different rates. The electromagnetic

component attenuates much faster as can be seen in Figure 5.3. The number of electromagnetic particles is vastly greater than the number of muons in an air shower but as the difference between X_{ground} and X_{max} becomes larger, the signals in the WCDs of the SD become more and more muon dominated. When this is the case, S_{1000} only depends on the muon component of the shower. The number of muons is predicted to be proportional to energy so a linear fit is expected in this situation.
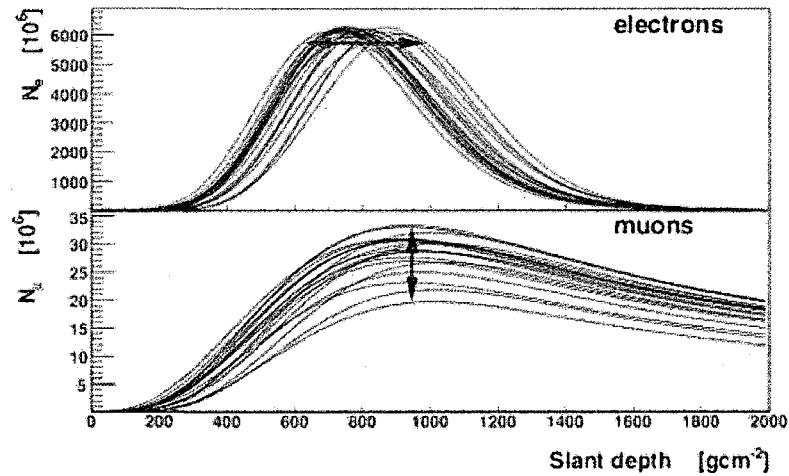
figure from [19]

Figure 5.3 The muonic and electromagnetic components attenuate at much different rates after shower maximum. Each of the 15 lines are for a simulated showers with energy of 10^{19} eV.

This situation occurs in two ways. For large zenith angles, $X_{ground} - X_{max} \gg 0$. Here only muons survive at ground level to contribute to S_{1000}. The Auger horizontal shower reconstruction task (60-90 degrees) exploits this by using the number of muons detected by the WCDs rather than S_{1000} [42]. Additionally, since low energy showers have a shallower X_{max}, the $X_{ground} - X_{max}$ difference can be large even for vertical showers. Observatories that look at lower energy cosmic rays can use muon counters at very deep depths to maintain sensitivity to air showers. Muon counters are not practical for the highest energy cosmic rays since they are more expensive than WCDs.

The expense makes deploying a large array cost prohibitive.

When $X_{ground} - X_{max} \approx 0$, electromagnetic particles completely dominate the signals in the WCDs closer than 1000 m to the shower core. The far away WCDs will mainly be triggered by muons since muons have a greater lateral spread. When S_{1000} is interpolated by fitting the signals in the WCDs, it will represent a combination of the electromagnetic and muon components. As the energy increases, the lateral spread of all components also increases such that the signals in the far away WCDs will also have significant electromagnetic contributions. Here, S_{1000} will be mostly dominated by electromagnetic particles. The total number of electromagnetic particles is predicted to be proportional to energy so a linear fit is also expected in this situation.

The above phenomenology illustrates that the average shower characteristic that S_{1000} represents may not be the same for all showers detected by the SD. The S_{1000} to E_{FD} relationship is expected to be approximately linear in limited ranges of both energy and zenith angle. In these limited ranges, S_{1000} represents a consistent shower attenuation of either the electromagnetic component, the muon component or the combination of components.

In a recent Auger publication [12], the relationship between S_{1000} and energy was shown to be linear for showers from approximately $10^{18.5} - 10^{19.5}$ eV. This was expected because the Auger Observatory was optimized for showers around 10^{19} eV. The altitude of the Auger surface detector was chosen so that air showers would be detected within a few hundred g/cm^2 of shower maximum for primary particle energies of 10^{19} eV. For energies higher than $10^{19.5}$ eV there are not enough hybrid data to verify if the relationship holds. At the very highest energies, the shower may be completely dominated by electromagnetic particles and the relationship may need to be modified. In the future ($\approx$ 5 years), the relationship can be tested up to 10^{20} eV with the greater number of golden hybrid events that will be recorded.

There is good reason to suspect that the relationship could be different below $10^{18.5}$ eV and there are ample statistics to measure the difference. Showers below $\approx 10^{18.5}$ eV generally only trigger three or four WCDs. This means that most of the WCDs in an event are relatively far away from the shower core compared to the total size of the footprint of the shower (more details on the shower footprint and number of WCDs in events is discussed in Chapter 6). This combined with the fact that the shower is detected at a depth well past X_{max}, means that the signals in the WCDs are becoming more and more muon dominated. A different relationship, or different linear fit is possible for the lower energy showers.

The apparent linear relationship seen in Figure 5.2 between S_{1000} and E_{FD} supports the idea that we are in a sufficiently limited range in energy and zenith angle. However, the phenomenological review given above suggests that the relationship will may not be the same for all energies or zenith angles. This is considered in the next section when the SD energy calibration is refined.

5.4 Refining the SD energy calibration

Using all the data from 30-45 degrees and the simple linear relationship of equation 5.1 is a valid first attempt at a surface detector calibration. The accuracy of this calibration is improved upon in this section with further refinements to better account for the phenomenology of air shower development.

5.4.1 Zenith angle bins

To take into account the differences in shower development at different observed slant depths, the data is binned in small zenith angle bins. Five bins of equal size in $cos^2(\theta)$ (≈ 3 degrees each) are used to make five calibration curve plots. The size of the bins

was limited by the available statistics. In the future smaller angular bins will be possible and eventually a 2-D fit function could be used. Figure 5.4 shows two of these plots. The values in the left plot correspond to a slant depth of $\approx 1020 \frac{g}{cm^2}$ and the right plot of $\approx 1210 \frac{g}{cm^2}$. The lower values of S_{1000} in the right plot reflect the lower number of particles that remain in the air shower at the later stage of development.
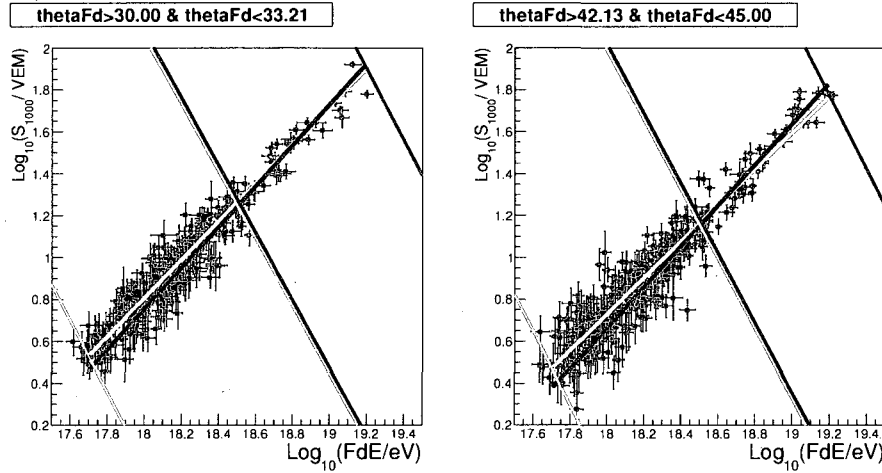
Figure 5.4 Different calibration equations are made for different zenith angle bins. The S_{1000} value for $Log_{10}(E_{FD}) = 18.4$ is lower for the larger zenith angle bin. This is due to measuring the shower at a later stage of development (here 150 $\frac{g}{cm^2}$)

The flux of primary cosmic rays has been shown to be isotropic except for the highest few handfuls of events [6]. If we keep this fact in mind, the impact of this refinement can be demonstrated by looking at the distribution of the number of events above a threshold energy versus the cosine squared of the zenith angle, θ. Since the integrated acceptance of the surface detector grows as $cos^2(\theta)$, the number of events in in each bin of $cos^2(\theta)$ should have the same. The distribution of events should be flat when plotted against $cos^2(\theta)$ for energies above 10^{18} eV. A flat $cos^2(\theta)$ distribution means that the detected flux of cosmic rays does not depend on the local coordinate system. This statement assumes that the efficiency of the surface detector

has a uniform zenith angle dependence (the efficiency does not have to be 100%, uniformity is the key). In general, for all angles and energies this is not true. For example, the efficiency goes to zero for angles greater than 90 degrees. However, in the 30-45 degree range for these energies the zenith angle dependence is uniform within 5-10%, so a flat $cos^2(\theta)$ is expected.

The impact of this refinement becomes clear if we compare the resulting SD energy calibration to a calibration without it. To do this, all the golden hybrid events from 30-45 degrees were initial used in one large zenith angle bin (like in Figure 5.2) to generate the SD energy calibration equation. This calibration yields an energy estimate, E_{SD}^{1-bin}, and was applied to the entire SD data set. On the left in Figure 5.5 the distribution of all the SD events versus $cos^2(\theta)$ is shown for energies above 10^{18} eV using the E_{SD}^{1-bin} energy assignment. The distribution is clearly not flat. There are fewer events at higher zenith angles $(cos^2(45)=0.5)$ than at lower angles $(cos^2(30)=0.75)$. Higher zenith angle events have an underestimated energy, so more events drop below the energy threshold compared to lower zenith angels.
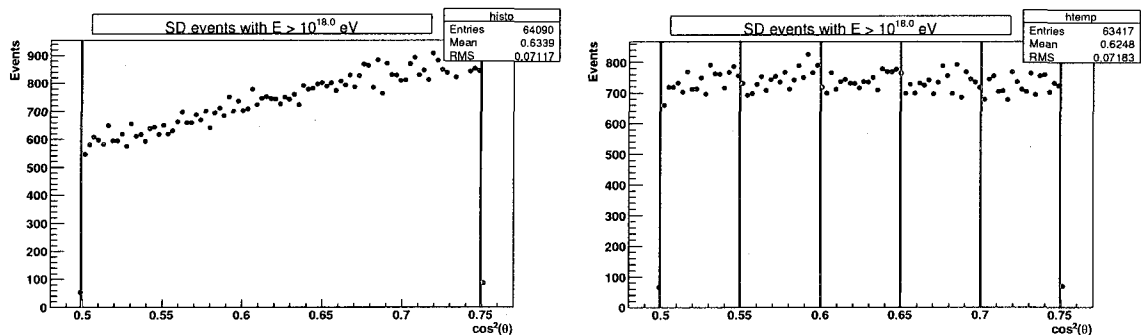
Figure 5.5 The number of SD events versus $cos^2(\theta)$. The left plot shows the bias introduced by using only one zenith bin in the calibration curve(i.e. assuming all showers have the same slant depth). The right plot uses 5 zenith bins of $\approx$3 degrees each for the showers from 30-45 degrees in zenith.

Using 5 smaller zenith angle bins in the range from 30-45 degrees and generating 5

different calibration equations results in an energy assignment, E_{SD}^{5-bins}. This energy assignment takes into account the different slant depths better than E_{SD}^{1-bin}. When E_{SD}^{5-bins} is applied to the entire SD data set the distribution is much flatter (the right of Figure 5.5). In each of the 5 zenith bins, the same basic effect remains, but with a much reduced magnitude.

5.4.2 Two fitting regions

For each zenith angle bin, the calibration plot is split into two regions. In each region a linear fit to the data points is made. In each region, the slope may be different because the air shower property that S_{1000} represents may be different. Figure 5.4 shows an example for two different zenith angle bins. As mentioned above, a previous Auger analysis demonstrated that the slopes in the upper regions are expected to be the same in all 5 zenith angle bins, with only the intercept changing. However, this was not the case when the slopes were fitted for initially. The low statistics of the upper region (due to the steeply falling flux of events) resulted in slopes that were not constant. To amend this, the value of the slope in the upper fitting region was calculated by a fit to the all the events into one plot (as in Figure 5.2 from section 5.2). Using this fixed slope, the intercept in each zenith bin was deduced by fitting the data in each zenith angle bin. In the lower regions, the higher statistics allowed for fitting of both the slope and intercept.

Figure 5.6 shows the fitting results in both regions from the 5 different zenith angle bins. The progression of the lines from lower to higher zenith bins can be seen. The lines in the two regions have not been constrained to run through the break point. The break point was somewhat arbitrarily defined to be $Log_{10}(E_{FD}/eV) = 18.5$. The choice of this value was motivated by the lower bound used in the analogous calibration curve from the recently published work from Auger [12]. The resulting
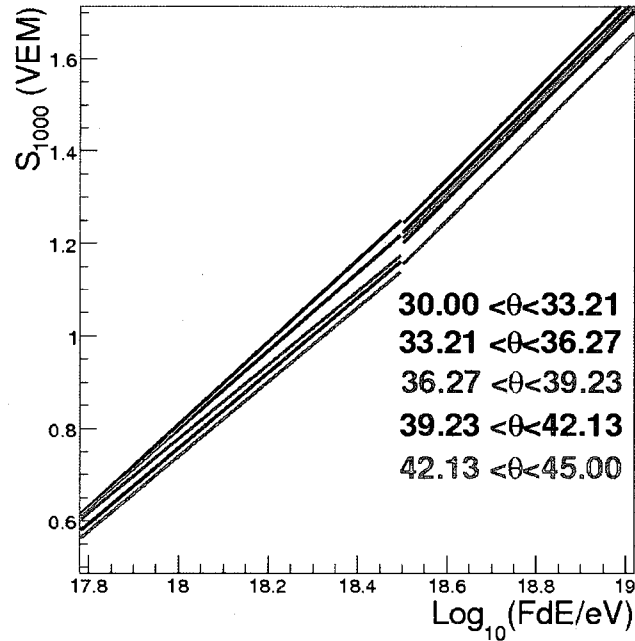
Figure 5.6 All the fitted calibration curves.

SD energies do not change significantly (less than 5% at 10^{18} eV) if the break point is changed by ± 0.2.

If the break point is lowered below $Log_{10}(E_{FD}/eV) = 18.3$, the upper region fit becomes dominated by the higher amount of statistics just above the new break point. The resulting upper region fit does not match well with the higher energy points on the plot. In this case, a third fitting region would could be used to obtain consistency for all energies. Three or more fitting regions and in the limit, a continuous function, would be ideal, but this is statistically impractical at this point. The functional form is not know a priori from models and simulations so the simple two region linear approximation is as well as we can do to accurately match the data over the entire energy range.

The reason for this change of slope in the two regions is most likely the result of

the phenomenological effects discussed above. The change in slope is in the direction expected if S_{1000} changes from a value that depends on both electromagnetic and muon components to one that is muon dominated. However, possible detector threshold effects could introduce a similar change in slope. For a given energy event, the detected S_{1000} value will fluctuate around the mean. At these low energies, with only three WCDs in the event, the statistical uncertainty on S_{1000} is large. The S_{1000} value can fluctuate downwards below the triggering threshold for the array. This can cause the data sample to be biased by upward fluctuations. The expected change in slope would be in the same direction as the phenomenological effects.

No phenomenological claims are being made, only the fact that a change in slope better represents the data. This two region fit is only an approximation of what is really happening in air shower phenomenology. With more statistics and the new enhancements to the surface detector (see section 7.4), it may be possible to better understand the reason for the change in slope. This understanding could help refine detector and air shower simulations.

Using an approximation will introduce systematic uncertainties. Specifically, for the S_{1000} values near the break point the choice of using the upper or lower line will systematically shift the energy in one direction or the other. This has a small effect on the energy assignment ($\approx 5\%$), when applied to the entire SD data set. However, the steeply falling flux of cosmic rays implies the number of events just below and just above the break point are not equal. For S_{1000} values near the break point the flux will be systematically shifted due to fluctuations. The measured flux at the break point energy value has a larger systematic uncertainty than at other energies. This will be mentioned again in Chapter 7 when the uncertainties on the cosmic ray flux are discussed.

5.4.3 Anti-bias cut

In each fitting region an anti-bias cut has been used as well. These cuts are demonstrated by the lines in Figure 5.4 and Figure 5.7. The lines are not exactly perpendicular to the linear fit lines. The reason for these cut lines is to reduce the impact of fluctuations on the fit. If the data were to be cut on simply one variable, say E_{FD}, then the fit would be biased by the upward fluctuating S_{1000} values. If the cut on the data was with S_{1000}, then the fit would be biased by upward fluctuating E_{FD} values. Upward fluctuations do not cancel with downward fluctuations because of the steep spectrum of cosmic rays. A perpendicular cut will reduce this bias (in the limit of infinite statistics the bias would be completely eliminated). To arrive at the perpendicular cut an iterative procedure is needed. A fit is done with no cut, then a first attempt at a perpendicular cut is tried and then a new perpendicular cut is defined. This procedure converges after only three or four iterations.

A perpendicular cut is only appropriate if the uncertainties on the both S_{1000} and E_{FD} are equal. Since they are not, 18% verses 8%, the cut line needs to be modified from the perpendicular slightly. The value of this modification was not optimized for each zenith bin, instead was chosen to be equal in all bins since the uncertainty in both S_{1000} and E_{FD} are approximately independent of zenith angle. The final calibration curves for each of the 5 zenith angle bins are shown in Figure 5.7.

5.5 Energy Resolution

Five different energy calibration equations from the plots in Figure 5.7 were made. The resulting SD energy, E_{SD}, was compared with the FD energy, E_{FD}. The histogram in Figure 5.8 shows the percent difference in energy assignment. A fit of a Gaussian distribution yields a width of $\approx 21\%$. This width is effectively the com-
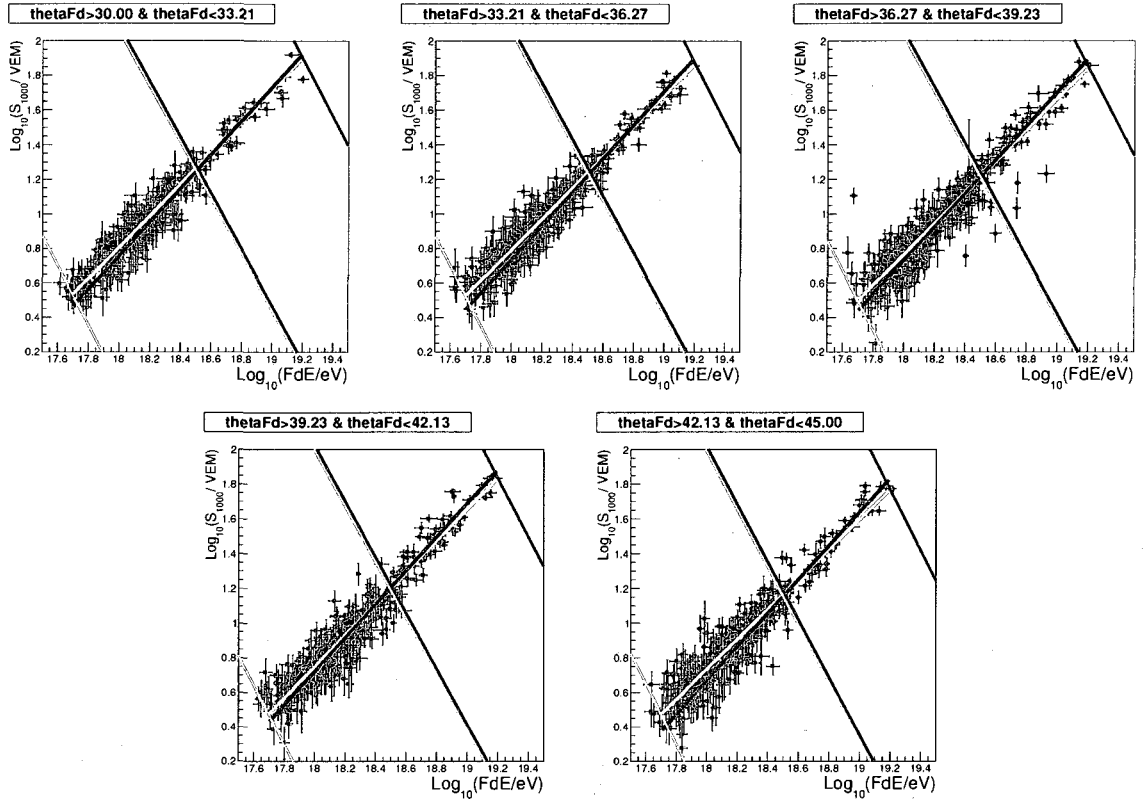
Figure 5.7 The final calibration curves for the five zenith angle bins.

bined scatter of the points around the fit lines in each of the calibration plots in Figure 5.7. The width is representative of the energy resolution of SD calibration and is consistent with what would be expected from the associated average event-by-event uncertainties on S_{1000}, 18%, and E_{FD}, 8%, are added in quadrature. This implies other uncertainties not considered (from the size of the zenith angle bins for example) are not biasing the SD energy calibration significantly.

The importance this SD energy calibration procedure can be seen by comparing the results with the SD energy assignment that was optimized for higher energy events and used in the recently published Auger results [12]. The plot in Figure 5.9 shows the percent difference between E_{SD} and E_{FD} as a function of energy. The solid points
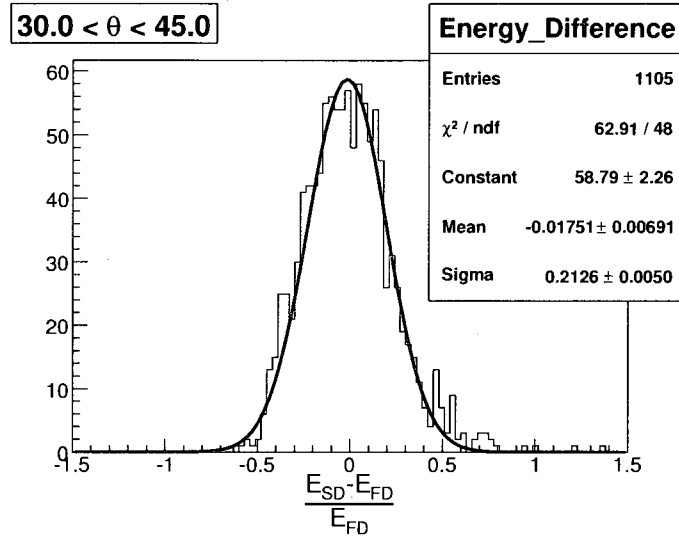
Figure 5.8 Combining the different calibration equations for each angle bin and comparing E_{SD} with the E_{FD}. The histogram has a width of 21%.
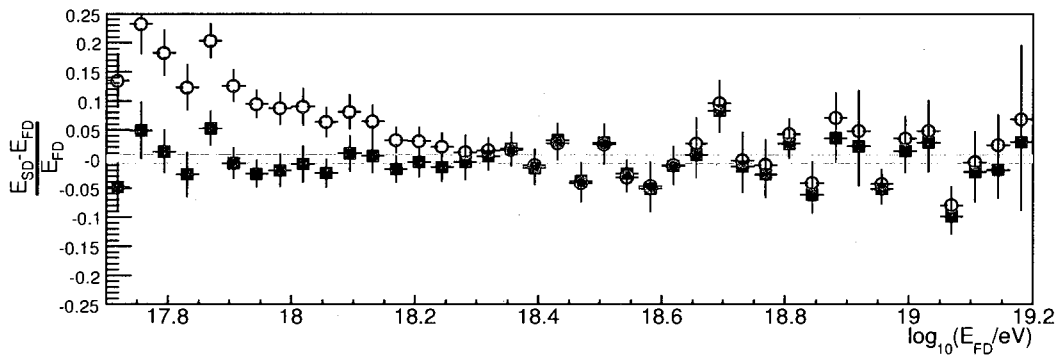
Figure 5.9 Comparing E_{SD} with the E_{FD}. The solid points are based on the SD energy calibration described in this chapter. The hollow points are based on the Auger high energy optimized SD energy calibration.

are the SD energy calibration described in this chapter and show very good agreement

with the FD energy over the entire energy range. The hollow points show a significant

bias at low energies.

Chapter 6

The efficiency of the surface

detector

The acceptance of the surface array is composed of the effective surface area and the solid angle window used. The efficiency and the uptime of the detector are then needed to calculate the time integrated exposure. This chapter explains the efficiency calculation for the surface detector for energies below $10^{18.5}$ eV. The next chapter uses this efficiency in the exposure calculation which is needed to plot the energy spectrum.

6.1 Introduction

The 1600 WCDs in the surface array compose roughly 3000 km^2 active detection area. Each point in this area can detect air showers for all azimuth angles and zenith angles from 0 to 60 degrees (showers greater than 60 degrees can be detected as well but different reconstruction algorithms are needed). The symmetry of the layout allows the array to be though of as a collection of individual unit cells. For the Auger

surface array the unit cell is a hexagon of WCDs (see Figure 6.1). The shower core coordinates from each event are translated to the ideal unit cell by a simple vector subtraction.

Understanding the detection efficiency of each unit cell is sufficient for understanding the efficiency of the whole array. It will be shown in later sections that the unit cell has further symmetries and a triangular cell is more appropriate for understanding the efficiency of the array at low energies.

The efficiency of the surface detector depends on the number of secondary particles in the EAS development at ground level. The number of particles depends on the primary particle energy, composition and the stage of a shower development when detected. As described in previous chapters, the zenith angle of shower relates strongly to the stage of shower development. Along with a zenith angle dependence, the efficiency will depend on the location of the shower core in the until cell. This is due to steeply declining lateral distribution of particles (as discussed in chapter 3).

The efficiency of the surface detector will saturate, or become 100% efficient, at high energies. For showers with zenith angles less than 60 degrees the saturation energy has been shown to be $10^{18.5}$ eV [43]. Below $10^{18.5}$ eV, the zenith and core location dependent efficiency needs to be known to measure the exposure and thus the flux of cosmic rays. There is an abundance of data at these low energies due to the power law flux of cosmic rays. Distributions of the core locations for this data show well defined average properties. Two models have been created to understand how a decreasing efficiency could cause these distributions. The models were used to understand the general trends in the data to see what can be deduced from the data alone. The models also show how the hybrid data can help in the measurement of the surface detector efficiency. All calculations of the efficiency that are discussed in the later sections were strictly made from the data alone without any dependence on
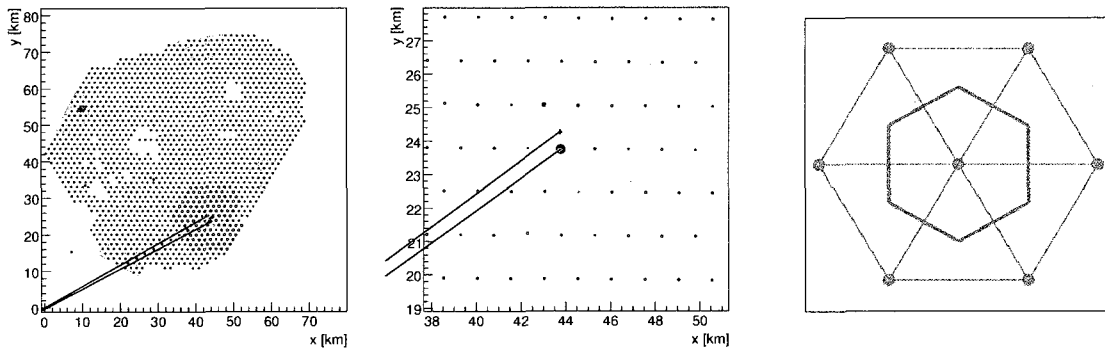
Figure 6.1 The ideal unit cell of the Auger surface detector is a hexagon of WCDs. In each event the WCD with the highest signal can be represented by the center station. All events in the array can be translated to the ideal hexagon by subtracting the hot tank vector from the shower core vector.

the specific parameters of the models. One of the models was developed by a fellow Auger collaboration member [44], and the other was developed specifically for this thesis.

The efficiency as a function of energy has been found for energies below $10^{18.5}$ eV with zenith angles between 30 and 45 degrees. Section 6.2 will give an overview of the trigger and event selection that will provide the needed background for the following discussions. Section 6.3 discusses a model of the zenith angle dependence. Section 6.4 introduces a toy model that accurately reflects the core location dependence seen in the data. Section 6.5 compares the models to the actual surface detector data and the hybrid data. Finally section 6.6 reviews the techniques we developed to measure the efficiency of the surface detector from the data itself. In the next chapter the associated uncertainties on the efficiency calculation are propagated to the exposure calculation.

6.2 The SD trigger and event selection

The high energy cosmic ray showers of interest are constantly incident on the surface detector. Along with these showers there is a constant bombardment of background secondary cosmic ray muons from the much higher flux of low energy cosmic ray showers. As described in Chapter 3 these muons provide a way to calibrate the individual WCDs, so they are useful, but a trigger system is needed to distinguish between the background muons and the cosmic ray showers with energy above 10^{18} eV.

Each WCD of the surface array is continually recording data and communicating with the central data acquisition system (CDAS). There are local triggering algorithms on each individual WCD as well as a central trigger that looks for geometrical patterns in groups of WCDs that have local triggers coincident in time. These time coincidences are signatures of large air showers.

6.2.1 Local station triggers

The first level trigger, T1, has two different algorithms used to scan the FADC trace. The first, called threshold (Thres), simply looks for a time bin (25 ns) with an amplitude greater then 1.75 VEM. The second, called time over threshold (ToT), looks for 13 consecutive time bins (out of a total of 120 in each trace) with amplitude over 0.2 VEM. The second level trigger, T2, also acts on the local stations. All the ToT triggers automatically pass the T2 trigger and the Thres triggers must have at least one time bin with signal above 3.2 VEM.

The rates of the Thres T2 triggers are uniform in all the WCDs in the surface array within a few percent. This is expected because each WCD is, in principle, equally sensitive to signals sufficiently above background (here background refers to the dark current in the PMTs). There are seasonal effects that cause modulations in

the T2 rate that are observed to be uniform throughout the array. The ToT trigger is sensitive to low amplitude signals spread in time. The ToT rates have more variation from WCD to WCD, ranging between 0.5 to 5 Hz. The sensitivity to these low amplitudes depends on the individual PMTs, the characteristics of the water in the tank and temperature variations. This lack of uniformity of the trigger response over the array must be factored into the calculation of the efficiency of the surface array.

The local station trigger is broken into two levels to reduce the technical specifications of the electronics of each WCD. The T2 trigger, reduces the local station event rate from ≈3 kHz (all signals) to 20 Hz (signals that pass the T2). The local station stores all the higher-rate T1 FADC traces in local memory. However only the signals from the T2 triggers are sent to CDAS because of communication bandwidth limitations. The thresholds have been chosen so that each properly functioning WCD should have a 20 Hz T2 rate. This rate is constantly monitored and is used for diagnostic purposes and to measure the working uptime of each WCD. A sudden drop off in the T2 rate from a station indicates the station has a problem and needs to be attended to. The problem could be in the communication or could be a different hardware failure. A sudden spike in the T2 rate is also indicative of a problem. Lighting from a storm over the array can cause large spikes in the T2 rate. In the early deployment stages, failures was common and average uptime for each WCD were only 85-90%. More recently the uptime for each WCD is greater than 95% [43].

The next level trigger, T3 is a central trigger that uses the timing and coordinate locations of the T2 triggered stations. If a T3 trigger is fulfilled then FADC traces from both the T1 and T2 triggered WCDs are transfered from local memory to CDAS. Once the traces are transfered the local memory is cleared. The memory is also cleared in the local stations if the station is not part of a T3 event. This allows for a small local memory requirement but also implies that data could be lost if for some reason

the WCD had communication troubles with CDAS.

6.2.2 The central level trigger

The third level of the trigger, T3, looks for time coincidences and geometrical configurations of the T2 triggers. Once a T3 trigger has been fulfilled, any nearby stations with T1 triggers are also included in the event provided that the signals are in time coincidence with the stations in the T3 trigger.

In the T3 algorithm, all of the T2 triggered stations that were sent to CDAS are scanned for signals that occur within a specific time window. The stations with signals inside the time window are then checked to see if the distances separating the signals are compatible with the speed of light. The stations are then searched for geometrical patterns. Two algorithms are used for the T3 trigger. The need for different algorithms arises from the vastly different topological patterns between near vertical (0-60 degrees) and near horizontal showers (greater than 60 degrees).

The first algorithm scans for three stations with ToT type T2 triggers in a compact triangle of stations. The rate of events that pass this trigger algorithm is currently $\approx$1800 per day and was approximately 1.1 $\times$ the number of working WCDs during the deployment of the surface array. This geometrical configuration selects a very pure sample of real showers from the background. For showers that have zenith angles less than 60 degrees, the samples are $\approx$90% pure.

The second T3 algorithm looks for a four-fold coincidence of stations with either type of T2 trigger within $\approx$6km of each other. This allows very loose geometrical configurations to pass the T3 trigger. This configuration has a low selection purity but is needed to have sensitivity to showers with zenith angles greater then 60 degrees. Of the events that pass this type of T3 trigger only a few percent are real showers. The T3 trigger was not designed for purity, but was instead designed to include as much
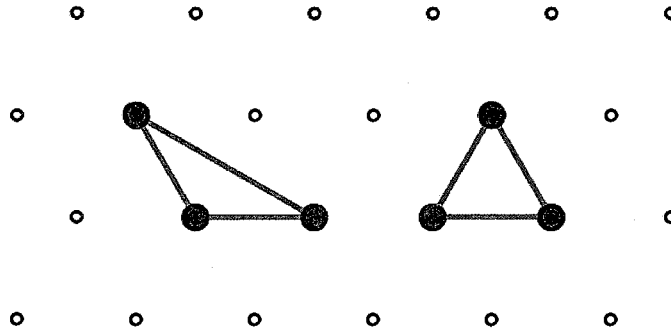
Figure 6.2 A compact triangle can be either equilateral (right) or isosceles (left). The three WCDs that define the triangle are tested for compatibility of a plane shower front moving at the speed of light.

data as was practically possible to make sure all real physics events were recorded. Off line event selection is then needed to reduce the T3 data stream to real physical showers from high energy cosmic rays.

6.2.3   Event selection

Different selection criteria are needed for events with zenith angle less than or greater than 60 degrees. Only the criteria for less than 60 degrees are discussed here. The first event selection cut, T4, filters the T3 data down to real physics events. The second cut, T5 is used to insure a non-biased reconstruction due to edge effects.

The criteria needed to fulfill the T4 cut are more strict than the loose T3 requirements. At least three stations in a compact triangle (see Figure 6.2) must survive for the T4 requirement. The compact triangle defines a plane that is used to check the other stations for time compatibility with a shower front in this plane. Any stations not compatible are removed and labeled as "accidental." These stations are not used in the event reconstruction.

If a shower lands near the border of the array the reconstructed core location will be pulled towards the array and the shower parameters will have a systematic bias. The T5 cut limits the events to the showers that are contained in a unit cell of WCDs that are actively working. Such a unit cell is called an active hexagon. The six WCDs surrounding the WCD in the event with the largest integrated signal must all be working properly. The events on the border of the array cannot pass the T5 requirement. This limits the usable size of the array by eliminating the events on the border of the array and events that land near holes in the array due to nonworking stations or stations that were not yet deployed. The activity level of each unit cell of WCDs is monitored by the T2 rates. The T5 cut reduces the number of events by $\approx 25\%$, where 20% come from edge events and holes from non-deployed stations and 5% from non active unit cells due to temporary hardware or communication failures.

The multileveled trigger and event selection (T1-T5) efficiently reduces the event rate to $\approx 3 \ 10^{-5}$ Hz from the ≈ 3 kHz rate of detected background plus signal rate at the local station without any significant loss of real showers [43]. The 20 Hz T2 rate is continually sent to CDAS and allows for constant monitoring of the uptime of each WCD. The next section will show how the efficiency of the array is modeled based on these trigger and event section criteria.

6.3 Zenith angle dependency of the SD efficiency

This thesis analysis is limited to the study of showers with zenith angles constrained to the zenith angular window of 30-45 degrees. This choice was motivated by the model that was developed by a fellow collaborator [45]. This model is based on a full air shower simulation and a full detector simulation. The reliance on air shower simulations inherently introduces hard to quantify systematic uncertainties (due to lack

of knowledge of the primary composition, unknown hadronic interaction parameters, etc.). For this reason, this model was only used to make relative predictions on the zenith angle dependencies of the surface detector efficiency.

The full details of this model will not be discussed, but the final results can be seen in Figure 6.3. Here different efficiency curves are shown for different zenith angles for simulated proton showers of different energies. The efficiency drops steeply for all angles below 3×10^{18} eV. The range from 30-45 degrees is predicted to be more efficient than the more vertical (0-25 degrees) or the more inclined (53-60 degree) showers. Similar plots were made for iron primaries and a mixed composition scenario. The general trends seen in Figure 6.3 hold for all composition scenarios.

The 30-45 degree zenith angle range has the majority of the well reconstructed hybrid events. As discussed in the previous chapter, the hybrid events are crucial to generate reliable calibration curves. It will be shown in the following sections how the hybrid events are also crucial to efficiency calculations. The choice to limit this analysis to this zenith angle between 30-45 degrees is somewhat arbitrary. The techniques developed later in the chapter could be generalized to other zenith angles. This generalization would result in a gain in statistics but it will be harder to understand the associated systematic uncertainties. This is mainly due to the lack of hybrid statistics at other zenith angles. The available SD statistics from 30-45 degrees are more than sufficient to generate a reliable energy spectrum. All the final results are also unchanged if the angular window is shifted by a few degrees.

6.4   Core location dependency of the SD efficiency

I developed a toy model of the acceptance to investigate the core location dependency of the SD efficiency. This model is based on the general properties of air shower
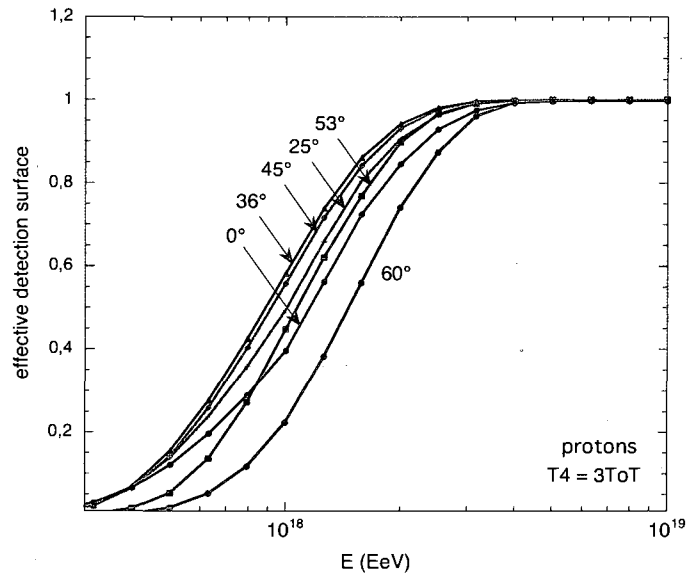
figure from [45]

Figure 6.3 The effective area of the surface detector as function of energy for different zenith angles. This work is based on a full Monte Carlo shower simulation and full detector simulation. The most efficient zenith angles are roughly between 30-45 degrees.

phenomenology and a simplified version of the T2 trigger used in the surface detector. For a shower of fixed energy and zenith angle, the lateral distribution function (LDF) predicts the size of the signal in each WCD based on the distance to the shower core location. The probability of the incident signal to record a T2 trigger in each WCD is based primarily on the signal size. The LDF used was derived from the data [46] for energies above $10^{18.5}$ eV and has been extrapolated to lower energies. The dependence of the T2 trigger probability as a function of signal, P(S), has been deduced from the data. The work to obtain P(S) was done in conjunction with colleagues and the details can be found in [43] and [45]. Figure 6.4 shows P(S) as well as an LDF fit to a specific low energy event.

For a fixed energy and zenith angle an event is simulated in the toy model by choosing a random core location in the unit cell of stations. The expected signal based
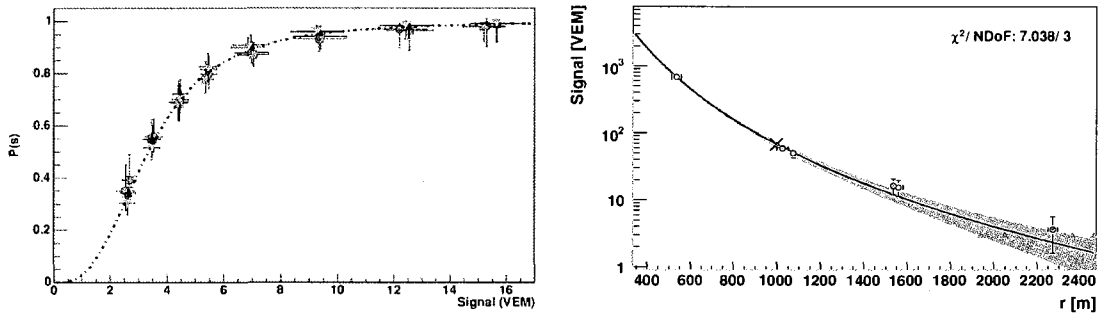
Figure 6.4 The probability of a single WCD to trigger as a function of signal, $P(S)$ (left) and the lateral distribution function, LDF (right), combine to estimate the probability of a WCD to trigger as a function of distance from the shower core..

on the LDF is then found in each WCD. Each WCD then either records a trigger or does not based on the local station trigger probability, P(S) (the T2 triggering condition). The T3 central trigger algorithm is then applied to the triggered stations in the hexagon unit cell. The resulting events that pass the T3 central trigger fill a two dimensional histogram. This histogram shows the density distribution of triggered events. The left histogram in Figure 6.5 is the density distributions of triggered events mapped on the unit hexagonal cell for the energy of $\approx 10^{18.0}$ eV.

6.4.1   The ideal triangular unit cell

While the hexagon unit cell is useful for event selection and monitoring, the symmetry of the distributions makes using a triangle unit cell more appropriate. This is demonstrated in Figure 6.5, a triangular unit cell exploits the symmetry of the density distributions better than a hexagon. The density in the center of the triangle is higher than near the verities. The density falls off smoothly and can be seen in a one dimensional representation by binning the core locations in equal density rings. The rings of constant density can be approximated by Reuleaux triangular

rings. Reuleaux triangles are geometric shapes that have equal width, examples are are shown on the left in Figure 6.6. Equal width means that any straight line drawn from one edge to another that passes through the center is of the same length. A circle is a familiar shape of equal width. Reuleaux triangles can be constructed by drawing circles around each of the stations on the vertexes of an equilateral triangle. The three circles will overlap, creating a region that is in the shape of a Reuleaux triangle.
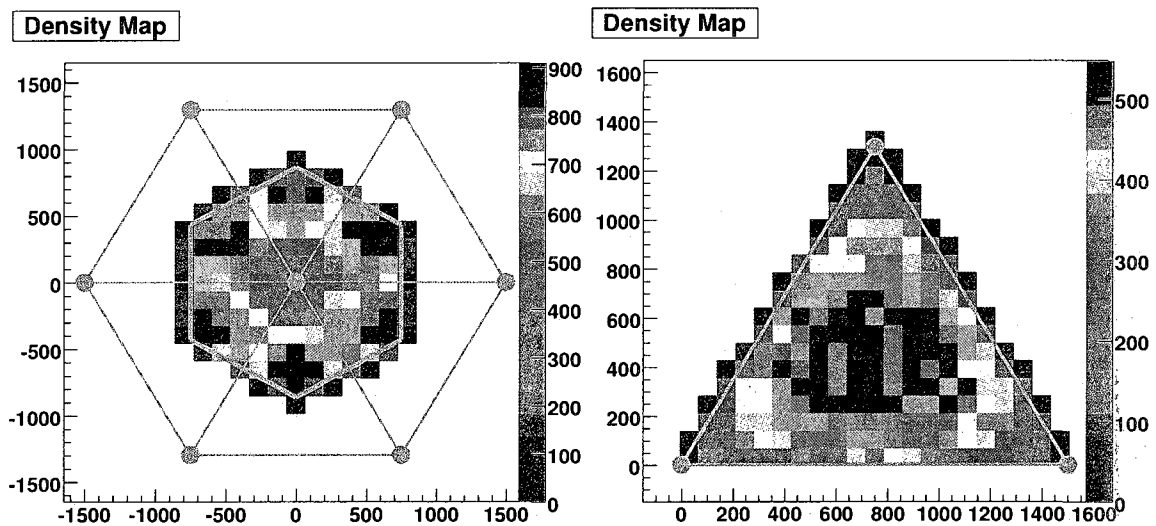
Figure 6.5 Toy Model results. The density distributions in the unit hexagon cell display a symmetry that makes using a triangle more appropriate. All the events first get translated to the ideal hexagon cell and then get mapped onto the ideal triangle of stations.

On the right of Figure 6.6 the density drop off in one dimension is plotted in what I call a "density plot". In the density plots the x-axis is a measure of the size of the Reuleaux ring. The lower the value the closer the ring is to the center. The value of 0.8 represents a ring that just touches the edge of the unit triangle. The Reuleaux rings are centered at the middle of the unit triangle. The size of the ring is not measured as the distance from the center, but is instead measured as the fractional

area the ring would represent if the ring were a solid disk rather than a thin ring. This is done for the ease of interpreting the density plots that will be made clear in the later sections. The number of events in each Reuleaux ring divided by the area of the ring is the calculated density shown in the density plots on the y-axis. The error bar for each density point is the density times one over the square root of the number of events (counting statistics). There are no values between .8 and 1 because those rings would include area outside of the triangle. The last point at 1.0 represents the density between the largest enclosed Reuleaux triangle and the edges of the triangle.

The trends in the model data show that for each energy, the density in the center rings is constant out to a certain point. After this point the density falls off smoothly. Since this is a model, we know the number of events that both pass and fail the triggering conditions. Comparing the number that pass with the number that fail is a measure of the simulated efficiency. Mapping the efficiency on the unit cell will then give the core location dependency of the efficiency. Hundreds of thousands of random core locations throughout the unit cell are chosen to identity the core location dependency. Figure 6.7 shows the triggered, non-triggered and total incident density distributions for the model at an energy of $\approx 10^{18}$ eV.

For other energies, the density distribution and density plots have similar features. The density is always highest in the center and falls off near the vertexes. This is expected because a shower that lands in the center will have signals in the three stations of approximately equal size since they are at similar distances from the shower core. If that same shower is moved off center then the distances will be different and the furthest station will have a much reduced signal (due to the steeply falling LDF) and thus have a lower probability of triggering. The local trigger saturates above a certain signal so if the off-center shower still produces signals above the saturation level then the event will still pass the central trigger. This can be seen as the density
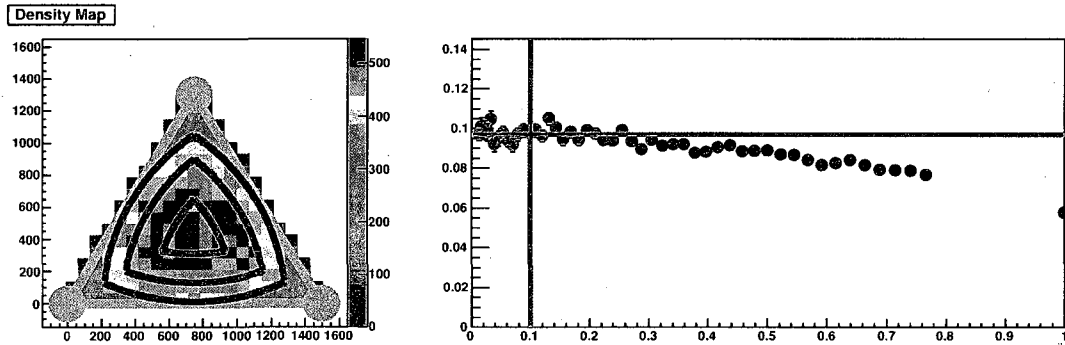
Figure 6.6 Toy Model results. The fall off in density from the center is smooth. Reuleaux rings moving from the center of the triangle outwards represent approximately equal density regions. These density plots have a characteristic flat region in the center and a steady fall off as the Reuleaux ring moves towards the edges of the unit triangular cell.
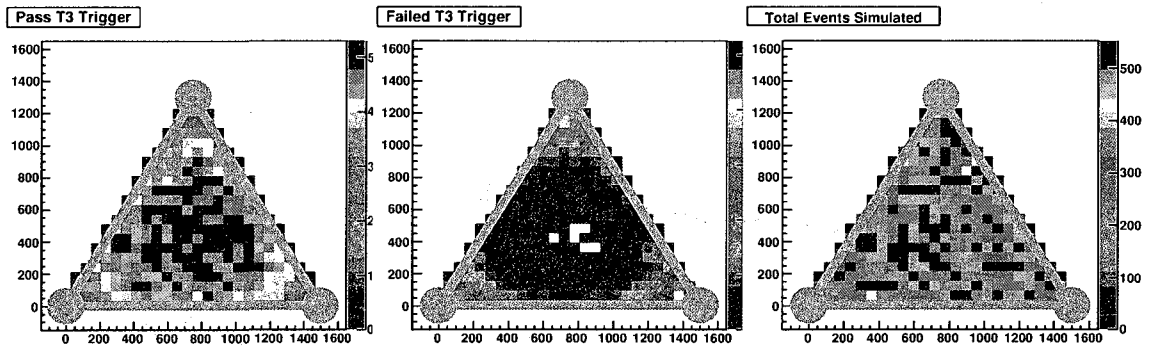
Figure 6.7 Toy Model results. All the simulated showers are in the histogram on the right. The ones that passed the T3 central trigger are on the left and the ones that failed are in the center histogram. The efficiency of the triangular unit cell is found by looking at the ratio of the left histogram (passed the T3) to the right histogram (total number of showers). The centers of the triangle can be seen to be the efficient.

remaining constant out to a certain Reuleaux ring and appears as a flat region on the density plots. Inside this flat region of constant density, the signals in all 3 stations (all at different distances), are all above the local station saturation level. This level is nominally ≈10 VEM as demonstrated in Figure 6.4 (left). By comparing histograms

like the ones in Figure 6.7 we find that the efficiency in the center regions is near full efficiency. If the whole triangle were fully efficient, the density would be uniform over the unit triangle. The resulting density plot would have all the points on the horizontal line in Figure 6.6.

6.4.2 Estimating the efficiency from the density plots

The integral of the points in the density plot in Figure 6.6 represent the total number of detected events in the unit triangle, $N^{\triangle}_{Detected}$. While the center regions are not exactly 100%, due to fluctuations, the density in the center regions is a measure of what the density across the unit cell would be if the cell were fully efficient. The integral of the horizontal line in Figure 6.6 is then an estimate of the total number of incident events on the unit triangle, $N^{\triangle}_{Incident}$. This is an easy integral to do and is just the center density in the flat region, ρ^{flat} times the area of the triangle, $A^{\triangle}$. The efficiency of the flat center region, ϵ^{flat}, can be used to adjust the integral and more accurately estimate $N^{\triangle}_{Incident}$.

$$N^{\triangle}_{Incident} = \frac{1}{\epsilon^{flat}} \rho^{flat} A^{\triangle} \tag{6.1}$$

The overall efficiency of the unit triangle, $\epsilon^{\triangle}$ can now be found,

$$\epsilon^{\triangle} = \frac{N^{\triangle}_{Detected}}{N^{\triangle}_{Incident}} = \frac{N^{\triangle}_{Detected}}{\frac{1}{\epsilon^{flat}} \rho^{flat} A^{\triangle}} \tag{6.2}$$

This method of estimating, $\epsilon^{\triangle}$, relies on two pieces of information. Firstly, the density plot for the selected energy needs to have large statistics so A^{flat} and the drop off in density can be easily determined. Secondly, a measure of ϵ^{flat} must be known or assumed (strictly speaking, one could assume that $\epsilon^{flat} \approx 95\%$ and assign a systematic uncertainty of $\pm 5\%$). With these pieces of information the efficiency

can be calculated directly from the density plot. This is advantageous because for the real data, all we have is the number of detected events, so a way to estimate the number of incident events is needed to calculate the efficiency.

Throughout the surface array, differences between detectors is expected. The ToT rate is indicative of this as well as the observed seasonal effects [40]. To properly model the entire surface detector all these differences must be weighted and averaged according to the specific properties of each individual WCDs. The parameters of the toy model were derived from the data but for practical reasons they could not be derived equally from all the individual WCDs [45]. To estimate the effect of having different WCDs, the parameters were changed to see the effect in the density plots and distributions. In Table 6.1 the estimation of the efficiency from the density plot is compared with the true efficiency obtained from the toy model for the same energy, but with varied model parameters.
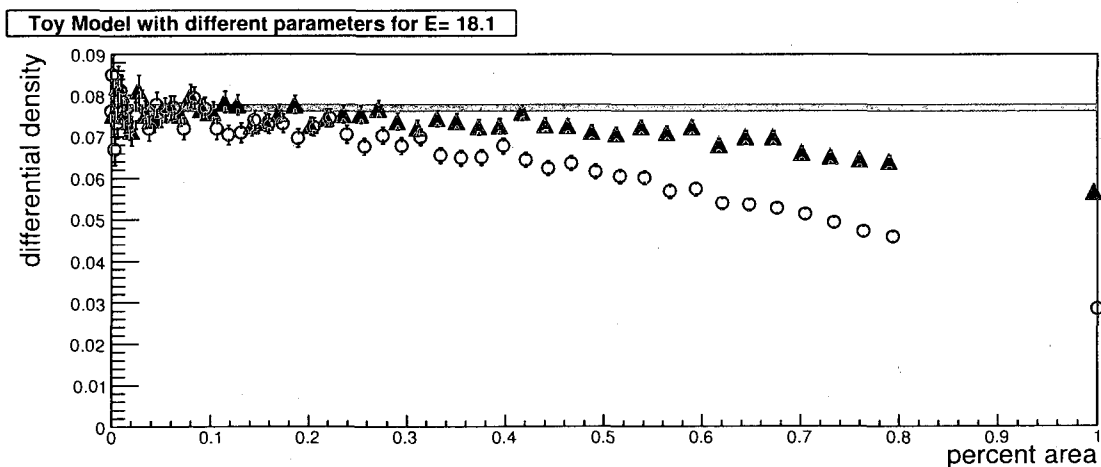
Figure 6.8 Toy Model results. If the parameters in the toy model are changed the density plot for the unit triangle changes. The center flat regions remain close to 100% efficient, but the drop off can vary considerably. The solid triangle points and the hollow circle points represent an efficiency for the whole unit cell of 91% and 74% respectively.

The parameters in the toy model were varied from their nominal values. The

A^{flat}	ϵ^{flat}	estimate $\epsilon^{\triangle}$	true $\epsilon^{\triangle}$
0.145	0.932	0.651	0.662
0.205	0.986	0.896	0.907
0.131	0.962	0.637	0.650
0.131	0.977	0.719	0.735

Table 6.1 An estimation of overall efficiency can be compared with the true efficiency for the simulated events with different parameters in the toy model.

resulting density plots changed but the same general features remained: flat constant density regions in the center and a drop off in density as the Reuleaux rings approached vertexes of the unit triangular cell. Figure 6.8 shows two cases with different parameters. Both plots have a flat center region but the hollow circles have a steeper drop off.

Different density distributions and plots were made for the same energy with a large range of model parameters. In each case the size of the constant region changed as did the steepness of the drop off. In Table 6.1, four cases are shown for rather extreme values of the parameters. The overall efficiency of the whole triangle changed quite significantly, up to 25% and the estimation method was able to account for these changes with a possible systematic bias of 2%.

An important point learned from this exercise is that the efficiency in the flat constant regions in each case, ϵ^{flat}, always remain near full efficiency ($95^{+4}_{-2}\%$). This is important because in general we do not have a way to measure the efficiency of the entire surface detector, $\epsilon^{\triangle}_{SD}$. What we can do, is to measure the efficiency of

specialized subsets of SD data, $\epsilon^{\triangle}_{subset-SD}$. In general,

$$\epsilon^{\triangle}_{SD} \neq \epsilon^{\triangle}_{subset-SD} \tag{6.3}$$

However, based on the work demonstrated in Table 6.1, we can conclude that,

$$\epsilon^{flat}_{SD} \approx \epsilon^{flat}_{subset-SD} \tag{6.4}$$

So, if we have a measure of $\epsilon^{flat}_{subset-SD}$, then we can use the approximation above to get ϵ^{flat}_{SD} and then use equation 6.2 to approximate $\epsilon^{\triangle}_{SD}$. By using this approximation for ϵ^{flat}_{SD}, we will introduce a systematic uncertainty but this will be smaller than the systematic if we equated the total triangle efficiencies.

The size of the systematic uncertainty introduced by assuming equation 6.4 is an exact equality will depend on how representative the subset is of the average properties of the entire set. If the subset is a completely random sample, the associated systematic will be minimized. If the subset has biases on properties that affect the triggering conditions (like temperature), the systematic will be larger. The study above changed the parameters in the toy model drastically to see what the extreme range in ϵ^{flat} could be.

As mentioned above this method may have a systematic bias of 2% when the true ϵ^{flat} is known for the specific subset. From this bias, plus the range of the ϵ^{flat} values from the study above, we arrive at a conservative estimate of a $\pm 4\%$ systematic uncertainty on the efficiency calculation. Remember that this uncertainty is needed because we only have $\epsilon^{flat}_{subset-SD}$ and we equate this with ϵ^{flat}_{SD}. In general the WCDs in the surface array are fairly uniform so that large efficiency differences in different subsets are not expected to be as large as the amounts shown in Table 6.1. Using a $\pm 4\%$ systematic uncertainty is possibly too conservative. Later in the chapter (section 6.5.1) we will see that this systematic of $\pm 4\%$ is on the same order as

the statistical uncertainties. Only when there are more statistics for the $\epsilon^{flat}_{subset-SD}$ measurement will a more careful treatment of this systematic will be appropriate.

6.5 Comparing the model with the data

Looking to the data, we see the general trends predicted by the toy model. Figure 6.9 shows the core location distributions and density plots for three different energies. The highest energy, $10^{18.5}$ eV, shows a constant density for all rings in the triangle as expected, since the surface detector has been shown to be 100% efficient above this energy [43]. The lowest energy of $10^{17.6}$ eV, does not show a constant density region in the density plot. The middle energy, $10^{18.05}$ eV, shows a constant region in the center and then a steady fall off, indicating the center of the triangle is still near 100% efficiency. Since this is the real SD data, we generally have no measure of the number of events that failed to trigger. However, to verify that the center regions are indeed near full efficiency, the hybrid data can be used.

6.5.1 Using the hybrids

The trigger and event selection algorithms for the surface detector cleanly separate the background cosmic rays from the high energy showers. However the strict requirement for at least three WCDs to trigger in time coincidence sets the energy threshold for the surface detector. For hybrid events, the triggering conditions are more relaxed. Only one WCD needs to have a trigger in time coincidence with the shower as defined by the fluorescence detector. This relaxed triggering condition allows for the detection of hybrid events with lower energy than the surface detector alone.

As discussed in chapter 3, when a hybrid event also passes the SD trigger requirement it is called "golden." The golden hybrids are a specialized subset of the
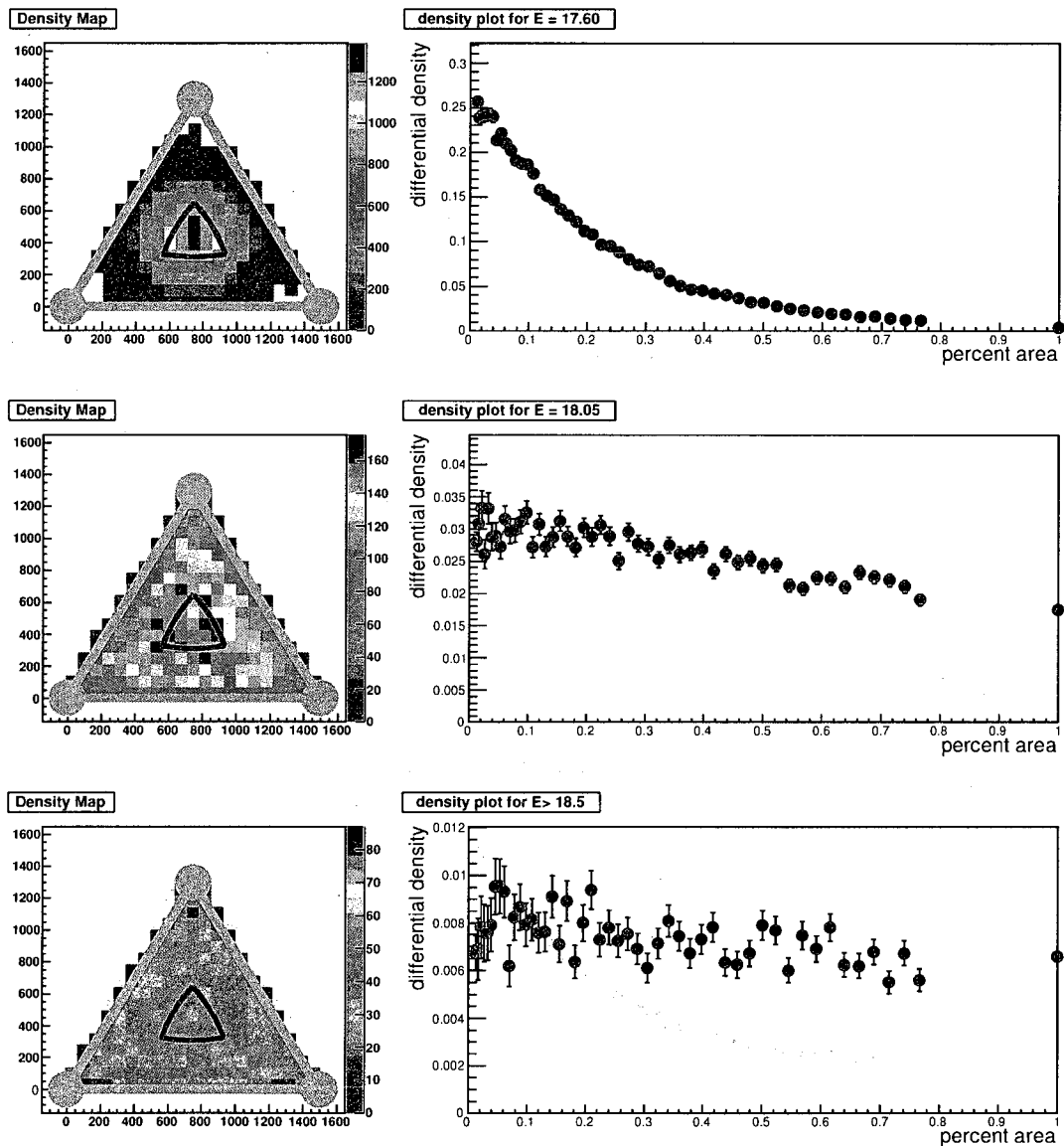
Figure 6.9 Real SD data. The fall off in density from is center is smooth. Reuleaux rings moving from the center of the triangle outwards represent approximately constant density region. The top plot are for energy $10^{17.6 \pm 0.05}$ eV, the middle is $10^{18.05 \pm 0.05}$ eV and the bottom is for $10^{18.5+}$ eV.

entire SD data set. In the previous chapter we saw how useful they are for the energy calibration of the SD events. The entire hybrid data set also includes non golden events, or events that did not have a distinct SD trigger (typically only 1 or 2 WCDs

with signal). These non-golden events are below the surface detector trigger threshold and they can be used to measure the efficiency of the surface detector for the golden hybrid data set, ϵ_{gold}.

This can be done only if the hybrid trigger and the surface detector trigger are independent, which is the case for the Auger Observatory. When the triggering conditions are independent the probability for a low energy event that has recorded a hybrid trigger to also record a SD trigger is then dependent only on the SD trigger probability. This trigger probability can be measured by looking at the ratio of golden hybrids (both SD and hybrid trigger) to the total number of hybrids (which is the non-golden plus the golden hybrids). This ratio is called the golden hybrid efficiency, ϵ_{gold}.

$$\epsilon_{gold} = \frac{N_{gold}}{N_{hybrid}} = \frac{N_{gold}}{N_{nongold} + N_{gold}} \qquad (6.5)$$

The golden hybrid data set is a subset of the entire surface detector data set. The subset is not a random sample however. Unfortunately, the golden hybrids are a very specialized subset of the SD data that includes many biases. Since the viewing aperture of a fluorescence telescope grows with energy, only nearby showers of low energy will trigger. This means that generally only the WCDs that are closest to the individual fluorescence buildings are included in the ϵ_{gold} measurement. The lack of ToT trigger uniformity discussed earlier will not be averaged out unless the efficiency calculation properly weights the contributions from the varying WCDs. The SD operates 24 hours a day all year round so average temperature differences between night and day will affect the triggering rate since all the hybrid data has to be recorded at night. Since the winter nights are also longer, more data will be from the colder winter months. These are the major reasons why the golden hybrid data set is not a random subset of the SD data with respect to low energy trigger efficiency (These differences do affect the S_{1000} measurement but are much less drastic, so the golden

hybrids can be used for surface detector energy calibration).

We can not directly relate ϵ_{gold} to the efficiency of the entire surface detector, ϵ_{SD}. Detailed independent studies are needed to accurately relate these two efficiency but that is not needed because of the technique developed in section 6.4.2. All we need is a measure of the efficiency in the center region. The hybrid data is used only to find ϵ_{gold}^{flat}. This efficiency is approximately equal to ϵ_{SD}^{flat} as described previously. Equating the efficiencies in the flat regions only, introduces a small uncertainty, 4%, on the SD efficiency for the whole SD data set.

6.5.2 Golden hybrid efficiency

The measurement of ϵ_{SD}^{flat} proceeds a follows. All the well reconstructed hybrid events (as described in chapter 5) are mapped onto the unit triangle based on their core locations. The number of golden and non-golden hybrid events are found for each Reuleaux ring. For each bin in energy the integral efficiency ϵ_{gold} is calculated from the center out to each Reuleaux ring. Integral efficiency is used rather then differential density because of the low number of hybrid events. The differential density plots that were made with the SD data are useful understanding the core locations dependency in detail but high number of events are needed. The number of hybrid events in each energy bin is smaller by a factor of 50-100, and differential density plots are hard to interpret with low statistics. Using integral efficiency plots is appropriate with the hybrid data because the flat center area, A^{flat}, is easy to identify and all that is needed is ϵ_{gold}^{flat}.

Figure 6.10 shows one such plot for the energy bin of $10^{18.25}$ eV. The point where the efficiency drops below near full efficiency ($\approx 96\%$ in this case) is identifiable in each energy bin. This point, at which a monotonic drop in efficiency starts, defines A^{flat}.
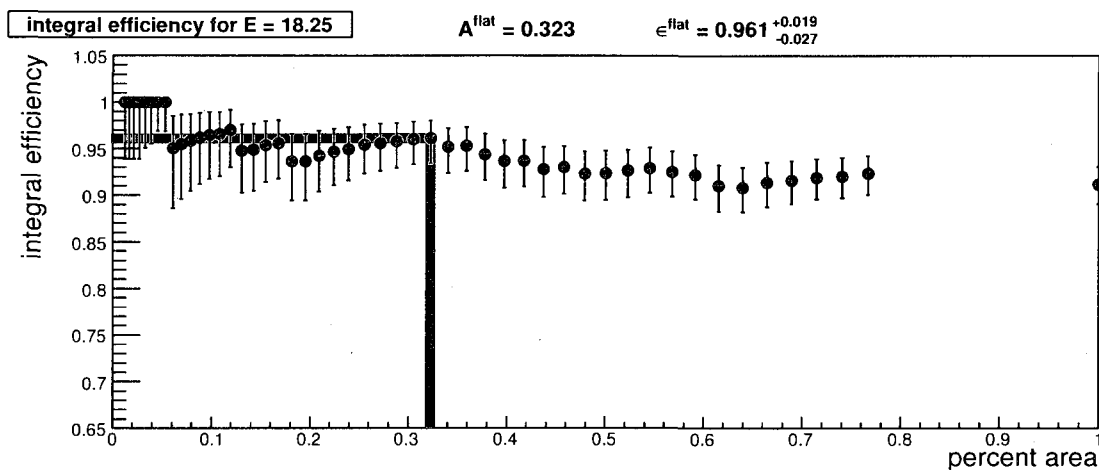
Figure 6.10 Hybrid results. The integral efficiency as a function of percent area as measured by the hybrid events for the energy of $10^{18.25}$ eV. The center 32% region has approximately constant efficiency greater than 96%. As the Reuleaux disc is expanded the efficiency drops off steadily. The value at percent area = 1.0 is roughly .91, indicating that the golden hybrid efficiency for the whole triangle is around 91% at this energy.

For each energy bin, the efficiency in the center is calculated from the hybrid data. A likelihood function was created that is based on the binomial probability of detecting a number of golden hybrid events out of a total number of hybrid events. The maximum of likelihood for ϵ_{gold}^{flat} is just $\frac{N_{golden}}{N_{hybrid}}$ as expected. The reason for creating the likelihood function is to estimate the uncertainty. The calculated efficiency, $\frac{N_{golden}}{N_{hybrid}}$, is varied around the maximum $\frac{N_{golden}}{N_{hybrid}}$ and the resulting likelihood is computed. The one sigma statistical uncertainty is defined when likelihood has dropped by 68%. As discussed above, there will be a ±4% systematic uncertainty due to the fact that the golden hybrid dataset is a specialized subset of the entire surface detector data.

The results of the ϵ_{gold} calculations and the associated statistical uncertainties are shown in Table 6.2. In Figure 6.11, the integral efficiency plots are displayed for energies from $10^{17.85}$ to $10^{18.5+}$ eV. The center regions of constant efficiency are seen in each energy bin down to $10^{17.95}$ eV. In the lowest energy bin of $10^{17.85}$ eV, no such
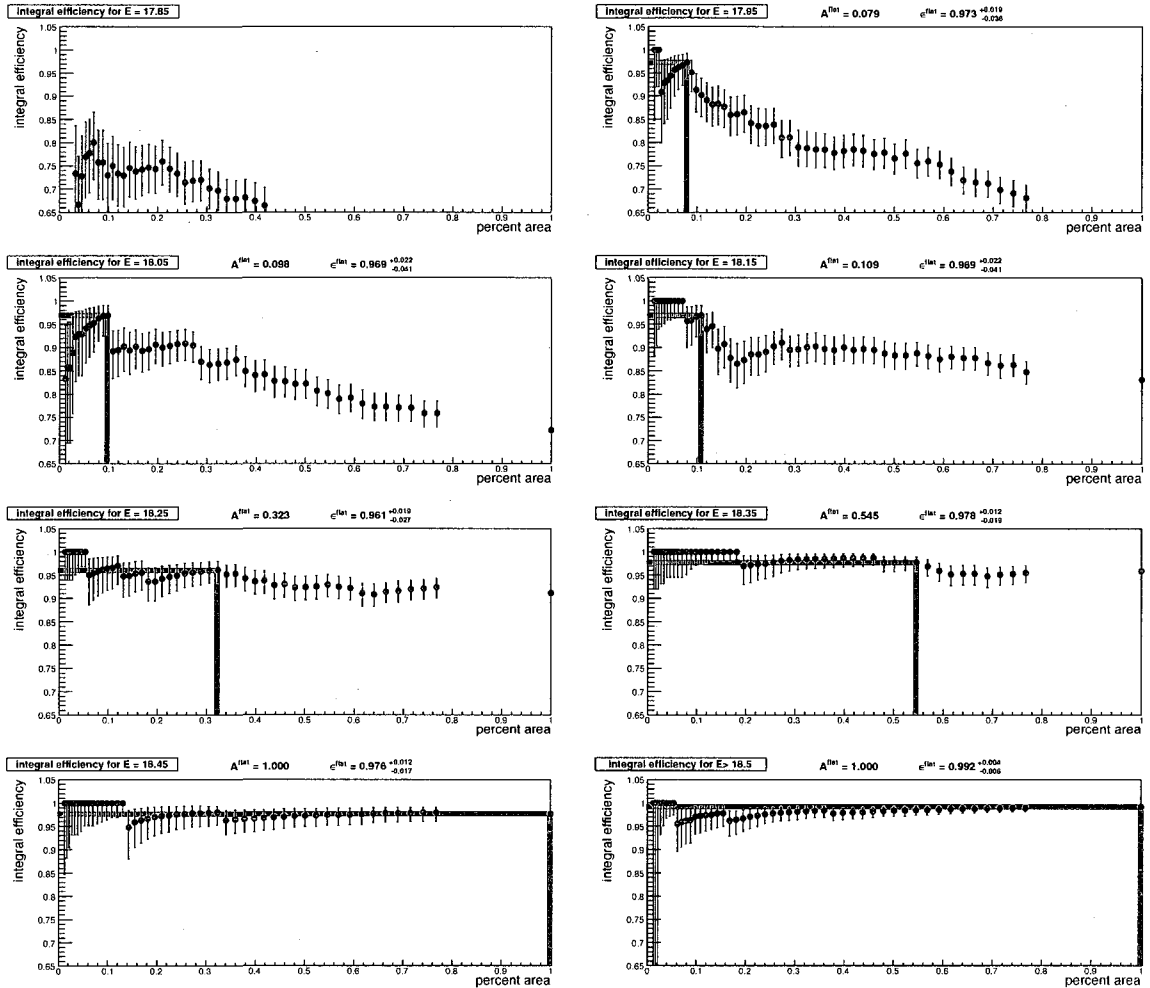
Figure 6.11 All the ϵ_{gold} plots.

region is found.

6.6   Calculating the Surface Array Efficiency

The surface detector data displays the general trends predicted by the toy model. For bins in energy below $10^{18.5}$ eV there exist regions of the surface detector that remain very close to full efficiency. The density of events plotted against Reuleaux ring can be used to define these regions and the hybrid events can be used to verify

$log_{10}(E/eV)$	A^{flat}	ϵ_{gold}	$\sigma_{\epsilon_{gold}}$
17.85	N/A	N/A	N/A
17.95	0.08	0.973	$^{+0.02}_{-0.04}$
18.05	0.10	0.969	$^{+0.02}_{-0.04}$
18.15	0.11.	0.969	$^{+0.02}_{-0.04}$
18.25	0.32	0.961	$^{+0.02}_{-0.03}$
18.35	0.55	0.978	$^{+0.01}_{-0.02}$
18.45	1.00	0.976	$^{+0.01}_{-0.02}$
18.5+	1.00	0.992	$^{+0.00}_{-0.01}$

Table 6.2 ϵ_{gold} is measured for each energy bin. An associated asymmetric uncertainty, $\sigma_{\epsilon_{gold}}$, is also calculated.

that the regions are indeed near 100% efficiency. By knowing the average density of events in a region of near fully efficiency, the total number of events incident on the surface detector can be estimated. The efficiency of the entire SD data set can then be estimated by the ratio of the number of detected events to the estimated number of incident events. By using all the SD data in the density plots, all the non-uniformities over the array are correctly averaged together.

6.6.1   General strategy summary

For each bin in $Log_{10}(E/eV)$ of size 0.1 the general strategy is as follows:

FD An integral efficiency plot is made from the hybrid data passing the quality cuts described in chapter 5;

FD The golden hybrid efficiency in the center region, ϵ_{gold}^{flat}, is calculated from the integral efficiency plot;

FD The center region area, A_{hybrid}^{flat} is defined by the integral efficiency plots with the hybrid data;

SD A differential density plot is made from all the SD data passing the T5 cut

SD The average density in the center constant region, called ρ_{SD}^{flat}, is found from the SD differential density plot with the center region defined by A_{hybrid}^{flat};

- The values of ϵ_{gold}^{flat} and ρ_{SD}^{flat} combine to estimate the total number of incident events, $N_{Incident}^{\triangle}$, on the whole triangle;

$$N_{Incident}^{\triangle} = \frac{1}{\epsilon_{gold}^{flat}} \times \rho_{SD}^{flat} \times A^{\triangle} \tag{6.6}$$

- The efficiency of the entire surface detector, $\epsilon_{SD}^{\triangle}$ is then calculated

$$\epsilon_{SD}^{\triangle} = \frac{N_{Detected}^{\triangle}}{N_{Incident}^{\triangle}}. \tag{6.7}$$

All the surface detector data from January 1st 2004 through June 31st 2008 were used in this analysis. The data set was reduced by limiting the analysis to zenith angles between 30-45 degrees. The only ambiguity in the process is defining the center flat regions of constant density and efficiency. It is possible to estimate $N_{Incident}^{\triangle}$ from the density and efficiency from any part of the center region. The maximal region has been used to reduce statistical uncertainties. However, a measure of this maximal region or area can be found from both the SD data and from the hybrid data. They typically do not match up, as shown in Table 6.3. The hybrid area, A_{hybrid}^{flat}, is typically significantly smaller. The areas are expected to be different because the hybrid data is not a random subset of the SD data. Additionally, integral plots were used for the hybrid data and differential plots for the SD data. The discrepancy may be due to the low statistics in each hybrid energy bin. In all cases the lower area, A_{hybrid}^{flat}, was used to be conservative. All the differential density plots are shown in Figure 6.12.

$log_{10}(E/eV)$	A^{flat}_{hybrid}	A^{flat}_{SD}
17.85	N/A	0.079
17.95	0.079	0.168
18.05	0.098	0.224
18.15	0.109	0.250
18.25	0.323	0.346
18.35	0.545	0.702
18.45	1.00	1.00
18.5+	1.00	1.00

Table 6.3 The flat area is estimated from both the hybrid and the SD data.

6.6.2   Results

Table 6.4 shows the results of estimating the number of incident events for different energy bins.

$log_{10}(E/eV)$	A^{flat}	ρ_{flat}	$\sigma_{\rho_{flat}}$	$N_{Incident}$	$N_{detected}$
17.95	0.079	4.799	0.079	48054	33014
18.05	0.079	2.94	0.061	29439	23119
18.15	0.109	1.837	0.041	18474	15484
18.25	0.33	1.077	0.018	10928	9741
18.35	0.545	0.611	0.01	6091	6055
18.45	0.702	0.407	0.007	4064	3930
18.5+	1	0.696	0.008	6839	6883

Table 6.4 Raw values of the center area, density and error on the density for each energy bin. The density and ϵ^{flat}_{gold} is then used to estimate the number of incident events.

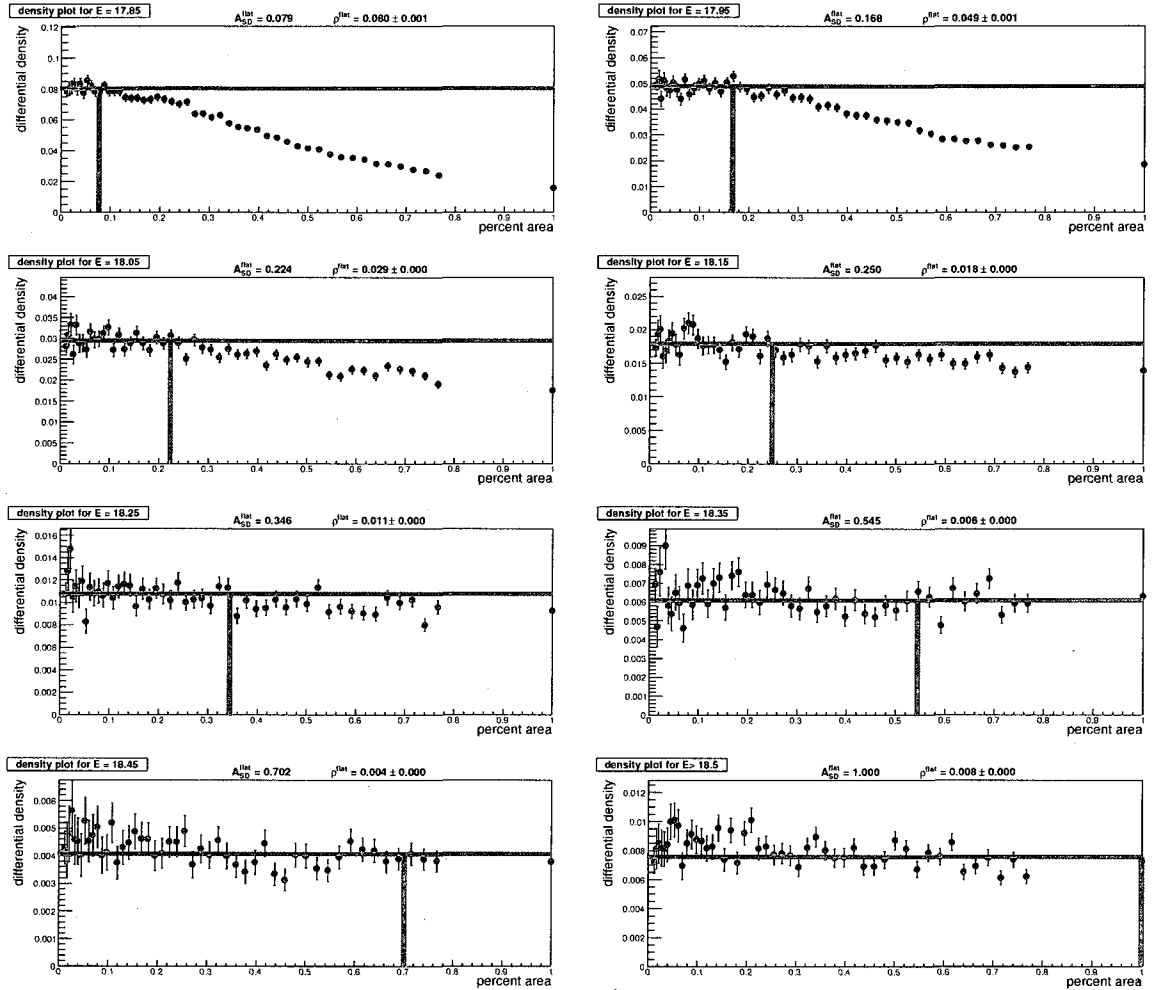**Figure 6.12** All the SD differential density plots.

Table 6.5 shows the final SD efficiency calculations and the estimated uncertainty.

6.6.3   Uncertainties

This technique has been developed to minimize the systematic uncertainties in the efficiency calculation. All the non-uniformities in the surface array arising from differences in individual WCDs, temperature variations, etc. are automatically correctly

$log_{10}(E/eV)$	ϵ_{SD}	$\sigma_{\epsilon_{SD}}$
17.95	0.687	$^{+0.05}_{-0.07}$
18.05	0.770	$^{+0.05}_{-0.07}$
18.15	0.838	$^{+0.05}_{-0.06}$
18.25	0.891	$^{+0.05}_{-0.08}$
18.35	0.994	$^{+0.05}_{-0.06}$
18.45	0.967	$^{+0.05}_{-0.05}$
18.5+	1.012	$^{+0.04}_{-0.05}$

Table 6.5 ϵ_{SD} and $\sigma_{\epsilon_{SD}}$ is measured for each energy bin.

averaged together. The uncertainties that remain are small and are dominated by the number of hybrid statistics. The final efficiency calculation in equation 6.7 has two terms that contribute to the overall uncertainty, $\sigma^2_{\epsilon^{\triangle}_{SD}}$.

$$\sigma^2_{\epsilon^{\triangle}_{SD}} = \sigma^2_{N^{\triangle}_{Detected}} + \sigma^2_{N^{\triangle}_{Incident}} \tag{6.8}$$

Where,

$$\frac{\sigma_{N^{\triangle}_{Detected}}}{N^{\triangle}_{Detected}} = \frac{\sqrt{N^{\triangle}_{Detected}}}{N^{\triangle}_{Detected}} \tag{6.9}$$

$$\left(\frac{\sigma_{N^{\triangle}_{Incident}}}{N^{\triangle}_{Incident}}\right)^2 = \left(\frac{\sigma^2_{\epsilon^{flat}_{gold}}}{\epsilon^{flat}_{gold}}\right)^2 + \left(\frac{\sigma^2_{\rho^{flat}_{SD}}}{\rho^{flat}_{SD}}\right)^2 \tag{6.10}$$

The uncertainty, $\sigma_{N^{\triangle}_{Detected}}$, is based on counting statistics and $\sigma_{N^{\triangle}_{Incident}}$ is a combination of uncertainty from the terms in equation 6.6 used to estimate $N^{\triangle}$. No uncertainty is assigned for the area of the triangle $A^{\triangle}$. The uncertainties for ϵ^{flat}_{gold} and ρ^{flat}_{SD} are in Table 6.2 and Table 6.4 respectively. These results cannot be thought of exactly as efficiencies since some of the values are greater than one. Instead they should be interpreted as just a scaling factor for the exposure calculation that will be discussed in the next chapter.

Chapter 7

The Cosmic Ray Energy Spectrum

The Pierre Auger Observatory was built to study the highest energy cosmic rays. The surface detector (SD) was designed to be fully efficient for air showers with energies greater than 10^{19} eV [47]. The actual implementation of the design proved to be more efficient and the SD efficiency is still 100% for energies down to $10^{18.45}$ eV for showers with zenith angle less than 60 degrees. While not fully efficient, the Auger surface detector (SD) has sensitivity to detect and accurately reconstruct showers with energies well below 10^{18} eV. The fluorescence detector (FD) is also sensitive to cosmic rays with these lower energies. The golden hybrid events, which have both a SD trigger and a FD trigger, have proved very useful in both the energy determination and the efficiency calculation of the SD events at these low energies. In chapter 6 we explained how the average efficiency for the entire SD was calculated for energies as low as $10^{17.95}$ eV. Chapter 5 explained how the energy of these low energy showers is calculated with a resulting energy resolution of 21%.

In the first section of this chapter the techniques developed in the previous chapters are used to build the energy spectrum with SD data from $10^{17.95} - 10^{19.05}$ eV. This is half a decade of energy lower than the recently published SD energy spectrum from

the Auger collaboration [12]. The lowering in energy allows for the full observation of the spectral feature known as the "ankle". The ankle has been observed by previous experiments and in the second section, comparisons are made. The difficulties in comparing measurements that are common to the cosmic ray field are also discussed. In the third section the astrophysical implications of the shape of the ankle, as measured by this thesis, are explored. In the final section the next steps the Auger collaboration are taking in measuring the spectrum at both higher and lower energies are briefly described.

7.1 The energy spectrum

The energy spectrum plots the flux of cosmic rays versus energy. In each bin of energy the number of events and the time integrated exposure are measured. The ratio of these quantities is the intensity of cosmic rays. To relate the intensity to the flux we follow the convention adopted by the Auger collaboration as first reported in [48].

7.1.1 Building the spectrum

The acceptance of the surface array is composed of the effective surface area (a) and the solid angle window used (Ω). The efficiency (ϵ_{SD}) and the uptime of the detector (τ) are then needed to calculate the time integrated exposure. Exposure is often measured in units of $[m^2 \ sr \ s]$. It is appropriate in ultra high energy cosmic ray experiments to measure exposure, A, in units of $[km^2 \ sr \ year]$.

$$A = \int \epsilon_{SD} \ d\tau \ d\Omega \ da \tag{7.1}$$

The regularity of the surface detector implies that the entire array can be thought of as a collection of "unit hexagon cells" of WCDs. The choice of the T5 quality trigger

restricts the data to events that have their shower cores in an active hexagon cell of WCDs. A hexagonal cell is considered active if the central and the six surrounding WCDs are functioning properly. If even one WCD is not working (see Figure 7.1) the cell is classified non-active and is not counted. The uptime of each WCD is monitored continuously. The number of active hexagon cells, N_{hex}, in the entire array is calculated on a second by second basis.
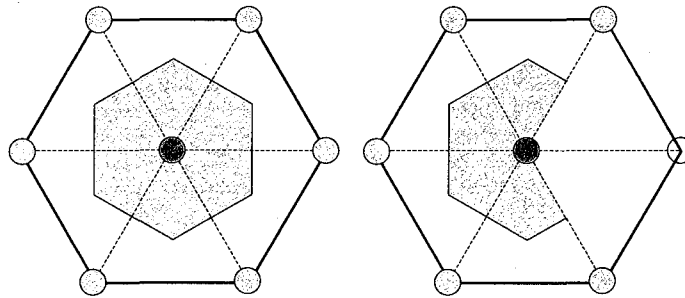

Figure 7.1 The unit cell of the surface detector is a hexagon. The active area is shaded. If an event falls outside the shaded region it would be counted as landing in a different unit cell. For the cell to be counted as active, the six surrounding WCDs must all be functioning properly (left). If even one of the six surrounding stations is non-working, then the cell is not considered active (right). Only active cells are counted in the exposure calculation.

For energy bins for which the efficiency is 100%, the acceptance calculation is purely geometrical. When this is true the maximal exposure will be obtained for the specific geometry used. To calculate the maximal integrated exposure, the number of active hexagons for each second are added up and then multiplied by the area (shaded area of Figure 7.1 (left)) and integrated solid angle for the zenith and azimuth range used. The techniques in this thesis were developed to work in the energy range where the efficiency is below 100%. The efficiency for each energy, as measured in Chapter 6, is used to scale the maximal exposure to the actual integrated exposure for each energy bin. Also, only the limited zenith angle range of 30-45 degrees is used (all

azimuth angles are used).

The uncertainty on the exposure calculation has contributions from the uptime and the efficiency calculations. Any uncertainty on the area or solid angle is not considered significant. The uptime calculation has an estimated 3% uncertainty [43]. The uncertainty on the SD efficiency has been estimated for each energy bin in the previous chapter.

Total Integrated Exposure from January 1, 2004 - June 30, 2008
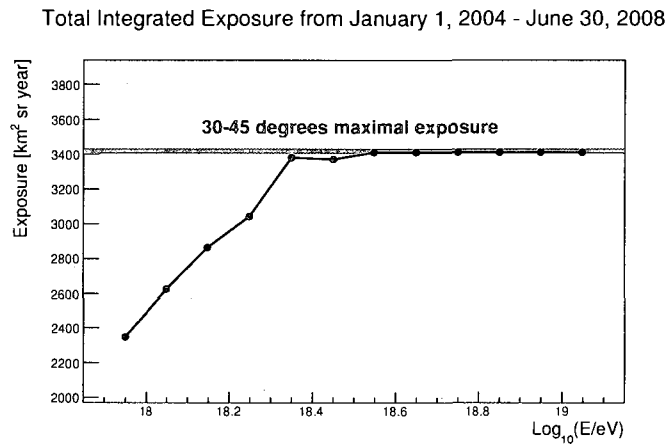
Figure 7.2 The integrated exposure is calculated summing the contributions of each unit hexagon cell in the surface detector. The SD efficiency lowers the exposure for energies below $10^{18.45}$.

Figure 7.2 shows the calculated exposure from January 1, 2004 through June 30, 2008 as a function of energy. The surface detector had been in various stages of construction throughout this time. The maximal exposure for zenith angles from 30-45 degrees during this time period was ≈ 3400 $[km^2 \ sr \ year]$. Now that the full array has been deployed the yearly exposure, in this zenith range, will be ≈ 1400 $[km^2 \ sr \ year]$, so the results in this thesis represent only two and a half years of full Auger statistics.

The number of events detected, N_{events}, per unit exposure, A, is called the intensity, I.

$$I = \frac{N_{events}}{A} \qquad [\frac{events}{km^2 \ sr \ year}] \qquad (7.2)$$

The intensity is related to the flux, J, by taking the differential with respect to energy,

$$J = \frac{dI}{dE} \qquad [\frac{events}{km^2 \ sr \ year \ eV}] \qquad (7.3)$$

The events are counted in $log_{10}(E/eV)$ bins so we actually measure the change in intensity per $log_{10}(E/eV)$,

$$\frac{dI}{d((log_{10}(E/eV)))} \qquad (7.4)$$

The $log_{10}(E/eV)$ bins of E are related to natural log E bins by,

$$\frac{dI}{d(log_{10}(E/eV))} = \frac{10}{ln(10)} \frac{dI}{d(ln(E/eV))} \qquad (7.5)$$

Where,

$$\frac{dI}{d(ln(E/eV))} = E \frac{dI}{dE} \qquad (7.6)$$

Now we have a way to relate the intensity we measure to the flux

$$\frac{dI}{d(log_{10}(E/eV))} = \frac{10}{ln(10)} E \times J \qquad (7.7)$$

So the flux in each $log_{10}(E/eV)$ bin of energy is,

$$J = \frac{1}{E} \times \frac{10}{ln(10)} \times \frac{N_{events}}{A} \qquad [\frac{events}{km^2 \ sr \ year \ eV}] \qquad (7.8)$$

The uncertainty on the flux has contributions from the number of events in each bin and the exposure calculation. These uncertainties are added in quadrature to

get the total flux uncertainty. Table 7.1 shows these uncertainties for each of the 12 energy bins used.

As mentioned in chapter 5, a systematic bias is introduced into the flux calculation for the energy bins near the break point in the calibration curves. The size of this systematic was estimated by recalculating the flux with different break points in the SD energy calibration curves. The flux in the energy bin just below and just above the break point changed. For the other bins the flux was unchanged. When the break point was changed by ± 0.2, the flux in the two bins changed by as much as 7%. Based on this study an extra uncertainty of $\pm 5\%$ has been added to the flux uncertainty for the energy bins of 18.45 and 18.55, since the break point of 18.5 was used. Details of this study are found in the appendices.

$log_{10}(E/eV)$	$\frac{\sigma_J}{J}$
17.95	$^{+0.05}_{-0.07}$
18.05	$^{+0.05}_{-0.07}$
18.15	$^{+0.05}_{-0.06}$
18.25	$^{+0.05}_{-0.08}$
18.35	$^{+0.05}_{-0.06}$
18.45	$^{+0.07}_{-0.07}$
18.55	$^{+0.06}_{-0.07}$
18.65	$^{+0.04}_{-0.05}$
18.75	$^{+0.04}_{-0.05}$
18.85	$^{+0.04}_{-0.05}$
18.95	$^{+0.04}_{-0.05}$
19.05	$^{+0.04}_{-0.05}$

Table 7.1 The uncertainty on the flux measurement.

7.1.2 Energy systematics

The energy calibration of the surface detector has all the systematic uncertainties of the fluorescence detector measurement. There are four contributions to the FD systematic uncertainty; the fluorescence yield of nitrogen, atmospheric attenuation, detector calibration and reconstruction methods. Groups within the Auger collaboration are working on understanding and minimizing each of these contributions.

Contributions from the fluorescence yield of nitrogen result from pressure, humidity and temperature dependencies, as well as the overall absolute normalization. New measurements of these effects are being made by the AIRFLY experiment [29]. For now, only the dependencies on pressure, temperature and humidity from AIRFLY are being used. For the absolute yield, the results from Nagano [28] are used. There is a 14% uncertainty on this value and it is the dominant source in the total uncertainty on the energy measurement from the Auger fluorescence detectors.

The calibration of the detector components of the FD was discussed in chapter 4. The contribution is 9.5% for the absolute calibration (which includes night to night variations) and 3% for possible wavelength dependences. The wavelength dependence comes from the uncertainties on the multi-wavelength curve (section 4.4) convoluted with the average light profile incident on the detector. The uncertainty was estimated by reconstructing showers with different multi-wavelength curves. The average difference between the reconstructed energy for different multi-wavelength curves was 3%.

The atmosphere is constantly monitored. Strict quality cuts are used to limit the data to clear nights with no clouds in the field of view of specific telescopes. When these cuts are used the systematic uncertainty on the energy from atmospheric attenuation is quite small, on the order of a 1-2% [26].

The reconstruction algorithm introduces some significant uncertainties. The sys-

tematic uncertainty associated with the reconstruction has two components, the profile reconstruction and the missing energy assignment. To estimate the uncertainty on the profile reconstruction, two independent methods were compared. The key differences between the methods was the treatment of Cerenkov light and the width of the light track on the camera (from optical effects and from the lateral shower spread). The details of the differences are found in [27]. The results from these two methods have a $\approx 10\%$ difference in energy on average.

The FD measures only the electromagnetic components of the shower. The amount of energy to assign to each event due to secondary particles that do not contribute to the nitrogen fluorescence (muons, hadrons and neutrinos) is estimated from air shower models and simulations (any Cerenkov contribution can be neglected due to the vastly greater contribution from the electromagnetic component). The number of muons predicted by the different models can vary by a factor of two or more. This is not a large problem since over 90% of the shower energy at detection depth has already been transfered from the initial cosmic ray to the electromagnetic component of the air shower. The differences in the various models only impact the remaining energy so the total effect on the energy is not great. The total energy assignment differences between the models from the missing energy contribution is 4% [49]. The total systematic uncertainty on the calorimetric energy measurement of the Auger fluorescence detectors is $\approx 22\%$, and is summarized in Table 7.2.

7.1.3 The energy spectrum

Figure 7.3 plots $log_{10}(J)$ verses $log_{10}(E/eV)$ from $10^{17.95}$ to $10^{19.05}$ eV. Spectral slopes have been fitted to the first four points and the last four points. This was done to only demonstrate that a change in spectral slope at around $10^{18.6}$ eV is supported by the data. This primitive fitting procedure is sufficient to state that the spectral

source of uncertainty	percent
Fluorescence yield:	
absolute normalization [28]	14
pressure [29]	1
temperature [29]	5
humidity [29]	1
atmospheric:	
Rayleigh [26]	1
aerosol phase function [26]	1
Detector calibration:	
absolute [chapter 4]	10
wavelength [50]	3
Reconstruction:	
algorithm [49]	10
missing energy [49]	4
Total	**22**

Table 7.2 Total systematic uncertainties on the measurement of the energy from the fluorescence detectors. Since the SD is calibrated by the FD, this uncertainty is inherent in the energy assignment given to all the surface detector events.

index change called the "ankle" has been observed. A more rigorous fitting treatment is not needed until the data is compared with other experiments or with theoretical models. This will be discussed in the later sections.

The ankle becomes more apparent when the fractional flux, rather then the raw flux, is plotted versus energy. Figure 7.4 shows the ratio of the measured flux, J, to an assumed flux with spectral index of 2.69 ($A \times E^{-2.69}$, where A is normalized to the
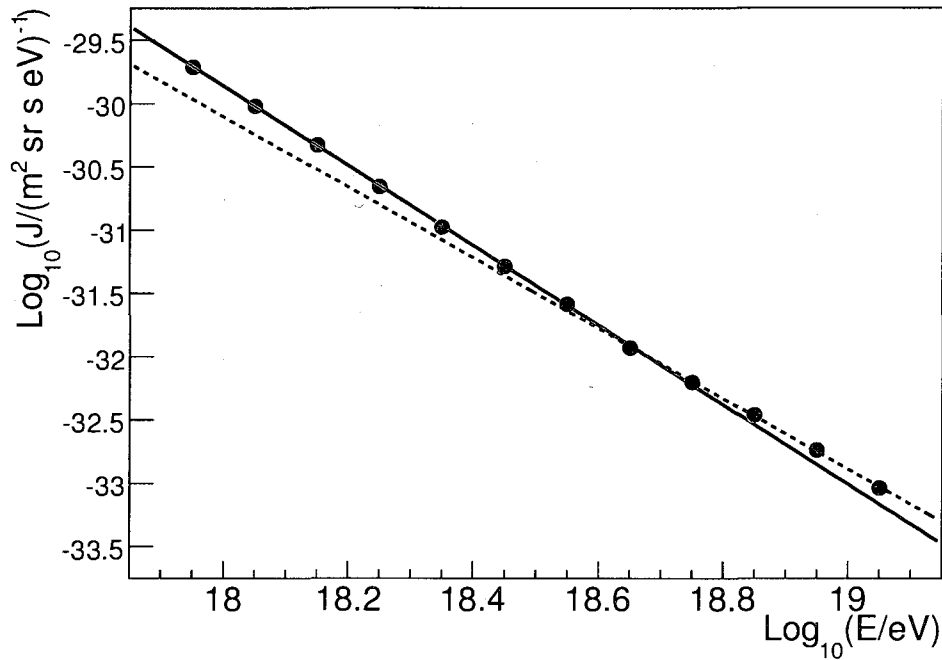
Figure 7.3 The SD energy spectrum as measured by this thesis work. A slope was fit to the first four points (solid line) and the last four points (dashed line). The data supports a change in the spectral slope that occurs at $\approx 10^{18.6}$ eV.

measured flux at $10^{18.65}$ eV). This ratio was subtracted by 1.0, so that the horizontal line at value zero represents equality (this is a convention used inside the Auger collaboration that is trying to be spread throughout the cosmic ray community). The data points from the published SD energy spectrum [12] are also shown. The seven energy bins of overlap between the two results have very good agreement. From $10^{18.6}$ to $10^{19.5}$ eV the measured flux is consistent with a spectral index of 2.69. The ankle is the change in spectral slope that is clearly seen at $\approx 10^{18.6}$ eV. The other spectral feature in this energy range, the suppression of the flux, is also readily apparent for energies above $10^{19.5}$ eV.
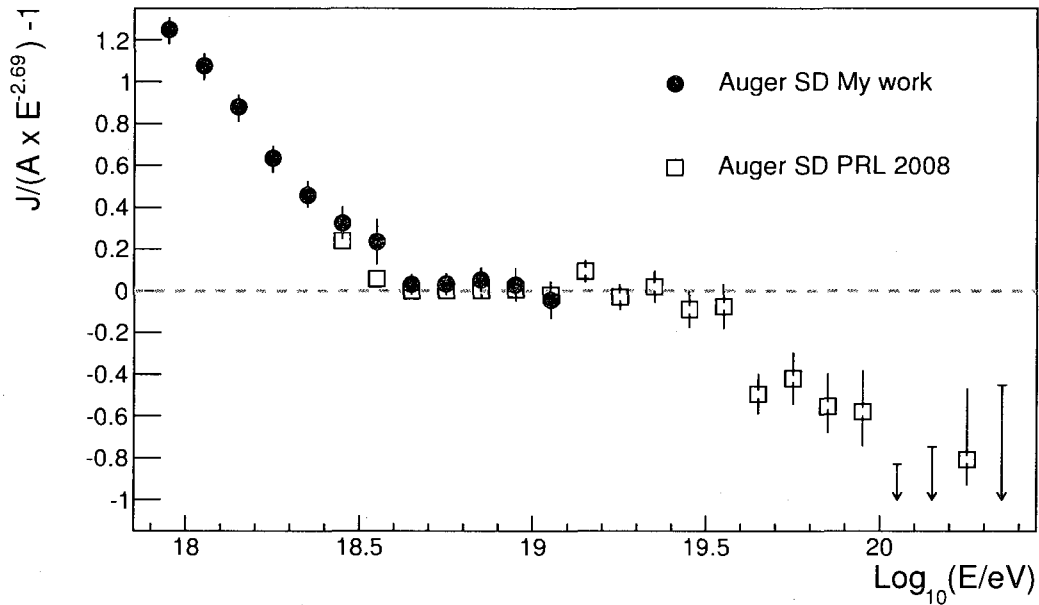
Figure 7.4 The full range of the Auger SD energy spectrum. The hollow points are based on the analysis from the PRL spectrum paper [12] and the solid points are from this thesis work.

7.2 Comparison with other measurements

The ultra high energy cosmic ray community has not agreed on many standards about published energy spectra. There are several common practices that are generally followed. Some of these practices are not very rigorous and can be confusing to a person not familiar with the field of high energy cosmic ray physics.

1. The energy resolution of the detector can significantly bias the measured flux and is typically not factored in when spectra are published. Experiments frequently publish their measured spectrum without taking the detector resolution into account. No standard attempt to unfold the true spectrum from the measured spectrum has been established for the field.

2. Generally, only statistical uncertainties on the flux measurement are shown (i.e. the uncertainty on the number of events). This practice is acceptable if each point in the spectrum has the same systematic uncertainty, but is not acceptable when the uncertainty has an energy dependence.

3. It is common practice to compare results between experiments with a scaled energy spectrum, $E^\gamma J(E)$ versus E. In these types of representation, the systematic uncertainties due to energy assignments from different experiments get magnified. Results can then appear to be in worse agreement than the quoted uncertainties.

When comparing measurements from one experiment to another the above practices need to be considered. Some of the shortcomings of these practices are being addressed by the Auger collaboration. Displaying energy spectra results as the fractional flux, instead of a scaled flux plot, was the first attempt at more rigor from Auger and was used in the recently published SD energy spectrum paper [12]. I mimicked the new practices for Figure 7.4 for comparison. In addition, the systematic flux uncertainties were added in quadrature with statistical uncertainties for each data point. However, there was no attempt to unfold the true spectrum from the raw measured values, which have detector resolution effects that bias the results.

Internally in the collaboration, we are trying to establish procedures and methods to deconvolve the detector resolution systematic effects from our data to unfold the true spectrum from the measured values. While we have not established a firm procedure as of yet, I have developed a preliminary procedure to conservatively estimate the magnitude of the effect. While trying to convert the entire field of cosmic ray research to adopt my procedure may be too ambitious for this thesis, the following subsection elaborates on the logic and the preliminary results of the unfolding

procedure (more details are found in a internal Auger document [51]).

7.2.1 Unfolding the energy spectrum

As explained in Section 7.1.1, to build the energy spectrum the data are binned in $log_{10}(E/eV)$ bins of energy. The energy measurement was shown in Chapter 5 to have resolution of $\approx 21\%$. This implies that the events that were measured in each $log_{10}(E/eV)$ bin of energy had some number of events that up fluctuated in from a lower energy and some that down fluctuated in from higher energy. Also, events fluctuated out of the bin in both directions. Properly treating binning effects due to fluctuations is a common problem in experiments in many fields, and typically there are established procedures how to account for this binning effect. However, for the cosmic ray community there are no such established procedures. The magnitude of the effect is greater in fields where the measured quantity is changing as steeply as the cosmic ray flux does.

The functional form of the cosmic ray flux is unknown because the true origins of high energy cosmic rays are still uncertain. Unfolding the measured spectrum into the true spectrum is hard to do a priori because both the flux and the detector resolution have to be known to estimate the magnitude of the binning effect. The goal of the Auger observatory is to discover the origins of cosmic rays so there is a bit of a conundrum. What I have done is to estimate the size of the possible correction by assuming a simple broken power law functional form of the cosmic ray flux.

I created a toy model that changed the number of events as a function of energy, $N(E)$, by following a power law distribution such that $N(E) \propto E^{-\gamma}$. Each event of this distribution could be "measured" by a detector with energy resolution σ_E so that each event in the toy model has both a true energy, E_{true} and a measured energy E_{meas}. In a real experiment all we have are the E_{meas} events and we bin the data in

E_{meas} energy bins for the spectrum. The idea is to use the toy model to estimate the average true energy of the events in each of the E_{meas} bins of energy. An adjustment to the E_{meas} values can then be deduced, for a given γ and σ_E.

In Figure 7.5, the histograms from the toy model for E_{meas} and E_{true} are shown for two different resolutions. When the toy data are binned by the E_{meas} value, the E_{true} distribution of those binned events will be much broader than the size of the bin. The worse the detector resolution (higher σ_E values) the more broad the E_{true} distribution in each E_{meas} bin. In Figure 7.5, the small E_{meas} bin of width 0.1 has a very broad E_{true} distribution for a $\sigma_E=0.40$ compared to the E_{true} distribution for a $\sigma_E=0.08$.
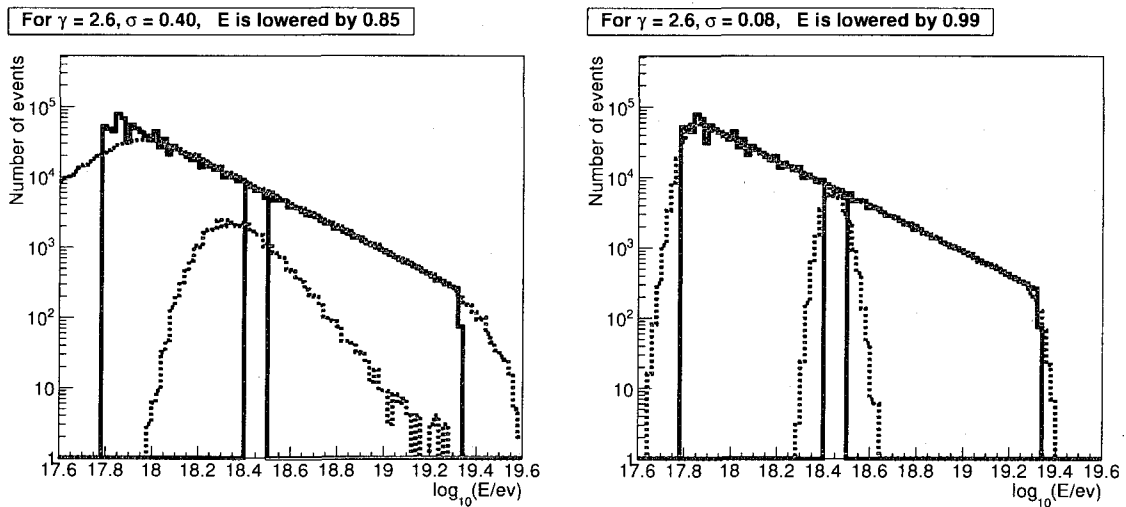

Figure 7.5 The toy model events are binned by the measured energy, E_{meas}. The true energies of those binned events have a much broader distribution that depends on the detector resolution, σ_E, and the steepness of the spectral index, γ. With a spectral index of $\gamma=2.6$, the correction on the energy varies from a 15% correction with $\sigma_E=0.40$, to only 1% when $\sigma_E=0.08$. The corrections are larger for a steeper spectral index.

The central value of the energy bin is typically used on energy spectrum plots. The E_{true} distributions show that the average true energy for events is not the center

of the E_{meas} bin. The correct energy to use for each E_{meas} bin is always a lower value than the center of the bin. With a spectral index of γ=2.6, the correction on the energy needed varies from a 15% correction with σ_E=0.40, to only 1% when σ_E=0.08. For a steeper spectral index of γ=3.3, the corrections needed are greater, 21% and 2% for σ_E=0.40 an σ_E=0.08 respectively.

The key items to highlight about the binning effect that are important for the cosmic ray community are summarized below.

- The unfolding of the measured spectrum into the true spectrum always shifts the energies in the measured spectrum to lower values because the flux is falling. The opposite would be true if the flux were increasing with energy.

- The larger the spectral index γ is, the larger the unfolding correction will be.

- Worse energy resolution (larger σ_E) leads to a larger correction.

- If the spectral index and the energy resolution are both constant then the unfolded spectrum will have the same spectral index as the measured spectrum.

The estimated magnitude of the correction needed for this thesis work was deduced by assuming realistic values for the parameters in the toy model. The γ value was approximated from the unfolded spectrum (Figure 7.3) by fitting the data with a broken power law, but the break point was not left free to float. The values of γ=3.2 for energies below $10^{18.6}$ eV and γ=2.75 for greater energies were found. Chapter 5 measured the average σ_E for all energies to be 21%. These values were used as approximate values in the toy model to estimate the size of the unfolding correction only. They are not meant to be the exact values. σ_E most likely has an energy dependence, since the fluctuations in S_{1000} go down as the energy increases due to the more stations in the event. Also the values for γ and the break point depend

on the fitting procedure. To account for these facts slightly different values for the parameters, σ_E and γ, were used in the toy model to estimate of the systematic uncertainty. σ_E was varied by $\pm 4\%$, based on the estimated S_{1000} uncertainty changes over the entire energy range and γ was varied by ± 0.15 based on the fitting error and a possible bias by leaving the break point fixed.
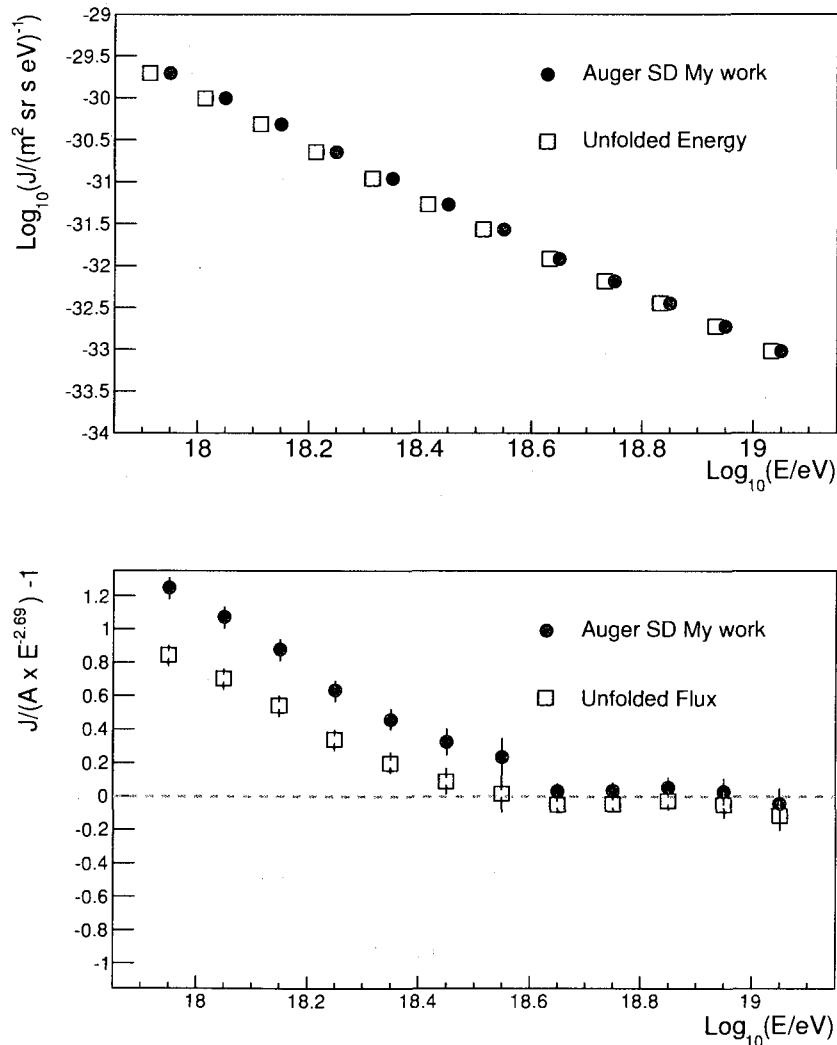
Figure 7.6 My unfolding procedure shifts the energy lower for each measured flux (top). The correction in energy could also be seen as a correction to the flux (bottom), since $J(E) \propto E^{-\gamma}$.

The unfolding correction lowers the energy for each bin of event by 8±6% below $10^{18.6}$ eV and by 3±4% above (see Figure 7.6 (top)). The flux is assumed to be equal to $E^{-\gamma}$, so unfolding correction could be approximated as a flux correction rather than an energy correction. By viewing the correction on the flux axis, the flux is lowered by 28±13% below $10^{18.6}$ eV and by 8±y% above (see Figure 7.6 (bottom)). The uncertainty on the unfolding can be translated to the flux as well which will add in quadrature with the other flux uncertainties. This is useful when comparing measurements that have different flux systematics, but the same energy scale.

Inside the Auger collaboration the flux of cosmic rays can be measured in three independent ways. The standard SD spectrum (including the extension from this thesis), the hybrid only spectrum [52] and the SD spectrum using horizontal showers [42]. Each of these methods produce a spectrum that has the same energy scale but a different energy resolution and an independent flux measurement. Currently, the horizontal spectrum is in preliminary stages of analysis and not yet ready to be published. The hybrid spectrum is in a more advanced state and should be published later this year. The following subsection compares my unfolded results with the preliminary unfolded hybrid spectrum.

7.2.2 Comparison to the hybrid energy spectrum

Both the exposure calculation and energy resolution in the hybrid spectrum are different from the standard SD spectrum and the low energy extension of this thesis. Figure 7.7 compares the unfolded spectra from the hybrid, the standard SD spectrum and the work from this thesis. The standard SD spectrum has been unfolded in the same manor as described in section 7.2.1 but using the appropriate different values for σ_E and γ as described in [12].

The unfolding correction has been applied as a flux correction in each spectrum.
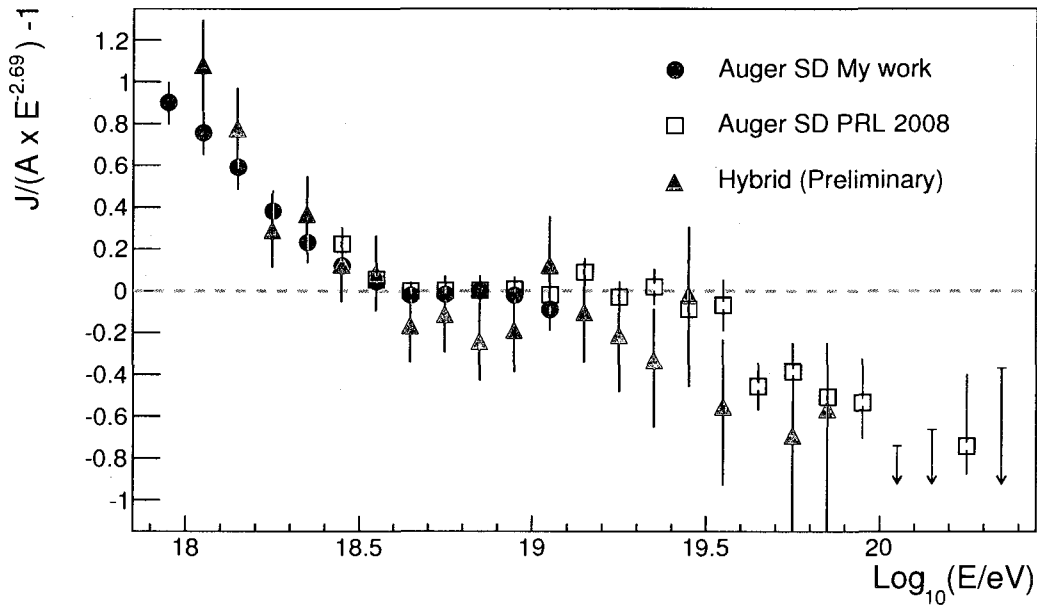
hybrid data from [53]

Figure 7.7 The spectrum from this thesis and higher energy SD spectrum
have been unfolded. They are compared to the preliminary unfolded hybrid
spectrum. All Auger spectra have the same energy systematic uncertainties.
The hybrid flux and the SD flux calculations are independent.

The larger uncertainties on each point in the SD spectrum reflect the contribution
from the unfolding correction. The energy scale and thus the energy systematic
uncertainties are the same for all Auger spectra, and are based on the FD energy
scale described in section 7.1.2. An energy systematic shift moves all points in each
spectrum the same amount so it is appropriate to not consider the energy scale in
this comparison. This is not the case when spectra from different experiments are
compared where the differing energy scales from experiment to experiment must be
considered.

The Auger collaboration is in a good position by being able to compare indepen-
dent measurements of cosmic ray flux with the same energy systematic. By doing this
a more rigorous and standard treatment of the unfolding correction can be decoupled

from the relatively large energy systematic uncertainties that are typical between different cosmic ray experiments. A more sophisticated unfolding procedure is currently being developed inside the Auger collaboration that exploits the relatively good energy resolution of the hybrid events (8%) as well as the horizontal energy spectrum measurement. A standard will develop that can be used in the whole community for past, current and future high energy cosmic ray experiments. Since no standard yet exists, it is difficult to compare results with other experiments. Only unfolded spectra will not be used for further comparison in the next section. An introduction to unfolding was discussed in this section for completeness. While the unfolding effect should not be forgotten in the following subsection, the large energy systematic differences between experiments is a more important issue when comparing energy spectra.

7.2.3 Comparison with previous experiments

While the Auger Observatory was still in the early planning stages in the late 90's, there were two ultra high energy cosmic ray experiments taking data, the Akeno Giant Air Shower Array (AGASA) and the High Resolution Fly's Eye (HiRes). The AGASA experiment was a 100 km^2 surface detector located in Japan that took data for over a decade. The HiRes experiment consisted of two fluorescence detector sites, HiRes I and HiRes II, that were located in Utah. Both the AGASA array and HiRes detectors have been decommissioned. The Auger Observatory was built as a successor to both these experiments. Inspired by Auger, many of the collaborators from both HiRes and AGASA are currently working on new cosmic ray observatory called the Telescope Array (TA). TA is a hybrid detector that is currently being built and when finished it will be about one third the size of Auger. The primary motivation for building TA was to understand the apparent discrepancy between the energy spectra

of AGASA and HiRes that is shown below.

In Figure 7.8, the energy spectra from different experiments are displayed. There are data from both fluorescence detectors (HiRes I, HiRes II, Auger) and surface detectors (AGASA, Auger). From this representation, the agreement or disagreement
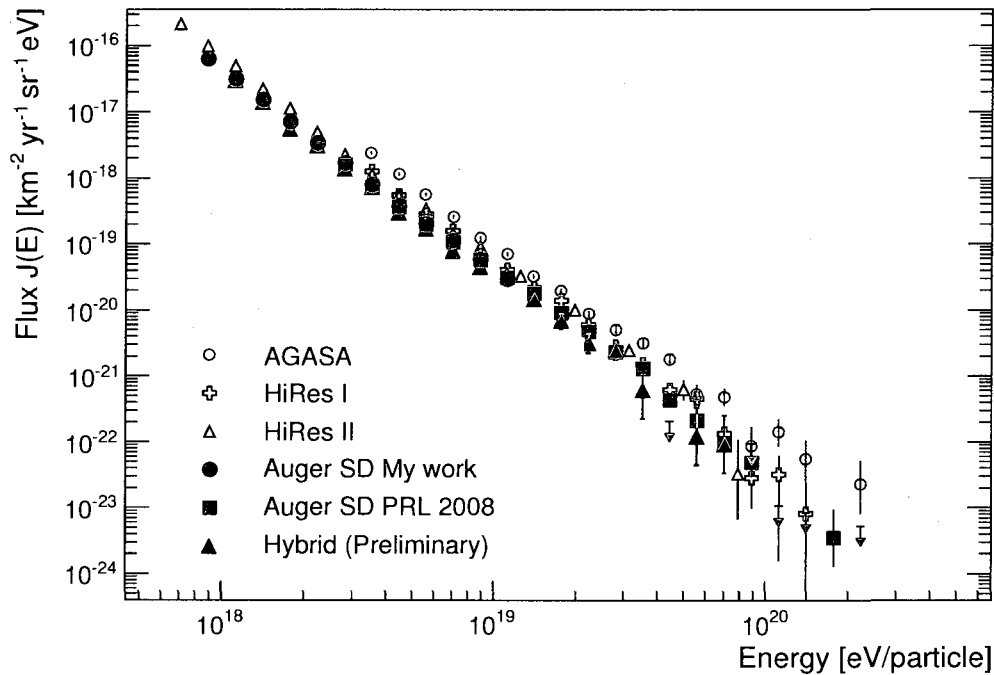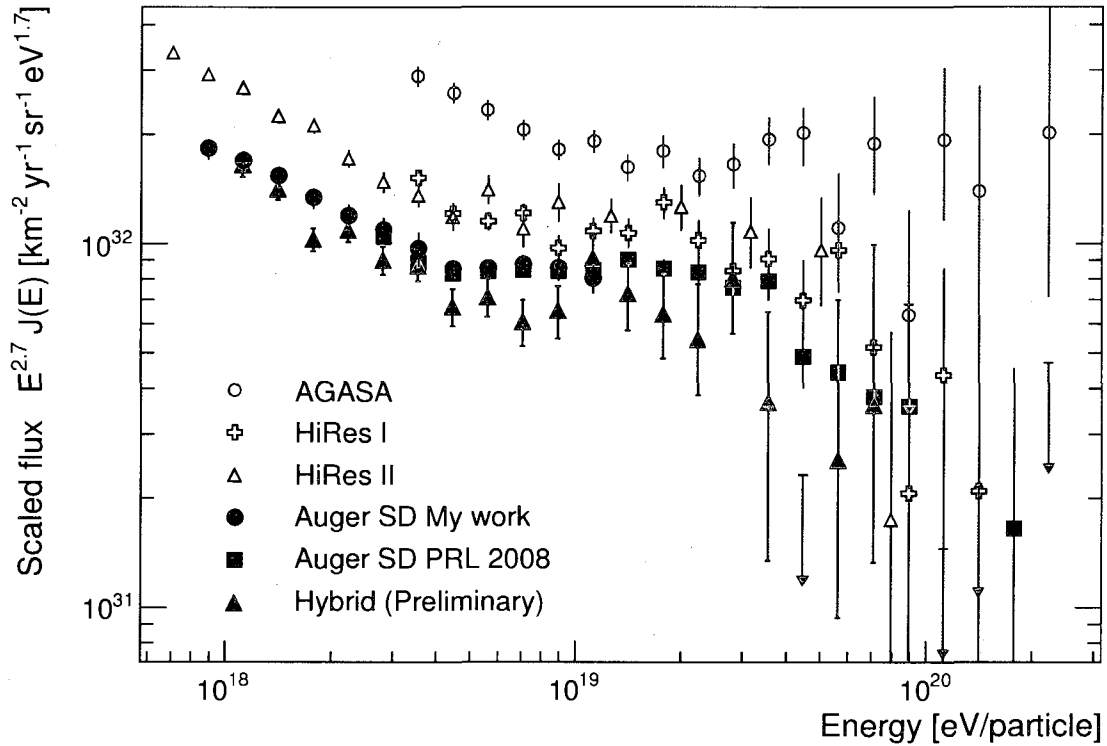

Data from other experiments collected by Ralf Engel as presented in [1]

Figure 7.8 The Auger SD energy spectrum compared with other experiments.

between experiments is hard to decipher. This is why scaled flux plots are typically used when comparing experiments. The energy spectrum is shown scaled, $E^{2.7}J(E)$ versus energy in Figure 7.9. The systematic differences between experiments become magnified in this representation. No systematic uncertainties are displayed, which as mentioned previously, is unfortunately typical in the community.

The effect of a systematic energy uncertainty of 15-30% is hard to visualize on
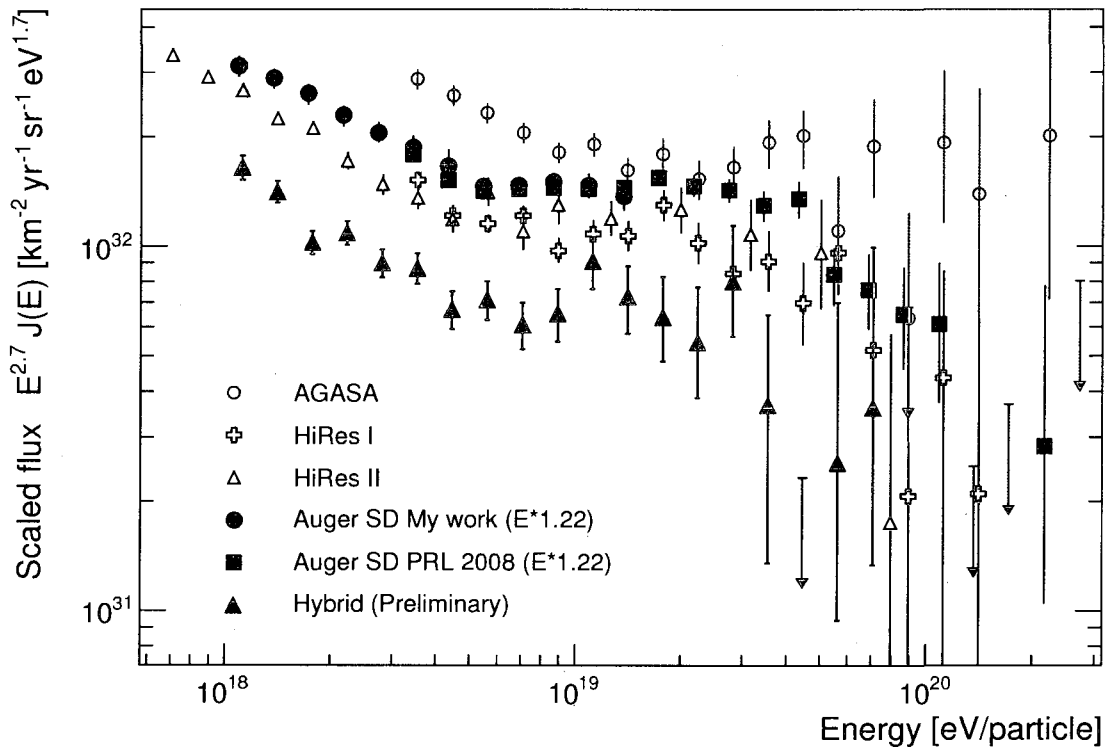
Data from other experiments collected by Ralf Engel as presented in [1]

Figure 7.9 The raw (meaning not unfolded) Auger SD energy spectrum compared with the raw Auger hybrid spectrum, Hires and AGASA. The unfolded Auger SD spectrum and hybrid spectrum do not agree as well as in Figure 7.7, this illustrates the importance of the unfolding.

scaled flux plots with a log-log scale. To illustrate how a systematic shift in energy moves the points in a $E^\gamma J(E)$ plot, Figure 7.10 shows the Auger data scaled up by 22%, which is the quoted systematic energy uncertainty discussed earlier in section 7.1.2 (not to be confused with the 21% energy resolution from the unfolding discussion).

The scaled flux plots (Figures 7.9 and 7.10) show the raw (meaning not unfolded) spectra from Auger and the published data from both HiRes and AGASA. This is because the unfolding is still in a premature stage for the Auger data and the details

Figure 7.10 The Auger SD energies are scaled up by 22% to show the effect of a systematic energy shift on a scaled flux plot. The Auger hybrid energies have not been scaled. What appeared to be very large differences between experiments in Figure 7.9 now seem to be within quoted energy systematics between experiments.

of any unfolding attempts from other experiments are unclear. The effect moves in the same direction for each experiment but by differing amounts depending on the energy resolution. The energy resolutions of these experiments ranges from 15-30% so the effect is of similar magnitude (except for the Auger hybrid spectrum which has an 8% energy resolution and an almost negligible unfolding effect). While this is obviously wrong, the unfolding effect is less important than the energy systematic differences between experiments. The unfolding correction in this thesis work resulted

in a flux correction of 28% or 8% (see Figure 7.6). While this may seem like a large effect compared to the other quoted flux uncertainties of around 5% (Table 7.1), this is a relatively small effect to worry about when comparing the spectra from different experiments where a systematic uncertainties on the energy measurement of $\approx 20\%$ can translate to a $\approx 80\%$ systematic flux uncertainty.

When the energy scale uncertainties are considered the apparent disagreement between the experiments is not nearly as significant. Before trying to make any definitive claims about the absolute energy or flux normalization values of the spectral features, some qualitative statements about the three experiments will be made.

Comparing HiRes and Auger:

- Both HiRes and Auger see an ankle (change in spectral slope between 10^{18} and 10^{19} eV) and a suppression of the flux at the highest energies.

- The relative energy difference, or the flux normalization differences, are within systematic uncertainties. If the HiRes energy scale is decreased by 10-12% or the Auger energies increased by 10-12%, the absolute flux normalizations agree.

Comparing AGASA and Auger:

- AGASA observes an ankle, but at a significantly higher energy. The AGASA ankle occurs above 10^{19} eV. Even when the Auger energies are scaled up by 22%, the Auger ankle is well below 10^{19} eV (roughly a 50-60% energy scale shift is needed to match the ankles between Auger and AGASA).

- AGASA does not see a suppression the flux. Even if the AGASA energies are lowered by 50-60%, no suppression is evident.

The Auger results qualitatively agree quite well with the HiRes results, especially when the energy systematics between experiments are considered. Both the HiRes and

Auger results do not appear to be compatible with AGASA. There are two plausible reasons for the disagreement. The AGASA energy scale is determined by models and Monte Carlo simulations, whereas Auger and HiRes use the fluorescence technique. It has been shown that the Auger energy scale is not compatible with an energy scale based on simulations. The simulations appear to be underestimating the number of muons in the shower at ground level by 50%. This incorrect characterization of the air shower results in an approximately 30% higher energy scale as predicted by the simulations [41]. In addition, a bug in the AGASA analysis code was recently found that incorrectly assigns the energy of some events. The results have not been updated or retracted yet. This is because most everyone in AGASA is currently working on a new project and nobody is actively working on the reanalysis. Up to 10-15% shifts in energy are anticipated and there is an energy dependence on the amount of the shift (based on personal communication with AGASA members). For these reasons, the apparent non-agreement of the AGASA data should not be emphasized strongly. No full comparison can be made until the analysis bug and the discrepancies in air shower simulations are worked out.

Unlike the AGASA data, the overall general agreement between the HiRes and Auger spectra is very encouraging. The overall shape of the energy spectrum observed by both Auger and Hires is very consistent. There does appear to be a systematic difference between the experiments. This difference is well within the quoted uncertainties. This can be seen by comparing the spectra in more detail with fractional flux plot. In Figure 7.11 the fractional flux is displayed. The energies from HiRes have been scaled by a factor of 0.9 (conversely the Auger energies could have been scaled up). This shift of energy is approximately half the 1-sigma quoted energy systematic uncertainty.

In Figure 7.11 the ankle spectral feature and the suppression of the flux seem to
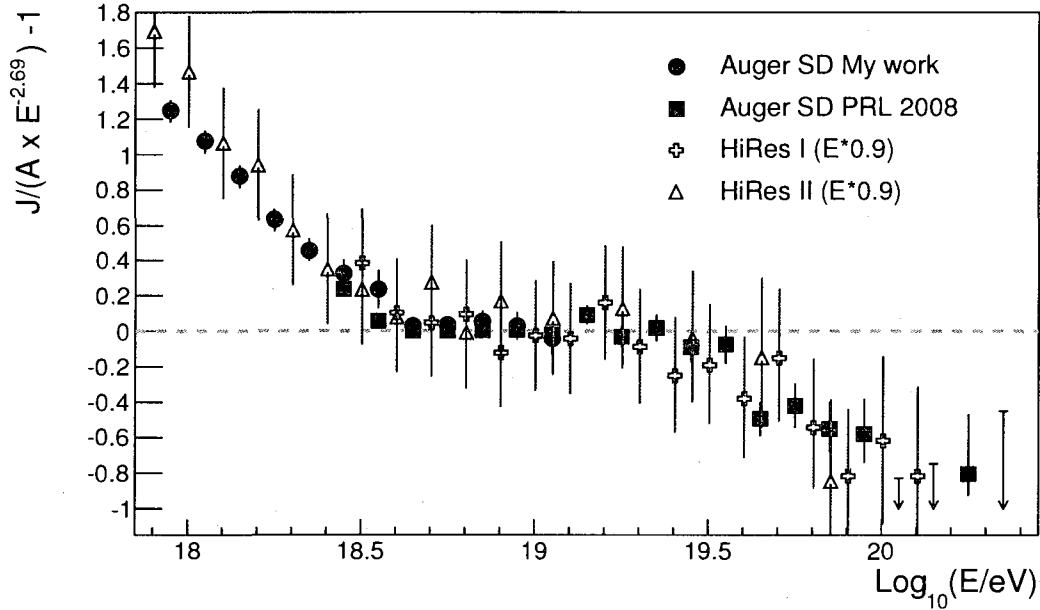
Figure 7.11 The full range of the Auger SD energy spectrum. The hollow points are based on the analysis from the PRL spectrum paper [12] and the solid points are from this thesis work. The Hires data are also shown.

coincide at the same energies for each experiment. More definitive conclusion can now be made about the cosmic ray spectrum:

- The spectral feature known as the "ankle" has been observed in this thesis analysis. The ankle occurs at $\approx 10^{18.6}$ eV which is consistent with the previous observation from HiRes.

- The spectral slope before the ankle is ≈ 3.2 and ≈ 2.7 afterwards.

- There is a suppression of the flux above $\approx 10^{19.6}$ eV.

These conclusions are still only approximate. The experimental uncertainties as well as the unfolding correction need to be considered when calculating the above values. The absolute values of the numbers quoted above do not have quoted uncer-

tainties because they are not precisely defined physical quantities. Where the ankle and the suppression are located depend on how the spectrum is fitted, which depends on the fitting function used. Different astrophysical scenarios use different functions to describe the spectrum, so the values must be computed in context of each scenario. The two leading source scenarios are discussed in the next section. Each scenario has free parameters, so predicting the exact values of the ankle and suppression can still be rather vague. However, when the approximate values from above are combined with the arrival direction and mass composition studies from Auger, strong constraints the source scenarios can be made. This is discussed in the following section. A more advanced fitting analysis of the spectrum will be more useful in the near future (2-3 years) when the reduction in the systematic uncertainties, a more standard unfolding procedure and a doubling of the statistics will constrain models even further.

7.3 Astrophysical interpretations

Recalling from chapter 1, there are certain pieces of information that need to be considered when trying to interpret the astrophysical significance of energy spectrum above 10^{18} eV:

1. The sources of the cosmic ray flux must be make a transition from galactic-dominant sources to extragalactic-dominate sources at some energy. Cosmic rays will have sufficiently high energy to escape confinement by magnetic fields in other galaxies at high energies. It is believed that no galactic sources are strong enough to accelerate cosmic rays to the very highest energies.

2. As extragalactic cosmic rays propagate from their sources they will interact with the cosmic microwave background (CMB). Extragalactic protons of energy

greater than 10^{18} eV (roughly the energy threshold for the following interaction), that travel large distances, will suffer from energy losses due to pair production.

$$p + \gamma_{CMB} \rightarrow p + e^+ + e^- \qquad (7.9)$$

Galactic cosmic rays would not travel distances far enough to have a significant fraction of their energy lost due to this process.

3. At an energies above $\approx 6 \times 10^{19}$ eV, extragalactic protons will start to loose energy via pion production, the so called GZK-effect.

$$p + \gamma_{CMB} \rightarrow N + \pi \qquad (7.10)$$

4. Nuclei will lose energy from spallation with the CMB for energies above 10^{19} eV. The energy threshold and magnitude of the energy loss depend on the specific nucleus.

5. The fraction of heavy nuclei in the flux may be changing, specifically the fraction before and after the ankle may be significantly different.

There are two leading scenarios to explain the ankle and the suppression of the flux in the observed energy spectrum. Each scenario emphasizes the above factors in different ways and has observable predictions for the energy spectrum, mass composition and arrival directions of cosmic rays. The scenarios will be gone over separately. The first scenario, what I am calling the "Pure proton extragalactic source model" makes the strongest predictions, and it does this with very few free parameters in the model. The other scenario, the "Two source mixed composition model" has many more free parameters and makes rather vague predictions. The current results from Auger, including this thesis work are highlighted throughout the discussions of these scenarios.

7.3.1 Pure proton extragalactic source model

The first scenario makes some reasonable assumptions and has a small number of free parameters.

- The cosmic ray flux starts to be dominated by extragalactic sources at energies above 10^{17} eV. The flux at the energies near the ankle and above would be from extragalactic sources alone.

- This extragalactic flux is assumed to have a constant spectral index, a source distribution that follows the matter density of the universe, and to be dominated by protons.

- The spectral index of the source, γ_g, governs the strength of the source. This is a free parameter in this model, but the results do not depend strongly on the value of γ_g.

This scenario is very straight forward with reasonable assumptions, but it makes strong predictions. Along with the energies of the spectral features, the mass composition and the arrival directions of cosmic rays have identifiable signatures.

The assumption that the sources of cosmic rays above 10^{18} eV are from extragalactic sources that follow the matter density of the universe is very reasonable. Most of the powerful astrophysical objects in the universe meet this criteria. If this is the case, the large distances the extragalactic cosmic rays travel will result in energy losses from pair and pion production. The farther the cosmic ray travels the more interactions with the CMB it will have. These interactions will impact how the pattern of arrival directions of cosmic rays, or the cosmic ray sky, looks.

The cosmic ray sky will appear isotropic at energies from 10^{18} to $\approx 6 \times 10^{19}$ eV. In this energy region, a cosmic rays can come from a very distant source (1000 Mpc)

and still retain a large fraction of its initial energy. The loss of energy due to pair production is not very drastic. At distance scales greater than a few hundred Mpc, the universe is very isotropic, thus the cosmic ray sky should observed to be isotropic as well. Above 6×10^{19} eV, the energy losses from pion production are very severe. Only cosmic rays from nearby sources ($\approx$100 Mpc) can retain an energy above this threshold. The universe within $\approx$100 Mpc is not isotropic, so the cosmic ray sky above this energy should begin show structure.
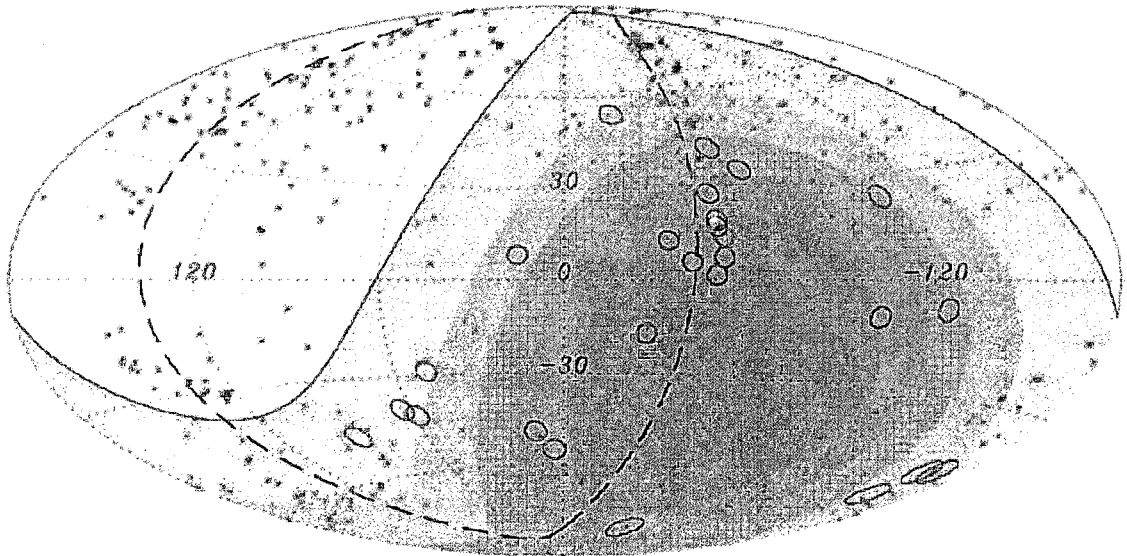

Figure from [6]

Figure 7.12 A projection of the celestial sphere in galactic coordinates with circles of radius 3.1 degree centered at the arrival directions of the 27 highest energy events detected by the Auger Observatory. The positions of the AGN within 75 Mpc are indicated by red asterisks(472 AGN total with 318 in the field of view of the Auger). The solid line represents the border of the field of view for zenith angles smaller than 60 degrees (Auger can only see the southern sky). Each colored band has equal integrated exposure with the darker colors representing larger relative exposure. The dashed line is the super galactic plane. 20 out of 27 events correlate with an AGN. This rejects the claim of isotropy of the arrival directions of cosmic rays with the highest energies with a 99% confidence level [6].

The arrival direction studies from the Auger Observatory support the above predictions. The 27 highest energy events have energy above 5.7×10^{19} eV. Below this energy nothing but an isotropic sky is seen (more on this below). When the arrival directions of the 27 highest energy events are compared with the locations of the nearby astrophysical objects a strong correlation is seen, 20 out of 27 events are within 3.1 degrees of an active galactic nucleus (AGN) (see Figure 7.12). The correlation rejects the claim of isotropy for these 27 events with a 99% confidence level [6]. AGN are approximately distributed uniformly in the local matter density of the universe. This implies that the sources of cosmic rays may not be due to AGN, but that AGN may be tracers to the real sources that come from some other type of object that is also distributed along with the local matter density. More events are needed to rule out or identify possible sources. This result of anisotropy in the cosmic ray sky is the first big step to eventually being able to perform cosmic ray astronomy!

So far the extragalactic proton scenario is looking pretty good, but the patterns in the cosmic ray sky are not the only prediction of this scenario. Let us look closer into what this scenario predicts about the ankle and the suppression of the flux in the energy spectrum.

The energy losses from interactions with the CMB will cause a modification to the observed flux on earth. Under the assumptions, if there were no energy loss mechanisms, the observed flux would have a constant spectral slope. The interactions will cause a modification of the source flux. The ratio of the expected flux on earth with no energy losses to the simulated flux with the energy losses from interactions with the CMB is called the modification factor, η. Figure 7.14 plots the modification factor as a function of energy, $\eta(E)$. This plot is similar to a fractional flux plot. The horizontal line at 1.0 represents no energy losses. The modification of the flux from pair production losses will cause a dip in the observed spectrum above 10^{18} eV (the

line marked e^+e^-). The line marked "total" is the combined losses from both pair and pion production. The severity of pion production mentioned above is evident in this plot.
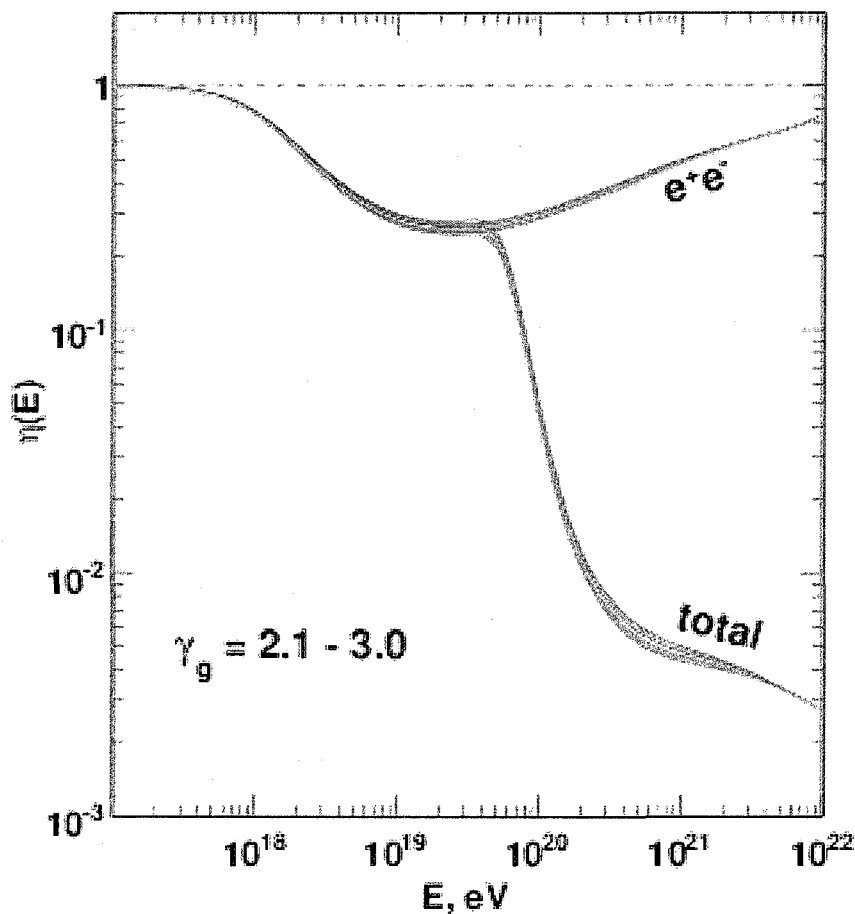
Figure 7.13 The modification factor, $\eta(E)$, is the ratio of the source flux to the expected observed flux. The flux is modified from interactions with the CMB. Displayed is $\eta(E)$ for a pure proton source model. Sources are assumed to be distributed with the matter density in the universe. The strength of the sources is denoted by, γ_g. The "ankle" spectral feature and the suppression of the flux are interpreted as propagation energy losses in this model for a wide range of source strengths.

The ankle is interpreted to occur as a result of pair production losses and the suppression of the flux results from pion production losses [54]. Both the ankle and the

suppression of the flux are explained by propagation effects alone. The interactions between the protons and the CMB are at center of mass energies that have been well studied by terrestrial accelerators. The absolute energies of the modifications to the flux can be defined rather precisely. The ankle, or where the flux begins to flatten will be just above 10^{19} eV and the suppression will begin at 6×10^{19} eV. The difference between the ankle and the suppression should be just a little over a half a decade in energy.

The final conclusions about the energy spectrum presented in the previous section (Figure 7.11 do not match very well with the predictions from this model. Both the absolute energy of the ankle and the suppression are lower than predicted, 4×10^{18} eV 4×10^{19} eV respectively. An upward global shift of energy of 50% would make the observed values more in line with the predicted values, but the energy difference between these values would remain too large. The measured energy difference is approximately one decade of energy, which is almost double what this model predicts.

Well this model is 1 for 2 now with its predictions. The third signature of this model is that the cosmic ray flux is dominated by protons. If there is more than a 15-20% contribution from heavier elements (like C, N, O and Fe) in the cosmic ray flux, the ankle can not be a result of pair production interactions during propagation [54]. The energy losses for heavy nuclei with these energies are much different than protons. When nuclei of these energies propagate, they undergo pair production at a lower rate and at an energy higher (depending on the nuclear binding energy) than protons. They also will undergo nuclear spallation with the CMB. The modification factor for iron is compared to the factor for protons in Figure 7.14.

Mass composition studies are very difficult in air shower observatories. On an event by event basis it is impossible to distinguish between a shower from an iron nucleus and from a proton. The inherent fluctuations of the depth of the first inter-
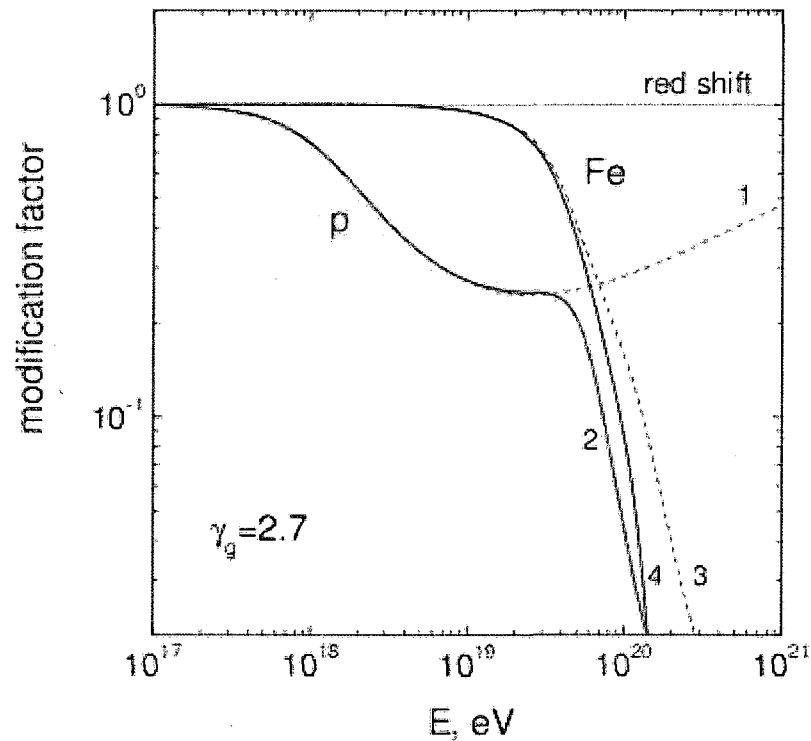
Figure 7.14 The modification factor for protons and iron nuclei. Line numbers 1 and 3 are for pair production losses for protons and iron respectively. Line numbers 2 and 4 are the total energy loss mechanisms. The energy losses have less effect on iron compared to protons at lower energies. If 15-20% percentage of the extragalactic cosmic ray flux is iron, then the "ankle" spectral feature can not be explained by propagation effects. The suppressions for both iron and protons are at almost the same energy, by a chance of nature.

action in the atmosphere and shower development are of the same magnitude as the difference in a proton shower and an iron shower. Only on a statistical basis can the average mass composition of the cosmic ray flux be determined.

The most sensitive parameter to mass composition is the average depth of shower maximum, $< X_{max} >$. In Figure 7.15 the elongation rate, or $< X_{max} >$ vs. Energy, as measured by Auger is displayed. The predicted elongation rates for protons and
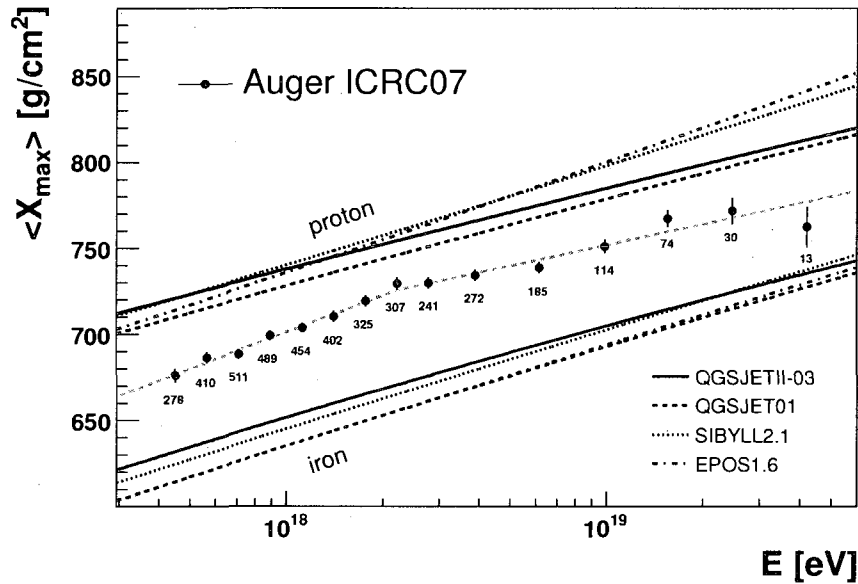
figure from [38]

Figure 7.15 The elongation rate, or $< X_{max} >$ vs. Energy, as measured by the Auger Observatory. The predicted rates for protons and iron from different simulations are displayed.

iron are also displayed. The prediction lines for both protons and iron have an approximately constant slope over the energy range displayed. The data falls in between the predictions and a constant slope is not supported. This is indicative of a probable mixed composition that is changing in this energy region.

While the cosmic ray sky predicted in this first scenario is supported by the data, the energy spectrum and mass composition studies do not. The next scenario makes less bold predictions and has more free parameters.

7.3.2 Two source mixed composition model

The second, and oldest scenario, interprets the ankle as the transition from a galactic dominated flux to an extragalactic dominated flux. This scenario assumes that the galactic component of the flux extends to energies up to and beyond the ankle. The

extragalactic flux, which is much flatter, starts to dominate the galactic flux at the exact location of the ankle. The ankle designates the energy at which the galactic and extragalactic flux components contribute equally to the total flux. After the ankle the flux is dominated by extragalactic sources and the galactic component steeply drops to zero [55].

There will be losses from pair production for the extragalactic components only. The galactic component would not suffer energy loses because of the smaller distances traveled. As a result, the total flux would not be modified as significantly, from pair production energy losses. Unlike the previous scenario the energy of the ankle does not have a predicted value, it is a free parameter in the model. The suppression of the flux is predicted by this scenario and thus the anisotropy in the cosmic ray sky above 6×10^{19} eV is predicted as well. The mass composition can be either only protons or a mixed composition, whatever fits the data better. The energy spectrum and the elongation rate of Figure 7.15, can be matched by this model by adjusting the source strength and composition.

This scenario can support all the observed data by reasonably modifying the free parameters. The only real signature of this scenario is that a significant galactic component is predicted, galactic sources or correlation of arrival directions with the galactic matter density, might be identifiable. Previous experiments had hints of a possible source near the galactic center. Both the AGASA experiment and the SUGAR [56] experiment observed an isotropic cosmic ray sky. However, these experiments observed small excesses of events in small regions of the sky near the galactic center for energies of around 10^{18} eV.

The Auger Observatory has collected more than three times the statistics as AGASA at these energies and over 10 times the SUGAR statistics. Figure 7.16 shows the sky map of Auger for a patch of sky around the galactic center. The colors

on the sky map represent the significance of the number of events detected above (or below) an isotropic expectation. The AGASA experiment found a 4.5σ excess, as compared to an isotropic prediction in a 20 degree radius region near the galactic center. The larger circle in Figure 7.16 shows this region. The SUGAR experiment found a 2.9σ excess in a 5.5 degree radius region also near the galactic center. The smaller circle in Figure 7.16 shows this region.
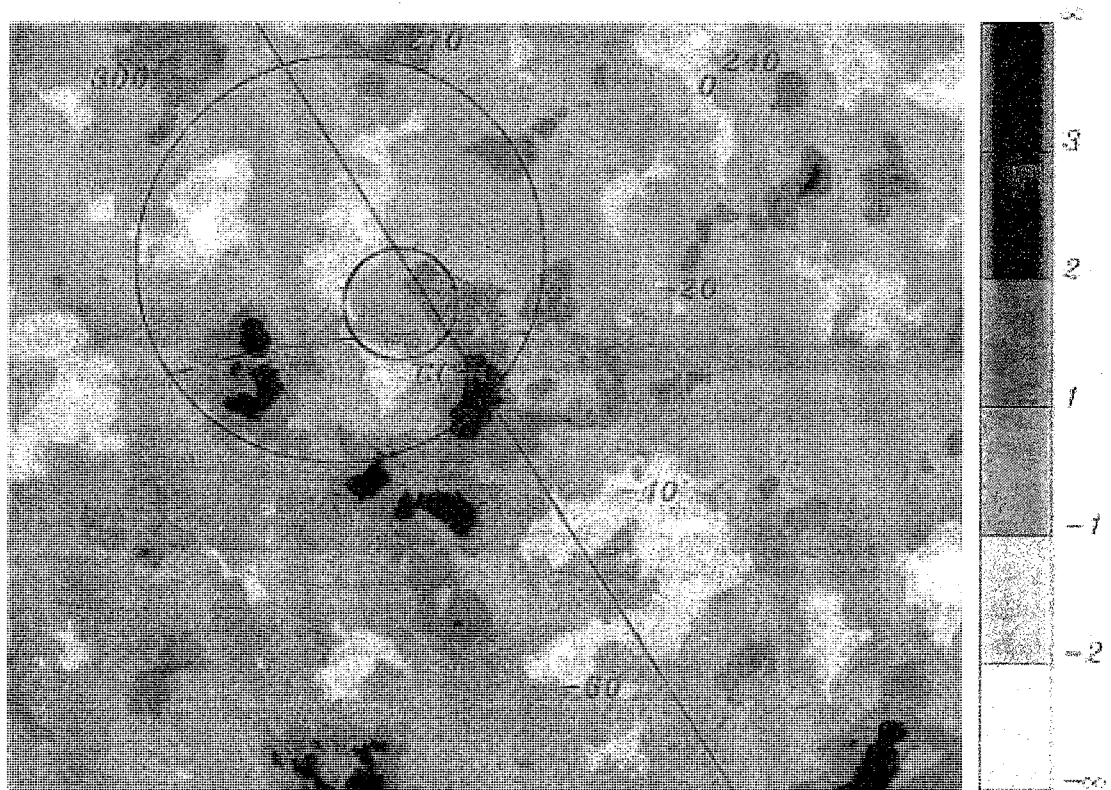

Figure from [5]

Figure 7.16 Sky map from the Auger Observatory showing the significance of the excess (or deficit) of cosmic rays compared to an isotropic expectation. The map is made from the arrival directions cosmic rays with energies of $10^{17.9} - 10^{18.5}$ eV. The region displayed is near the galactic center which is indicated with a cross, lying along the galactic plane (solid line). The regions where the AGASA experiment found their largest excess as well as the region of the SUGAR excess are indicated by the large and small circles respectively.

Based on the sky map in Figure 7.16, the previous claims of excesses are ruled

out. No significant excess or deficit is seen by Auger near the galactic center. If there are sources of galactic cosmic rays at energies around 10^{18} eV, the galactic center is a very likely location for a candidate. Even though no excess is seen, this does not rule out the region yet. The galactic magnetic fields could be strong enough that the deflections could smooth out the arrival directions. With more statistics (5+ years), the Auger Observatory will be able to set very strict limits on excesses at the galactic center and most nearby possible galactic sources (such as X-ray binary stars).

If galactic sources are ruled out, this gives support to the first scenario which assumed that all sources were extragalactic at these energies. Even so, the "pure proton extragalactic source model" while elegant and bold in its predictions is strongly disfavored based on the elongation rate and energy spectrum results from the Auger Observatory. Even if galactic sources at 10^{18} eV are eventually ruled out, the "two source mixed composition model" scenario could adjust its assumptions by saying the ankle is the result of two different extragalactic sources. This scenario has too many free parameters and does not make any definitive claims that can be ruled out. However, it is the best theory we have left. More precise measurements of all cosmic ray observables are needed to place strict limits on the free parameters in this model. When this is done, the model, while vague at this point, will be useful for identifying the sources of ultra high energy cosmic rays

7.4 The next steps for the Auger Observatory

The techniques developed in Chapters 5 and 6 were both statistically limited in some aspects. In two years the statistics will double and the corresponding uncertainties will be reduced. More stringent cuts will be usable and more sophisticated fitting routines will become practical. Besides just waiting for more statistics, the Auger

collaboration is taking steps to upgrade the Observatory.

The original design report from the Auger Observatory was to have both a northern and southern hemisphere location. For mainly political reasons inside the cosmic ray community, only the southern site was funded initially. The northern site is currently in the research and development stages. The original motivation for both a northern and southern site stems from having full sky coverage and more statistics. Based on the results from the southern site, the idea has been to build the northern site much larger to gain more statistics at the highest energies. The surface detector array will be more sparse (a $\sqrt{2}$ mile square grid compared to a 1.5 km hexagonal grid). This will allow more surface area to be covered for less money. The sparseness will raise the surface detector 100% efficiency energy to a much higher energy than in the southern site. This implies that techniques, like the ones developed in Chapter 6 on measuring the efficiency, will be very important to extend the northern energy spectrum to lower energies.

Currently there is a proposed infill array to be built in the southern site. Deployment is already underway on 85 more water Cerenkov detectors more densely spaced in one specific surface array location. This effective surface area of this infill array will be approximately 25 km^2. The design is to be fully efficient for certain zenith angles down to 10^{17} eV. Along with the standard water Cerenkov detectors, muon counters will be deployed in each infill location. The infill is important for three reasons. Firstly, lowering the energy threshold the Auger SD will allow for greater overlap with other experiments. Secondly, the systematics on the energy calibration and reconstruction of low energy SD events can be cross checked. Lastly, the muon counters will provide useful information for mass composition studies.

Overlooking the new infill array, three high elevation fluorescence telescopes are being constructed. These telescopes will be able to observe air showers developing

higher in the atmosphere so that lower energy showers can be observed. These new telescopes are also serving the dual purpose of R&D for the proposed northern Auger Observatory site.

The combined effect of having a surface detector infill array, high elevation fluorescence telescopes and a large northern observatory site will be a tremendous upgrade to the already great detector in Argentina. Seemingly overnight, (well actually 20 years in the making...), the four decades of energy from 10^{17} to 10^{21} will be measurable with state of the art measurements from one observatory. It will be an exciting time, as many theories are ruled out by the new data as it rolls in. Eventually the true nature of the sources, propagation effects and hadronic interactions at high energies will have bounds that were previously thought impossible from a cosmic ray detector. We are all lucky to be around to see these events unravel!

Bibliography

[1] R. Engel, Nucl. Phys. Proc. Suppl. **151**, 437 (2006).

[2] T. K. Giasser, *Cosmic Rays and Particle Physics* (Cambridge University Press, 1990).

[3] R. Aloisio et al., Astropart. Phys. **27**, 76 (2007).

[4] A. M. Hillas, "Cosmic Rays: Recent Progress and some Current Questions," arXiv:astro-ph/0607109v2 (2006).

[5] J. Abrams et al., Astropart. Phys. **27**, 244 (2007).

[6] J. Abrams et al., Science **318**, 939 (2007).

[7] J. R. Horandel, "Galactic cosmic rays and the knee- Results from the KASCADE experiment," arXiv:astro-ph/0509253v1 (2005).

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

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

[10] D. Allard et al., Astronomy and Astrophysics **473**, 59 (2007).

[11] The Telescope Array, "http://www.telescopearray.org/," (2008).

[12] J. Abrams et al., Phys. Rev. Lett. **101**, 061101 (2008).

[13] J. Abrams et al., Astropart. Phys. **29,** 243 (2008).

[14] J. Abrams et al., Phys. Rev. Lett. **100(21),** 211101 (2008).

[15] P. Sokolsky, *Introduction to Ultrahigh Energy Cosmic Ray Physics* (Westview Press, 2004).

[16] I. Maris, Internal Auger Note GAP-2008-026 (2008).

[17] R. Sreekantan, *Extensive Air Showers* (World Scientific, 1998).

[18] P. Lefeuvre, G. Gorodetzky and P. Salin, "Measurement of the absolute fluorescence yield in Nitrogen between 0.5 and 2.3 MeV," *Proc. 29th ICRC (Pune)*, (2005).

[19] R. Ulrich, Internal Auger Note GAP-2008-004 (2008).

[20] A. M. van den Berg, "Radio detection of high-energy cosmic rays at the Pierre Auger Observatory," *Proc. 30th ICRC (Merida)*, (2007).

[21] X. Bertou et al., Nucl. Instr. Meth. **A568,** 839 (2006).

[22] A. M. Hillas, Acta Phys. Acad. Sci. Hung. **29,** 355 (1970).

[23] D. Newton, Astropart. Phys. **26,** 414 (2007).

[24] J. Abrams et al., "The Fluorescence Detector of the Pierre Auger Observatory," Nucl. Instr. Meth. (To be submitted 2009).

[25] M. Mostafa et al., "Hybrid performance of the Pierre Auger Observatory," *Proc. 29th ICRC (Pune)*, (2005).

[26] M. Proza, "Systematic study of atmosphere-induced influences and uncertainties on shower reconstruction at the Pierre Auger Observatory," *Proc. 30th ICRC (Merida)*, (2007).

[27] M. Unger, B. R. Dawson, R. Engel, F. Schussler, and R. Ulrich, "Reconstruction of Longitudinal Profiles of Ultra-High Energy Cosmic Ray Showers from Fluorescence and Cherenkov Light Measurements," arXiv:0801.4309v1 [astro-ph] (2008).

[28] F. Kakimoto et al., Nucl. Instr. Meth. **A372,** 527 (1996).

[29] M. Iarlor et al., Astropart. Phys. **28,** 41 (2007).

[30] M-UG 6 a special filter glass 3.25 nm thick from Schott DESAG .

[31] J. A. J. Matthews, "Optical Calibration of the Auger Fluorescence Telescopes," *SPIE Astronomical Telescopes and Instrumentation,* (2003).

[32] J. T. Brack, R. Meyhandan, G. J. Hoffman, and J. A. J. Matthews, Astropart. Phys. **20,** 653 (2004).

[33] A. Tamashiro et al., Internal Auger Note GAP-2004-068 (2004).

[34] National Institute of Standards and Technology NIST Special Publication 250-41 (1998).

[35] C. Aramo et al., "Optical relative calibration and stability monitoring for the Auger fluorescence detector," *Proc. 29th ICRC (Pune),* (2005).

[36] C. D. Fratte et al., Internal Auger Note GAP-2008-129 (2008).

[37] M. Roberts et al., Internal Auger Note GAP-2005-089 (2005).

[38] M. Unger, "Study of the Cosmic Ray Composition above 0.4 EeV using the Longitudinal Profiles of Showers observed at the Pierre Auger Observatory," *Proc. 30th ICRC (Merida),* (2007).

[39] M. Ave, "Reconstruction accuracy of the surface detector array of the Pierre Auger Observatory," *Proc. 30th ICRC (Merida),* (2007).

[40] C. Bleve, "Weather induced effects on extensive air showers observed with the surface detector of the Pierre Auger Observatory," *Proc. 30th ICRC (Merida)*, (2007).

[41] R. Engel, "Test of hadronic interaction models with data from the Pierre Auger Observatory," *Proc. 30th ICRC (Merida)*, (2007).

[42] P. F. San Luis, "Measurement of the UHECR spectrum above 10^{19} eV at the Pierre Auger Observatory using showers with zenith angles greater than $60°$," *Proc. 30th ICRC (Merida)*, (2007).

[43] J. Abrams et al., "The aperture of the Pierre Auger Observatory surface detector for extensive air showers below $60°$: from the trigger system to exposure calculation," Nucl. Instr. Meth. (To be submitted 2009).

[44] I. Lhenry et al., Internal Auger Note GAP-2003-088 (2003).

[45] D. Allard et al., "Determination of the aperture of the PAO surface detector," *Proc. 29th ICRC (Pune)*, (2005).

[46] B. Bauleo et al., "Measurement of the lateral distribution function of UHECR air showers with the Auger Observatory," *Proc. 29th ICRC (Pune)*, (2005).

[47] J. Abrams et al., Nucl. Instr. Meth. **A523**, 50 (2004).

[48] P. Sommers et al., "First estimate of the primary cosmic ray energy spectrum above 3 EeV from the Pierre Auger Observatory," *Proc. 29th ICRC (Pune)*, (2005).

[49] B. Dawson, "Hybrid Performance of the Pierre Auger Observatory," *Proc. 30th ICRC (Merida)*, (2007).

[50] P. Privitera, Internal Auger presentation (2007).

[51] R. Knapik, Internal Auger Note GAP-2009-000 (2009).

[52] L. Perrone, "Measurement of the UHECR energy spectrum from hybrid data of the Pierre Auger Observatory," *Proc. 30th ICRC (Merida)*, (2007).

[53] F. Schuessler, Internal Auger Note GAP-2008-155 (2008).

[54] V. Berezinsky:, "Transition from galactic to extragalactic cosmic rays," *Proc. 30th ICRC (Merida)*, (2007).

[55] T. Gaisser and T. Stanev, "High-energy Cosmic Rays," arXiv:astro-ph/0510321v1 (2005).

[56] J. A. Bellido, Astropart. Phys. **15**, 167 (2001).