Annual Modulation of the Cosmic Ray Anisotropy: A Time Series Analysis of the 20 TeV Sky with the IceCube Neutrino Telescope

by

Henry Tillman Corbett IV

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Bachelor of Science in Physics

at

GUILFORD COLLEGE

April 2014

Author . . . .
Department of Physics
April 10, 2014

Certified by . . . .
Steven S. Shapiro
Professor of Physics
Thesis Supervisor

Certified by . . . .
Paolo Desiati
Associate Scientist, University of Wisconsin - Madison
Thesis Supervisor

Accepted by . . . .
Thomas P. Espinola
Chairman and Glaxo Wellcome Professor of Physics,
Department of Physics

Annual Modulation of the Cosmic Ray Anisotropy: A Time Series Analysis of the 20 TeV Sky with the IceCube Neutrino Telescope

by

Henry Tillman Corbett IV

Submitted to the Department of Physics
on April 10, 2014, in partial fulfillment of the
requirements for the degree of
Bachelor of Science in Physics

Abstract

Data from the IceCube South Pole Neutrino Observatory has be used to characterize the anisotropy in the arrival directions of muons produced in cosmic ray air showers. The anisotropy consists of a superposition of a solar dipole associated with the Compton-Getting effect and a sidereal dipole and quadrupole of unknown origin. We utilized IceCube data from the 40, 59, 79, and 86 string configurations of the detector, comprising four years of data and 3.69×10^{10} events with a median energy of 20 TeV. We limited our analysis to data from four azimuthal regions, allowing the rotation of the earth to trace out a periodic signal. We used a Lomb-Scargle periodogram to approximate a frequency spectrum from the event rates. The frequency spectrum contained four peaks with a significance level greater than 5σ, including a peak at 0.997 day^{-1} that is consistent with a sideband caused by modulation of the solar dipole. If further analysis confirms this modulation, interference between the solar and sidereal time frames will need to be considered in future analyses of the anisotropy.

Thesis Supervisor: Steven S. Shapiro
Title: Professor of Physics

Thesis Supervisor: Paolo Desiati
Title: Associate Scientist, University of Wisconsin - Madison

This departmental honors thesis has been examined by a Committee of the Department of Physics as follows:

Paolo Desiati, Ph.D. . . . .
Project Advisor, Thesis Committee
Associate Scientist (UW-Madison)

Thomas Espinola, Ph.D. . . . .
Member, Thesis Committee
Glaxo Wellcome Professor of Physics

Kami Rowan, D.M.A. . . . .
Member, Thesis Committee
Associate Professor of Music

Steven S. Shapiro, Ph.D. . . . .
Chairman, Thesis Committee
Professor of Physics

Robert M. Whitnell, Ph.D. . . . .
Member, Thesis Committee
Professor of Chemistry

Contents

List of Figures

Chapter 1

Cosmic Rays

> The subject [of cosmic rays] is unique in modern physics for the minuteness if the phenomena, the delicacy of the observations, the adventurous excursions of the observers, the subtlety of the analysis, and the grandeur of the inferences.[1]

1.1 Discovery of Cosmic Rays and Motivations for Study

1.1.1 The Electroscope

An electroscope is a simple device, consisting of a two thin sheets (called “leaves") of metal, typically gold, connected on one end. Figure 1-1 shows an example of a simple gold-leaf electroscope. When the device is electrically connected to a source of charge, both leaves will acquire a like charge (i.e., both positive or both negative) and repel one another to some distance related to the charge’s magnitude. However, when left in a charged state, all electroscopes will lose charge and return to ground, seemingly with no dependence on how well isolated electroscope was from its environment. Understandably, this phenomenon caused considerable confusion, and led to the theory that some unknown process was causing a seemingly random decay of stable atoms to create the ions responsible for leeching charge.

---

[1] Karl K. Darrow, as quoted in the front matter of ?.

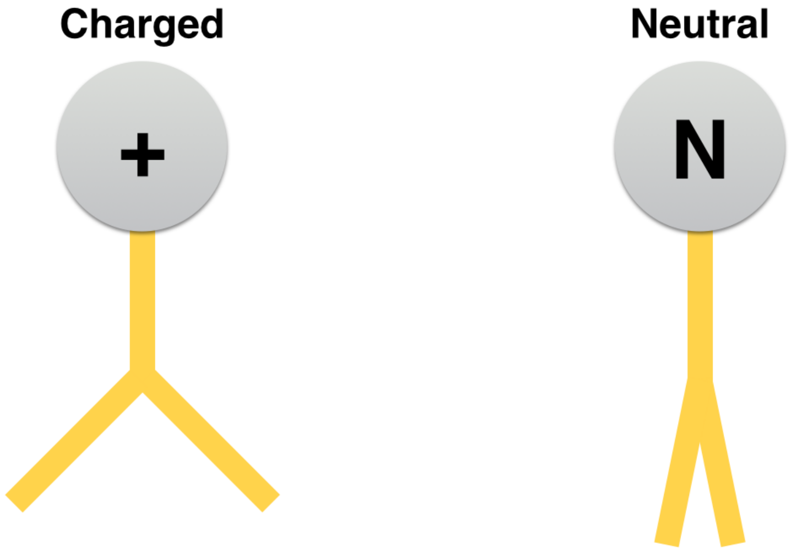

Figure 1-1: A gold-leaf electroscope. When charged (*left*), the leaves repel one another by a distance related to the magnitude of the charge; however, even a perfectly isolated electroscope will return to its neutral state (*right*) due to the ionization of the surrounding gas molecules by cosmic rays. Electroscopes were the primary tool of early cosmic ray physicists.

In the early twentieth century, science was well aware of the atom, at least insofar as that it contained equal quantities of positive and negative charge, and contained some quantized electron structure, and thus concluded that the inevitable loss of charge measured on the electroscope was related to a breakdown of the gas surrounding the leaves of the device, creating charged particles called *ions*. (?, p. 3). Additionally, thanks to the contributions of Wilhelm Roentgen towards pioneering x-ray physics in 1895, they were also aware that an electroscope placed in the path of an x-ray tube will discharge very rapidly. Finally, came from the 1898 isolation of radium by Marie and Pierre Curie, showing that charged particles – like those possibly responsible for electroscope discharge – can be emitted by radioactive materials (?, p. 5).

From these basic tenets, it was concluded that if charged particles resulting from radioactive materials could cause ionization of air and thus cause an isolated electroscope to discharge, there must be some dilute source of radioactivity in the Earth's crust creating a constant flux of charged particles. To test this, Thomas Wulf carried an electroscope to the top of the Eiffel tower (325 m; 1,063 ft) in 1909 and measured its rate of discharge to not be different from the rate at ground level. Three years later in early 1912, Albert Gockel took a similar device up to to an altitude of 600 m (2,000 ft) on a balloon flight and again observed no change. Finally, Austrian Victor F. Hess attempted a high altitude balloon test later in 1912 and observed the predicted steep drop off in the rate of discharge just beyond the altitude achieved by Gockel. Shockingly, as the balloon continued to climb, the rate of discharge began to increase again. By the time Hess reached an altitude of 5,000 m (16,500 ft), the rate had blow up to four times what he had observed at sea level. Hess published later that same year, claiming to have detected the presence of a measurable and penetrating "radiation from above" and was award a Nobel prize for his discovery 24 years later in 1936. Hess's observations were extended by Wilhelm Kohlhöroster between 1913 and 1919 in a series of risky high altitude balloon flights, who demonstrated that the increase in the rate of discharge continued up to an altitude of 8,500 km (28,000 ft) (?, p. 2-6). Unmanned balloons were used to further confirm the phenomenon at higher and high altitudes by Robert Millikan and Erich Regener in the 1920s, but still two possibilities remained: increase in ionization at higher altitudes could be explained by increasing concentrations of radioactive gasses or associated with electrical storms in the upper atmosphere. However, no evidence of significant concentrations of radioactive atmospheric components has ever been found, nor are discharge rates dependent on weather conditions.[2]

1.1.2 Evolving Motivations for Study

Since the time of Hess and his contemporaries, cosmic ray astrophysics has grown into a diverse field founded on close collaboration between astrophysicists, high energy particle physicists, plasma physicists, engineers, and – in recent decades – computer scientists. Initially, this collaborative interest came from the fact that cosmic rays provided the only source of high energy particles for experimentation up until the 1950s (?, p. 2). Eventually, particle accelerators became more commonplace and brought the possibility for

---

[2] At least, not short term weather conditions. There is measurably seasonal dependence on cosmic ray production due to the changing density of the atmosphere. See Section

more precise arrays of sensors and greater reproducibility through finely tunable energies and particle types; however, particle accelerators have always been limited to a certain energy threshold, a restriction not found in cosmic ray physics, where energy is limited only by the size of the telescope and the amount of time one waits.

1.2 Cosmic Ray Air Showers

In 1938, a group led by Frenchman Pierre Auger discovered that the particles detected as cosmic rays were coincident in time. From this, they concluded that the particles were actually part of a larger phenomenon, called an air shower (sometimes called an *extensive air shower* or simply a cascade) (?). The cosmic rays observed were consistent with particle events up to 20 PeV (?). In this phenomenon, an incoming cosmic ray particle, called the *primary*, enters Earth's atmosphere and collides with a molecule in the upper atmosphere, creating a cascade of secondary particles and electromagnetic radiation. High-energy particles, with energies greater than 10 TeV, will create a large number of secondary particles, as the secondary particles produced by the first molecular collision will continue moving and colliding with other molecules, creating even more particles (?, p. 3)

1.2.1 Composition of Air Showers

Particle cascades can be initiated by gamma rays, neutrinos, or electrons, though the incident particle is either a proton or a light nucleus in the overwhelming majority of cases. Upon interacting with an atmospheric molecule, the incident particle will break apart into short-lived pions and kaons, which decay into neutrinos, muons, electrons, and gamma rays. Figure 1-2 shows three types of air showers, produced by a gamma ray, a proton, or an iron nucleus. For the proton and iron nucleus, the majority of the energy of the primary particle is diverted into muons which penetrate downwards through the atmosphere to the ground. From computer simulations, we know that these muons tend deviate less than 0.2° from the arrival axis as they propagate (?). The remaining energy of the incident particle is dissipated into the gamma rays and electrons via repetitive

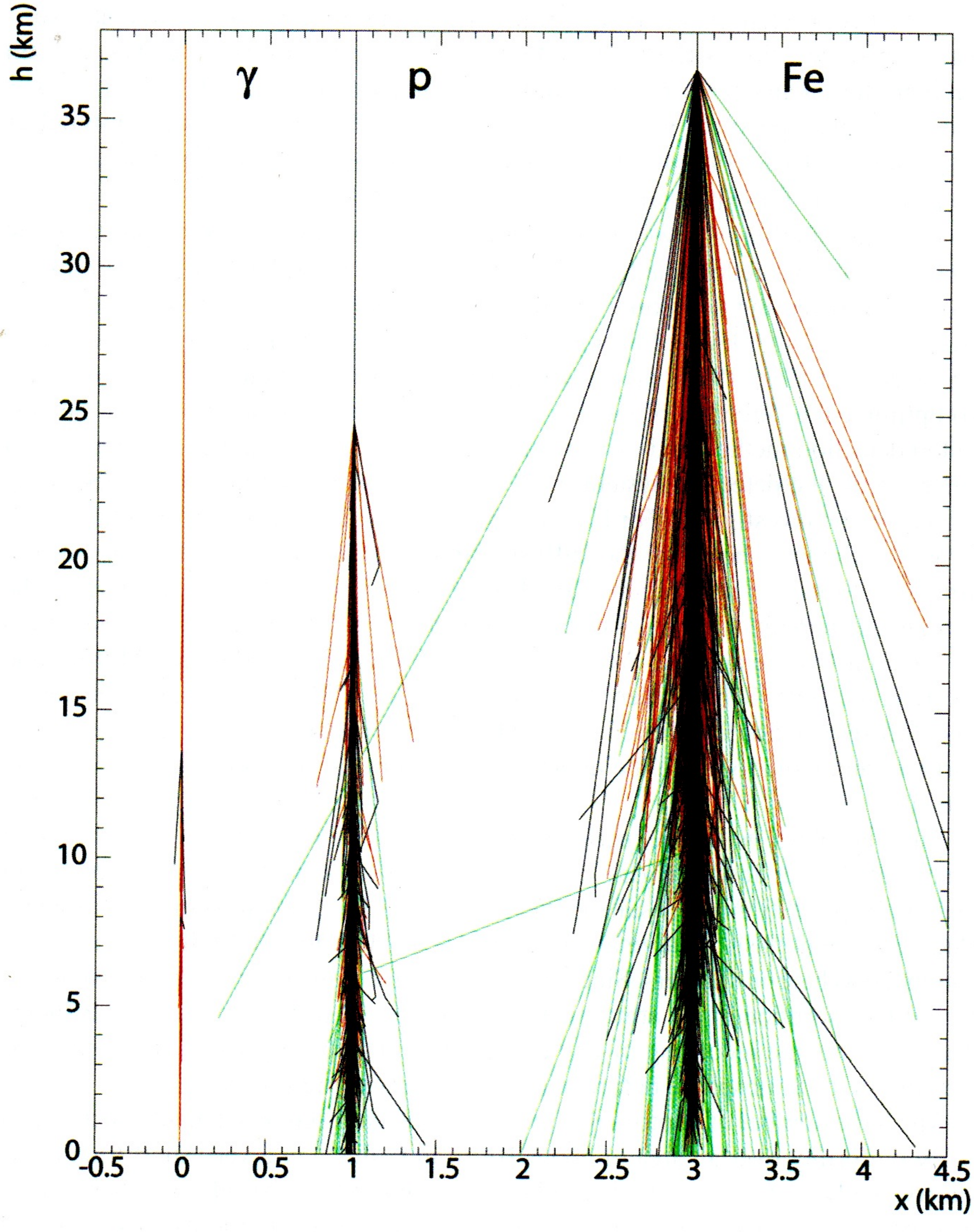

Figure 1-2: Side view of the particle trajectories resulting from three types 10 GeV cosmic rays from data produced by the KASCADE experiment. Electromagnetic products of the cascade are shown in red, hadrons are black, and muons are shown in green. The portion of the shower is predominantly made up of muons, which can carry sufficient energy to penetrate into the Earth's crust (?, pg. 21).

pair creation (gamma rays interact with atmospheric nuclei, producing pairs of electrons and anti-electrons – called positrons) and bremmstrahlung (literally, "braking radiation," wherein a particle decelerates via interaction with the atmosphere and converts its kinetic energy into high-energy light) processes. Effectively, we have superimposed cascades of photons (via bremmstrahlung), electron-positron pairs (via repetitive pair creation), and neutrinos. Depending on the arrival angle of the initial particle, the neutrino components of a high-energy cascade can penetrate completely through the planet and continue out indefinitely into deep space. The total composition of the cascade is approximately 90% electron-positron pairs, 1% hadrons (meaning, charged particles such as protons, pions, kaons), and 10% muons (?, p. 4-8) . Of these, only muons are detectable by the detector used in this thesis, as they carry the majority of the mass and energy of the late shower.

1.2.2 Longitudinal and Lateral Development

As the particle cascade moves through the atmosphere, the scattering of the secondaries causes the cascade to spread out perpendicular to the arrival axis. Figure 1-3 shows this spreading. The front of the shower shows a slight curvature, with a roughly disc-shaped cross section that is predominantly symmetrical about the arrival axis of the primary particle (with the exception of secondary particles carrying a sufficiently small energy as to be captured or influenced by the magnetic fields of the planet). By the time the cascade reaches the sea level, the disk of the cascade front can encompass several square kilometers. The disk thickens along the edges due to the lower energies of the particles farther from the primary axis creating closer scattering (??). The development of the shower, both parallel and perpendicular to the arrival axis, is dependent on the energy of the primary cosmic ray particle, the height at which the primary first interacted with the atmosphere, and the zenith angle (effectively, the path taken by the particle between the top and the bottom of the atmosphere). Additionally, the height of maximum development of the cascade is dependent on the vertical column density of the atmosphere (in mass per area), which is a function of barometric pressure and temperature in the atmosphere. In general, the higher the energy of the primary, the more deeply into the

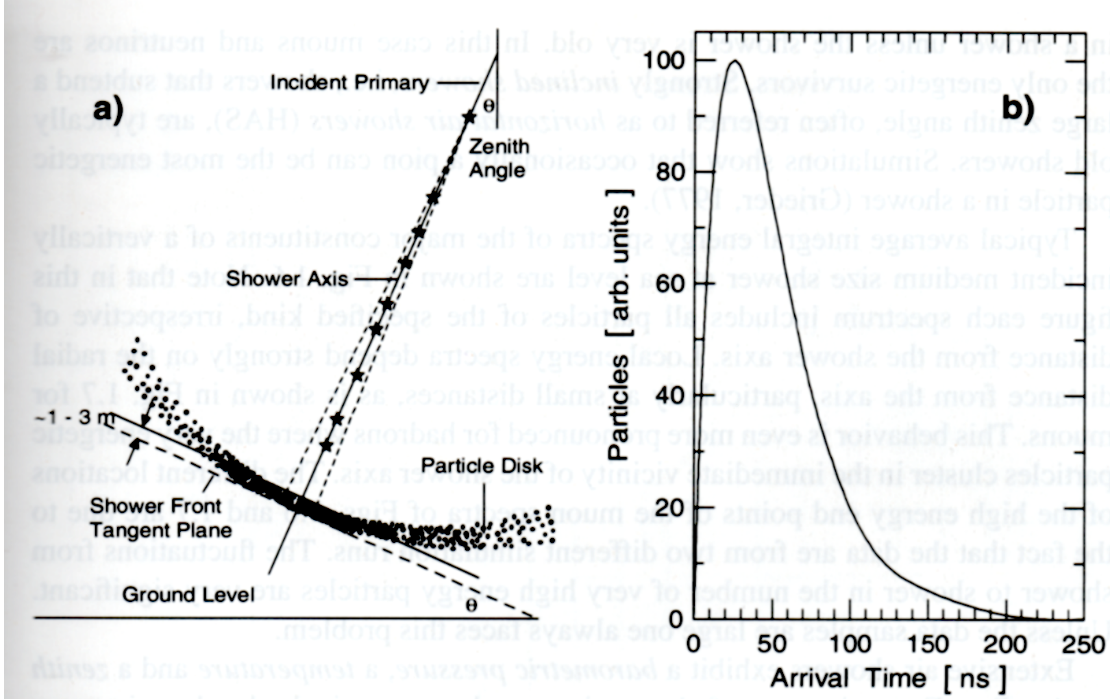

Figure 1-3: **(a)** Longitudinal profile of a cosmic ray air shower, showing the arrival axis and the profile of the leading disk. Note the thickening of the disk edges due to fall off in energy dispersion far from the central axis. **(b)** Cascade size (in arbitrary units) as a function of time in the reference frame of the detector. (?, pg. 9).

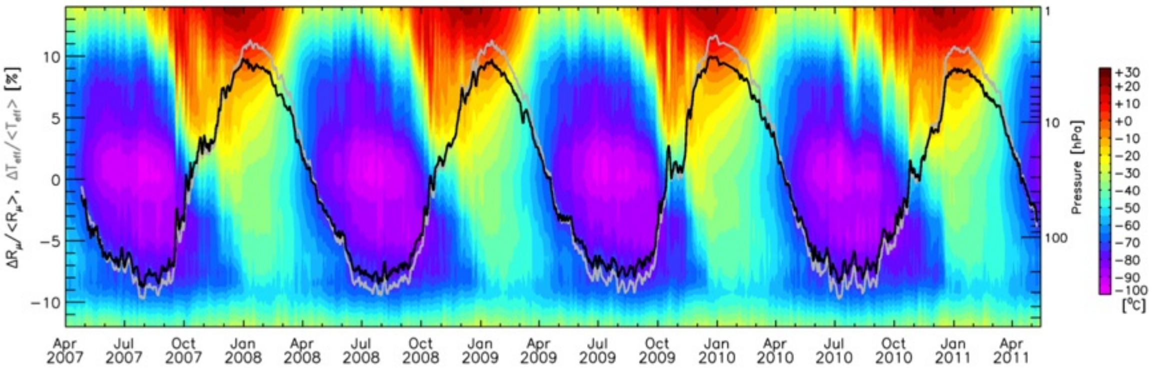

Figure 1-4: Relative event rates and effective atmospheric temperature for four years of IceCube data (2007-2011). Color scale indicates atmospheric pressure. Event rates increase with atmospheric pressure and temperature in the warmer austral summer months (?).

atmosphere the shower will penetrate. For a given particle energy, the height of the first collision will increase with increasing primary particle mass, which makes some degree of intuitive sense: Larger particles are more likely to collide earlier in their path through the atmosphere (?, p. 5).

1.2.3 Seasonal Variation

Cosmic ray interactions in the atmosphere are highly correlated with atmospheric temperature and pressure for both daily variations and the ±8% yearly modulation observed in cosmic ray muon event rates (?). This variation, as seen with the IceCube Neutrino Telescope, is presented in Figure ??. Event rates increase with temperature into the austral summer months, with a distinct ledge forming in spring due to yearly climatological turbulence at the South Pole. In a warmer atmosphere, the column density of the atmosphere is greater, causing incoming cosmic ray particles a greater likelihood of interacting.

1.3 Origins of Cosmic Rays

1.3.1 Supernovae and Gamma Ray Bursts

Supernovae have long been the primary candidate for cosmic ray production, dating back to a 1934 publication by Fritz Zwicky (?). The famous Crab Nebula, a supernova remnant, is a known point source of cosmic rays (?). Most recently, data from the Fermi Space Telescope revealed a neutral pion decay through a signature in the gamma-ray spectra of supernovae remnants. Neutral pions will decay into gamma rays with a characteristic energy following interactions between accelerating protons and the interstellar medium. Thus, detecting this signal in association with a supernova indicates that the supernova producing large numbers of high-velocity particles, carrying up to 10% of the total energy of the supernova (10^{43} J or 10^{61} eV) (?). Attempts have been made to correlate cosmic rays with gamma-ray bursts, though efforts today have only managed to set lower limits on the cosmic ray flux expected from a burst.

1.3.2 Relative Dust

Lyman Spitzer suggested that grains of interstellar dust could be accelerated to relativistic velocities approaching the speed light in 1949 by the radiation pressure of a supernova (?). A similar effect was predicted in 1950, due to the radiation pressure coming from Seyfert galaxies, which are highly luminous, quasar-like objects which produce large amounts of radiation from their active cores, which are powered by supermassive black holes powered by an accretion disk of inflowing matter (?). The luminosity of just the cores of these galaxies is typically comparable to the entire luminosity of an entire average-sized galaxy like our own Milky Way. Four years later, it was theorized that cosmic dust would typically have a non-zero charge, and could be accelerated by the same cyclotron effects experienced by any charged particle (?). Dust would acquire a charge due to interactions with starlight, the interstellar medium, or relic 2.7 K photons from the Big Bang (?).

The dust grain theory could not explain all cosmic rays, but could explain some otherwise anomalous events with a very sparse and flat distribution, which could be explained by the superimposed cascades produced as the many nuclei making up the dust particle break apart. To produce an air shower, the dust particle would need to have an energy of at least 100 TeV and begin interacting with the atmosphere at an altitude of at least 100 km (?). There have been multiple observations of air showers with parameters consistent with a dust-grain event, but only at the ultra-high end of the energy spectrum, with most candidates in the EeV range. More recent insights have shown that magnetic fields and solar wind within our solar system may be sufficient to break apart dust grains into individual nucei, making them indistinguishable from free-nucleus events (??)

1.4 Anisotropy and the Cosmic Ray Distribution

1.4.1 Motivations

The distribution of cosmic ray arrival directions could be a result of the distribution of Galactic sources of cosmic rays, most significantly close or recent supernovae. Though cosmic rays arrive at Earth filtered through the interstellar medium, the discrete distribustion of sources would still create an ordered distribution as observed from Earth (?). In the broadest sense, analysis of the distribution of cosmic rays provides us with a tool to identify cosmic ray sources (?, pg. 552). A better understanding of the anisotropic distribution (contrast with an isotropic, or evenly disordered, distribution) can also reflect the general pattern of cosmic ray propagation through the galactic (and extra-galactic, for sufficiently high energy events) magnetic fields (?).

1.4.2 Anisotropy Measurements

Comprehensive measurements of the anisotropy of cosmic ray arrival directions have become available in the past few decades, beginning in the Northern Hemisphere and at the lower end of the energy spectrum, in the range of tens to hundreds of GeV (?). Later investigations have verified that the anisotropy is also present in the TeV energy range (?). The TeV anisotropy is also present in the Southern Hemisphere sky, as confirmed by the IceCube Neutrino Observatory. Figure 1-5 shows the distribution of cosmic ray

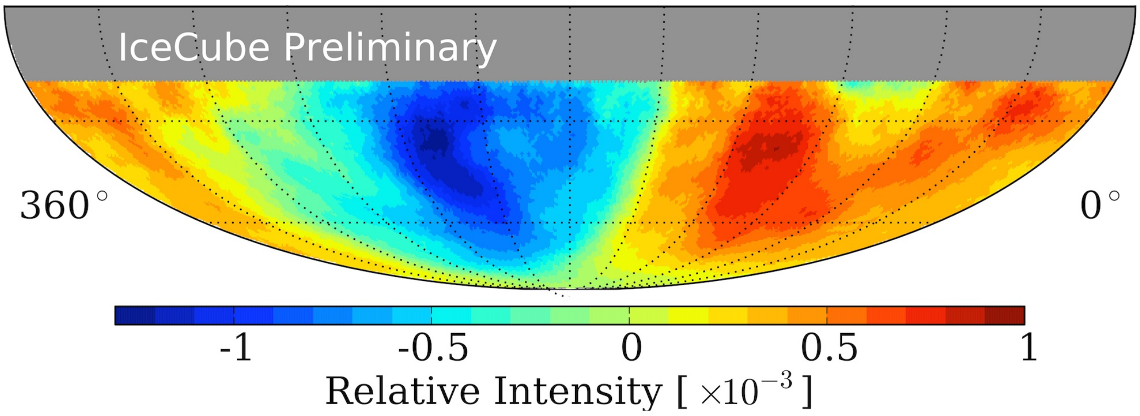

Figure 1-5: Skymap in celestial coordinates showing the relative intensity of 20 TeV cosmic rays as a function of arrival direction as seen by IceCube in the southern hemisphere sky. Large-scale dipole effects are visible at a 10^{-3} level. (?).

flux in the Southern Hemisphere sky. A clear dipole is present, with a deviation from an isotopic distribution of 10^{-3}, which is consistent with observations of the anisotropy made in the Northern Hemisphere (?). However, the distribution is more complex: Past investigations of the arrival directions of cosmic rays with IceCube (Abbasi et al., 2012) and other detectors have demonstrated a robust anisotropy in the arrival directions of cosmic rays, with respect to both the position of the sun (the solar component) and the position of the stars (the sidereal component).

1.4.3 Solar Component

Figure 1-6 shows the relative intensity of the cosmic ray event rate as a function of arrival direction in a coordinate system where the position of the sun is fixed. The data in Figure 1-6 consist of an integer number of sidereal years (longer than a solar year by about 20.4 minutes) of cosmic ray events. If the sidereal component is symmetric over the, the sidereal contribution will sum to zero over an integer number of years. If the data were not organized in this way, the observed distribution would be an interference pattern result from the superposition of the two effects. The data are fairly well fit by a simple dipole with an amplitude modulation of 10^{-4}, which is consistent with the predictions made by **?** based on the Doppler effect caused by the Earth's relative motion around the sun. Similar to having the front of one's coat become soaked while the back remains dry

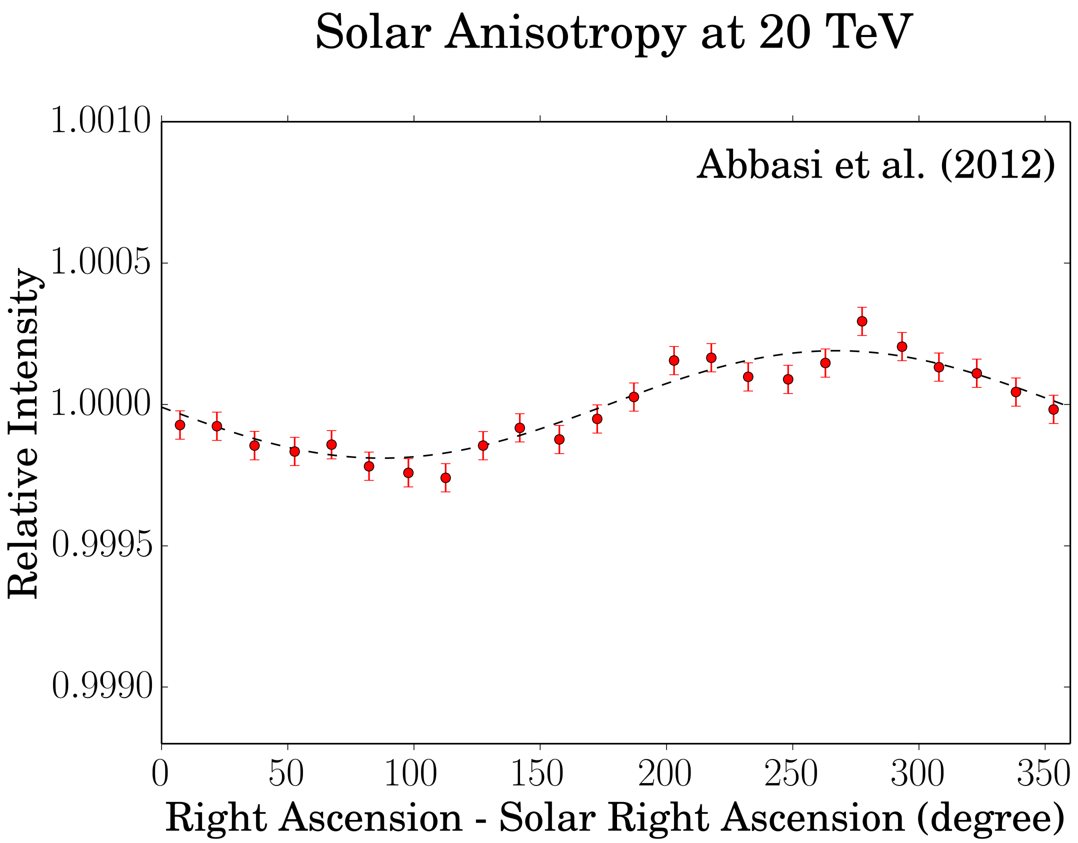

Figure 1-6: Relative intensity of cosmic rays vs. position in a coordinate system that is fixed with respect to the position of the sun, showing the solar dipole effect observed in the arrival directions of cosmic rays with IceCube. The 10^{-4} variation is attributable to the Compton-Getting effect (?).

when running through a rainstorm, we observe an excess of cosmic rays in the direction of the motion of the Earth along the ecliptic plane and a deficit directly opposite. This is referred to as the Compton-Getting effect. Due to its tractable physical explanation and low intensity, the solar component of the anisotropy provides a sensitivity test for investigations of the less-understood sidereal component (?).

1.4.4 Sidereal Component

Figure 1-7 shows the relative intensity of the cosmic ray event rate as a function of right ascension (a coordinate system that is fixed with respect to the stars). The origins of the sidereal anisotropy are not yet well understood, but the observed features can be described as a superposition of a dipole and a quadrupole. Additionally, significant small scale anisotropies in the sidereal reference frame (also unexplained) have been identified and described in data from the 59-string configuration of the detector (?). Similar to Figure 1-6, Figure 1-7 uses an integer number of years of data to isolate the component of interest; however, here we are using an integer number of solar (rather than sidereal) years, such that the solar contribution sums to zero.

One possible explanation of this effect again comes from the work of Compton and Getting, who also anticipated that a dipole structure would be observed due to the motion of the solar system around the center of the galaxy. This would correspond to a 0.35% excess of cosmic rays around a right ascension of 45°, which is not observed (?). The Earth's motion through space is likely more complicated than a simple dipole would suggest, meaning that the proposed galactic-scale Compton-Getting effect is at most one of several factors influencing the observed sidereal distribution.

1.5 IceCube South Pole Neutrino Telescope

The IceCube Neutrino Observatory was designed to identify high-energy neutral particles called neutrinos, but is also a valuable tool for studying the muons produced in cosmic ray air showers. IceCube is located at the Amundsen-Scott South Pole Station, and is currently the only instrument capable of characterizing the cosmic ray distribution of the Southern Hemisphere sky in the TeV range. In its completed configuration, IceCube observes tens of billions of events per year at a rate of about 2,500 events per second. Figure 1-8 is a diagram of the detector and gives a good sense of the scale. The complete configuration of the telescope uses 86 strings of 60 photomultiplier tubes (5,160 sensors in total), placed between 1.5 and 2.5 km beneath the Earth's Surface. The analysis presented in this thesis uses data collected from May 2008 to May 2012. During this time, the detector was still under construction, expanding from 40 strings in 2008 to 86 strings in 2012. Each configuration has a different sensitivity profile and noise characteristics.

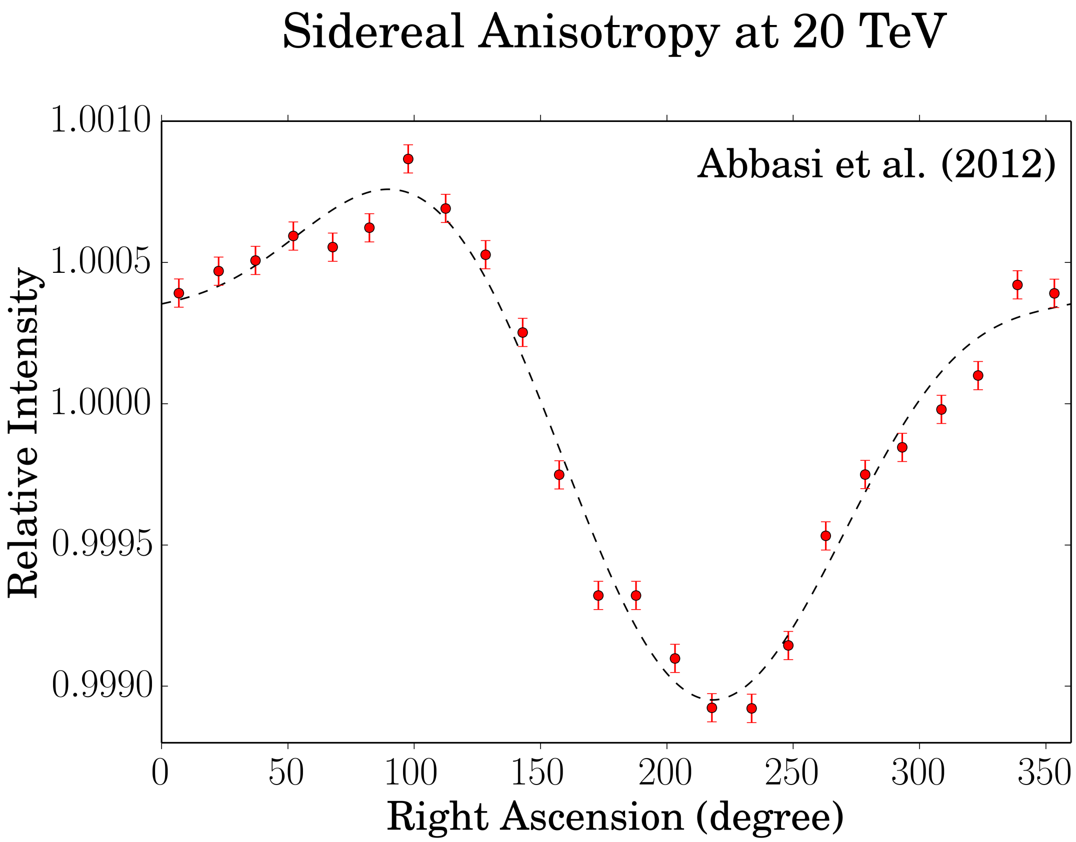

Figure 1-7: Relative intensity of cosmic rays vs. position in a coordinate system that is fixed with respect to the position of the stars. The anisotropy is of unknown origin and is fit here with a combination of dipole and quadrupole terms from a multipole expansion (?).

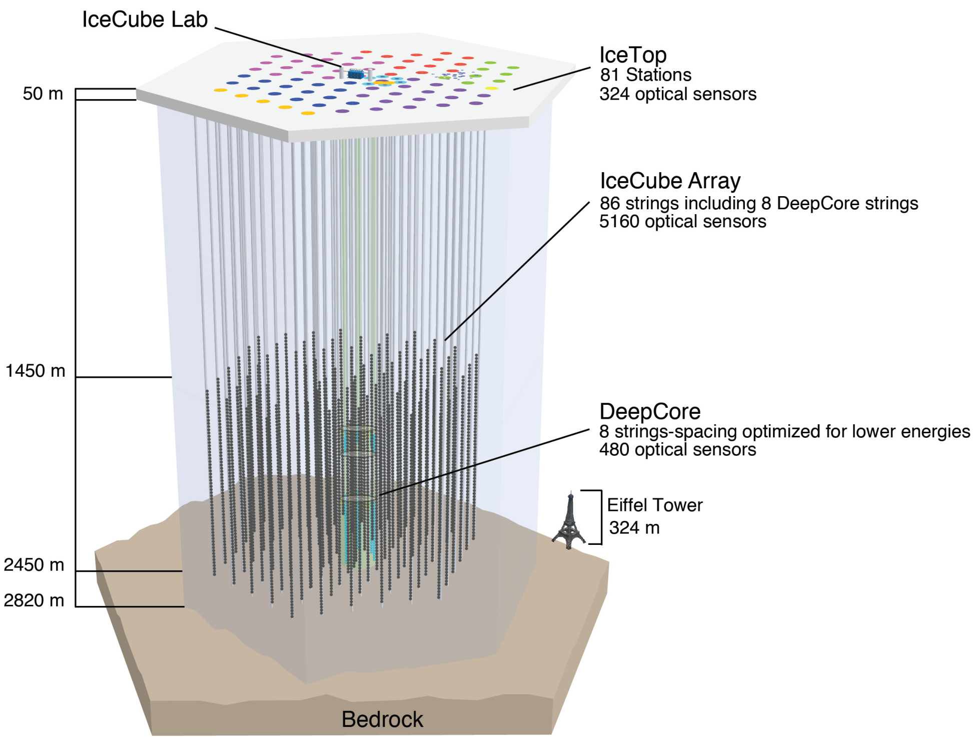

Figure 1-8: The IceCube detector, consisting of 5,160 photomultiplier tubes organized into 86 strings. Colored dots at the top of the strings indicate the year in which the strings were installed. The sensors detect the light produced during interactions between the ice and incoming particles via the Cherenkov effect (see **?**).

Chapter 2

Procedure

2.1 IceCube Detector

2.1.1 Filtering and Reconstruction

The IceCube detector consists of 5,160 photomultiplier tubes (PMTs) buried between 1.5 and 2.5 km below the South Pole in the Antarctic ice sheet. These PMTs detect light produced via the Cherenkov effect. As a charged particle moves through the detector, it will briefly polarize the ice molecules, which then revert to their ground state and release a photon as described in ?. An event is recorded when eight or more contiguous PMTs are triggered within a 5,000 ns time window. Based on the order in which the PMTs are triggered and the strength of the signal, the path of the particle can be reconstructed with a mean angular resolution of 3°. Specifically, the data set used in this analysis has a median cosmic ray parent particle energy of 20 TeV.

2.1.2 Drift Scanning

For this analysis, we isolated eighteen twenty-degree bins in geocentric azimuth and looked only at cosmic ray events with arrival directions within those bins. Each of these regions are stationary with respect to Earth, so that the rotation of the planet would move the entire Southern Hemisphere sky past each region once per day. We then expect a periodic variation in the event rates for each analysis region as the rotation scans through the anisotropic features of the cosmic ray distribution. Our analysis is greatly simplified by the IceCube detector's South Pole location; the sky visible above is constant throughout the year. Figure 2-1 is a diagram of this arrangement, with an example of a twenty-degree region shown in gray on the top of the detector. Figure 2-2 shows the same arrangement from above. This approach allows us to move analysis of the anisotropy from the spatial domain (as described in the introduction) to time series analysis, the techniques of which are already well established in the more conventional astronomical pursuits of variable star and radio observing. Instead of appearing as a spatial component, each of the components of the anisotropy will appear as a characteristic frequency in the spectrum of the event rate time series.

2.2 Event Rates

I calculated hourly event rates for each of the eighteen regions of the detector for five years of IceCube data taken in the 22, 40, 59, 79, and 86 string configurations of the detector. The time distribution of cosmic ray events follows a Poissonian distribution. One convenient property of the Poisson distribution is that the distribution of the time intervals between events is exponential. To calculate an hourly event rate for the detector, I developed code to bin the high-level DST[1] data by hour and calculate the time intervals-between events. I then made a linear, least-squares fit to each distribution. An example of on of these fits is shown in Figure 2-3.

The code was written using the Python program and is included in Appendix B. The slopes calculated in this fit, along with their corresponding uncertainties, χ^2-value, and corresponding hour (stored in Modified Julian Days, see Appendix A).

2.2.1 Computation

The process of calculating event rates involves using all of the available IceCube data for all configurations of the detector from 2008-2012, consisting of several terabytes of data

---

[1] Data Storage Format (DST) refers to the format in which data collected at the South Pole are packaged before transfer to the North. I worked with a reduced version of this data, stored in the common ROOT format.

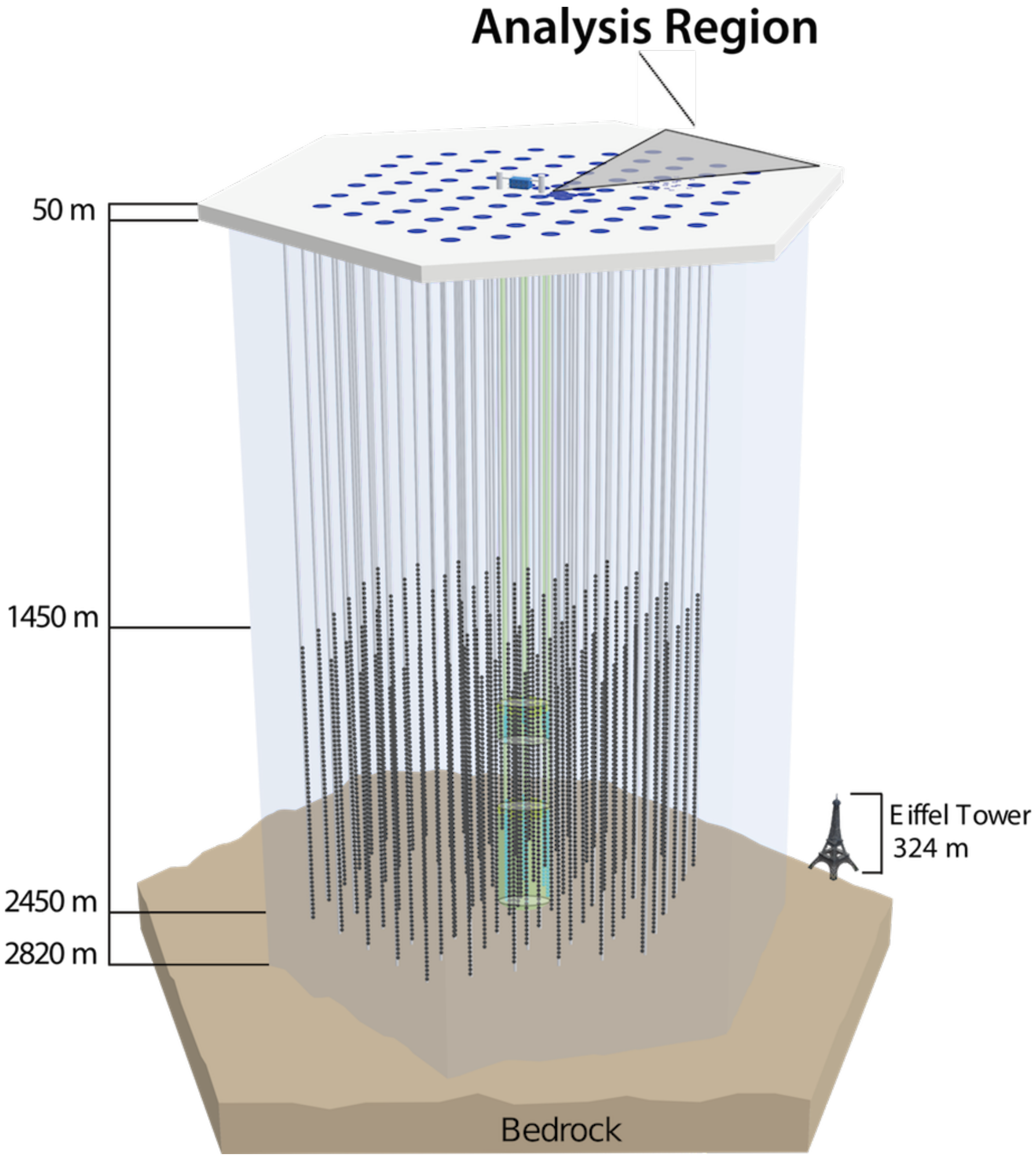

Figure 2-1: The IceCube South Pole Telescope with twenty-degree azimuthal bin outlined in gray. Each region is stationary with respect to the Earth. Every solar day, the rotation of the planet scans the entire Southern Hemisphere sky past each region. Anisotropic features of the distribution can then be detected as periodic signals in the regional event rate.

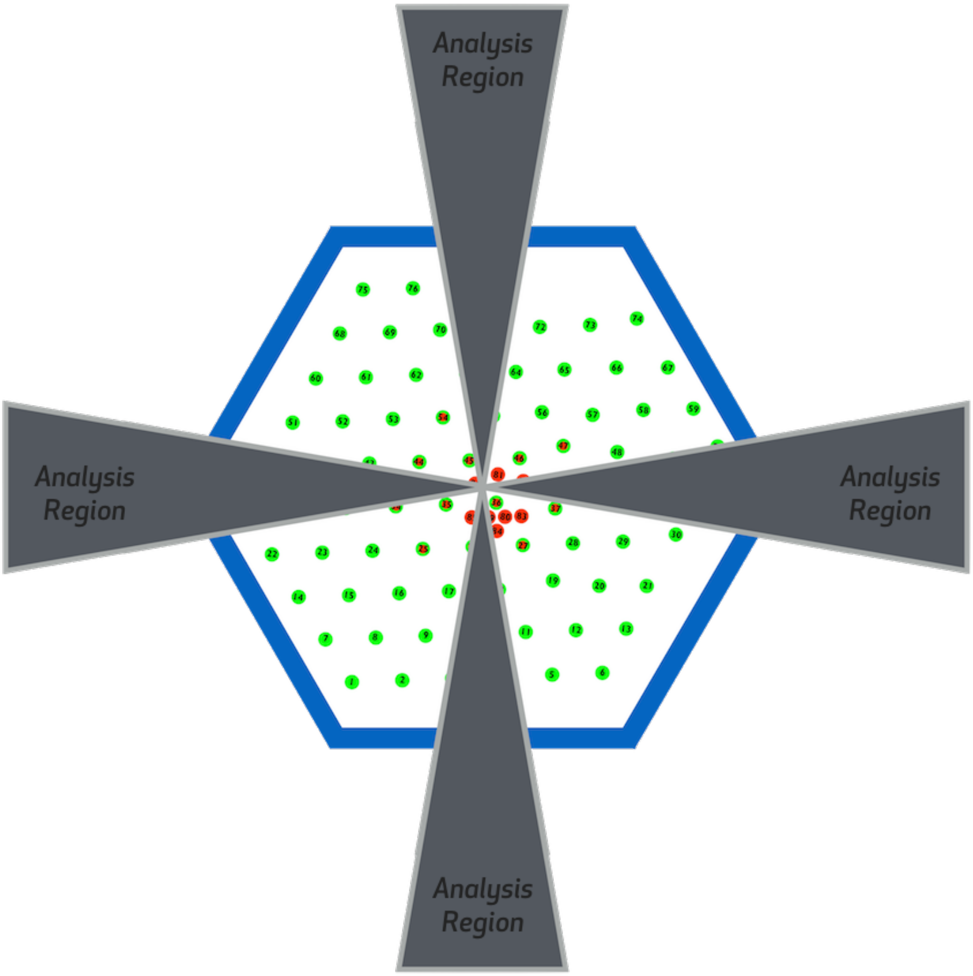

Figure 2-2: IceCube Telescope from above, showing four of the eighteen analysis regions. Four of the eighteen analysis regions are shown in gray. Note the uneven distribution of strings (indicated by numbered green dots) within each region. Strings represented by red dots are part of DeepCore, an IceCube extension that is not used in this analysis.

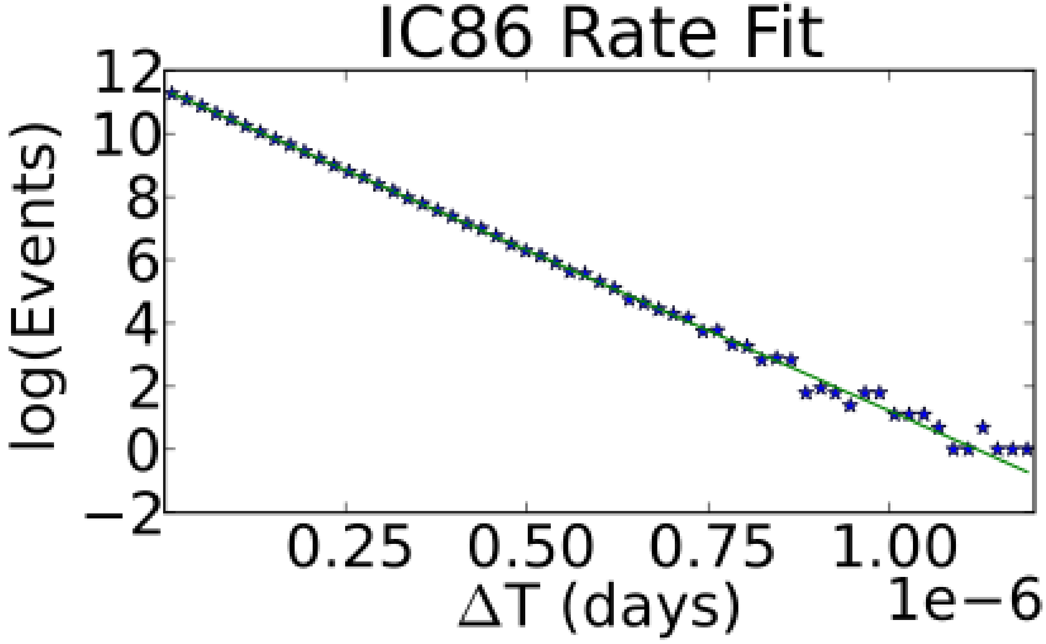

Figure 2-3: Natural log of the number of events vs. time interval between events. A histogram of the number of events vs. time is a Poisson distribution, which has that the distribution of the time intervals between events is exponential. The slope of a linear fit to the log of the exponential can then be interpreted as an event rate.

and 1.7 × 10^{11} events. It is a computationally intensive process. Data from the pipeline are stored in a binary format (used by the ROOT data analysis tool) and organized by day. Execution time for the rate calculation script depended on the configuration of the detector being considered; the 22-string configuration had a much lower sensitivity, and fewer events (mean event rate of 500 Hz), so one day of data for one region could be processed in roughly 15 minutes, whereas the current – and complete – 86-string configuration (mean event rate of 2,500 Hz) took up to two hours. All computation was done using the high performance computing cluster (NPX4) belonging to the Wisconsin IceCube Particle Astrophysics Center, which consists of 512 nodes, each with performance capabilities comparable to a modern desktop machine (64-bit, XX GHz, 2GB RAM). In total, calculating all of the necessary rates used approximately 30,000 hours of CPU time.

2.2.2 Normalization

My analysis used data from multiple configurations of the IceCube telescope, which was still under construction while these data were being collected. As the detector expanded, the noise rates increased and the noise in the signal decreased. Additionally, due to the hexagonal detector geometry of IceCube, the sensitivity of the regions varies due to an uneven coverage of photomultiplier tubes in all azimuthal directions. To allow the years to be combined into one data set without introducing low frequency noise into the spectrum, each of the years (meaning, each stage in the construction of the detector) was normalized to a variance of 1 and mean-subtracted to adjust for the increasing event rate in with every expansion.

2.2.3 Concatenation

Each region of the detector will observe the same time series, offset by some time interval as determined below. Because each region constitutes an independent measurement, we can combine them to better constrain the periodic variations observed. To do this, we add a constant to the time stamps of each region such that the time series from each region differ by an integer number of solar days. The constant added to each region is calculated as

$$dY = n * 365.24219647 - 6.24 \times 10^{-6} T, \qquad (2.1)$$

where dY is the time offset, n is the number of years of data used, and T is the number of Julian centuries since the beginning of the epoch.

2.2.4 Filtering

Analysis of the time series was done using the Lomb-Scargle algorithm for approximating a frequency spectrum. This method is equivalent to a least-squares fit, which is necessary due to the gaps and unequal spacing present in the data. The algorithm also has some convenient statistical properties. Namely, when normalized properly, the power values in the normalized Lomb-Scargle periodogram are exponentially distributed, allowing us to calculate the probability that a given peak arises from a random gaussian fluctuation. For more information regarding the algorithm, please see Appendix C.

However, the Lomb-Scargle algorithm is only an approximation; if lacks the orthogonal decomposition found with a Fourier transform. If the periodicities found in a series differ in amplitude by too great a factor, the largest effect will saturate the periodogram and make other periodicities undetectable. We found the spectrum of event rates to be completely dominated by the 8% modulation associated with the seasonal variation, so we applied a Butterworth notch filter to our data prior to calculating the periodogram so that the 10^{-3} amplitude oscillations associated with the anisotropy could be detected. The Butterworth filter was chosen do to its rapid roll-off and flat The notch parameters were defined to be arbitrary at the high end (100 oscillations per day - which would not be detectable with the hourly sampling rate we used) and slightly higher than the highest frequency contribution in the spectrum of the seasonal variation (1 oscillation every 2.2 days).

Chapter 3

Results

We calculated hourly cosmic ray event rates for eighteen twenty-degree regions of the IceCube detector using three years of data. The rotation of the Earth causes the anisotropic features of the Southern Hemisphere cosmic ray distribution to move past each region periodically, creating a periodic time series in the event rates. The rates from each region were concatenated to one another to produce an extended time series based on the independent measurements from each region of the detector. A Butterworth notch filter was applied to the data and a 3σ cut was applied the filtered event rates to eliminate outliers. Figure 4-1 shows the normalized event rates vs. time and a frequency spectrum before and after the filter was applied. Note that the high-frequency periodicities associated with the anisotropy are completely blown out without filtering.

Figure 3-2 shows the Lomb-Scargle periodogram of the concatenated and filtered event rate time series. We observe four peaks at at least the 5σ significance level in Figure 3-2, corresponding to frequencies of approximately 0.997, 1.00, 1.0027, 1.0054 and 2.0054 day^{-1} (reported peak frequencies correspond to the maximum value within a range of significant values). There are also peaks present in the periodogram which correspond to modulations of the other peaks over the span of the number of years of data used.

The Lomb-Scargle algorithm used here is normalized as described in **?** such that the the calculated power is proportional to the probability of making a type I error (a "false alarm" probability for identifying a frequency that is not actually present). Power maps

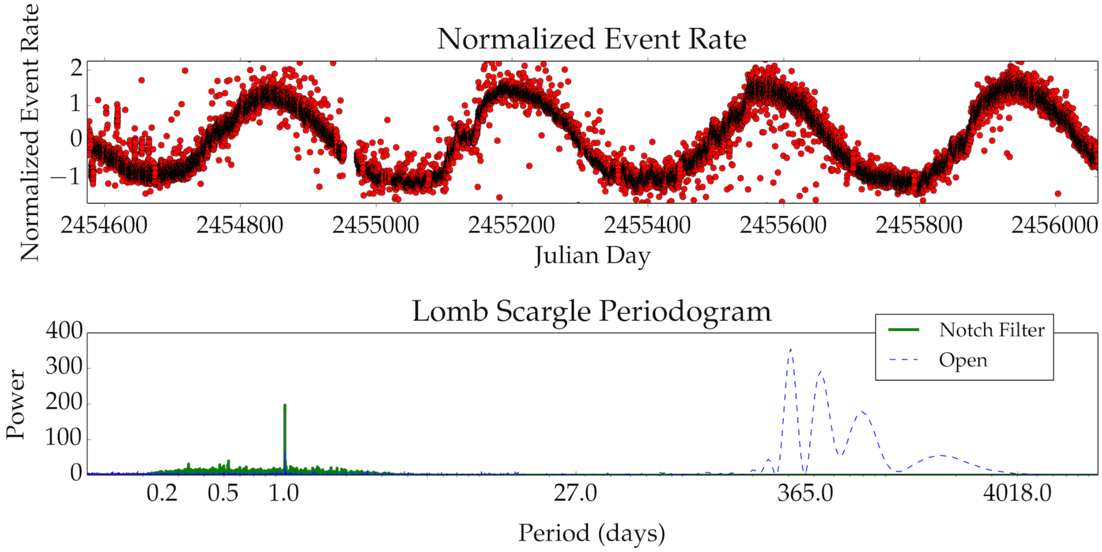

Figure 3-1: **Top:** Normalized event rates vs. time. The 8% modulation of the seasonal variation dominates the spectrum. The three weeks of missing data around Julian Day 2454900 are caused by downtime during the construction process.**Bottom:** Approximated frequency spectrum before and after the application of a Butterworth notch filter. Low frequency contributions are removed with filtering, allowing greater resolution in the higher frequencies.

to probability as

$$P(>z) = 1 - (1 - e^{-z})^{M}, \tag{3.1}$$

where z is the p-value of a type I error and M is the number of independently sampled frequencies (**?**). For more information regarding the Lomb-Scargle algorithm, see Appendix C.

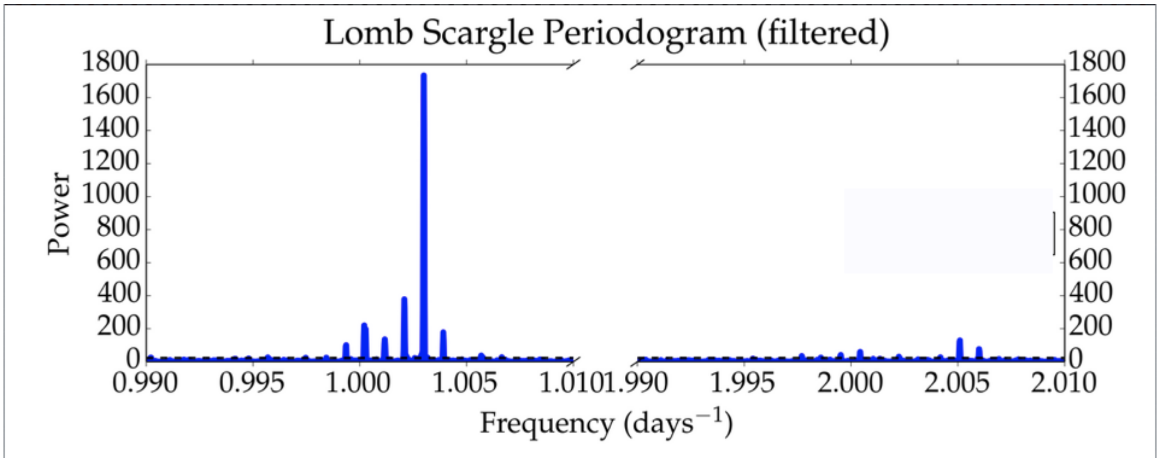

Figure 3-2: Approximated frequency spectrum for a time series produced by concatenating data from all eighteen analysis regions. Peaks are present at a minimum 5σ level for frequencies of 0.997, 1.00, 1.0027, 1.0054 and 2.0054 day^{-1}. a power of 72 corresponds to a significance value of 5σ.

Chapter 4

Discussion

4.1 Peaks Corresponding to the Anisotropy

We used the Lomb-Scargle algorithm to approximate a frequency spectrum for event rates with 20 degree azimuthal bins of the IceCube detector. We observed peaks in the frequency spectrum corresponding to frequencies of 1.00, 1.0027, and 2.0054 day^{-1}. These frequencies are interpreted as evidence of anisotropic features moving past each region with the rotation of the earth. The 1.0027 day^{-1} frequency corresponds to a feature moving past each region once per sidereal day. The 2.0054 day^{-1} frequency similarly corresponds to a frequency of twice per sidereal day. These two sidereal peaks are consistent with the quadrupole nature of the sidereal anisotropy as reported in **?**. The solar dipole is visible in the 1.00 day^{-1} peak.

4.2 Peaks Corresponding to Modulations of the Anisotropy

We also detected peaks in the periodogram corresponding to frequencies of 0.997 and 1.0054 day^{-1}. These peaks would correspond to long-term modulations of the components of the anisotropy. For the peak at 1.0054 day^{-1}, we can report a modulation of the sidereal anisotropy over the span of one solar year. Likewise, the peak at 0.997 day^{-1} indicates a modulation of the solar anisotropy over the span of one sidereal year. However, the power calculated in the periodogram is proportional to the square of the amplitude,[1] so we can conclude that while these long term modulations are present in the data, they appear to represent small modulations.

4.3 Spurious Peaks

Additionally, there are many other significant peaks present in the spectrum that cannot correspond to any physical phenomenon. These peaks represent modulations of the components of the anisotropy (and modulations of the modulations) over the course of the number of years of data used. We can dismiss these peaks from having a physical interpretation because if we are using only three years worth of data and thus cannot observe a modulation with a period longer than 1.5 years. These peaks must be an artifact of the analysis.

We are aware of two possibilities for the origins of these spurious peaks, both of which are related to varying noise levels. First, the IceCube detector was undergoing construction while these data were collected, and each configuration of the detector has a different noise level due to increasing size and the telescope apparatus itself going through radioactive decay and stabilizing over time. When the data from the eighteen analysis regions are concatenated end-to-end, there is thus a repeating pattern of decreasing noise levels, which could introduce a low-intensity and low-frequency contribution to the approximated periodogram that could not be filtered out by conventional means. Second, the telescope is not equally sensitive in all directions; there is an azimuthal sensitivity due to the unequal coverage of photomultiplier tubes. Figure **??** shows the distribution of photomultiplier tubes in the telescope and the corresponding relative sensitivity of the detector as a function of azimuth.

---

[1] For example, if the solar anisotropy is one-third the intensity of the sidereal anisotropy, we would expect to see the power corresponding to the sidereal peak to be nine times the power corresponding to the solar peak.

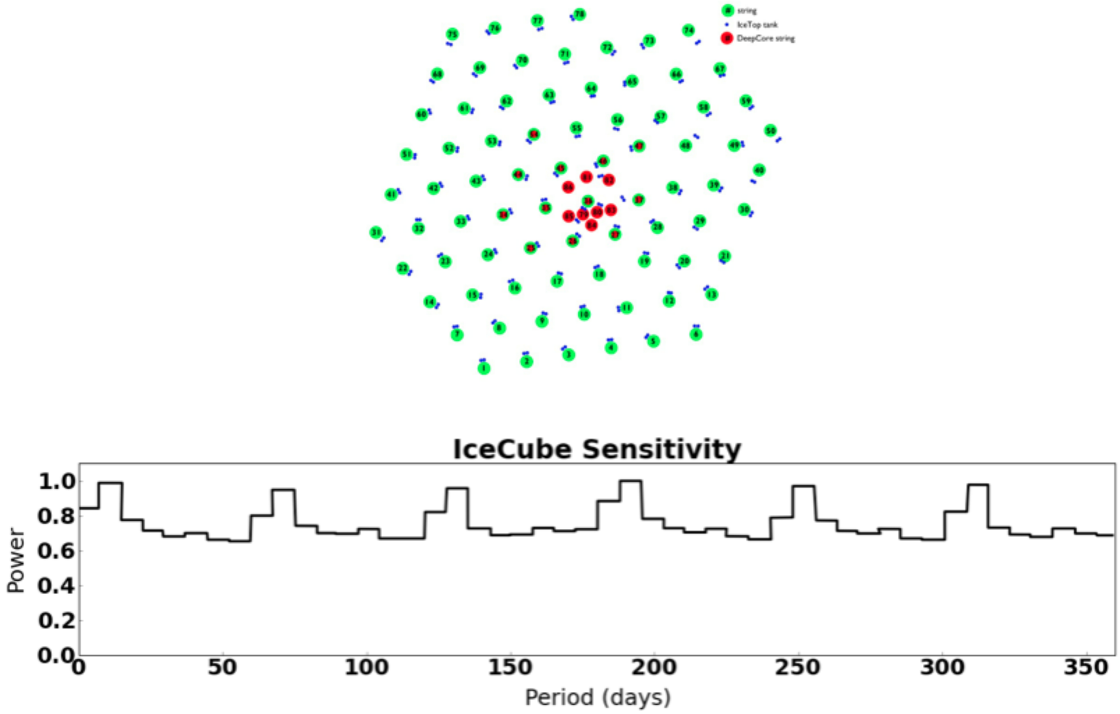

Figure 4-1: **Top:** Distribution of photomultiplier tubes in the telescope as viewed from above. **Bottom:** The number of sensors visible depends on the path taken through the detector, thus there is an azimuthal dependence for the relative sensitivity of the detector.

Appendix A

Overview of Units

A.1 Cosmic Ray Particle Energy

The analysis presented within this thesis uses the electron volt (eV) as the unit for cosmic ray particle energy. While this unit is potentially alien to some readers, the idea behind it is deceptively simple: 1 eV is the energy of an electron that has been accelerated by a 1-volt potential difference. For a sense of the scale of this unit, please refer to table **??**. An electron volt is a small unit, equivalent to $4.45 \times 10^{-26}\,\mathrm{kW \cdot h}$ or The median cosmic ray particle energy in this analysis is 20 TeV (tera-electron volts, or 10^{12} eV), which is comparable to the kinetic energy of a housefly.[1]

A.2 Julian Day

---

[1] Assuming a mass of 0.01 g and a speed of 3 $\frac{\mathrm{km}}{\mathrm{hr}}$.

Appendix B

Code

B.1 Event Rate Calculation

```
#!/usr/bin/env python

"""
    Get rates and dT distribution from a simple_dst .root file.

    This code reads in a DST file or chain of DST files containing
    particleevent records and calculates rates (in Hz) for events
    within the azimuthal range specified by the tuple range.

INPUTS:
    lowLim -- The lower limit on azimuth for the region.
    upLim  -- The upper limit on azimuth for the region.
    files  -- A file or list of .root files containing the
                CutDST structure.

FLAGS:
    -v     -- Include delta-time distribution in output pickle
                file (very large)
    -f     -- Include complete fit data (bin centers, log of
                observed counts, and fit values) in output
                pickle file.

OUTPUTS:
    A pickle file containing the rates and statistics for the
    four 20 degree regions for one day. This will contain modified
    julian dates with hourly rates, uncertainty in the rates, and
    the chi-square values used to arrive at those rates by default.
    With option -v, the file will also contain the exponential
    delta-time distribution for all of event in each hourly bin.
    Note that this is many megabytes of data and will be very
    resource intensive to open and manipulate. With option -f, the
    file will also contain the fit values, bin centers and the logs
```

```
    of the hourly event counts – this is probably what you want to
    use if you're examining the fits.


ALGORITHM:
    The rates are determined by a linear fit to the logarithm of
    the distribution the times between events.

USAGE:
    Typically called via HTCondor as described in cluster submit
    file. For testing, use:
        % ./regRate.py 0 20 /in/put/file /out/put/file

AUTHORS:
    Created by hank corbett (2013:6:28) : hcorbett391@gmail.com

    Modified by hank corbett (2014:1:30):
            — Linear fits are now calculated using np.polyfit()
            — Azimuth bin edges now come from arguments lowLim
                and upLim
            — dT histogram bins are now determined by the
                statistics of the  dT distribution (fits made
                from 0 to 3*std(dT)
            — Results are now stored in the more standardized
                pickle format
"""

try:
    from ROOT import TChain
    import argparse
    import os.path
    import numpy as np
    import scipy.stats as stats
    import pickle
except ImportError, e:
    print e
    raise SystemExit

parser = argparse.ArgumentParser(description=
    "Azimuth␣Limited␣Hourly␣Rate␣Calculator␣for␣the␣IceCube␣SPNT")
parser.add_argument("-v", action='store_true',
        help="Store␣dT␣distributions␣in␣jsons.")
parser.add_argument("-f", action='store_true',
        help="Store␣fit␣info␣in␣json␣file.")

parser.add_argument("lowLim", type=int,
        help="Lower␣azimuthal␣limit.")
parser.add_argument("upLim", type=int,
        help="Upper␣azimuthal␣limit.")
parser.add_argument("file",
        help="A␣simplified␣DST␣file␣(CutDST)␣in␣ROOT␣format")
parser.add_argument("outfile", help="Output␣file")
args = parser.parse_args()
```

```
lowLim = args.lowLim
upLim = args.upLim
file = args.file
outfile = args.outfile

chain = TChain("CutDST")
chain.Add(file)

unBinnedTimes = []
hours = []
rates = []
sRates = []
rChi = []
region = []
if args.v:
    dTDists = []
if args.f:
    centers = []
    counts = []
    fits = []

for event in chain:
    if (event.LLHAzimuthDeg > lowLim \
      and event.LLHAzimuthDeg < upLim):
        unBinnedTimes.append(event.ModJulDay)


unBinnedTimes = np.array(unBinnedTimes)
unBinnedTimes = np.sort(unBinnedTimes)

# Define hourly bins for the range of MJDs covered in the input files
bins = np.arange(np.floor(unBinnedTimes[0]),
                 np.ceil(unBinnedTimes[-1]), (1./24.))

# Assign each event to a bin
bIndices = np.digitize(unBinnedTimes, bins, right=False)

for hour in np.arange(1,len(bins)+1):
    hours.append(bins[hour-1])



    dT = []
    eventTimes = unBinnedTimes[np.where(bIndices == hour)]
    #for i in range(len(eventTimes)-1):
        #dT.append(eventTimes[i+1]-eventTimes[i])
    dT = [eventTimes[i+1]-eventTimes[i] for i in range(len(eventTimes)-1)]
    # Avoid empty bins skewing the fit — interval distribution up to 3sigma
    upperBinLimDT = np.mean(dT)+3*np.std(dT)
    binEdgesDT = np.linspace(0, upperBinLimDT, 10)
    countsDT, binsDT = np.histogram(dT, bins=binEdgesDT)
    binCentersDT = 0.5*(binsDT[1:]+binsDT[:-1])
    logCountsDT = np.where(countsDT > 0, np.log(countsDT), np.nan)

    fitParams, cov = np.polyfit(binCentersDT, logCountsDT, 1, cov=True)
```

```
    fit = fitParams[0]*binCentersDT + fitParams[1]
    var = np.diag(cov)
    r_chi = stats.chisquare(logCountsDT, fit, ddof=2)

    rates.append(-fitParams[0]/86400)
    sRates.append(var[0]/86400)
    rChi.append(r_chi)
    if args.v:
        dTDists.append(dT)
    if args.f:
        centers.append(binCentersDT.tolist())
        counts.append(logCountsDT.tolist())
        fits.append(fit.tolist())



if args.v:
    region = [hours, rates, sRates, rChi, dTDists]
if args.f:
    region = [hours, rates, sRates, rChi, centers, counts, fits]
else:
    region = [hours, rates, sRates, rChi]

filename = outfile + "/MJD"+str(int(np.floor(unBinnedTimes[0])))+"_reg_" \
                   + str(lowLim)+"_"+str(upLim)+".p"

pickle.dump( region, open(filename, "wb"))
```

B.2 Data Normalization and Filtering

B.3 Lomb Scargle Periodogram

Appendix C

The Lomb-Scargle Periodogram

The Lomb-Scargle periodogram is a tool for approximating a frequency spectrum from a time series and is in effect similar to a simple Fourier transform. However, the Lomb-Scargle periodogram differs from the Fourier transform in that it does not represent an exact spectral decomposition (i.e., the sampled frequencies are not necessarily orthogonal and the periodogram is not invertible). The periodogram is based on a least-squares fit of sinusoidal components to the given data, which makes the algorithm well-suited to analyzing unevenly-sampled data with large gaps, for which Fourier analysis would artificially introduce high levels of low-frequency noise. The algorithm is implemented as follows:

$$P_N = \frac{1}{2\sigma^2}\left\{\frac{[\Sigma_j(h_j-\bar{h})\cos(t_j-\tau)]^2}{\Sigma_j\cos^2\omega(t_j-\tau)}+\frac{[\Sigma_j(h_j-\bar{h})\cos(t_j-\tau)]^2}{\Sigma_j\cos^2\omega(t_j-\tau)}\right\}, \tag{C.1}$$

where $h(t)$ is a time series, ω is a sampled frequency in the spectrum, and τ is a time offset defined by the relationship

$$\tan(2\omega\tau)=\frac{\Sigma_j\sin 2\omega t_j}{\Sigma_j\cos 2\omega t_j}. \tag{C.2}$$

To calculate a complete periodogram, this algorithm is then calculated independently for each frequency in the range of frequencies in the range of interest. Further information about the formulation of the periodogram can be found in **?**. Specifically, my analysis used the faster, but equivalent, formulation of the algorithm described in **??**.

In addition to its utility with unevenly sampled data, the periodogram also has statistical properties that make it ideal for our investigation of periodicities in IceCube event rates. The periodogram power PN corresponds to a significance statistic that can be converted to a false-alarm probability as

$$P(>z)=1-(1-e^{-z})^M, \tag{C.3}$$

where z is the p-value of a type I error and M is the number of independently sampled frequencies (?). Using these significance statistics, we can calculate the probability that any given peak in the spectrum is a random fluctuation resulting from gaussian noise.