Measuring the Cosmic-Ray Energy Spectrum, Composition, and Anisotropy at PeV Scales Using the IceCube Observatory

by

James Bourbeau

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Physics)

at the

University of Wisconsin – Madison

2019

Defended on April 26, 2019

Dissertation approved by the following members of the Final Oral Committee:

Justin Vandenbroucke · Assistant Professor, Physics

Francis Halzen · Professor, Physics

Ellen Zweibel · Professor, Astronomy and Physics

Paolo Desiati · Senior Scientist, Physics

Abstract

Since the discovery of cosmic rays over one hundred years ago, many experiments have studied their properties. However, a definitive answer to the questions of where cosmic rays originate and how they are produced is still not known. Over the last several decades, a much more detailed understanding of high energy cosmic rays has begun to materialize. In particular, the cosmic-ray energy spectrum, with its transitions at 3 PeV (the "knee") and 3 EeV (the "ankle"), has been extensively investigated. Based on magnetic confinement arguments, it's generally believed that the energy range between the knee and ankle is where the transition from galactic to extragalactic sources of cosmic rays. The ability to distinguish between high energy cosmic rays of different composition and study the relative mass abundances of cosmic rays in this transition region can provide invaluable insight in answering the open questions surrounding the origins of cosmic rays.

This work focuses on measuring the composition-resolved cosmic-ray energy spectrum at and above the all-particle knee using one year of data collected by the IceCube Observatory. Sepcifically, we focus on making a two mass group spectrum measurement from $10^{6.4}$ GeV to $10^{7.8}$ GeV. The first mass group, referred to as the "light" mass group, is modeled using proton and helium cosmic rays, while the second, "heavy" mass group, is modeled using oxygen and iron cosmic rays. We observe a clear softening of the light spectrum near 3 PeV, while the energy spectrum for the heavy mass group follows a power-law like structure with a spectral index of ~ 2.7 throughout the entire energy range considered. The observed transition from a primarily light to a heavy-dominant spectrum takes place near $10^{7.1}$ GeV. This feature is characteristic of a potential rigidity-dependent cutoff, or Peters cycle. The change in relative mass abundance could also indicate a possible transition in the source population of cosmic rays.

In addition, a study to determine whether or not the light, heavy, or all-particle cosmic-ray energy spectra vary as a function of arrival direction is also presented. This marks the first time an analysis of this kind has been conducted using the IceCube Observatory. No statistically significant spectrum deviations were observed. The results from this analysis can be used to set a limit on the range of possible spectral deviations.

Acknowledgements

The work presented in this thesis would not be possible without the help of many people. First and foremost, I'd like to thank my wife Caitlin for her support, patience, and companionship. To the friends I've made during my time in graduate school, while some are moving away and some are staying the Madison, I look forward to many years of friendship and enjoyment together. I'd like to thank the entire WIPAC community for providing such an engaging and welcoming environment to work in. Being a member of WIPAC has made for a unique graduate school experience, and I'm all the better for it. In particular, this work would not have been possible without the help of Josh Wood, Paolo Desiati, Zach Griffith, and Justin Vandenbroucke. Thank you all. Lastly, I'd like to thank Stefan Westerhoff for his mentorship and guidance both in physics and in life. You are missed.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This thesis presents an analysis using data collected by the IceCube Observatory to study the cosmic-ray energy spectrum and composition at and above the all-particle spectrum knee. Specifically, a two-mass-group flux measurement from $6.4 \leq \log_{10}(E/\text{GeV}) \leq 7.8$ is presented. This energy region is particularly interesting as it's generally believed to mark the beginning of the transition from Galactic to extragalactic sources of cosmic rays. The two mass groups, referred to as "light" and "heavy", are modeled by proton/helium and oxygen/iron cosmic-ray primaries, respectively. In addition to a robust two-component flux measurement, a search for variations of the energy spectrum as a function of arrival direction is conducted for both the light and heavy mass groups, as well as for the all-particle spectrum. This provides information about any potential local sources of cosmic rays and, because the observed arrival direction is influenced by magnetic fields, the environment through which they propagate.

This chapter begins with a brief historical overview of cosmic-ray physics, then summarizes experimental results and our current understanding of the cosmic-ray energy spectrum and mass composition, followed by a discussion of extensive air showers. In Chapter 2, we present an overview of the detector components of the IceCube Observatory and the low-level signals recorded by the detector. Then in Chapter 3 we discuss the reconstruction algorithms used to map the signals measured by the detector to high-level parameters that are characteristic of primary cosmic ray particles. Chapter 4 focuses on using the reconstructed parameters from Chapter 3 to determine

a primary particle composition and energy for each air shower detected with IceCube. Next, an unfolding technique used to correct for the known detector response is presented in Chapter 5. Finally, the results for the light, heavy, and all-particle flux measurement are presented in Chapter 6, while the study of spectral variations as a function of arrival direction are shown in Chapter 7.

It should be noted that while previous and current work done by the IceCube collaboration [1] have focused on maximizing the number of composition groups analyzed, this work has a slightly different motivation. Here, we are concerned not only with producing a precise cosmic-ray composition measurement, but also studying how the energy spectrum varies as a function of sky position. To this end, in order to maximize the available statistics, while still providing composition separation, two mass groups were used.

1.1 Background

The history of studying cosmic rays dates back well over a hundred years. The discovery of ionizing radiation and the advent of the electroscope prompted the observation of an ambient level of background radiation. This radiation was originally thought to be produced only by known sources of radiation in the environment (e.g. radon). However, in a series of hot air balloon flights in 1911 and 1912, Victor Hess used electroscopes to measure the ambient radiation intensity as a function of altitude [2]. His findings, shown in Figure 1.1, demonstrated that while radiation levels initially decreased until an altitude of roughly 1 km was reached, the level of ionizing radiation then began to continually increase with altitude. This indicated that there was also a cosmic radiation component originating beyond Earth's atmosphere in addition to terrestrial radiation. Higher altitude balloon experiments in 1913 and 1914 by Werner Kolhörster confirmed Hess' findings [3]. Hess' balloon experiments are regarded as the discovery flight of cosmic rays and he was awarded the Nobel Prize in Physics in 1936 for his work.

Robert Millikan, who coined the term "cosmic rays", believed that cosmic rays were gamma rays, while others believed the radiation was composed of charged particles. In 1929, Walther Bothe and Kolhörster, using a coincident detection technique with two Geiger-Müller tubes, discovered that cosmic rays are indeed primarily charged particles [4], although the name cosmic *rays* remained. Then, in 1930 Bruno Rossi predicted an anisotropy in the intensity of cosmic rays which depended on their charge. The "East-West effect" was measured in the same year, providing evidence that cosmic rays are mostly positively charged particles.

In 1934, Rossi observed coincident detection of cosmic rays between detectors that were widely separated from one another [5]. Then, in 1939, Pierre Auger discovered large, extensive collections of correlated particles, called air showers, with estimated energies up to 10^{15} eV [6]. These showers were in fact secondary particles from highly energetic primary particles interacting in Earth's atmosphere. Ever since the discovery of extensive air showers, scientists have constructed detectors with larger and larger areas, probing incredibly high energy scales.

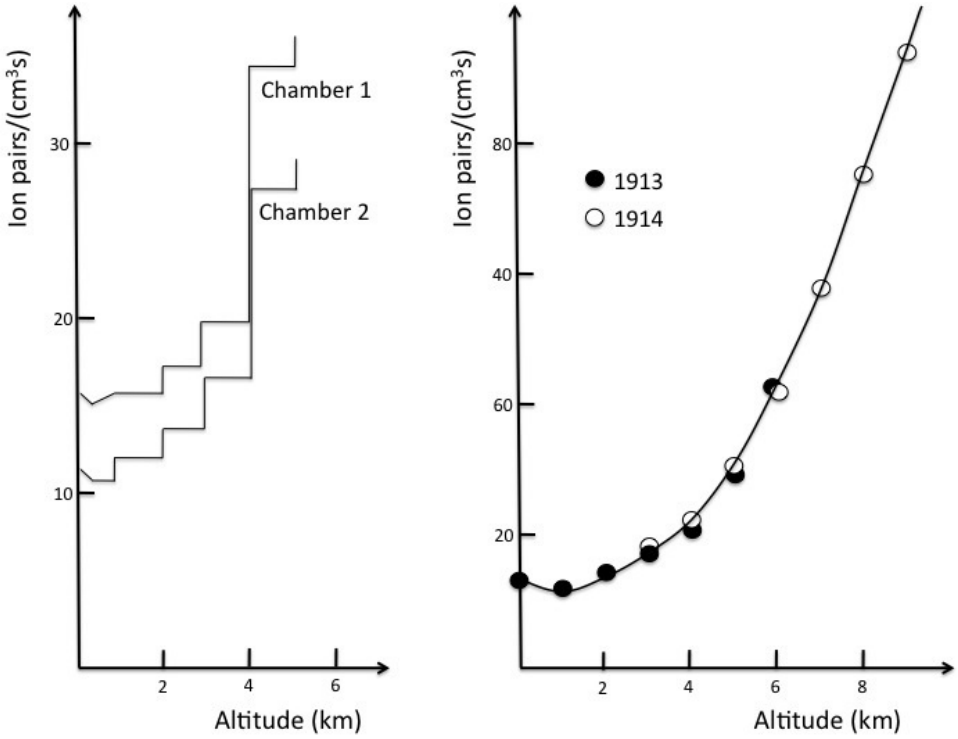


Figure 1.1: Ionization rate as a function of altitude measurements by (left) Hess in 1912 [2] and (right) Kolhörster from 1913-1914 [3]. The rise in rate with increasing altitude is attributed to the extraterrestrial origin of the penetrating radiation.

1.2 Energy Spectrum

Scientists have studied cosmic rays over an incredibly large energy range. The highest energy cosmic ray ever recorded was detected in 1991 by the Fly's Eye detector in Utah. The detected event had an estimated energy of 3.2×10^{20} eV [7], orders of magnitude larger than the highest man-made energies reached at the Large Hadron Collider (LHC).

The cosmic-ray energy spectrum, defined to be the number of particles at a given energy E per unit area, time, solid angle, and energy

$$J(E) = \frac{dN}{dE\ dA\ d\Omega\ dt} \tag{1.1}$$

has been measured over ten orders of magnitude in particle energy. Figure 1.2 shows the cosmic-ray energy spectrum as measured by several different experiments. One of the most remarkable

characteristics of the cosmic ray spectrum is that it is *approximately* governed by a power law:

$$J(E) \propto E^{-\gamma} \tag{1.2}$$

where E is the cosmic-ray energy and γ is the spectral index. However, there are a couple of notable transitions in the spectrum where the spectral index changes.

$$\gamma = \begin{cases} 2.7 & \text{for } E < 3 \times 10^{15}\text{eV} \\ 3.1 & \text{for } 3 \times 10^{15}\text{eV} < E < 3 \times 10^{18}\text{eV} \\ 2.6 & \text{for } E > 3 \times 10^{18}\text{eV} \end{cases}$$

The first of these transitions occurs at roughly 3×10^{15} eV where there is a slight steepening of the spectrum and the spectral index changes from ~ 2.7 to ~ 3.1 [8]. This downward bend is fittingly referred to as the "knee" of the cosmic-ray spectrum. The second transition occurs near 3×10^{18} eV where there is a hardening in the spectrum and the spectral index transitions from ~ 3.1 to ~ 2.6. In keeping with the same analogy, this feature is referred to as the "ankle" of the cosmic ray spectrum.

It's worth noting that there are hints of additional spectral index changes between the knee and ankle. Figure 1.3 shows the cosmic-ray energy spectrum which has been scaled by a factor of $E^{2.6}$ to compress the spectrum and emphasize changes in the spectral slope. A faint spectral hardening at approximately 10^{17} eV, the so-called "second knee", has been observed by both the KASCADE-Grande [9] and IceTop [10] experiments. However, additional high-precision measurements are needed for further study of these finer structures.

Cosmic rays, as they move through magnetic fields throughout the universe, are subject to deflections by the Lorentz force. The gyroradius, or Larmor radius, for a charged particles in a magnetic field is given by:

$$R_L = \frac{E}{c|Z|eB} = 1.1 \times 10^{-3}\text{pc}\Big(\frac{E}{\text{TeV}}\Big)\frac{1}{|Z|}\Big(\frac{B}{\mu G}\Big)^{-1} \tag{1.3}$$

where E is the energy of the particle, Z is the charge of the particle in units of e, c is the speed of light in vacuum, and B is the magnetic field strength. Additionally, the *rigidity* $R = R_L Bc$ of a

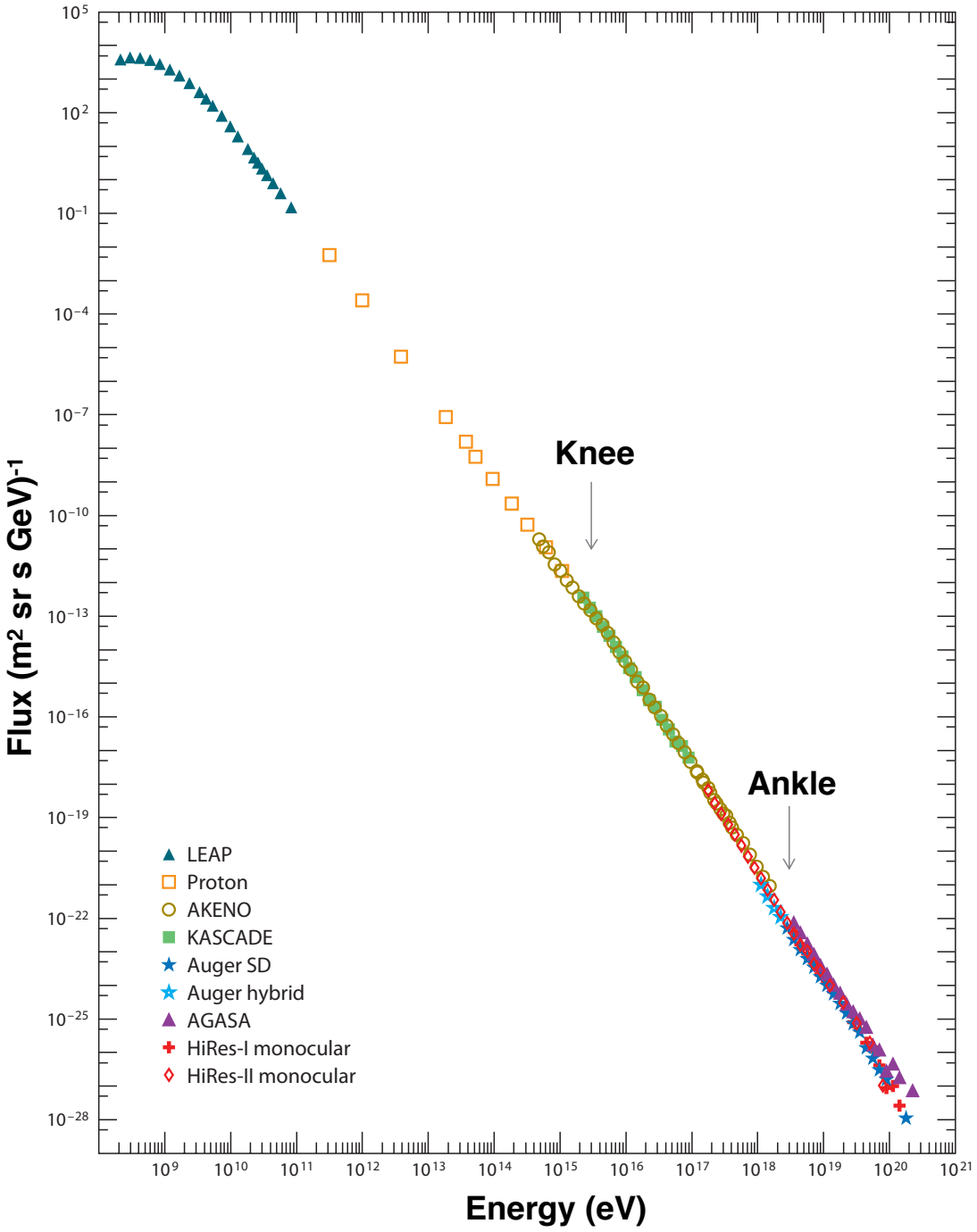


Figure 1.2: Overview of the cosmic-ray energy spectrum. Transitions in the power law spectral index at 3×10^{15} eV (the knee) and 3×10^{18} eV (the ankle) are indicated by arrows. Figure from [11].

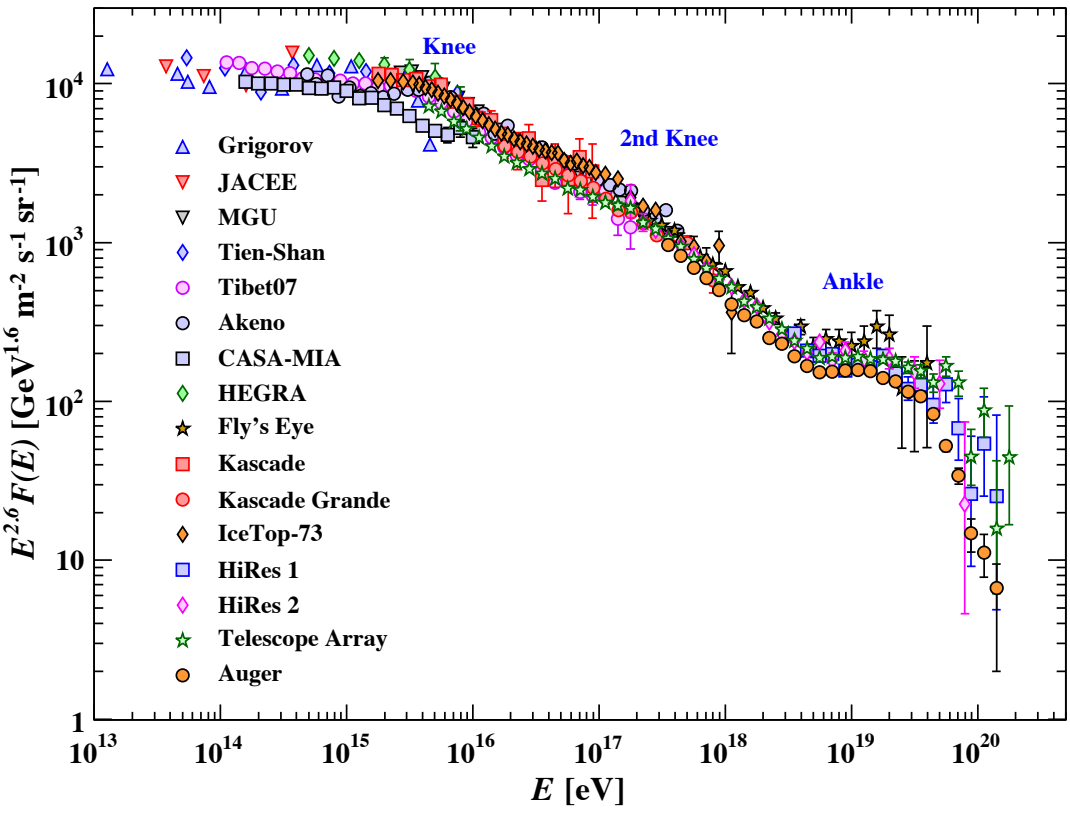


Figure 1.3: Weighted energy spectrum to compress spectrum and highlight spectrum differences. Figure from [12].

particle is often used when describing the characteristics of cosmic rays. The Milky Way is ~ 300 pc thick and is permeated with a magnetic field $B \sim 3$ μG [13]. Particles with PeV scale energies around the knee have a Larmor radius $R_L(E = \text{PeV}) \sim 0.3$ pc and can be magnetically confined within our galaxy. While particles at EeV scale energies near the ankle have a much larger Larmor radius $R_L(E = \text{EeV}) \sim 300$ pc and begin to escape from our galaxy. Based on these magnetic confinement arguments, cosmic rays at and below the knee are generally believed to originate from within our galaxy, while cosmic rays above the ankle are thought to be sourced from extragalctic origins. Note that there is additional evidence, e.g. recent EeV-scale anisotropy measurements by the Pierre Auger Observatory [14], that are consistent with an extragalactic origin of high-energy cosmic rays. This makes the energy range between the knee and ankle, where the transition from to extra sources of cosmic rays is believed to happen, an extremely interesting region for scientific

exploration and is a motivation for the work presented in this thesis.

At the highest energies, there is an observed cutoff in the cosmic-ray energy spectrum that could arise from either a fall in intensity of the extragalactic source population or from high energy cosmic rays interacting with cosmic microwave background (CMB) photons. Above $\sim 5 \times 10^{19}$ eV, protons can lose energy through interacting with CMB photons:

$$p + \gamma_{\mathrm{CMB}} \rightarrow \Delta^{+} \rightarrow n + \pi^{+} \tag{1.4}$$

$$p + \gamma_{\mathrm{CMB}} \rightarrow \Delta^{+} \rightarrow p + \pi^{0} \tag{1.5}$$

Above this energy scale, known as the Greisen-Zatsepin-Kuzmin (GZK) limit, the mean free path of protons is limited to a few tens of Mpc and a suppression in the observed cosmic-ray flux is predicted.

Our conventional understanding of the origin of Galactic cosmic rays is that they may be produced by supernova remnants through an acceleration mechanism first proposed by Fermi in 1949 [15]. With this acceleration mechanism, in which cosmic rays interact with moving magnetic fields, the maximum energy a particle may be accelerated to is proportional to its electric charge Z. For a maximum achievable energy for protons E_{p}, helium is expect to reach a maximum energy of $E_{\mathrm{He}} = 2E_{\mathrm{p}}$, for lithium $E_{\mathrm{Li}} = 3E_{\mathrm{p}}$, and so on. This scenario, first proposed by Peters in [16], predicts a rigidity-dependent cutoff sequence of cosmic rays known as a Peters cycle. These sequential cutoffs result in an increase in the average cosmic-ray mass with energy. Likewise, many models of the energy spectrum describe structures in the spectrum, such as the knee and ankle, as the interplay of various source populations which result in distinct energy spectra for different cosmic ray species. See [17] and [18] for two instances of such models. It's worth noting there are other proposed mechanisms, e.g. a Galactic wind termination shock, by which cosmic rays may be accelerated to energies between the knee and ankle [19] and [20].

Generally, the cosmic-ray energy spectrum and composition are intimately intertwined in theories related to cosmic-ray production and propagation. Thus knowledge of the mass composition of cosmic rays is crucial in furthering our understanding of the origin of high-energy cosmic rays.

1.3 Composition

The full cosmic-ray spectrum is comprised of ∼90.8% protons, ∼9.1% helium nuclei, and ∼0.12% heavier nuclei with $Z > 2$ [21]. The relative abundance of hadronic cosmic rays follows the chemical composition of the solar system, shown in Figure 1.4, apart from elements with Z = 3-5, 9, and 21-25 where the cosmic ray abundances are larger. This mismatch originates from spallation processes in which cosmic-ray nuclei fragment as they propagate through the interstellar medium.

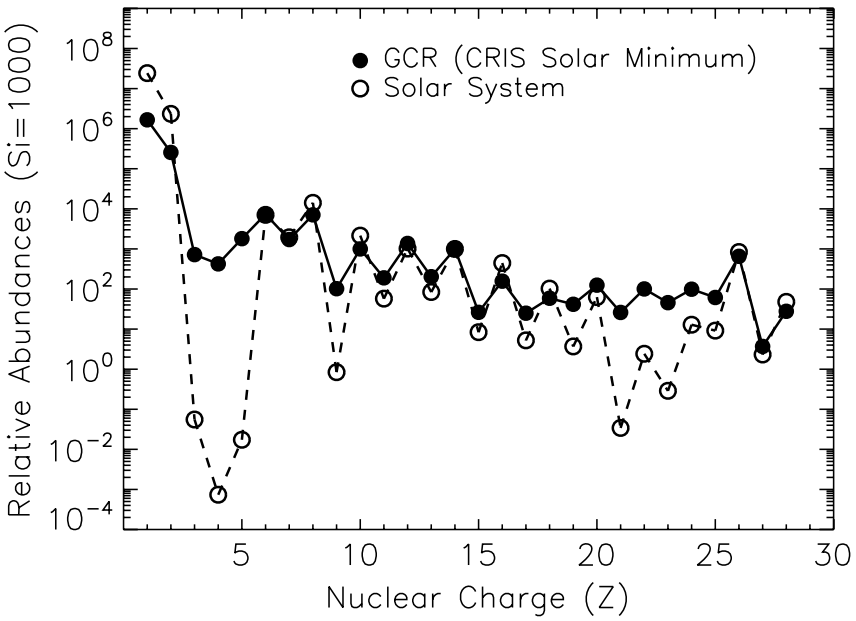


Figure 1.4: Relative abundances of cosmic rays as measured by the CRIS instrument. The relative abundance of cosmic rays matches the chemical composition of the solar system, with a few exceptions. The mismatch between elements with Z = 3-5, 9, and 21-25 arises from spallation processes in which cosmic-ray nuclei fragment as they propagate through the interstellar medium. Figure from [22].

Cosmic-ray composition from ∼1 GeV to ∼100 TeV has been measured with great precision by several satellite (e.g. AMS) and balloon-based detectors (e.g. CREAM). A summary of these measurements is shown in Figure 1.5. The spectrum is proton-dominated until roughly 10 TeV [23], at which point helium becomes dominant. These direct detection experiments have been incredibly successful, however they are limited in the energy ranges they can investigate.

Due to the steeply falling nature of the cosmic-ray spectrum, only ∼1 particle per m^2 per year

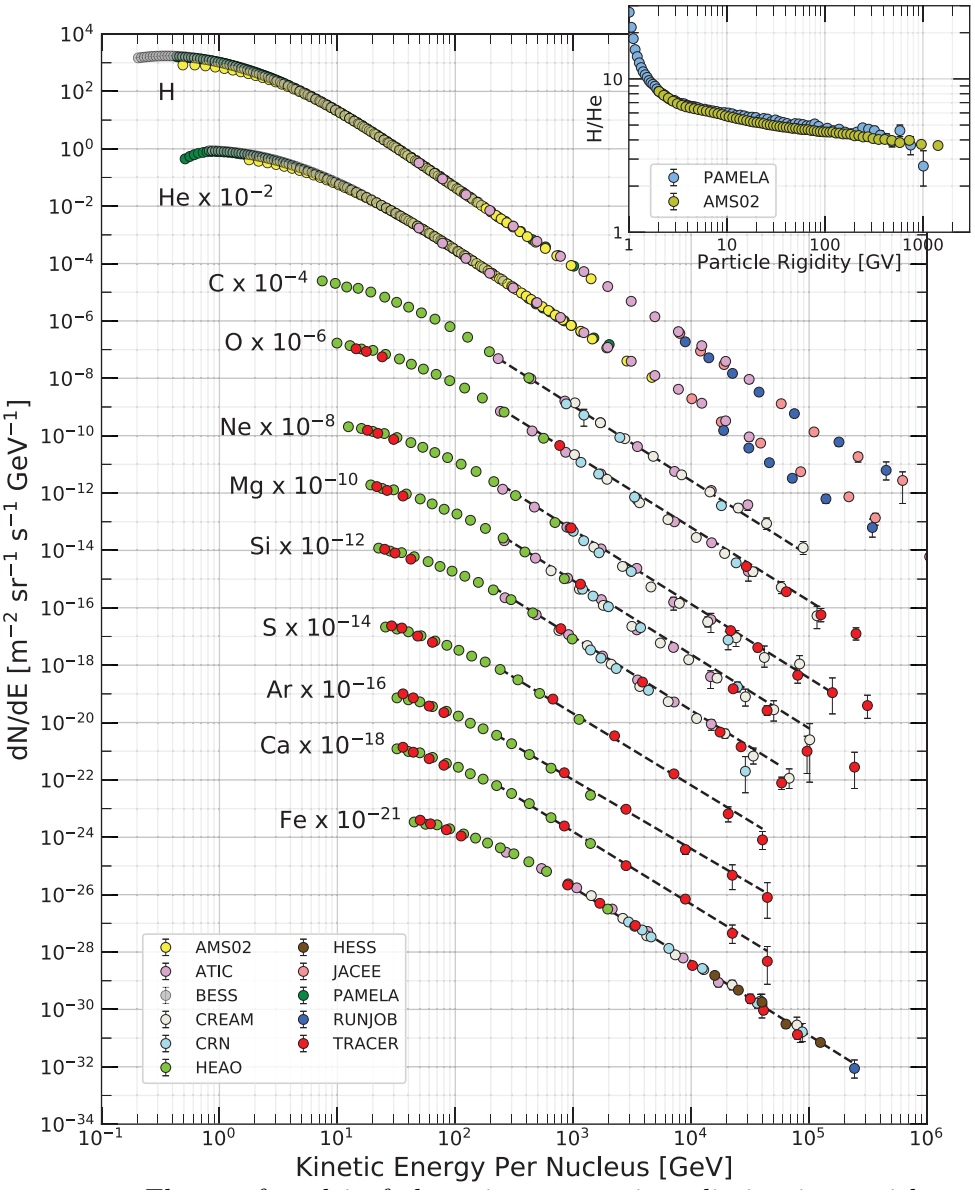


Figure 1.5: Summary of direct detection energy spectra measurements for various cosmic ray primary nuclei. For ease of visual comparison, the magnitude of each spectrum is scaled by the indicated factor. The inset shows the H/He ratio at constant rigidity. Figure from [12].

is expected to be detected at energies near the knee. Such a small event rate makes it unfeasible for satellite and balloon-based detectors to probe the energy regime above the knee. For detection of high-energy cosmic rays, large ground-based detectors are needed in order to have a statistically meaningful event rate.

However, cosmic-ray composition measurements with ground-based detectors are inherently difficult as these detectors do not directly measure cosmic ray particles, but instead measure secondary particles produced from cosmic rays interacting with air nuclei in Earth's atmosphere (air showers are discussed in more detail in Section 1.5). Due to the indirect nature of these experiments, large fluctuations from shower to shower tend to smear out the observable differences between cosmic ray particles of different masses. This leads to a degradation in the composition resolution of ground-based experiments. Only the average mass, or at most a few mass groups, can be measured with any confidence, as compared to the energy spectrum measurements for individual cosmic-ray species made by direct detection experiments.

Results from ground-based detectors that are sensitive to the mean logarithmic mass $\langle \ln A \rangle$ of cosmic rays are shown in Figure 1.6. These measurements show a gradual increase of the average logarithmic mass of cosmic rays between $10^{15} - 10^{17}$ eV, followed by a transition to a lighter composition. At high energies near the ankle ($E_{\text{ankle}} \sim 3 \times 10^{18}$ eV), the observed composition is consistent with a primarily hydrogen flux. This decrease in the average cosmic-ray mass may be indicative of the emergence of extragalactic sources of cosmic rays. Above energies of 10^{19} eV results become inconclusive, with the Telescope Array Surface Detector (TA) observing a proton dominated flux [24], while the Pierre Auger Observatory reports a spectrum consistent with iron [25].

Recent analyses from ground-based experiments have made possible the measurement of energy spectra for distinct mass groups. Observations from KASCADE [26] and KASCADE-Grande [9] [27] indicate a mass-dependent cutoff sequence at energies at and above the knee. In particular, the KASCADE-Grande group found evidence just below 10^{17} eV of a knee-like structure for heavier cosmic rays, shown in Figure 1.7, opening the possibility of a transition to different sources near this energy.

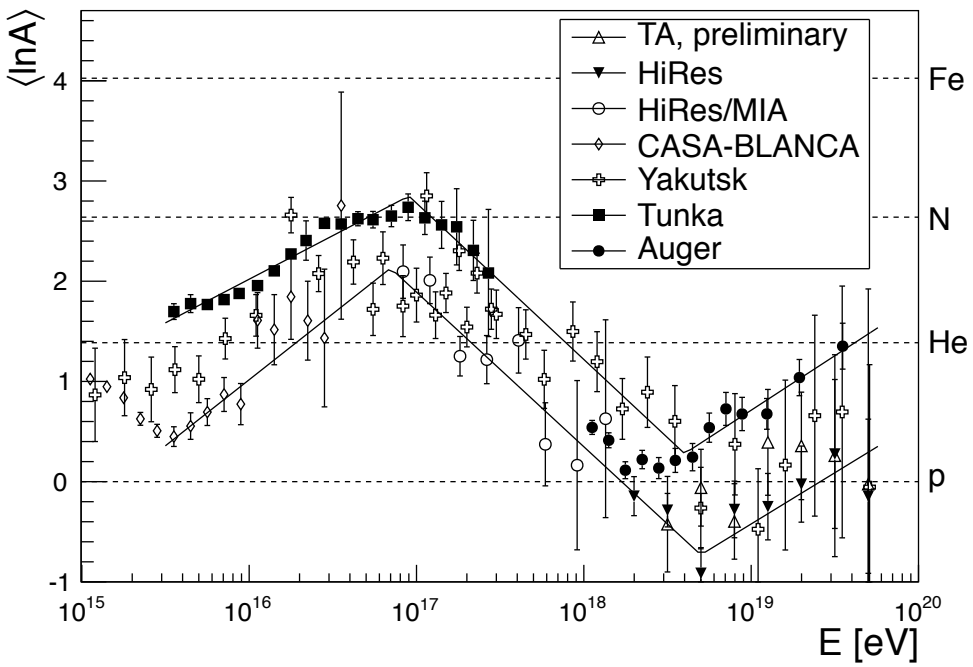


Figure 1.6: Average cosmic ray primary mass ⟨lnA⟩ as a function of energy as measured by several experiments. Error bands are the upper and lower boundaries of the data presented and serve as estimates on the experimental systematic uncertainty. Figure from [25].

While these mass group spectrum measurements are an achievement in their own right, there are large systematic uncertainties originating from both the manifestly indirect detection using air showers, as well as our limited understanding of the interactions which occur in air shower generation. This is one of the motivating factors for performing the composition analysis which is presented in this thesis.

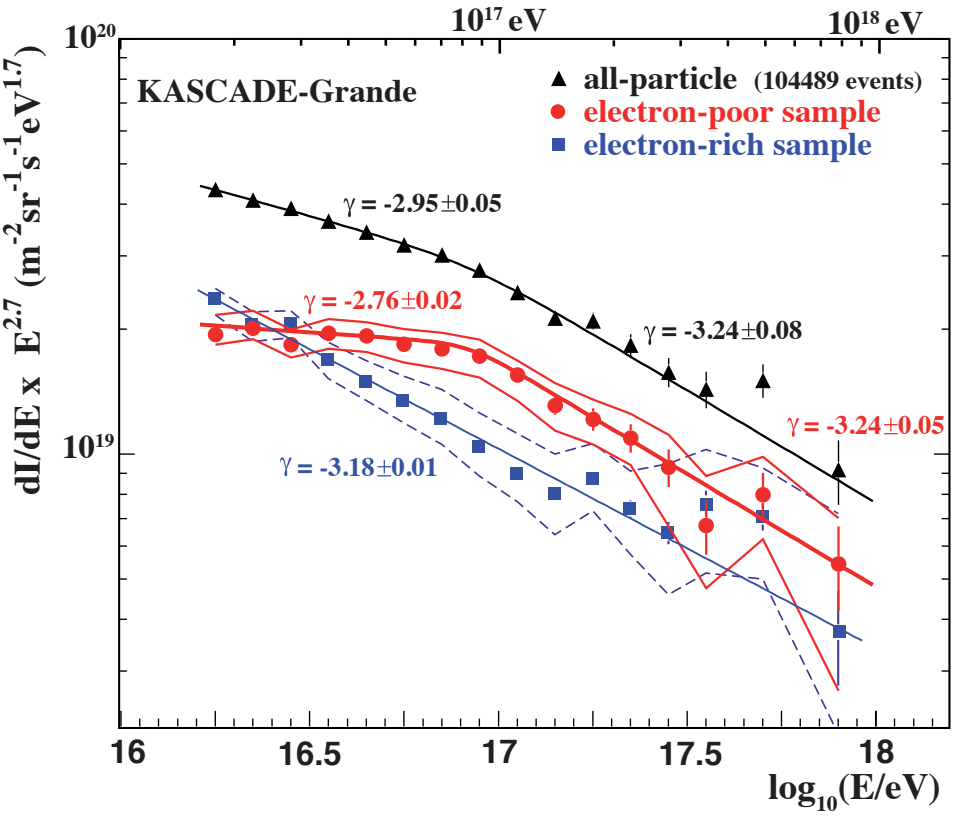


Figure 1.7: Energy spectrum for electron-poor (heavy) and electron-rich (light) components from KASCADE-Grande. Fits for spectral indices are also indicated. The error bars show statistical uncertainties, while systematic uncertainties are represented by bands. Figure from [9].

1.4 Spectrum Anisotropy

In addition to measuring the energy spectrum and relative abundances of cosmic rays, scientists have begun to explore how the energy spectrum varies as a function of arrival direction. Using seven years of detector data, the Telescope Array recently reported an observed hotspot in which the energy spectrum deviated from the rest of the sky with a post-trial statistical significance of 3.74σ [28]. This was determined by comparing energy distributions for events that lie inside spherical caps to those outside, across different points in their field of view. The observed spectrum deviation pre-trial significance is shown in Figure 1.8.

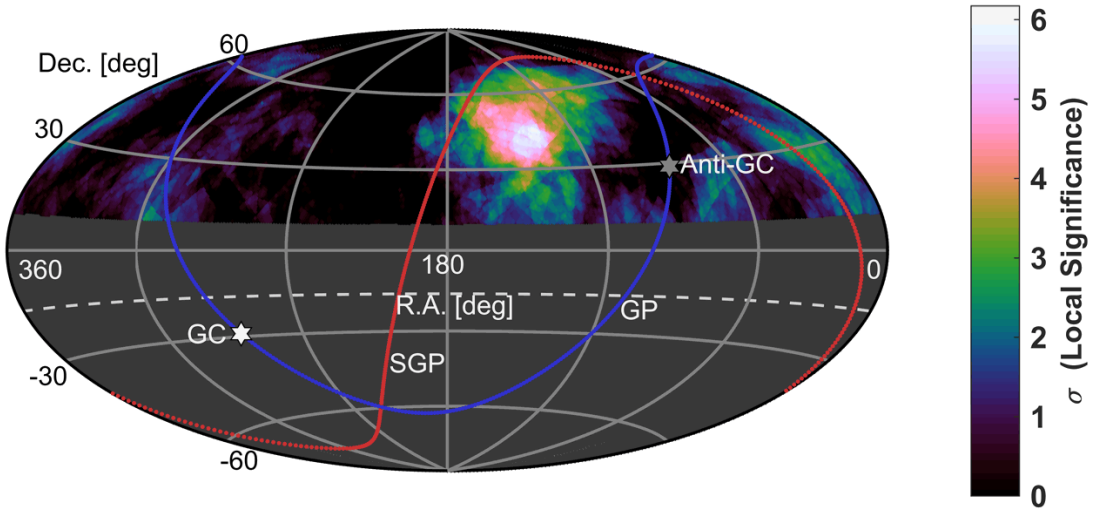


Figure 1.8: Energy spectrum deviation pre-trial significance from the Telescope Array. The maximum is 6.17σ and occurs at 9h16m, 45° in equatorial coordinates and is 7° from the the arrival direction hotspot presented in [29]. The dashed curve at a declination of -16° defines the available field of view. Solid curves indicate the Galactic plane and supergalactic plane. Figure from [28].

The spherical cap with the largest energy spectrum deviation, shown in Figure 1.9, was interpreted as having a deficit of events with energies $10^{19.2} \leq E/\text{eV} < 10^{19.75}$ and an excess for $E \geq 10^{19.75}$ eV. In addition, the quoted hotspot is only 7° from the the arrival direction hotspot location presented in [29]. The TA group claims this feature is suggestive of an energy-dependent magnetic deflection of ultra high-energy cosmic rays.

This work by the TA collaboration serves as motivation for the complementary analysis presented in this work. Here, we use a similar method, described in detail in Section 7.2, to study how

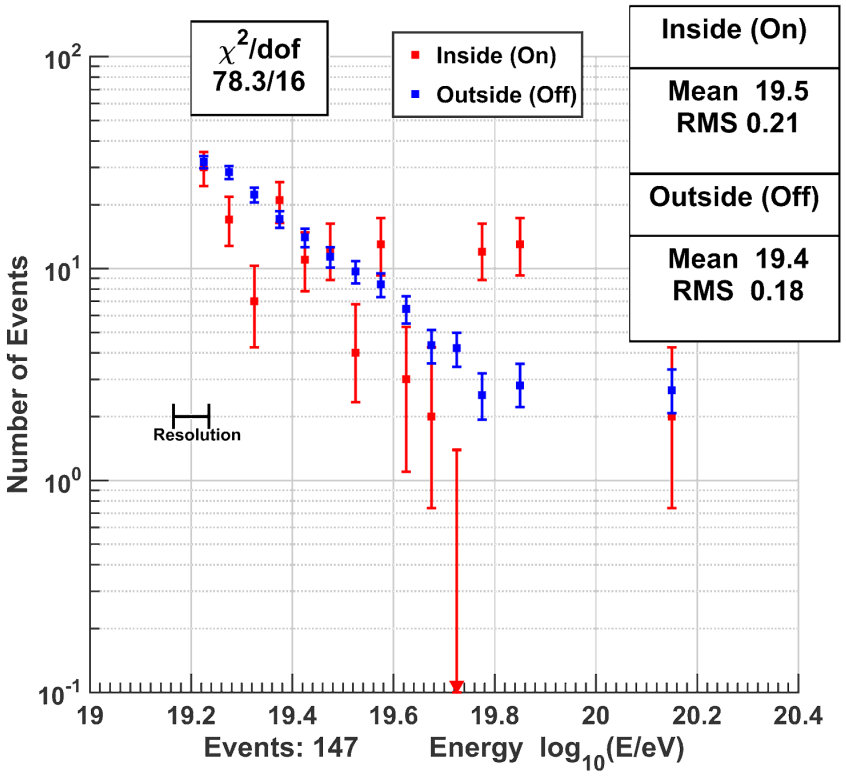


Figure 1.9: Most significant energy spectrum deviation from the Telescope Array. The energy distribution for events that are within the "on" region is shown in red, while the blue markers indicate the "off" region energy distribution. The mean and root mean square (RMS) for the on and off region energy histograms are also indicated. Figure from [28].

the cosmic-ray energy spectrum varies with arrival direction at and above the all-particle knee. In addition, composition information is used to determine if spectral variations differ for separate mass groups. One major benefit of this technique, since it is concerned with measuring only *relative* differences between spectra, is that it is largely insulated against systematic errors, most of which influence spectra in both regions considered in the same manner.

1.5 Extensive Air Showers

Cosmic rays that enter Earth's atmosphere interact with air nuclei and produce a cascade of secondary particles through sequences of interactions [30]. This interaction can be represented as:

$$\mathrm{cr} + A \rightarrow \mathrm{cr} + X + \pi^{\pm,0} + \mathrm{secondaries} \tag{1.6}$$

where cr is an incident cosmic ray, A is an atmospheric nucleus, X is the fragmented nucleus, and $\pi^{\pm,0}$ are charged and neutral pions. While other secondary particles such as $K^{\pm,0}$ and hyperons are also produced, the production cross section for these particles is an order of magnitude smaller than that of pions. Thus, air shower characteristics are largely dominated by pions and other secondaries can generally be ignored [31]. Depending on the energy and type of the particles generated in air shower interaction, they will either interact again or decay.

At each hadronic interaction, the resulting neutral pions receive, on average, roughly one-third of the incident energy [32]. These neutral pions then immediately decay to photons:

$$\pi^0 \rightarrow \ \gamma + \gamma$$

and the resultant photons subsequently produce electron-positron pairs $\gamma \rightarrow \ e^+ + e^-$. The $e^\pm$ pairs can produce additional photons via bremsstrahlung or positron annihilation with atmospheric electrons. The photon, electron, and positron secondaries produced in air shower interactions are often called the electromagnetic (EM) component of the air shower.

The remaining two-thirds of energy goes into charged pions. Depending on the energy of these charged pions, they will either initiate a subsequent hadronic interaction which produces more pions, or decay leptonically to $\mu^\pm$, $e^\pm$, and neutrinos.

$$\begin{aligned} \pi^+ &\rightarrow \mu^+ + \nu_\mu \\ &\rightarrow e^+ + \nu_e \\ \pi^- &\rightarrow \mu^- + \bar{\nu}_\mu \\ &\rightarrow e^- + \bar{\nu}_e \end{aligned}$$

The muons resulting from charged pion decay comprise what's called the hard, or penetrating, component of air showers. The origin of this term is discussed in the next section. Detection of

these secondary particles that constitute an extensive air shower requires a detailed knowledge of how energetic particles lose energy as they travel through matter.

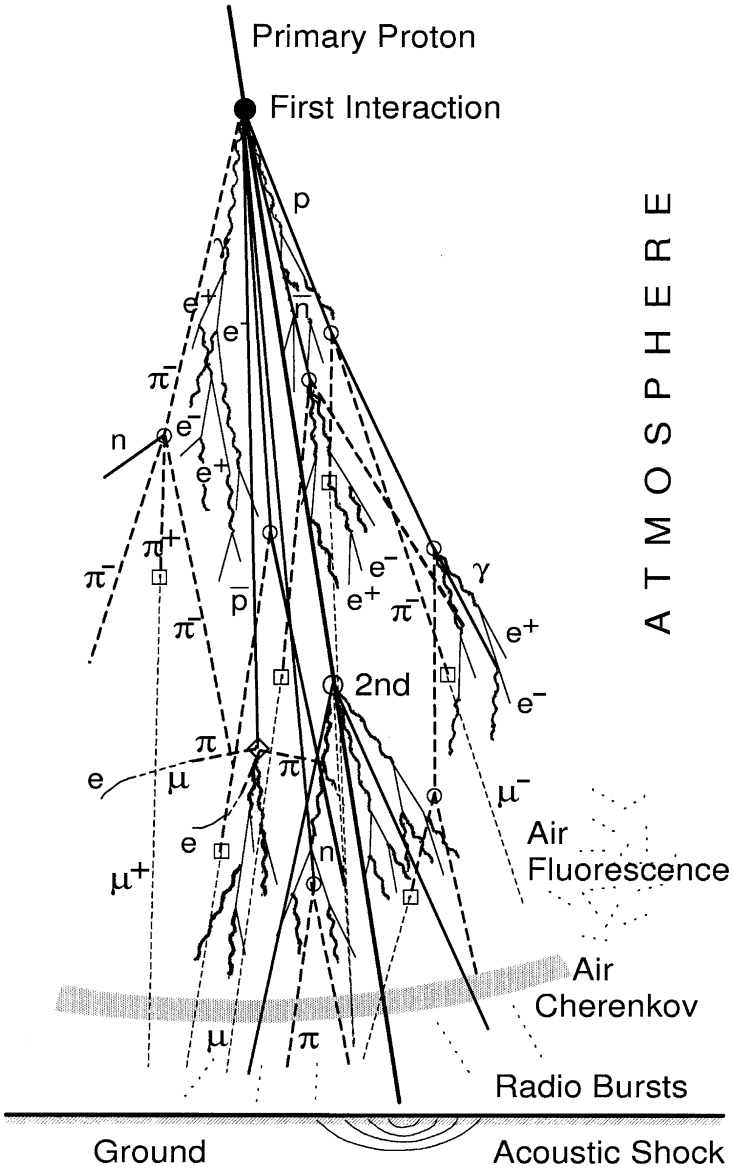


Figure 1.10: Simplified schematic illustrating the development of an air shower induced by a proton cosmic ray primary interacting with the Earth's upper atmosphere. Note that neutrinos are not shown. Figure from [30].

1.5.1 Energy Losses in Matter

Both electrons and muons lose energy as they propagate through matter. There are several processes by which energy is lost: ionization of the surrounding medium, e^+e^- pair production, bremsstrahlung, and photonuclear interactions. The intrinsic properties of a particle (e.g. mass), as well as the energy of the particle, determine the relative magnitudes for each of these processes. Typically ionization losses are dominant below some characteristic energy scale, while stochastic losses from e^+e^- pair production, bremsstrahlung, and photonuclear interactions dominate at higher energies [33].

Electrons above 1 GeV lose energy propagating through matter primarily from bremsstrahlung. The power radiated by bremsstrahlung emission is inversely proportional to mass squared, so electrons lose energy at a rapid rate and the range they can traverse through matter is limited to a few meters of water equivalent (mwe). Thus they are only detected at or above ground level.

Muons, being ∼200 times more massive than electrons, can penetrate much further through matter. For muons below approximately 1 TeV, energy losses due to ionizing the surrounding material are the dominant source of energy loss. However, above these energies stochastic losses become the dominant source of energy loss. Figure 1.11 shows the contribution from each of these processes as a function of energy for muons passing through ice.

The average muon energy loss in matter is typically parameterized via [33]:

$$-\frac{dE}{dX} \approx a(E) + b(E)E \tag{1.7}$$

where $-dE/dX$ is the energy loss per unit path length traveled though a material, and $a(E)$ and $b(E)$ are approximately constant terms accounting for the muon continuous ionization energy losses and stochastic losses, respectively. For ice at the South Pole $a = 0.26$ GeV mwe^{-1} and $b = 3.60 \times 10^{-4}$ mwe^{-1} [35]. Using Equation 1.7, we can determine the average range in ice for a muon with starting energy, E_0, to be [33]:

$$x_0 \approx \frac{1}{b}\left[1 + \frac{b}{a}E_0\right] \tag{1.8}$$

As an example, a 1 TeV muon will, on average, travel a distance of 2.4 km in ice.

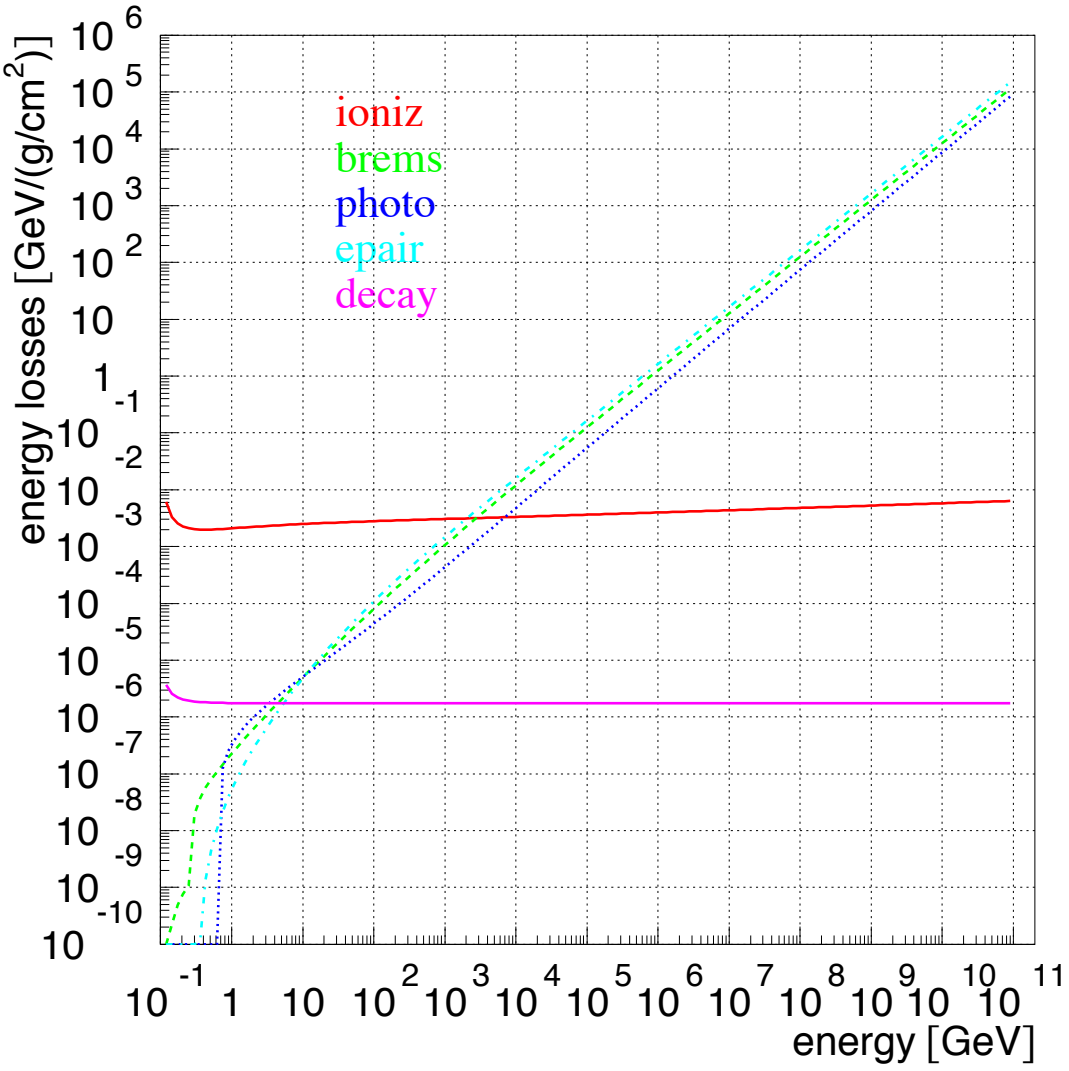


Figure 1.11: Muon energy losses due to ionization, e^+e^- pair production, bremsstrahlung, and photonuclear interactions in ice. Below ∼1 TeV continuous losses to ionization are dominant, however at higher energies stochastic losses begin to take over. Figure from [34].

1.5.2 Experimental Technique

Due to the indirect nature of air shower detection, no single observable can be used to determine both cosmic ray primary particle energy and mass. Instead, multiple parameters are used in conjunction to extract energy and mass estimates. In the case of this work, measuring the number of muons and number of electrons produced in an air shower allows for both primary energy estimation along with composition sensitivity.

Using the Heitler model [32], depicted in Figure 1.12, the number of electrons in an air shower at shower maximum $N^A_{e,\max}$ initiated by a cosmic ray with nuclear mass number A can be approximated as:

$$N^A_{e,\max} \approx N^p_{e,\max} \propto E \tag{1.9}$$

where $N^p_{e,\max}$ is the number of electrons in a proton-initiated shower and E is the primary particle energy. So the number of electrons at shower maximum yields an estimate for the primary particle energy, independent of the the composition. Whereas the number of muons in an air shower initiated by a cosmic ray with nuclear mass number A is roughly:

$$N^A_\mu \approx N^p_\mu A^{1-\beta} \tag{1.10}$$

where N^p_μ is the number of muons in a proton-initiated air shower and β is, from air shower simulations, predicted to be between 0.88 to 0.92 [32]. It's important to note that the number of muons increases with both energy E and mass number A. This can be understood using the superposition principle in which a cosmic ray with total energy E and mass number A can be approximated as A independent nucleons, each with an energy of E/A. These A nucleons will undergo the same interaction in Equation 1.6, but the resultant pions will have, on average, lower energies. Due to a smaller Lorentz boost factor, these lower energy pions have a higher leptonic decay probability which results in a larger number of muons for air showers induced by heavier cosmic rays. The number of muons and electrons observed at ground level for simulated proton and iron air showers at various energies is shown in Figure 1.13.

Note in Figure 1.13 the relative rotation between the iron and proton shower contours as the primary particle energy increases. This non-linearity originates from the rate of increase in number

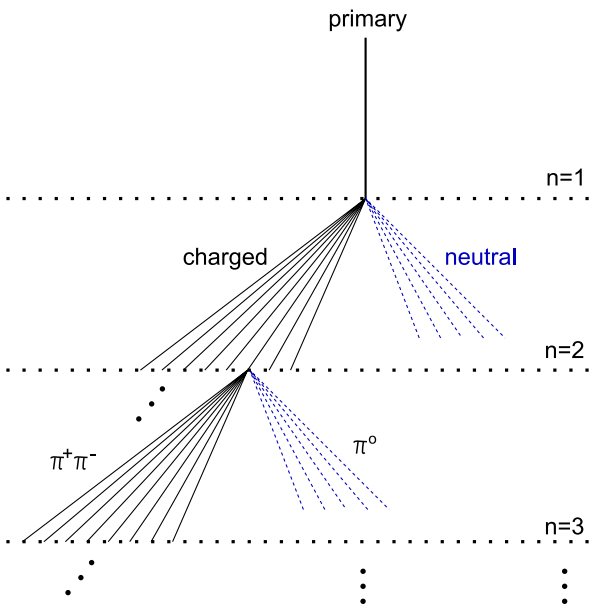


Figure 1.12: Heitler model in which air shower development is modeled as a sequential branching process. Each line segment indicates a particle, with solid lines representing charged pions and dashed lines denoting neutral pions. Note that not all pion lines are shown after the $n = 2$ level. Figure adapted from [32].

of electrons being larger for iron showers as compared to proton initiated air showers. At higher energies, proton initiated showers have not yet reached their shower maximum, which reduces the rate of increase of the number of electrons in the shower compared to iron induced showers that have already passed their shower maximum [25].

In this work, we use the IceCube Observatory, discussed in Chapter 2, to measure both the electromagnetic and muonic components of air showers. This simultaneous estimation of both the number of electrons and number of muons provides energy and composition sensitivity.

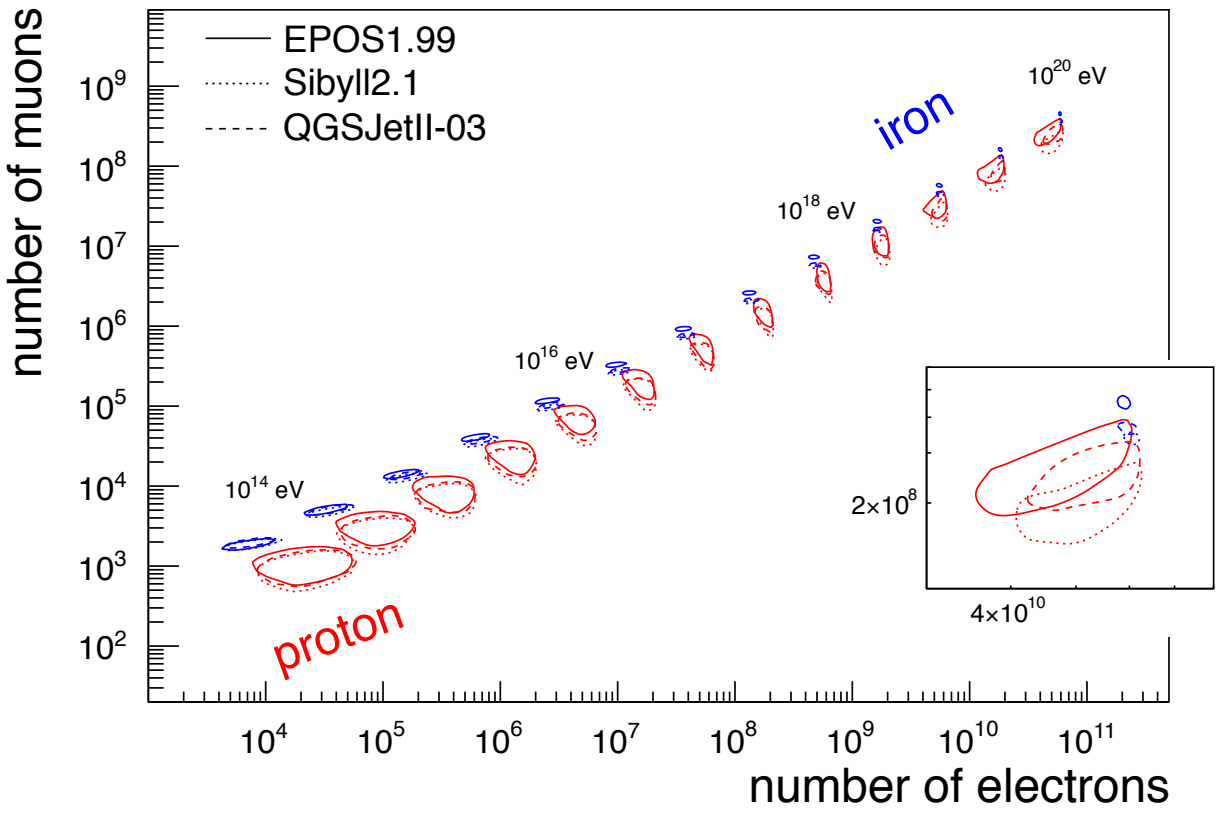


Figure 1.13: Number of muons and electrons for simulated proton and iron air showers at various energies. The particles are observed at ground level for a vertical shower at 800 g/cm^2 (note that IceTop is at 692 g/cm^2 [36]). Each contour indicates the 90% containment of air shower events, while the various line styles show differences in the hadronic interaction model used for producing the simulated showers. Inset shows a detailed view at 10^{20} eV. Figure from [25].

1.5.3 Air Shower Simulation

Paramount for the science mission of any air shower experiment is the ability to develop a mapping between the observed detector's signals and quantities of interest related to primary cosmic ray particles (e.g. direction, location, energy, etc). In order to develop such mappings, Monte Carlo (MC) simulation is used to model both extensive air shower development and the corresponding detector response. Air shower simulation is generated using the CORSIKA [37] software, which is the standard simulation package used throughout the cosmic-ray community. CORSIKA tracks the development of air shower particles that are induced from atomic nuclei primary particles interacting with Earth's atmosphere. For a given incident primary particle, the particle is traced until it interacts with an air nuclei as determined by the nucleus-air cross-section. Then, all subsequent interactions of secondary particles are modeled according to the decay and air nuclei interaction probabilities for the resultant particles. The output from CORSIKA is a set of particle type, location, direction, energy, and arrival time relative to the primary particle first interaction for all the secondary particles at a specified observation altitude.

CORSIKA allows users to choose between several different hadronic interaction models when simulating air showers. For low energy hadronic interactions below 80 GeV we use the FLUKA interaction model [38], while above 80 GeV Sibyll 2.1 [39] was chosen as the nominal hadronic interaction model used to generate air shower simulation. Effects related to choosing other interaction models, including Sibyll 2.3 [39], QGSJet-II-04 [40], and EPOS-LHC [41], are discussed in Section 6.3.3.

Air showers initiated by proton (p), helium (He), oxygen (O), and iron (Fe) primaries were generated according to an E^{-1} energy spectrum with minimum and maximum primary energies of $E_{\text{min}} = 10^{5.0}$ GeV and $E_{\text{max}} = 10^{8.0}$ GeV, respectively. The simulated energy spectrum of E^{-1}, as opposed to a more realistic E^{-3} spectrum, was chosen in order to generate sufficient numbers of high-energy air showers in a timely manner given the available computing resources. In total, the simulated detector response to 200,000 air showers was generated for each of the four primary particle compositions. The incident zenith angles θ for the generated showers were sampled from a $\cos\theta$ distribution between $0^\circ < \theta < 65^\circ$, while the primary azimuth angles were uniformly sampled

over the full 2π range. Table 1.1 summarizes the generated air shower simulation used in this work.

Composition	Energy Range $[\log_{10}(E/\text{GeV})]$	Number of Events
P	5-8	200,000
He	5-8	200,000
O	5-8	200,000
Fe	5-8	200,000

Table 1.1: Summary of simulated air showers. Air showers for proton (P), helium (He), oxygen (O), and iron (Fe) primaries were generated according to an E^{-1} energy spectrum between minimum and maximum primary energies of $10^{5.0}$ GeV and $10^{8.0}$ GeV, respectively. The shower zenith angles θ were sampled from a $\cos\theta$ distribution between $0^\circ < \theta < 65^\circ$. The Sibyll 2.1 hadronic interaction model was used in production of secondary air shower particles.

Chapter 2

The IceCube Observatory

The IceCube Neutrino Observatory is a neutrino detector located at the Amundsen-Scott South Pole Station in Antarctica. It's comprised of two components: an air shower surface array, also known as IceTop, and an in-ice array of detectors deep in the Antarctic ice. Each of these components is designed to detect Cherenkov radiation produced from relativistic particles propagating through ice. A schematic of the observatory is shown in Figure 2.1. While the primary science mission for IceCube is to discover and understand the astrophysical origins of high-energy neutrinos, it is also an excellent detector of PeV scale cosmic rays. This chapter discusses the basic detection principle used by IceCube as well as the two detector components of the observatory.

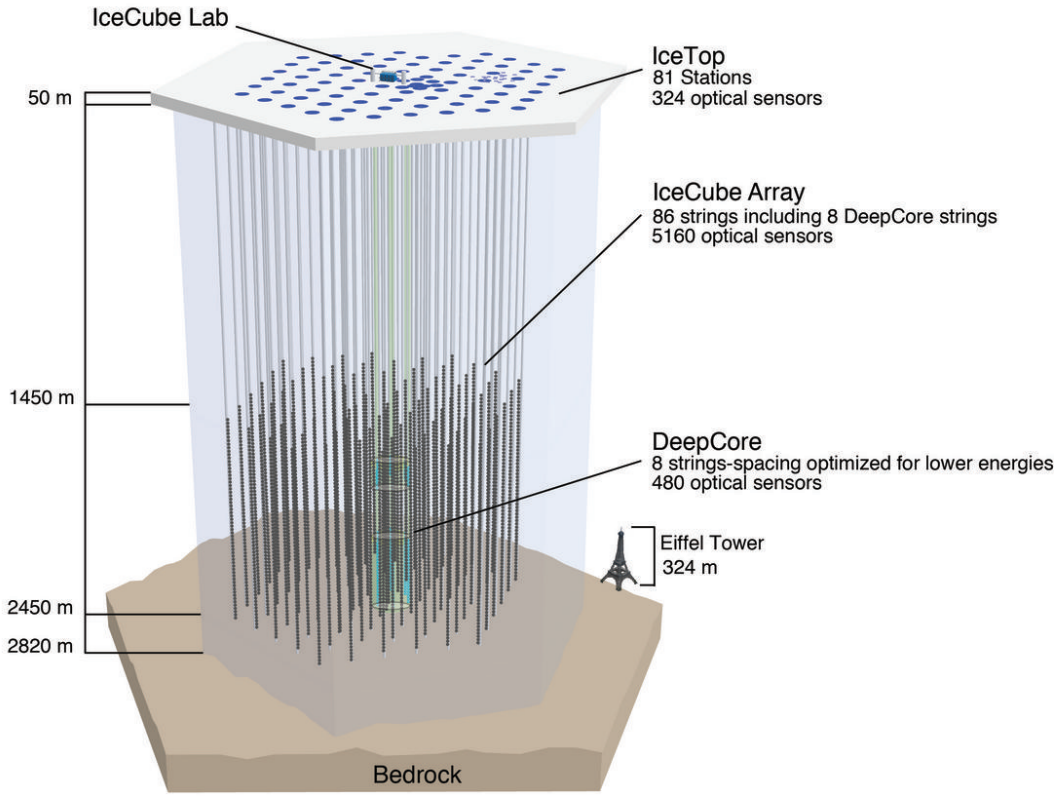


Figure 2.1: The IceCube Neutrino Observatory. The observatory is comprised of an air shower array at the surface and an in-ice array deep in the Antarctic ice. The Eiffel tower is shown for scale.

2.1 Cherenkov Radiation

As discussed in Section 1.5, cosmic rays interact with atoms in Earth's atmosphere, creating air showers of particles. Those particles are detected via the Cherenkov radiation they produce. Cherenkov radiation is the electromagnetic radiation produced when relativistic, charged particles travel through a dielectric medium faster than the local speed of light in that medium [42]. The local speed of light in a material is given by c/n, where c is the speed of light in vacuum and n is the index of refraction for the medium. As a relativistic particle moves through a medium, the emitted Cherenkov radiation forms a coherent wavefront of light (depicted in Figure 2.2), much like the wake formed by a duck moving through water in a pond.

The Cherenkov radiation wavefront, often referred to as the Cherenkov cone, is characterized by the opening angle:

$$\cos\theta_c = \frac{c}{vn} \tag{2.1}$$

where θ_c is the opening angle of the Cherenkov cone, v is the speed of the particle, c is the speed of light in vacuum, and n if the refractive index of the medium. For a highly relativistic particle (i.e. $v \approx c$) moving through ice ($n = 1.3$) the Cherenkov angle is $\theta_c \approx 39.72°$.

The IceCube Observatory, along with other particle detectors, capitalizes on this light generation process by instrumenting a transparent medium with light sensors to detect Cherenkov radiation produced by relativistic air shower particles.

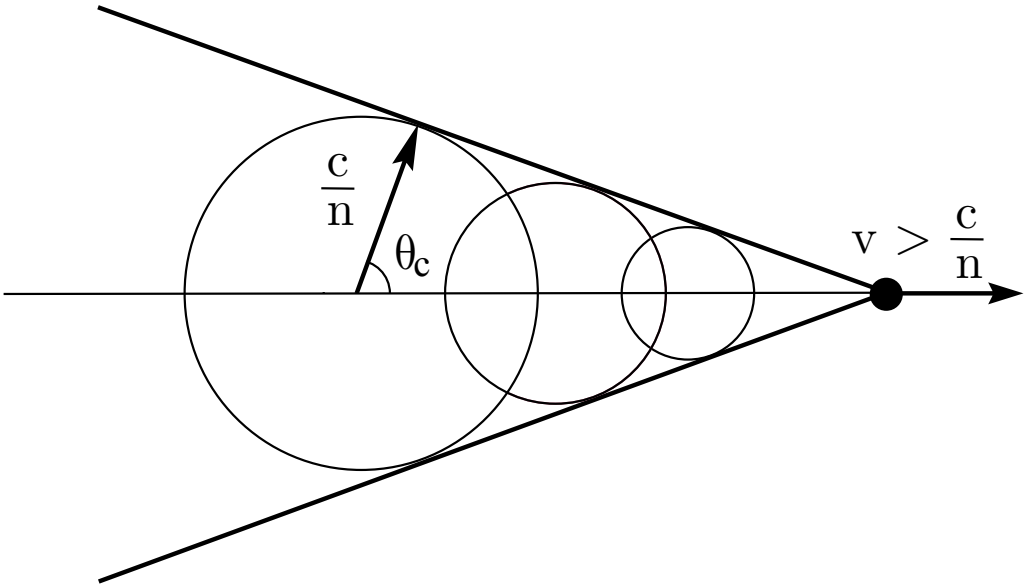


Figure 2.2: Cherenkov radiation wavefront diagram for a relativistic particle moving at speed $v > c/n$. Cherenkov light is emitted at an angle θ_c (see Equation 2.1) relative to the particle's velocity vector. The emitted radiation constructively interferes to form a cone-like wavefront with the vertex of the cone located at the position of the particle.

2.2 Detector components

2.2.1 Surface Array (IceTop)

IceTop is an ice Cherenkov air shower array located at the surface of the South Pole which is sensitive to cosmic ray induced extensive air showers with primary energies from roughly 100 TeV to 1 EeV. It consists of 81 stations arranged in a grid (shown in Figure 2.3) over a 1 km^2 area with a nominal spacing of 125 m between stations. Each station is comprised of two ice Cherenkov tanks that are separated by a distance of 10 m.

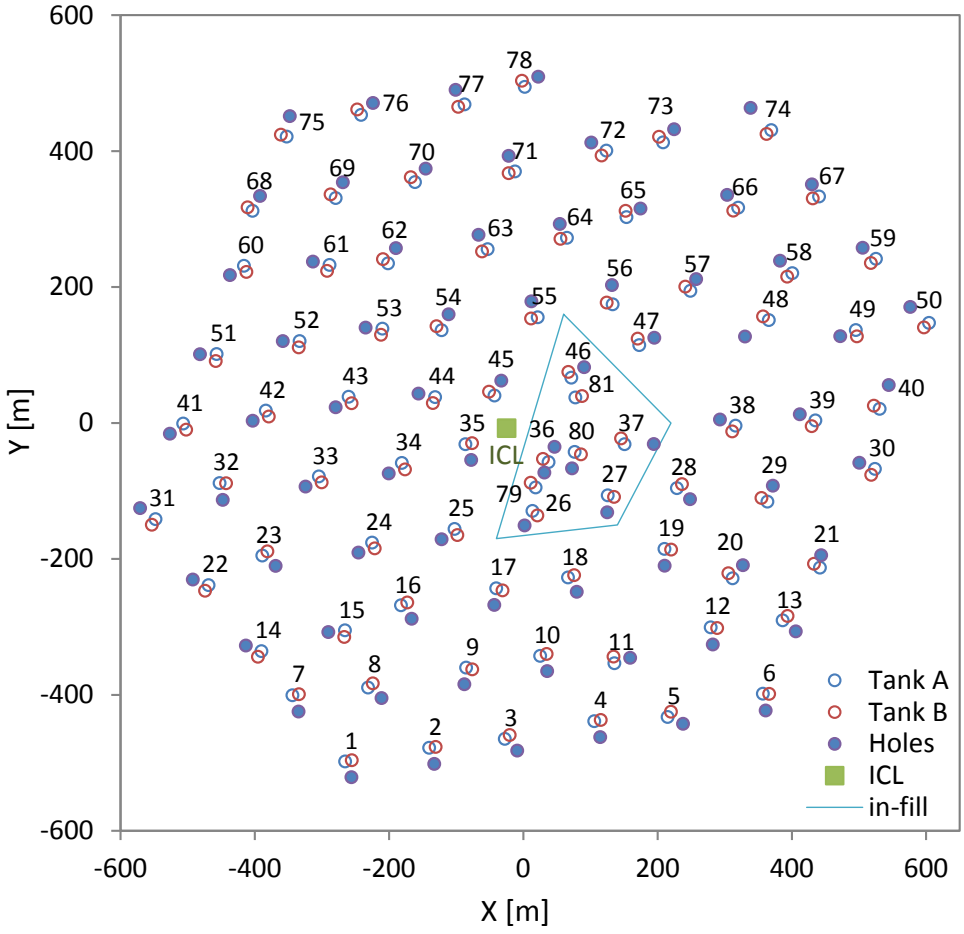


Figure 2.3: Diagram showing IceTop tank and in-ice array string locations. The nominal spacing between in-ice strings is 125 m, while the spacing between IceTop tanks is 10 m. The IceCube Lab (ICL) location is indicated by the green square. Figure from [36].

A cross section schematic of an IceTop tank is shown in Figure 2.4. Each tank is 1.3 m tall

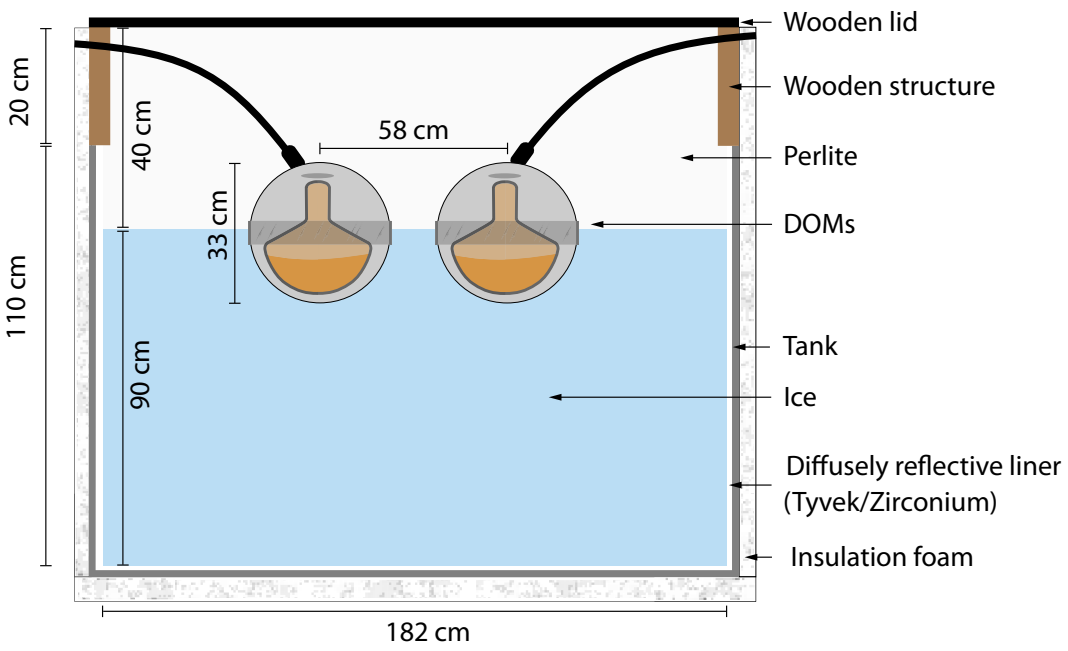


Figure 2.4: Cross-section of an IceTop tank. Figure from [36].

and filled with 0.9 m of clear ice. Two digital optical modules (DOMs) for detecting Cherenkov radiation from air shower particles are placed in each tank. The DOMs in each tank are facing downwards and located along a diameter of the tank symmetrically about its center with a distance of 58 cm between the two DOM centers. The space between the tank ice and lid (0.4 m) is filled with perlite for thermal insulation and protection against light exposure. A procedure in which the water in the surface tanks was made to freeze from the top to the bottom of the tank was used to ensure clear and crack-free ice [36].

2.2.2 In-Ice Array

The in-ice array consists of 86 cable assemblies (called "strings") that are drilled deep into the Antarctic glacial ice. The strings are spaced approximately 125 m apart from one another and are each instrumented with 60 DOMs between 1450 m to 2450 m beneath the surface. The DOMs, which are identical to those used in IceTop, are evenly spaced along each string in intervals of 17 m and each string is located next to a corresponding IceTop station (see Figure 2.3 for the locations of

the in-ice strings and IceTop stations). Altogether, the in-ice array instruments a cubic kilometer of Antarctic ice.

The in-ice strings were deployed into boreholes, each approximately 60 cm in diameter to a depth of 2500 m, that were drilled using a high-pressure hot-water drill [43]. Due to the extreme weather conditions at the South Pole, drilling was able to be done only during the months of December and January (the two warmest months in Antarctica). Construction of both the surface and in-ice arrays began in 2004 and was completed in December 2010.

2.3 Digital Optical Modules

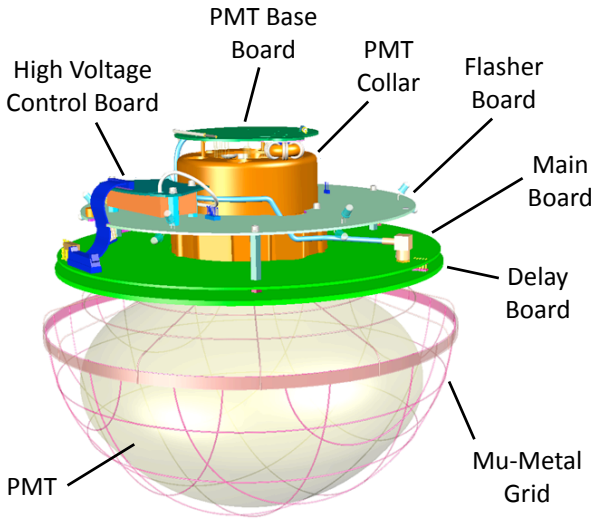


Figure 2.5: Schematic of a digital optical module. Note that the glass pressure housing unit is not shown in this diagram. Figure from [44].

The digital optical module, or DOM, is the central light sensor device used throughout IceCube. DOMs are used to detect Cherenkov radiation given off by relativistic particles interacting with ice. Each DOM is comprised of a 10 inch diameter Hamamatsu R7081-02 photomultiplier tube (PMT) [45], power supply, light sources for calibration, and onboard electronics that allow for the capture and digitization of PMT signals [46]. The PMT and electronics are housed in a glass pressure sphere with an outer diameter of 13 inches and thickness of 0.5 inch. In order to shield against

effects due to the Earth's magnetic field, the PMT is enclosed by a mu-metal wire grid. A DOM schematic diagram is shown in Figure 2.5.

To fully capture the dynamic range of signals produced by air showers, the two DOMs in each IceTop tank are operated at different gains. The high gain (HG) DOM operates with a gain of 5×10^6 and the low gain (LG) DOM has a gain of 1×10^5. The in-ice array DOMs operate with a gain of 1×10^7.

Several operations, including data acquisition, calibration, and communications, are integrated into a single circuit board (called the "main board") at the neck of the PMT [46]. PMT signals are captured and digitized by the DOM electronics when the PMT pulse voltage crosses a specified threshold. This threshold corresponds to a signal of 23 PE and 270 PE for the HG and LG IceTop DOMs, respectively, and 0.25 PE for the in-ice array DOMs. PMT signals are digitized using both an Analog Transient Waveform Digitizer (ATWD) and a 10-bit fast Analog-to-Digital Converter (fADC). The ATWD is a custom integrated circuit that samples the PMT output signal in 3.33 ns wide time bins with 128 total bins for a total sampling time of 427 ns. The fADC is a commercial product used to continuously digitize PMT waveforms at a sampling rate of 40 MHz for a total of 6.4 μs. The ATWD yields a digitized signal with a small timing resolution over a short time window, while the fADC provides time coverage for longer signals. Figure 2.6 shows an example PMT signal that has been digitized by both the ATWD and fADC.

2.4 Calibration

Each DOM in IceTop can have a different response to a given signal due to variations in the ice properties, tank materials, or snow coverage atop the individual tanks. In order to calibrate for these various effects, each DOM is calibrated using a single-muon spectrum.

The average muon energy at IceTop is $\mathcal{O}(2$ GeV) and muons at this energy are minimum ionizing in ice [36]. Thus the energy deposited in an IceTop tank by these muons is proportional to the traveled path length through the tank. This is evident in the charge spectrum for each tank. An example charge spectrum is shown in Figure 2.7. The peak at $\sim$134 PE is produced by vertical muons passing through the tank, while the background below the vertical muon peak is

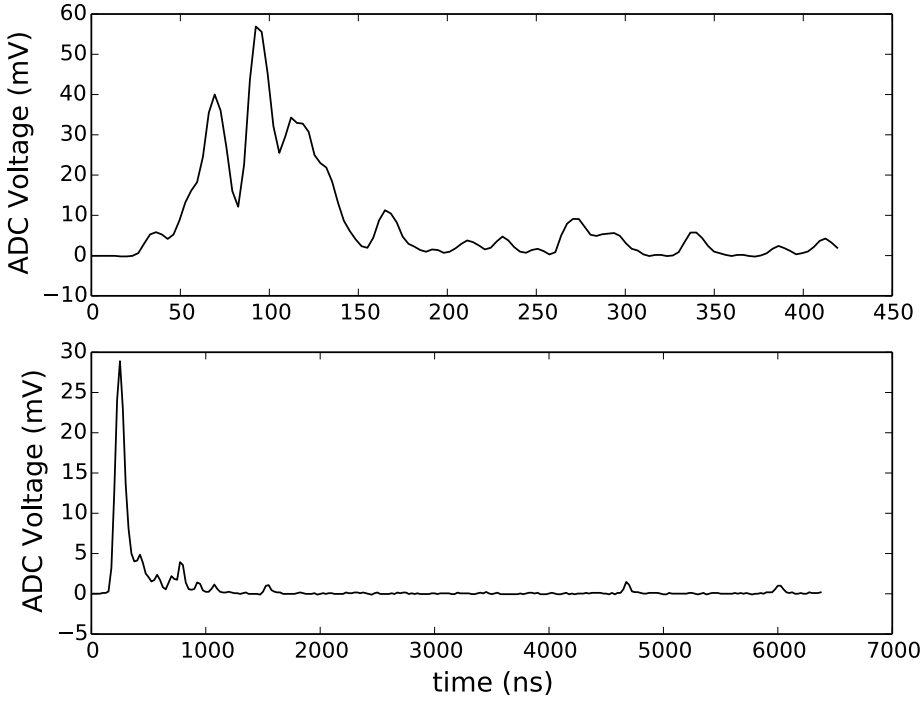


Figure 2.6: The same PMT signal sampled by the ATWD (top) and fADC (bottom). Figure from [44].

predominately from the electromagnetic background.

The muon spectrum for each tank is fit using the empirically derived formula [36]:

$$f(x) = \underbrace{p_0 \left[L(x; p_1, p_2) + \frac{1.85}{p_1} \frac{1}{\exp\Big((x - p_1)/(p_2)\Big) + 1} \right]}_{f_\mu} + \underbrace{p_3 \exp\Big(p_2 x\Big)}_{f_{em}} \tag{2.2}$$

where $L(x; p_1, p_2)$ is a normalized Landau distribution. The f_μ term is the contribution to the charge spectrum due to muons and the f_{em} term accounts for electromagnetic background. The fitted muon term is then used to calibrate each IceTop DOM to the same input signal (vertical muons). A vertical equivalent muon (VEM) is defined to be the charge value at 95% of the muon peak which is the maximum of the muon contribution in the fit in Equation 2.2. A VEM is the standard unit of deposited signal used in a IceTop tank. The charge value of 95% was chosen in response to a measurement of tagged vertical muons which resulted in a 5% lower muon peak [36].

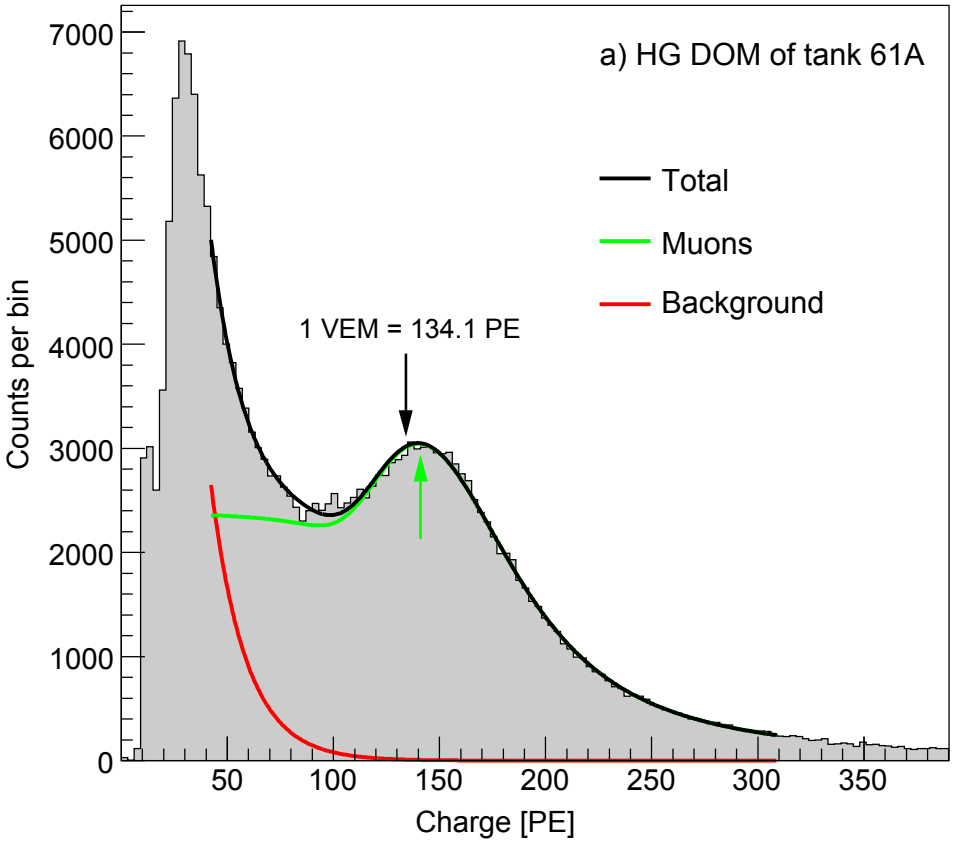


Figure 2.7: Charge spectrum for an IceTop DOM during calibration monitoring on April 15, 2012. The black curve is the fitted function from Equation 2.2. The green and red curves are the fitted muon peak and electromagnetic background components, respectively. For this particular DOM, a VEM is equivalent to a deposited charge of 134.1 PE. Figure from [36].

Chapter 3

Air Shower Reconstruction

Given the charge, timing, and position information from an air shower detected by the surface and in-ice arrays, we are left with the task of using this information to infer, or reconstruct, parameters associated with the primary cosmic ray particle that induced the air shower. This chapter discusses the reconstruction techniques used to engineer the high-level features used in this work. Specifically, the reconstructed parameters which will be used in later steps of this analysis are the air shower energy proxy, which correlates with the number of air shower electrons, and direction, as determined using IceTop, and the composition-sensitive muon number proxy, as determined using the in-ice array.

3.1 General Reconstruction Principles

In Chapter 2, we discussed the low-level DOM signals in the surface and in-ice arrays. Using the timing distribution of these signals, we can group signals across the arrays into different events. Each event of interest in this analysis represents the detected signals associated with a single air shower passing through IceCube. An example event is shown in Figure 3.1. This chapter will review the event reconstruction methods used for this analysis. In particular, it is important to understand the main goals of our reconstructions:

1. With IceTop: the best possible measure of the air shower core position, direction, and an energy proxy parameter

2. With the in-ice array: a muon number proxy

The reconstructed core position and direction will be crucial for the data quality cuts discussed in Section 3.4, while the energy and muon number proxies will be used to determine primary particle composition.

The reconstruction techniques used for both the surface and in-ice arrays are maximum likelihood methods. These reconstructions are described by a likelihood function, $\mathcal{L}(x|\theta)$, that contains information about the observed detector signals for an event, x, and is parameterized by a set of free parameters $\theta = \{\theta_j | j = 1, ..., m\}$ that we wish to know. In our case, the free parameters are related to characteristics of detected air shower, e.g. air shower direction, core position, etc. The method of maximum likelihood estimation states that the estimated values for the θ parameters are those that maximize the likelihood function. Intuitively, this is choosing the values for the likelihood parameters that maximize the probability of measuring the observed detector signals.

Maximum likelihood-based algorithms perform better when the seed, or starting point for the optimization procedure, used for the reconstruction is close to the true solution. To this end, an initial, analytical plane wave reconstruction is performed for IceTop signals which then serves as a good seed for the more advanced IceTop likelihood reconstruction. The result of the IceTop likelihood reconstruction is then used as a seed for the in-ice likelihood reconstruction.

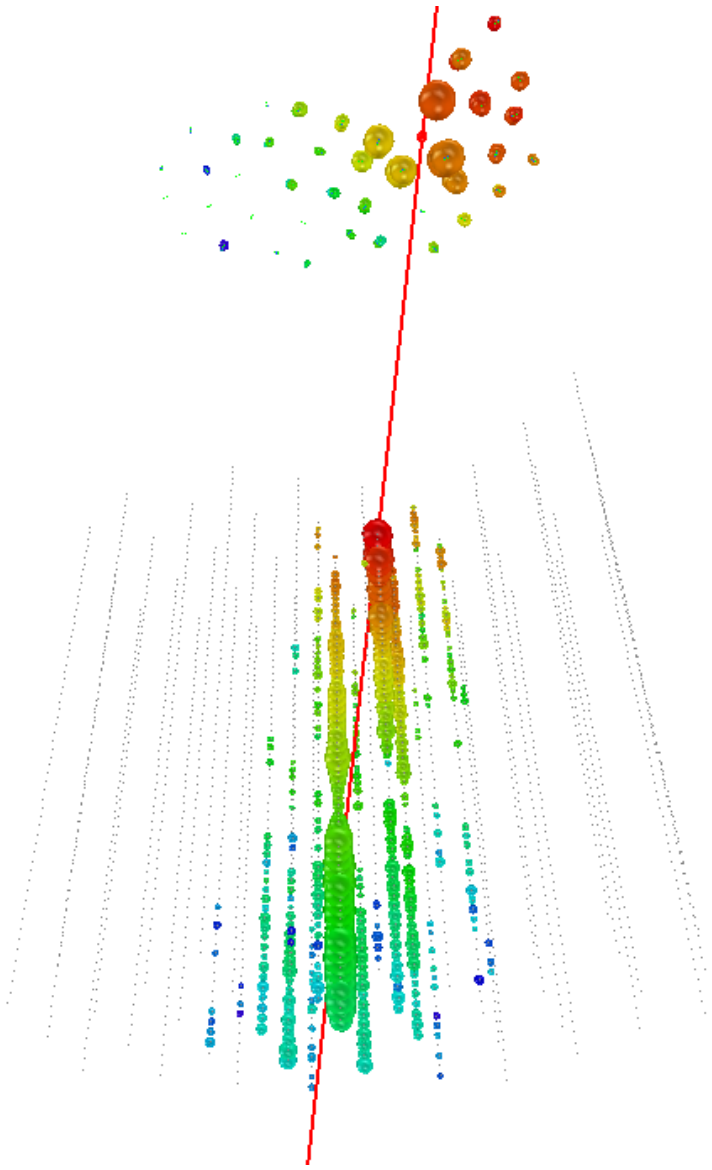


Figure 3.1: Coincident air shower event. IceTop can be seen towards the top of the diagram, while the in-ice array is shown below. The signal deposited in each DOM is shown by a sphere surrounding the corresponding DOM. The larger the measured signal, the larger the size of the sphere. Timing information is encoded in the color of the sphere, with earlier times in red and later times in blue. The red line indicates the reconstructed air shower track.

3.2 IceTop Air Shower Reconstruction

An initial quick, analytical reconstruction is performed in which cosmic ray air showers are approximated as plane waves moving at the speed of light c. The shower core position estimate for

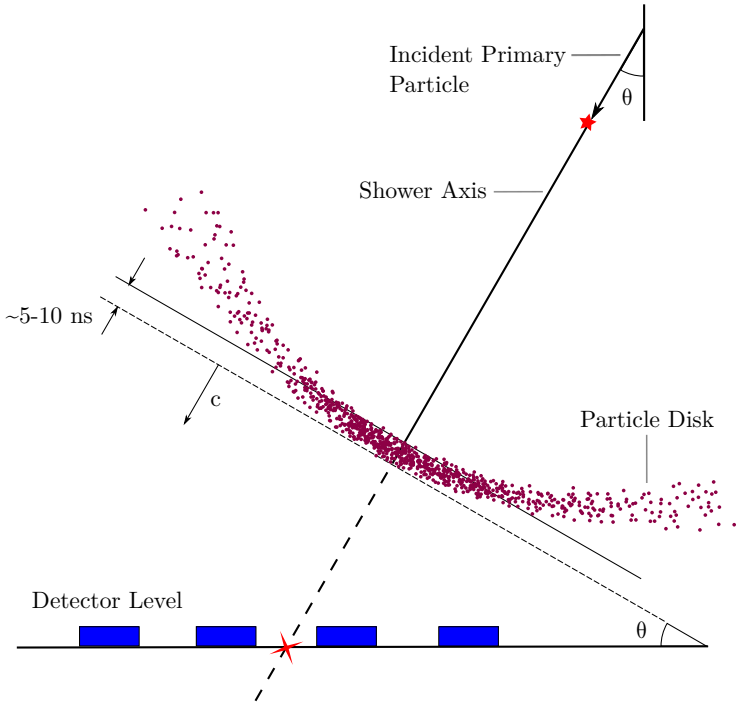


Figure 3.2: Spherical particle disk of an extensive air shower.

the plane wave reconstruction is given by the centroid of the air shower. This is defined to be the charge weighted average of the tank positions:

$$\vec{r}_0 = \frac{\sum_i^{n_{\text{tanks}}} q_i^w \ \vec{r}_i}{\sum_i^{n_{\text{tanks}}} q_i^w} \tag{3.1}$$

where $w = 0.5$ is the charge weighting power, q_i is the charge measured in the i-th tank, and $\vec{r}_i$ is the position vector for the i-th tank. Note that the weighted average is not taken over all tanks, but instead restricted only to the $n_{\text{tanks}} = 7$ tanks with the largest measured signal. The values for the charge weighting power w and n_{tanks} were chosen to optimize the performance of the plane wave reconstruction [36].

Using the shower core position and timing information from the tanks, we can determine the direction of the air shower. Under the shower plane wave approximation, the time of each tank hit can be written as a function of the tank position relative to the shower core. That is,

$$t_i^{\text{plane}} = t_0 + \frac{(\vec{r}_i - \vec{r}_0) \cdot \vec{n}}{c} \tag{3.2}$$

where $\vec{n}$ is a unit vector pointing in the direction of motion of the shower plane wave, $\vec{r}_0$ is the shower core position, and t_0 is the time at which the shower core passes through ground level. The shower direction is determined by minimizing a χ^2 function between the fit and measured times

$$\chi^2 = \sum_i \frac{(t_i^{\text{plane}} - t_i)^2}{\sigma^2} \tag{3.3}$$

where $\sigma = 5$ ns is the trigger time uncertainty, t_i^{plane} and t_i are the plane wave fit and measured times of the i-th tank hit, respectively. Expanding equation 3.3 in terms of the components of $\vec{n}$ and $\vec{r}_i$ we obtain

$$\chi^2 = \frac{1}{\sigma^2} \sum_i \left(t_0 - t_i + \frac{1}{c}\big(n_x(r_{i,x} - r_{0,x}) + n_y(r_{i,y} - r_{0,y}) + n_z(r_{i,z} - r_{0,z})\big) \right)^2 \tag{3.4}$$

Given that $\vec{n}$ is a unit vector, i.e. $\vec{n} = \left(n_x, n_y, -\sqrt{1 - n_x^2 - n_y^2}\right)$, minimizing equation 3.4 requires solving a system of non-linear equations. To simplify this procedure, we make the assumption that the z-component of the tank positions is constant and equal to the z-component of the shower core position, i.e. $r_{i,z} = r_{0,z}$. This is a reasonable assumption to make, given that the tank altitudes vary by only 6 meters over the entire $\sim$1km^2 scale IceTop detector [36]. Under this constant tank altitude assumption, the values of n_x and n_y which minimize Equation 3.4 can be found analytically through matrix inversion.

To help mitigate any biases introduced by the constant tank altitude assumption, a second iteration of the χ^2 minimization is performed using times that have been corrected to account for the tank's relative height differences. The corrected times used for this second iteration, in place of the original measured times, are:

$$t_i^{\text{corr}} = t_i + \frac{(z_i - z_0)\cos(\theta)}{c} \tag{3.5}$$

where t_i^{corr}, t_i, and z_i are the corrected hit time, the measured time, and z-coordinate for the i-th tank, respectively, and θ is the reconstructed zenith angle from the first iteration. The direction

found from this second iteration of χ^2 minimization is the shower direction returned by the plane wave reconstruction.

The estimated shower core position and direction from the plane wave reconstruction are used as a seed for the more advanced IceTop likelihood reconstruction. This likelihood reconstruction, which includes factors for shower front curvature, takes into account information related to the deposited tank charge spatial distribution, timing information, and charge thresholds.

3.2.1 Spatial Distribution

The expectation value of the deposited tank charge as a function of distance from the shower core, called the shower lateral distribution (LD), is characteristic of the shower primary energy, composition, and state of shower development (often referred to as the shower "age"). To extract this information, the lateral distribution is parameterized by the empirically derived lateral distribution function (LDF)[47]:

$$S_{\mathrm{LDF}}(R) = S_{\mathrm{ref}} \cdot \left(\frac{R}{R_{\mathrm{ref}}}\right)^{-\beta-\kappa\log_{10}(R/R_{\mathrm{ref}})} \tag{3.6}$$

where $S_{\mathrm{LDF}}(R)$ is the expected signal deposited a distance R away from the shower core, S_{ref} is the charge deposited at a reference distance R_{ref} away, and β and κ are additional slope and curvature parameters. Previous studies have found that fixing the value of κ has little impact on the likelihood reconstruction. So while β is a free parameter, κ is set fixed to 0.303 [36]. A reference distance of $R_{\mathrm{ref}} = 125$ m was chosen in order to minimize the correlation between the S_{ref} and β fit parameters. The expected signal at this distance is denoted by S_{125} and is highly correlated with the primary particle energy. Thus, S_{125} serves as a good energy proxy and is used so throughout this analysis. Figure 3.6 shows reconstructed S_{125} values as a function of primary particle energy. An example shower lateral distribution and corresponding fitted LDF is shown in Figure 3.3.

The charges measured by IceTop are affected by snow accumulation on top of the tanks. On average, the snow height on the tanks increases 20 cm every year [36]. This causes a reduction in the measured signal in the tanks over time as air shower particles lose energy in, or are fully absorbed by, the snow atop the surface tanks. To correct for this reduction in the measured signal,

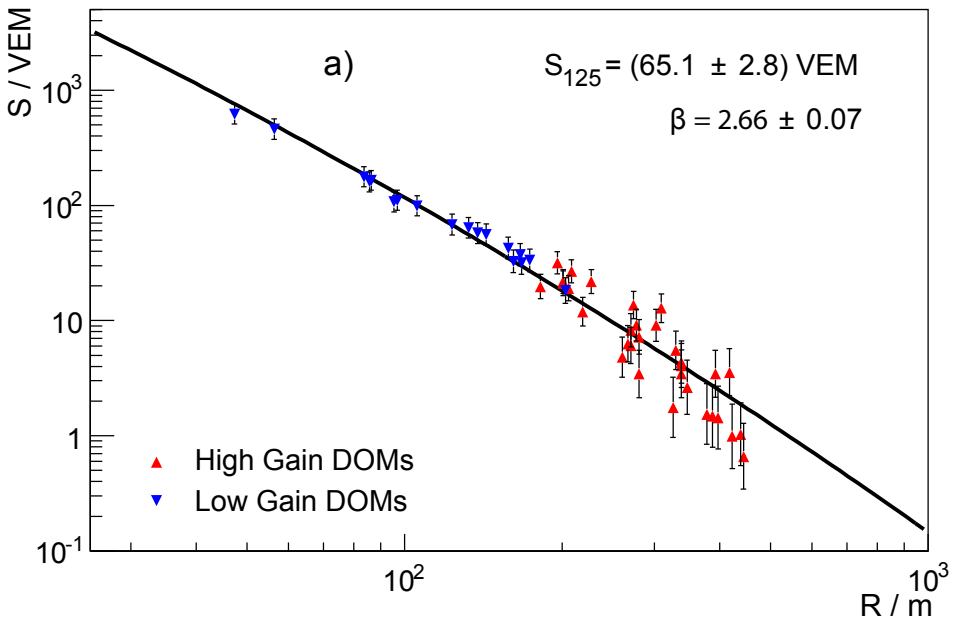


Figure 3.3: Fitted lateral distribution for an air shower detected by IceTop. The data points show the deposited tank charges and tank distance for this event. The black curve indicates the fitted LDF from Equation 3.6. The fitted LDF parameters for this air shower are $S_{125} = 65.1$ VEM and $\beta = 2.66$. Figure from [36].

an exponential absorption model is used to modify the expected charge in each tank [36]. The corrected tank signals are given by:

$$S_{\mathrm{corr},i} = S_i \cdot \exp\left(\frac{d_{\mathrm{snow},i} \sec\theta}{\lambda_{\mathrm{eff}}}\right) \tag{3.7}$$

where S_i is the measured charge in the i-th tank, $S_{\mathrm{corr},i}$ is the corrected charge, $d_{\mathrm{snow},i}$ is the depth of the snow above the i-th tank, θ is the shower zenith angle, and λ_{eff} is an effective absorption length in snow. The effective absorption length of $\lambda_{\mathrm{eff}} = 2.25$ m was found to give a consistent S_{125} distribution across different portions of IceTop with varying snow coverage, and is used in this work.

In addition to accounting for tank snow coverage, two additional factors related to the measured signals in each tank are also considered in the likelihood reconstruction: un-triggered tanks and saturated tanks. When large amounts of Cherenkov light are produced in IceTop tanks, the tank PMTs begin to exhibit a non-linear response to input light signal. This effect, known as PMT

saturation, results in an underestimate of the true signal size. In contrast, for small signals there is a probability that a tank with a certain expected signal will have a measured signal less than the tank threshold and not trigger. Thus, the distribution of tanks that do not trigger also contains meaningful information. In order to account for these effects, tanks which do not trigger, or are saturated, are accounted for separately in the likelihood reconstruction.

3.2.2 Timing Distribution

In addition to the spatial charge distribution, the tank hit timing distribution is also included in the likelihood fit. The time of each tank hit can be written as

$$t_{\text{fit},i} = t_0 + \frac{(\vec{r}_i - \vec{r}_0) \cdot \vec{n}}{c} + \Delta t\Big(R\Big) \tag{3.8}$$

where t_0 is the time the shower core reaches the ground, $\vec{r}_0$ is the shower core position at ground level, $\vec{r}_i$ is the position of the i-th tank, and $\Delta t(R)$ is an additional term related to the curvature of the shower front as a function of the distance to the shower axis. For the case of $\Delta t(R) = 0$, we recover the plane wave approximation in Equation 3.2. The additional curvature term is given by the empirically derived formula:

$$\Delta t(R) = aR^2 + b\left(1 - \exp\left(-\frac{R^2}{2\sigma_{\text{curve}}}\right)\right) \tag{3.9}$$

where $a = 4.823 \times 10^{-4}$ ns/m^2, $b = 19.41$ ns, and $\sigma_{\text{curve}} = 83.5$ m [36]. Based on experimental data, fluctuations in the arrival times are given by

$$\sigma_{t,i} = 2.92 \text{ ns} + 3.77 \times 10^{-4} \text{ ns} \cdot (R_i/m)^2 \tag{3.10}$$

where R_i is the distance of the i-th tank to the shower axis [48]. An example fitted tank hit time distribution is shown in Figure 3.4.

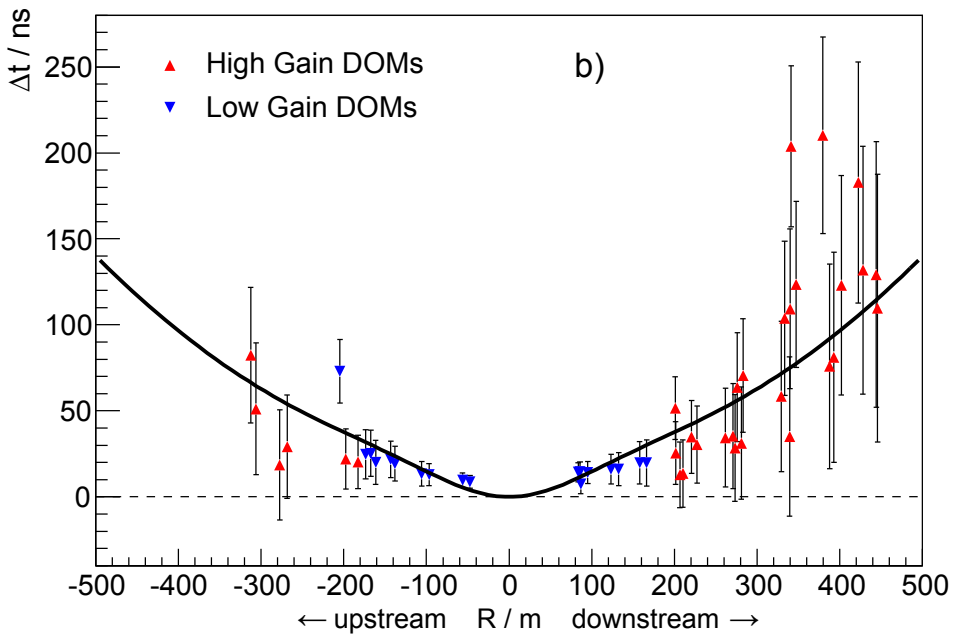


Figure 3.4: Fitted tank hit timing distribution for an air shower detected by IceTop. The data points show the tank hit times as a function of tank distance from the shower core for this event. The black curve indicates the fitted timing distribution from Equation 3.8. The terms "upstream" and "downstream" refer to tanks being hit before and after the shower core reaches the ground. Figure from [36].

3.2.3 Likelihood Fit Procedure

The full air shower reconstruction likelihood, $\mathcal{L}$, is the product of four individual terms:

$$\mathcal{L}(\vec{r}_0, t_0, \vec{n}, S_{125}, \beta) = \mathcal{L}_q \times \mathcal{L}_t \times \mathcal{L}_{\text{nohit}} \times \mathcal{L}_{\text{sat}} \tag{3.11}$$

where the $\mathcal{L}_q$ term relates to the probability for the triggered tanks to measure the observed charge, $\mathcal{L}_t$ accounts for the expected tank hit timing distribution, and $\mathcal{L}_{\text{nohit}}$ and $\mathcal{L}_{\text{sat}}$ are terms for the un-trigger and saturated tanks, respectively. For a detailed discussion regarding these individual likelihood terms, see [36].

The full IceTop likelihood is a function of seven free parameters: the shower core position $\vec{r}_0 = (x_0, y_0)$, time the shower passes through the shower core position t_0, shower direction $\vec{n} = (\theta, \phi)$, and LDF parameters S_{125} and β. In general, high-dimensional likelihood spaces can be complicated and have multiple local extrema. To help convergence towards global minimum, an iterative fitting procedure is used. At each step, only a subset of the full parameter space is explored and the fitted

parameters from the previous step are used as the seed for the following step to ensure the best possible starting point for the minimizer at a given step. Specifically, the fitting procedure steps are:

1. The shower core and direction are seeded with the results from the plane wave reconstruction. S_{125}, β, and the shower core position are fit, while the shower direction is kept fixed. The shower core position is restricted to a box with 400 m sides about the plane wave core position.

2. Use the fitted parameters from the previous step as a seed. All seven likelihood parameters are fit. The shower core position is restricted to a box with 30 m sides about the seeded core position. β is bound between 2.0 and 4.0.

3. Use the fitted parameters from the previous step as a seed. The shower direction is kept fixed, while S_{125}, β, and the shower core position are refit. The shower core position is restricted to a box with 70 m sides about the seeded core position.

The output of this maximization are estimates for the parameters $\vec{r}_0 = (x_0, y_0)$, t_0, $\vec{n} = (\theta, \phi)$, S_{125}, and β.

3.2.4 Reconstruction Performance

The resolution of the reconstructed shower core position and direction from the IceTop likelihood reconstruction are shown in Figure 3.5. Note that the reconstruction procedure yields a shower core position resolution between 3-9 meters, and a shower direction angular resolution between 0.2-0.45 degrees.

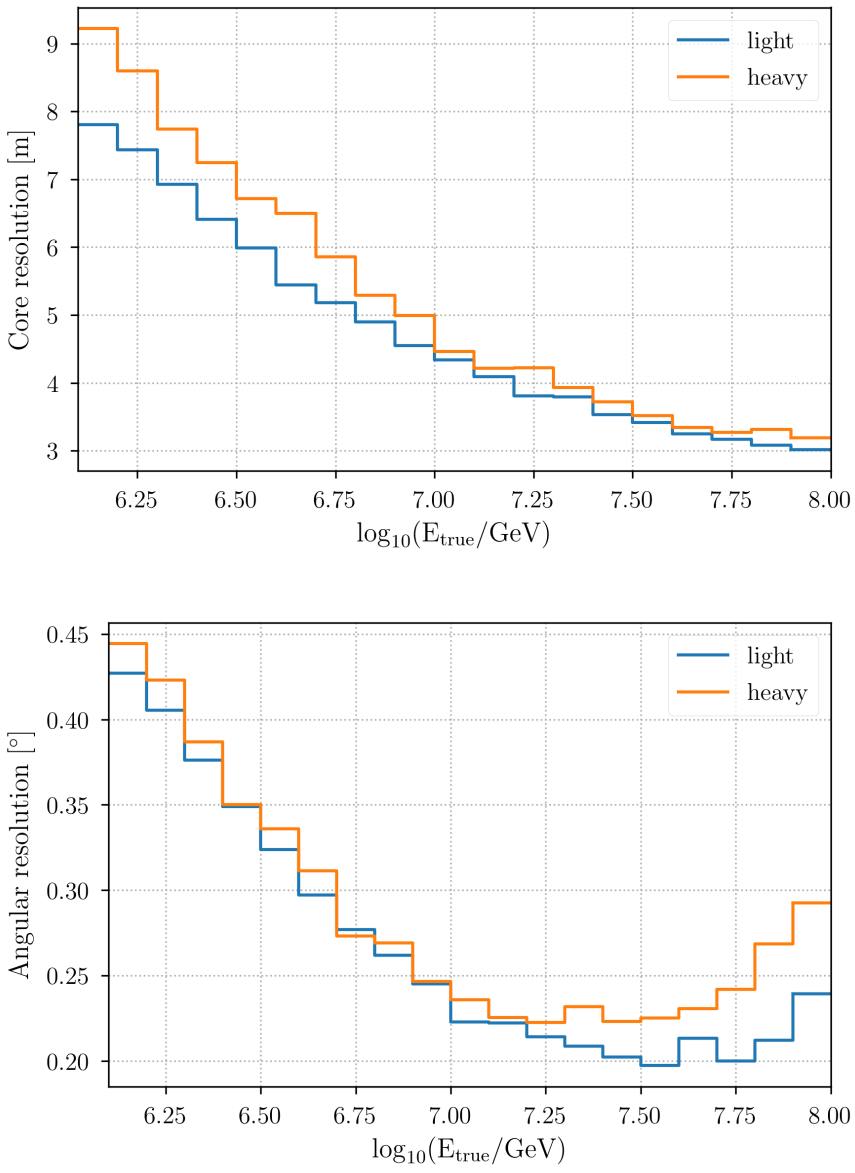


Figure 3.5: Air shower reconstruction performance. Top: Reconstructed shower core position resolution. Bottom: Reconstructed shower direction angular resolution.

The reconstructed S_{125} LDF parameter is highly correlated with the primary cosmic ray energy and serves as a good energy proxy. Figure 3.6 shows the distribution of reconstructed S_{125} values versus the true primary particle energy.

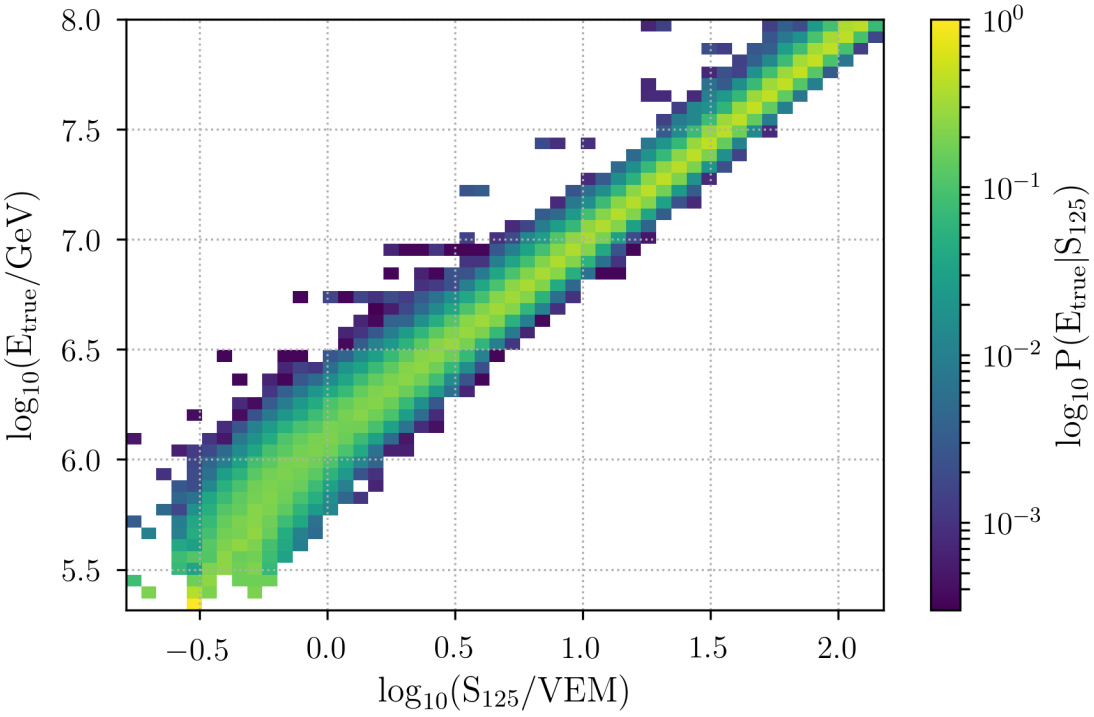


Figure 3.6: Histogram of the reconstructed $\log_{10}(S_{125})$ vs. true primary energy for simulated air shower events. The histogram is normalized such that each column sums to one. S_{125} is highly correlated with energy and thus serves as a good energy proxy.

3.3 In-Ice Energy Loss Reconstruction

The in-ice array detects Cherenkov radiation emitted by high energy muons as they make their way through the Antarctic ice. The amount of light that each DOM in the array detects is related to the energy loss per unit length of the muons traveling through the ice (often called simply the "energy loss" and denoted by dE/dX), which is a function of the muon multiplicity of the bundle. Specifically,

$$\frac{dE_{\mu,B}(X)}{dX} = \int_{E_{min}}^{E_{max}} \frac{dE_{\mu}(X)}{dX} \frac{dN_{\mu}(X)}{dE_{\mu}(X)} dE_{\mu}(X) \tag{3.12}$$

where $\frac{dE_{\mu,B}(X)}{dX}$ is the muon bundle energy loss at a slant depth of X, E_{min} is the minimum energy at ground level required for a muon to reach a depth of X, and E_{max} is the maximum muon energy that can be produced from a primary particle with energy E_0 and mass A. The muon bundle energy loss is the primary composition-sensitive variable used in this work.

A maximum-likelihood based technique is used to reconstruct the muon bundle energy loss, given the observed signals from the in-ice array DOMs. As muons pass through the ice, the Cherenkov radiation emitted at each point along their trajectory can contribute to the observed signals in each of the in-ice DOMs. The reconstruction algorithm partitions the muon bundle track into segments, each 20 meters long, and treats each segments as a cascade that contributes to the overall signal observed by the in-ice DOMs. Figure 3.7 depicts the segmented track contributions to the observed DOM signals. Adding up the contributions from each segment, the observed signal in the in-ice DOMs is given by:

$$Q_{\mathrm{exp},i} = \sum_j \Lambda_{ij} x_j + \rho_i \tag{3.13}$$

where $Q_{\mathrm{exp},i}$ is the expected observed charge series in the i-th DOM, x_j is the energy loss of the j-th track segment, ρ_i is the average charge expected from noise in the i-th DOM, and Λ_{ij} is the expected charge at the i-th DOM from a 1 GeV cascade in the j-th track segment. The transmission coefficients, Λ_{ij}, are obtained from using spline smoothed simulation [49].

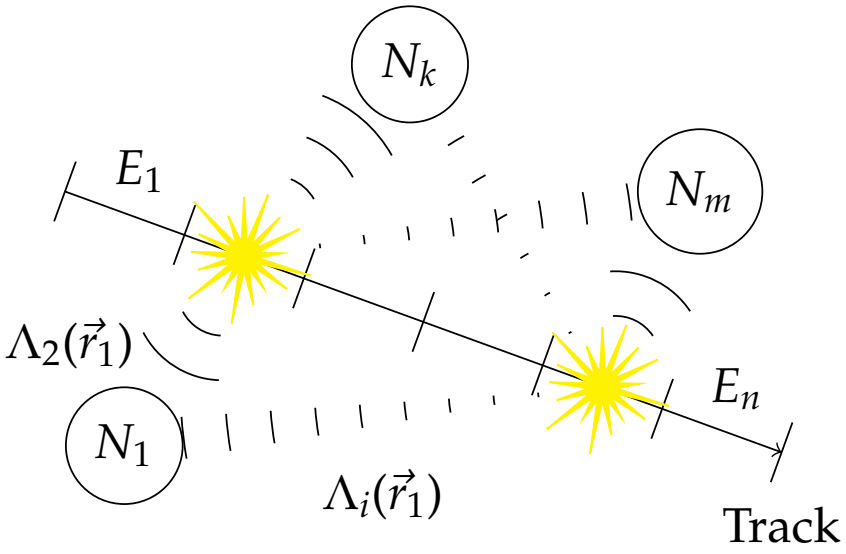


Figure 3.7: Diagram depicting the segmented energy loss reconstruction for the in-ice array. Circles represent DOMs that measured N_k photoelectrons. The muon bundle track is segmented into bins of 20 meters in length and are indicated by the periodically spaced lines perpendicular to the track direction. The energy deposited in the n-th bin is given by E_n.

The aim of the reconstruction algorithm is to determine the energy losses, x_j, in Equation 3.13, given the observed DOM signals, Q_i. This is achieved by maximizing the Poisson likelihood:

$$\mathcal{L} = \prod_i \frac{\left(Q_{\text{exp},i}\right)^{Q_i}}{Q_i!} \exp(-Q_{\text{exp},i}) \tag{3.14}$$

where Q_i and $Q_{\text{exp},i}$ are the observed and expected signals in the i-th DOM from Equation 3.13. The above likelihood is maximized using the pre-conditioned conjugate gradient (PCG) method.

The segmented in-ice muon energy losses from the likelihood reconstruction are then fit to produce an energy loss profile as a function of slant depth. Previous studies have demonstrated that the composition resolving power is better when assessed at smaller slant depths [35]. So for the purposes of this analysis, we take the muon energy loss to be evaluated near the top of the in-ice array, for vertical showers, at a slant depth of $X = 1500$ m.

Figure 3.8 shows the reconstructed energy loss as a function of true energy for simulated proton and iron induced showers. We can see, for a given true energy, iron induced showers have, on average, a larger muon energy loss than proton showers. In Chapter 4, we will discuss how dE/dX

is used to provide air shower primary mass identification.

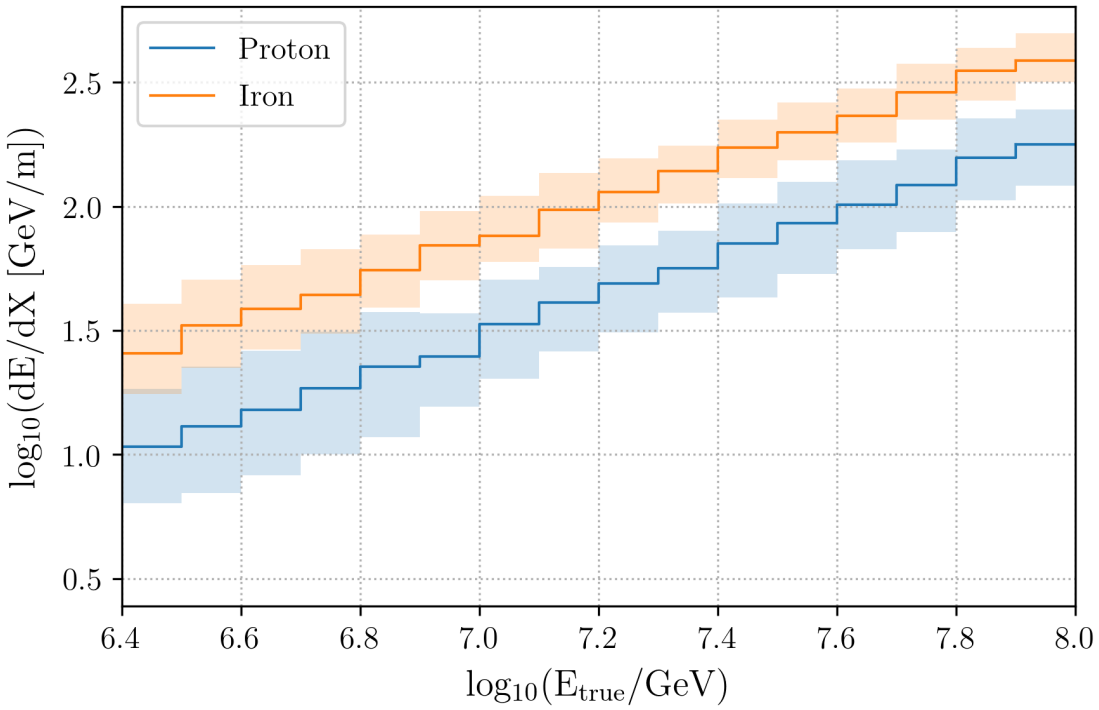


Figure 3.8: The reconstructed in-ice muon bundle energy loss as a function of true energy for proton and iron induced air shower simulation. For each energy bin the solid line indicates the median energy loss, while the shaded region shows the 68% containment of the energy loss distribution in that bin.

3.4 Data Quality Cuts

Quality cuts are used to filter out events for which either the surface or in-ice reconstructions are known to perform poorly. In particular, the reconstruction algorithms perform best with coincident events in which the shower trajectory passes through the instrumented area and volume of the surface and in-ice arrays. Several of the quality cuts applied are to ensure the rejection of un-contained events [10]. The following set of data quality cuts are used for this analysis.

1. Minimum number of hits to perform quality reconstructions:
    - At least 5 IceTop stations must trigger
    - At least 8 in-ice array DOMs must trigger
2. Containment cuts to ensure that air showers pass through both the surface and in-ice arrays:
    - The reconstructed shower core position must lie within the IceTop area
    - The reconstructed shower track must pass through the in-ice array volume
    - At least one IceTop tank must have a snow corrected signal greater than 6 VEM
    - The neighboring tank in the same station as the largest signal tank must have a snow corrected signal of at least 4 VEM
    - The Icetop station with the largest deposited charge cannot be at the edge of the detector
3. Event topology cut:
    - The fraction of triggered stations within a circle about the center-of-gravity of the shower with radius to the farthest away hit station must be greater than 0.2
4. General reconstruction quality cuts:
    - The IceTop maximum likelihood reconstruction is required to converge
    - The LDF parameter β is required to be between 1.4 and 9.5
    - The in-ice array reconstruction is required to converge with an $\log_{10}(r \log l) < 2.0$

- The total predicted charge from the in-ice reconstruction must be at least 90% of the measured charge

Chapter 4

Composition & Energy Estimation

A crucial component of this analysis is the ability to assign a primary particle energy and composition from cosmic ray air showers detected with IceCube. Given the availability of air shower simulation and the corresponding simulated detector response, supervised machine learning is a useful technique for automating these tasks. This chapter gives an overview of supervised machine learning and discusses its application and use within this analysis. Specifically, we'll cover how reconstructed parameters from Chapter 3 are mapped to primary particle energy and composition.

4.1 Supervised Machine Learning

Machine learning is the subset of artificial intelligence in which algorithms learn from data, rather than being explicitly programmed. Given a set of data relevant to a specific task, machine learning algorithms find statistical structure in the dataset to help accomplish the task. This is in contrast with non-machine learning approaches in which a program with hand crafted rules is used to accomplish the task.

Supervised machine learning is the subset of machine learning where a dataset X of labeled examples is used to develop an algorithm that maps between an input example and the corresponding label (sometimes referred to as the target). The goal of supervised machine learning is to develop a mapping f that yields predicted labels $\hat{y}$ that generalizes well to future, yet unseen data.

$$\hat{y} = f(X) \tag{4.1}$$

The dataset X, called the training dataset, is formatted such that each example in the dataset is a row, while each column consists of a property that describe the examples. The columns are often referred to as the *features* of the training dataset.

Supervised machine learning problems are typically divided into two categories: classification problems and regression problems. The distinction between these two categories is based on the label which is predicted. Classification tasks are those in which the label we wish to predict is discrete (e.g. labeling primary particles as light or heavy), while regression tasks are those for which the label is continuous (e.g. predicting primary particle energy). Figure 4.1 illustrates the difference between typical classification and regression problems.

4.2 Model Validation

The goal of developing a model that generalizes well to unseen data in supervised machine learning is different than in other optimization problem, e.g. maximum likelihood methods, where the objective is typically to find the best fit solution, given a dataset. In supervised machine learning, we want the model to perform well not only on the training dataset, but also generalize well to

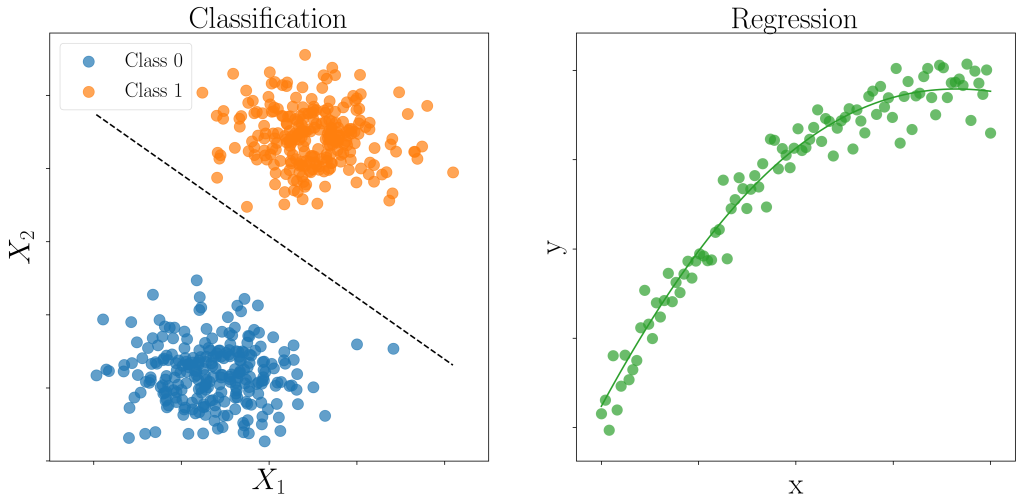


Figure 4.1: Diagram illustrating classification and regression tasks. Left: Example dataset where each sample belongs to one of two classes (data points are shown in different colors to distinguish their class label). The dashed line shows a decision boundary for a linear model that is used to identify the two classes. Everything below the line is classified as class 0, while points above the line are assigned to class 1. Right: Example regression problem where given (x, y) data points, a model finds a polynomial function to best fit the data (shown by the solid line).

new data. In order to ensure the predictions from a trained model generalize well, and don't just describe the training dataset, model validation is paramount.

4.2.1 Cross Validation

One of the most commonly used ways to quantify how well a model generalizes is the k-fold cross validation (CV) technique [50]. In k-fold cross validation, a training dataset is randomly partitioned into k different subsets, or folds. This partitioning is such that each fold has approximately the same number of samples. Next, $k-1$ of the folds are used as a dataset to train a model. The one fold that was not used in training (called the test or validation fold) can then be used to calculate a model performance metric for the model trained on the $k-1$ other folds. In this manner, the performance of the model on unseen data can be estimated. This process is then repeated $k-1$ times, each time choosing a different fold as the test fold. Finally, an average model generalization performance and performance uncertainty are obtained by calculating the mean and standard deviation of the model performance metrics for each iteration of the cross validation. Figure 4.2 illustrates the k-fold cross

validation algorithm for $k = 10$.

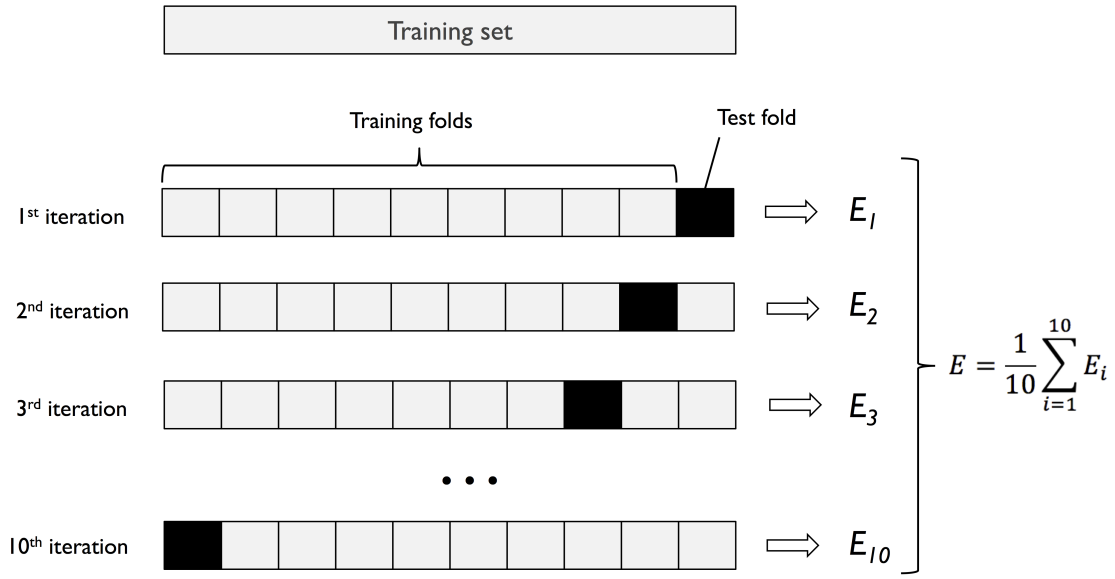


Figure 4.2: Illustration of the k-fold cross validation algorithm (diagram depicts the $k = 10$ case). Figure from [50].

For classification problems, a slight variant of the k-fold cross validation algorithm is often used. This variant, the stratified k-fold cross validation algorithm, differs from the standard k-fold algorithm only in how the training dataset is partitioned into folds. While the k-fold algorithm performs this partitioning randomly, the splitting in the stratified k-fold algorithm is done randomly while also preserving the percentage of samples for each class. This way, each fold has approximately the same number of samples belonging to each class as the entire training dataset. Classification problems in this work are validated using stratified 10-fold cross validation.

4.2.2 Hyperparameter Optimization

Machine learning algorithms have two distinct types of parameters: those that are learned during the training process from data and those that are free parameters which parametrize the algorithm itself. Parameters that aren't learned directly from data are referred to as hyperparameters. For

example, a one-dimensional polynomial regression model takes the form:

$$\hat{y}(X) = \sum_{i=0}^{n} w_i x^i \tag{4.2}$$

where x is the independent variable, w_i is the coefficient for the i-th power term, and n is the largest power in the polynomial to consider. During the learning process, the polynomial coefficients w_i are determined from training data, but the maximum order n is a free parameter that must be specified ahead of time. In this example case, the maximum polynomial degree n is a hyperparameter for the model.

In many cases hyperparameters are related to the capacity of a model. In the polynomial regression case, the larger the maximum polynomial degree n, the more complex a relationship can be modeled (shown in Figure 4.3). For smaller model complexities, the model cannot accurately represent the underlying structure of the data. This results in both a low training and testing cross validation score and the model is said to *underfit* the training dataset. As the model complexity increases, so to does the average training and testing performance. However, as the model capacity becomes very large, the model is able to capture particularly complex structure that is specific to the dataset used to train the model. In this case, the model performs exceedingly well on the training dataset, but does not generalize well, resulting in a poor testing cross validation score. Here the model is said to *overfit* the training dataset. Figure 4.4 illustrates the typical training and testing performance as a function of model complexity. The ideal values for model hyperparameters are those that avoid both under and overfitting.

Hyperparameter optimization is often done by choosing the combination of parameter values that optimizes the testing cross validation score for the problem at hand. For example, in regression tasks a common scoring function is the average mean squared error (MSE), while the fraction of correctly labeled examples (the accuracy score) is often used in classification tasks.

This optimization is done numerically by sampling the space of possible hyperparameter combinations. There are several different ways to go about sampling the hyperparameter space: grid search, random search [52], and model-based optimizations. For the grid search method, the user selects a finite set of hyperparameter values to explore. A model is then trained for every combina-

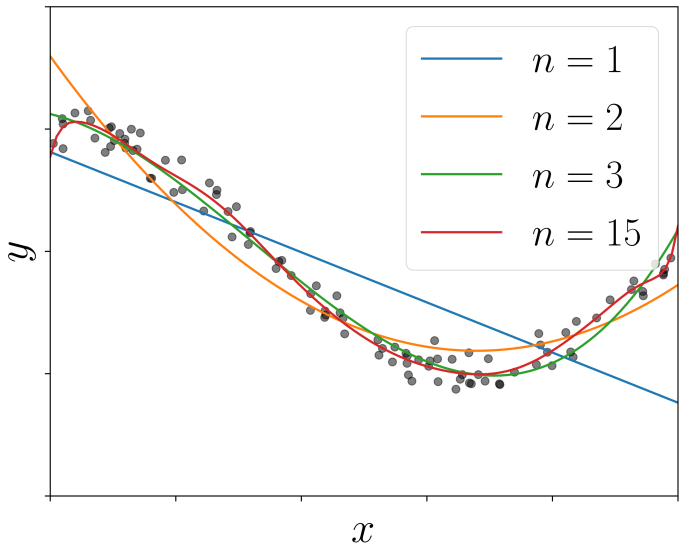


Figure 4.3: Example polynomial regression models for various values for the maximum polynomial degree n. The same training data is used for each model and indicated by the gray data points. For smaller values like $n = 1, 2$ the model has too low a capacity and cannot fully capture the variation in the data, i.e. underfitting the training data. For larger values like $n = 15$, the model becomes very complex and starts to fit fluctuations in the training, i.e. overfit the training data. An optimal value of $n = 3$ avoids both under and overfitting.

tion of hyperparameter values. The combination that has the best validation score is then selected as the optimal set of parameters. The downside to the grid search method is that the computational cost of performing the search grows exponentially with the number of hyperparameters. The random search method randomly samples hyperparameter values according to user-defined probability distributions (e.g. uniform or log-uniform). The random search method will converge faster when only a few hyperparameters have a significant influence on the model performance, which is typical [52]. Figure 4.5 illustrates how the grid and random search methods compare when a hyperparameter has little overall effect on model performance. For this work, we make use of the random search method. We'll now discuss the training features, models, and model validation used to assign cosmic ray air showers a primary particle energy and composition.

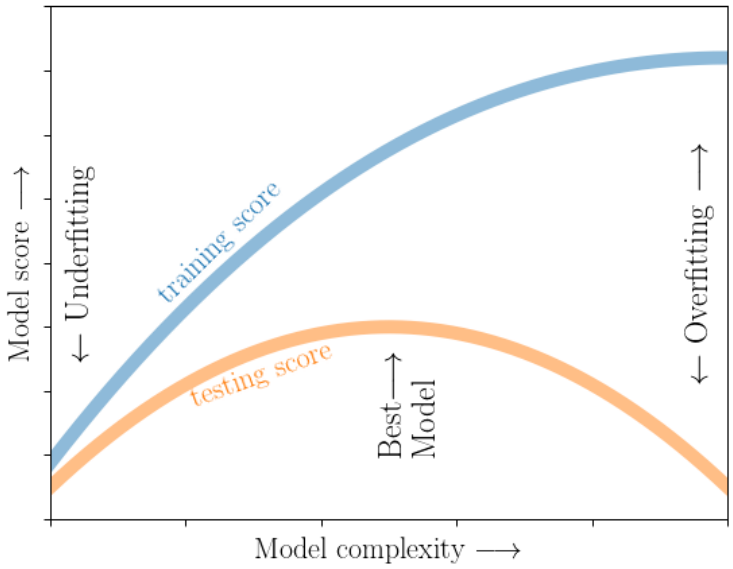


Figure 4.4: Diagram depicting typical model performance as a function of model complexity. For low model complexity, both the training and testing validation scores are low (underfitting). For high model complexity, the training score continues to increase while the validation score reaches a maximum and then begins to decrease (overfitting). Optimal model complexity avoids both under and overfitting. Figure adapted from [51].

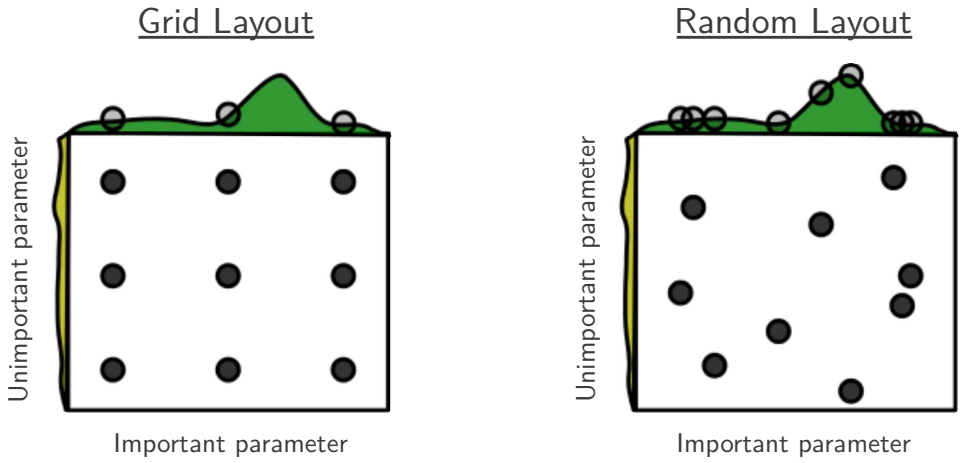


Figure 4.5: Comparison of grid and random search methods for hyperparameter optimization. (Left) grid search is run for every combination of specified hyperparameter values. (Right) Random search is run over randomly chosen hyperparameter values in accordance with a user-defined probability distribution. Figure from [52].

4.3 Composition Classification

On average, the more massive a cosmic ray particle, the more muon-rich the induced air shower will be for a given energy. Thus, the ability to distinguish cosmic ray composition comes from mass separation in the energy-muon number space. In Chapter 3 we discussed the reconstruction algorithms used to engineer both the energy proxy parameter, S_{125}, and muon number proxy, dE/dX, for detected air showers. Figure 4.6 shows the reconstructed $\log_{10}(S_{125})$ and $\log_{10}(\mathrm{dE/dX})$ distributions for light and heavy induced air shower simulation for various primary energies. We

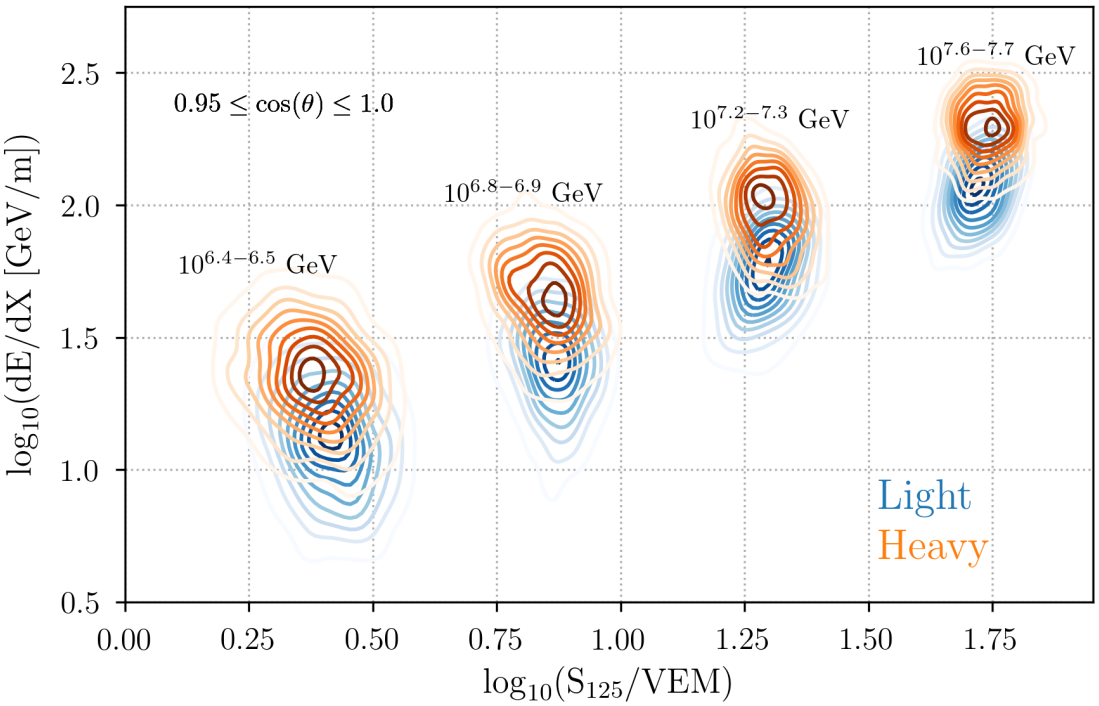


Figure 4.6: The reconstructed $\log_{10}(S_{125})$ and $\log_{10}(\mathrm{dE/dX})$ for light (blue) and heavy (orange) induced air shower simulation for various ranges of primary energies. There is composition separation power between the light and heavy air showers in the $\log_{10}(S_{125})$-$\log_{10}(\mathrm{dE/dX})$ space. The reconstructed $\log_{10}(S_{125})$ and $\log_{10}(\mathrm{dE/dX})$ values vary slightly with the true zenith angle, θ, of the cosmic ray primary. Here we show air showers with $0.95 \leq \cos(\theta_{\mathrm{true}}) \leq 1.0$.

can see that for a given primary energy, air showers from heavier cosmic ray primary tend to have a larger value for dE/dX. However, the same value for dE/dX can be found for both light and heavy air showers at different energies. The muon content, in terms of the reconstructed in-ice energy loss, of air showers induced by heavier cosmic ray primaries is similar to that of air showers

induced by a lighter primary at higher energies. This illustrate the need to measure both energy and muon number proxies *simultaneously*.

To distinguish between light and heavy induced air showers, reconstructed variables from air shower simulation are used as a training dataset for a classification model to predict whether a detected air shower event originated from a light or heavy primary. The model used for this task is the boosted decision tree (BDT) classification algorithm. The BDT implementation in the XGBoost Python library is used[1]. The training features of the BDT classification dataset are: the energy proxy $\log_{10}(S_{125})$ from the IceTop reconstruction, the muon number proxy $\log_{10}(\mathrm{dE/dX})$ from the in-ice array reconstruction, and the cosine of the reconstructed shower zenith angle $\cos(\theta)$ from the IceTop reconstruction. It's worth noting that several additional candidate training features (e.g. total deposited in-ice charge, charge deposited at different depths in the in-ice array, etc.) were also evaluated, however none were found to improve the composition classification accuracy and so were not included in the final set of training features.

The BDT algorithm is an ensemble learning method. In ensemble methods, a collection of models are each trained and their predictions are aggregated together to obtain a single output that can have better performance than any of the individual models that make up the ensemble. In the case of a BDT, the ensemble is composed of many decision trees.

Decision tree classifiers are trained by sequentially partitioning a dataset into nodes such that samples with the same labels are grouped together. To begin, all samples in the training dataset begin in a single node, the root node. The root node is then split into two child nodes according to a splitting condition on a feature in the dataset. This splitting condition is chosen such that it minimizes the impurity of the two resulting child nodes. The metric used in this analysis to quantify the impurity of a node is the Gini impurity [53]:

$$I_G = 1 - \sum_{i=1} p_i^2 \tag{4.3}$$

where p_i is the fraction of samples belonging to class i in the node. The Gini impurity is a measure of how mixed the classes of a node are. For example, a node with samples belonging to a single class is completely pure and has an impurity of $I_G = 0$. A node with an equal mixture of two

[1] Version 0.72 of XGBoost was used for this work

classes is maximally impure and has $I_G = 0.5$. There are other potential impurity metrics one could choose to use, however this choice generally has a negligible impact on model performance.

The process of node splitting is then applied to each child node until a stopping condition is met. There are several different stopping conditions that may be met: a minimum required decrease in the impurity to split a node, or requiring a minimum number of samples in a node to perform a splitting. However, the stopping condition most closely associated with the decision tree model complexity is the maximum number of splits that are allowed. This maximum number of splits, or maximum depth of the tree, is a hyperparameter of all tree-based learning models.

The series of decision trees that are sequentially trained in the BDT have their predictions aggregated in an additive fashion. That is

$$f(x) = \sum_{m=1} \gamma_m f_m(x) \tag{4.4}$$

where $f(x)$ is the full BDT, $f_m(x)$ are the individual decision trees, and γ_m are expansion coefficients that are chosen to minimize a loss function that quantifies the error in the BDT model predictions. A logistic loss function is used in this analysis. The process of sequentially training models and aggregating their prediction is called boosting. For a thorough discussion of boosted decision trees, and additive models more generally, please see [53].

In order to determine the optimal hyperparameters for the composition classification BDT, a random search is performed. Figure 4.7 shows the results of the hyperparameter optimization. Also shown in the figure is a comparison of the cross-validated training and testing accuracies as a function of the BDT tree maximum depth. While the random search yields an optimal value for the maximum depth of 6, to avoid overfitting a maximum depth of 4 is used instead. The final chosen hyperparameter values for the BDT are shown in Table 4.1.

The classification accuracy for the BDT composition model as a function of true primary energy is shown in Figure 4.8. The solid line and error bands for each energy bin indicate the mean and standard deviation of the 10-fold cross validation testing accuracies, respectively. Note that the model has a roughly uniform classification accuracy of $\sim$75% across the energy range relevant for this analysis.

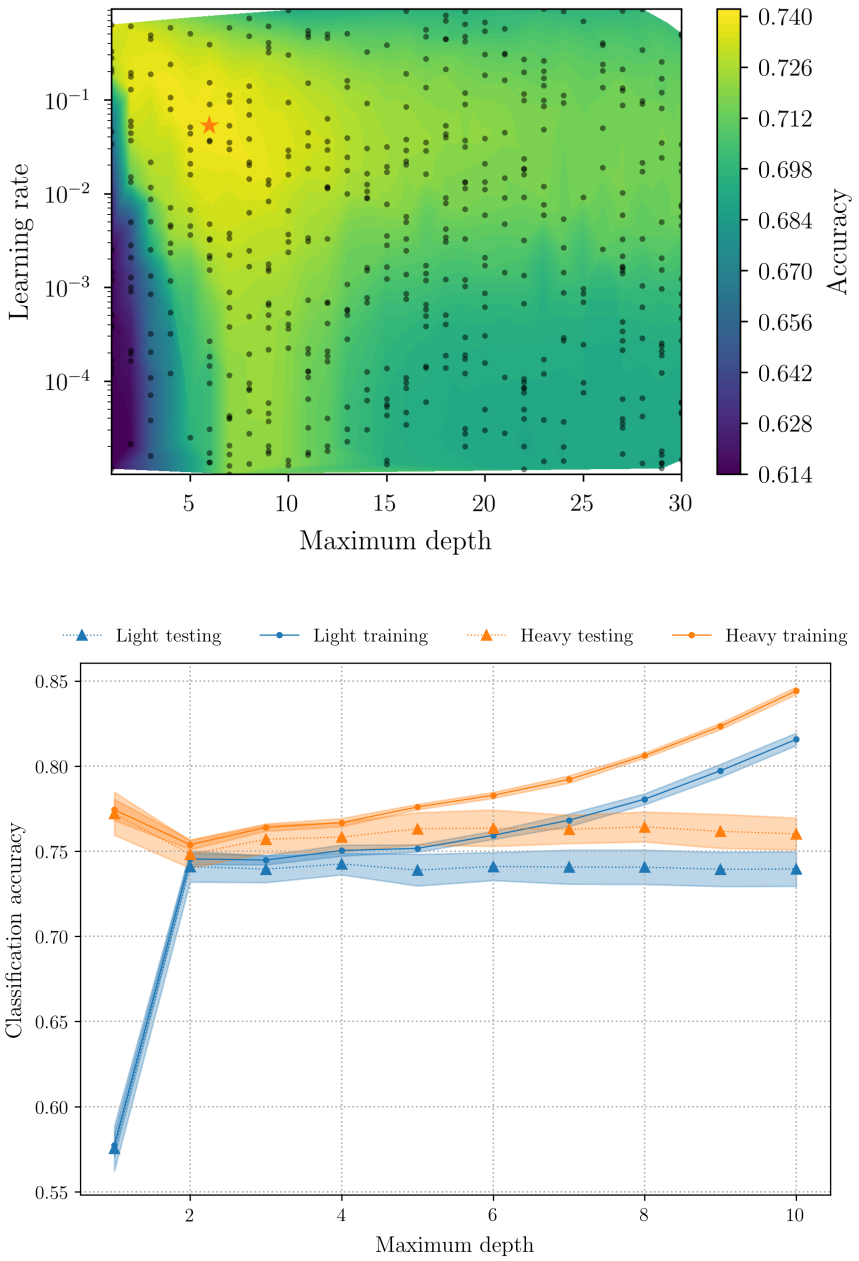


Figure 4.7: Top: Random search hyperparameter optimization for composition classifier. The black data points show individual hyperparameter combinations tested. The color scale indicates the average cross-validation accuracy score on the test fold. The orange star indicates the hyperparameter combination that maximizes the testing accuracy. Bottom: Validation curve displaying the average accuracy score for the composition BDT classifier on the testing and training CV folds as a function of the maximum tree depth. While the random grid search optimization yields an optimal maximum depth of 6, a maximum depth of 4 was chosen in order to reduce overfitting. A learning rate of 0.05 is used here.

Parameter Name	Value
Number of decision trees in BDT	150
Node splitting criterion	Minimize Gini impurity
Maximum depth of trees	4
Minimum samples required in a node	1
Boosting loss function to minimize	Logistic loss
Boosting learning rate	0.05
Minimum required loss reduction	0

Table 4.1: Tuned hyperparameters for boosted decision tree used for composition classification. Model implementation from XGBoost.

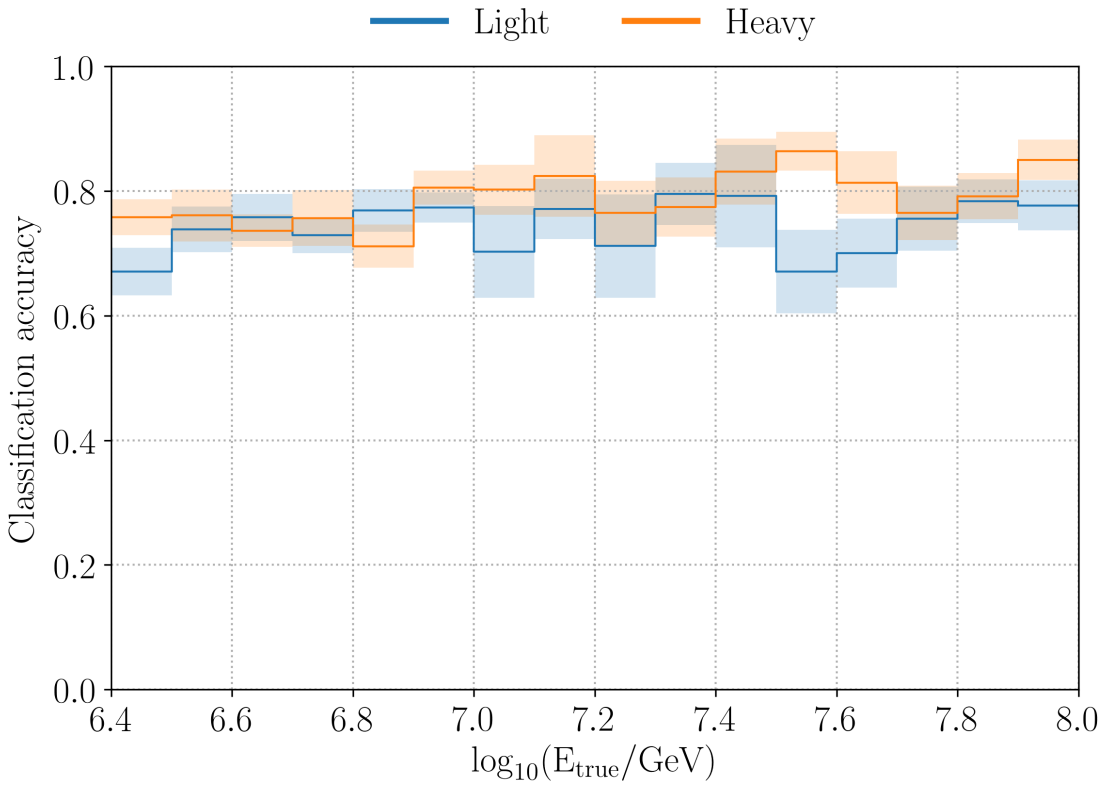


Figure 4.8: Composition classification accuracy as a function of energy. Classification accuracy is defined to be the fraction of examples which are correctly labeled. The solid lines indicate the average testing cross validation accuracy for each energy bin, while the shaded bands show the standard deviation of the accuracy scores.

4.4 Energy Reconstruction

A multiple linear regression model is used to assign each air shower event a reconstructed energy. In such a model, the target variable, in this case the primary particle energy, is represented as a combination of the dataset features. The training dataset features used for the energy regression model are the same as those used for composition classification in Section 4.3. The mapping between these features and the reconstructed primary energy is given by:

$$\log_{10}(E_{\text{reco}}) = \beta_0 x_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{01} x_0 x_1 + \beta_{02} x_0 x_2 + \beta_{12} x_1 x_2 \tag{4.5}$$

where x_0, x_1, and x_2 are $\cos(\theta)$, $\log_{10}(S_{125})$, and $\log_{10}(\text{dE/dX})$, respectively. The set of coefficients $\{\beta\}$ are determined from the training dataset using the method of ordinary least squares[2]. The optimized values for the coefficients are shown in Table 4.2. Inclusion of interaction terms between the features (e.g. the $x_0 x_1$ term in Equation 4.5) allows us to model when the effect a training feature has on the output variable is influenced by another feature.

Coefficient	β_0	β_1	β_2	β_{01}	β_{02}	β_{12}
Value	-0.64421306	0.3867754	0.38175117	0.43598762	-0.30033984	0.00855433

Table 4.2: Multiple linear regression model coefficients for energy reconstruction.

The performance of the trained energy regression model is illustrated in Figure 4.9. In addition, the energy reconstruction bias and resolution is shown in Figure 4.10. The energy resolution decreases with energy and ranges between $0.125 \leq \log_{10}(\text{E}_{\text{reco}}/\text{E}_{\text{true}}) \leq 0.05$, while the bias increases with energy. Note that both the energy bias and resolution are of order of the energy bin size used in this analysis, $\Delta \log_{10}(\text{E/GeV}) = 0.1$. Inclusion of additional feature interaction terms in the multiple linear energy regression model were shown to decrease the energy reconstruction bias at higher energies. However, this reconstruction bias reduction has a negligible affect on the results of this analysis. Therefore, additional feature interaction terms were not included. How the reconstructed energy bias and resolution are mitigated is the topic of Chapter 5.

Several more complex regression models, such as tree-based, support vector machine, and neural network models, where also evaluated as candidate models for primary particle energy regression.

[2]Version 0.19.1 of Scikit-learn [54] was used for the linear regression model

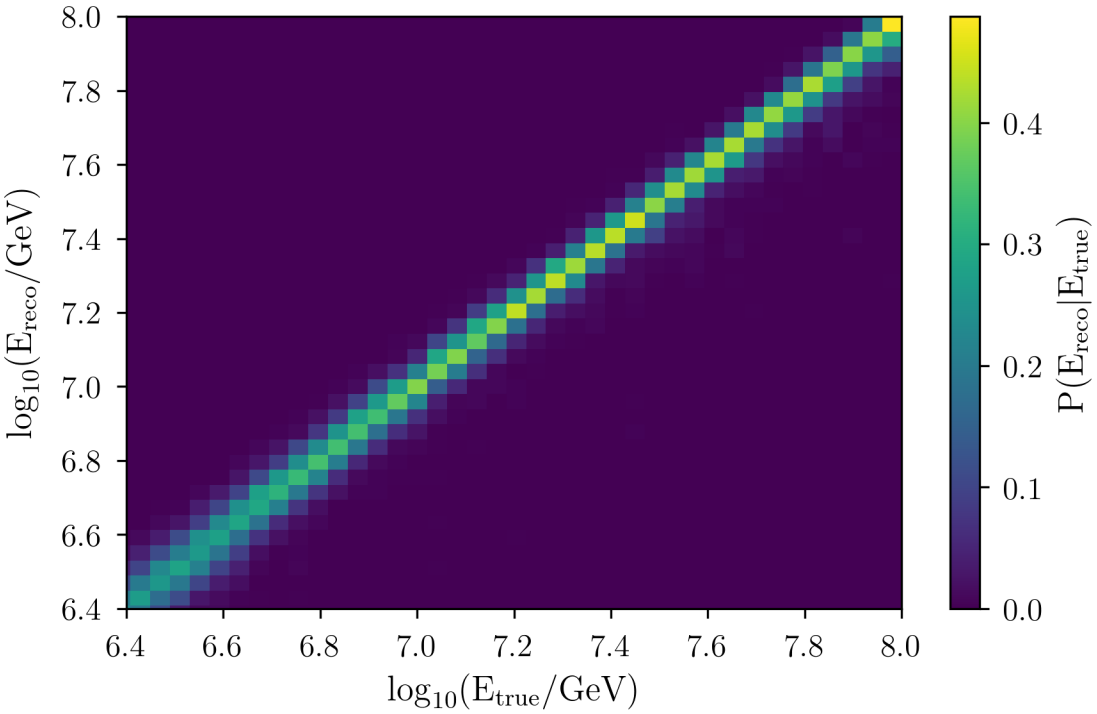


Figure 4.9: Distribution of true and reconstructed primary particle energies. The histogram is normalized such that each column sums to one

However, these models were found to have a similar performance to the more simple multiple linear regression model. This lead us to select the linear regression model for energy reconstruction in this work.

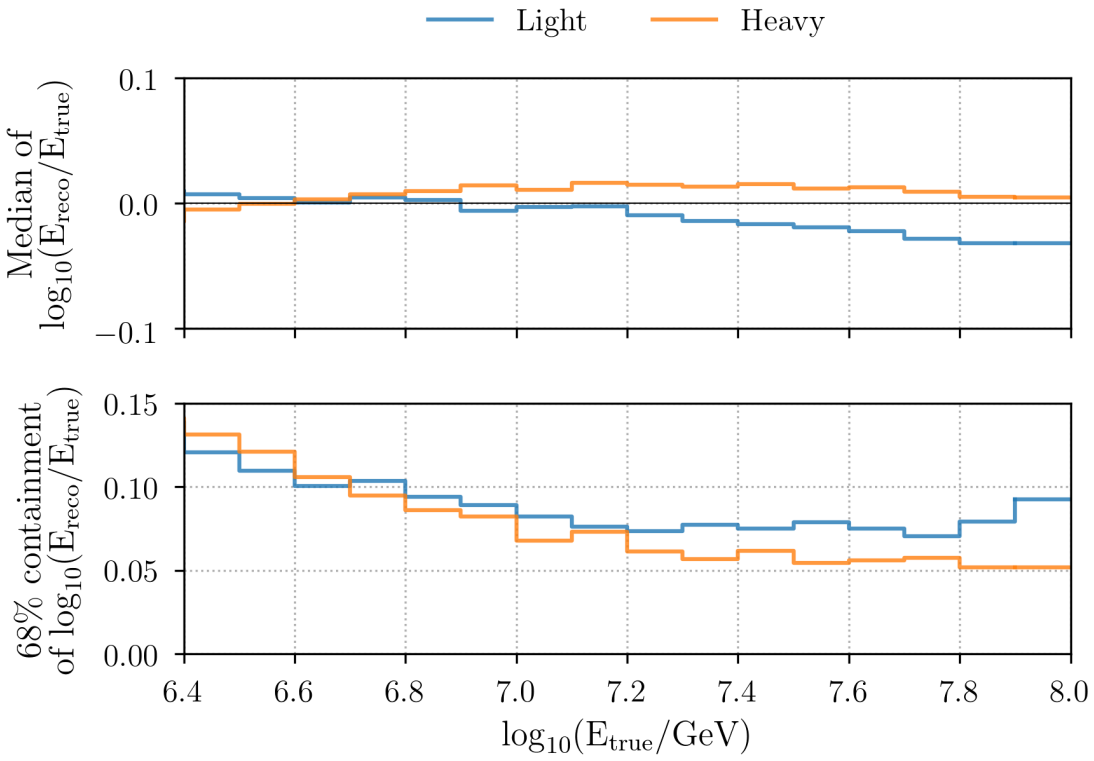


Figure 4.10: Reconstructed energy bias and resolution. For each true energy bin the reconstructed energy bias is the median of the $\log_{10}(E_{reco}/E_{true})$ distribution, while the reconstructed resolution is the 68% containment of the $\log_{10}(E_{reco}/E_{true})$ distribution for that true energy bin.

Chapter 5

Iterative Unfolding

In an ideal world, experimentalists would have access to a perfect detector that makes no error in measuring a desired quantity. However, real detectors have finite resolutions, characteristic biases that cannot be eliminated, and less than 100% detection efficiencies. Indeed, Figures 4.8 and 4.10 show we assign air showers the correct primary particle composition approximately 75% of the time and estimate the energy with an resolution of roughly 0.1 in $\log_{10}(E/\text{GeV})$. Unfolding is a method that corrects for these biases and resolutions. This chapter discusses the unfolding technique used to incorporate our knowledge of how our detector distorts our measurements to help mitigate these effects.

5.1 The Inverse Problem

When scientists make measurements there is typically some underlying true distribution of interest (e.g. the energy spectrum of cosmic rays). However, distortions to this true underlying distribution are introduced by the imperfections of measuring devices. Therefore, experimental observation can be characterized by two quantities: the true distribution we wish to measure and the observed distribution we actually measure. An important task for experimentalists is to be able to infer the true distribution from an observed distribution, i.e. to correct for the distortions caused by the measurement process. This correction, or deconvolution, is known as unfolding.

Typically, through simulation, effects from a measurement process can be encoded into a matrix that relates how true distributions, or causes, are smeared into observed distributions, or effects. Namely, a mapping between causes C_μ and effects E_j is developed:

$$n(E) = R\ \phi(C) \tag{5.1}$$

where $n(E)$ is the counts distribution for the observed effects, $\phi(C)$ is the true cause counts distribution, and R is the response matrix that encodes how a detector smears causes amongst the various observed effects. Note that both $n(E)$ and $\phi(C)$ are simply histograms of counts; $n(E_i)$ and $\phi(C_\mu)$ are the counts in the i-th effect bin and μ-th cause bin, respectively.

Ideally, one could simply invert the response matrix in Equation 5.1 to obtain the mapping from observed effects back to true causes. However, in practice this inversion is often either not possible, i.e. R is a non-invertable matrix, or problematic. For example, due to statistical fluctuations in the simulation used to construct R, often times inverting R can lead to non-physical, negative counts in the true cause distribution. When these problems arise, one must instead construct a pseudo-inverse, or unfolding, matrix that is used to map from observed effect back to true causes:

$$\phi(C) = M\ n(E) \tag{5.2}$$

where $M \approx R^{-1}$. There are several different techniques for constructing an unfolding matrix, each with its own benefits and tradeoffs. For this work, we use the D'Agostini iterative unfolding method as outlined in [55] and discussed in the following section. This method has the benefit of being able to converge to a result independent from the starting condition of the unfolding.

5.2 D'Agostini Iterative Unfolding

D'Agostini iterative unfolding applies Bayes' theorem to the inverse problem. Bayes' theorem states that the probability that a given observed effect E_j originates from a cause C_μ is given by

$$P(C_\mu|E_j) = \frac{P(E_j|C_\mu)P(C_\mu)}{\sum_{\nu}^{n_C} P(E_j|C_\nu)P(C_\nu)} \tag{5.3}$$

where n_C is the number of possible causes and $P(C_\mu)$ is the probability of the μ-th cause. $P(E_j|C_\nu)$ is the probability that a given cause C_ν results in an observed effect E_j, i.e. the response matrix R. We recognize the probabilities of the possible causes $P(C_\mu)$ as a prior distribution that encodes our current knowledge of the true underlying distribution of causes. Being a probability distribution, the prior distribution is normalized such that $\sum_{\nu} P(C_\nu) = 1$.

The construction of $P(C_\mu|E_j)$ in Equation 5.3 allows us to write the true cause distribution as:

$$\phi(C_\mu) = \sum_{i=1}^{n_E} P(C_\mu|E_i)n(E_i) \tag{5.4}$$

where n_E is the number of possible effects. Given that real detectors generally have less that 100% detection efficiencies, we allow for the possibility that a cause will not result in any measured effect. That is, the response matrix is normalized to our detection efficiencies ϵ_μ for the cause C_μ:

$$0 \leq \epsilon_\mu = \sum_{j=1}^{n_E} P(E_j|C_\mu) \leq 1 \tag{5.5}$$

Using this normalization condition, we can rewrite Equation 5.4 as:

$$\phi(C_\mu) = \frac{1}{\epsilon_\mu} \sum_{i=1}^{n_E} P(C_\mu|E_i)n(E_i) \tag{5.6}$$

$$= \sum_{i=1}^{n_E} M_{\mu i}\ n(E_i) \tag{5.7}$$

where M, the pseudo-inverse matrix for the unfolding, is defined to be:

$$M_{\mu j} = \frac{P(E_j|C_\mu)P(C_\mu)}{\Big[\sum_{k}^{n_E} P(E_k|C_\mu)\Big]\Big[\sum_{\nu}^{n_C} P(E_j|C_\nu)P(C_\nu)\Big]} \tag{5.8}$$

The unfolding matrix is composed of the response matrix $P(E|C)$, which is constructed from simulation, the measured effects distribution $n(E_i)$, as collected from a detector, and our best prior

knowledge of the underlying cause distribution $P(C_\mu)$. Using these quantities, along with Equations 5.7 and 5.8, we can obtain an unfolded true cause distribution. This unfolded distribution now represents our best estimate of the cause distribution. We can now iteratively apply this unfolding procedure, at each step using the unfolded distribution from the previous iteration as the prior distribution for the next iteration of unfolding. In this manner, the unfolding matrix for the i-th iteration

$$M^i_{\mu j} = \frac{P(E_j|C_\mu)\phi^i(C_\mu)}{\epsilon_\mu \sum_\rho P(E_j|C_\rho)\phi^i(C_\rho)} \tag{5.9}$$

is used to produce the unfolded distribution at the $i+1$ iteration

$$\phi^{i+1}_\mu = \sum_j M^i_{\mu j} n(E_j) \tag{5.10}$$

Unfolding iterations continue until some specified stopping condition is met. For example, a statistical test comparing the unfolded distribution between subsequent iterations is used as a convergence criterion for this work.

While any prior may be chosen as the starting point for unfolding, it is common practice for experimenters to choose either a uniform or a non-informative Jeffreys prior [56] to avoid introducing any bias into the unfolding procedure. Jeffreys prior, which is given by:

$$P_{\mathrm{J}}(C_\mu) = \frac{1}{\log(C_{\max}/C_{\min})C_\mu} \tag{5.11}$$

where $C_{\min}$ and $C_{\max}$ are the minimum and maximum allowed values the true causes C_μ can take, is often an appropriate choice when the cause distribution spans several orders of magnitude.

5.3 Unfolding Uncertainties

While typical Poisson counting errors can be used to estimate the uncertainty of the observed counts distribution $n(E)$, this is not the case for the unfolded counts distribution $\phi(C)$. During deconvolution, each bin in the unfolded distribution receives contributions from several different $n(E_j)$ in a correlated manner. From Equation 5.4, we can see that two sources of uncertainty for the unfolded distribution need to be considered: those from the Poisson uncertainties of the

observed counts distribution $n(E)$, and those originating from the uncertainties in the response matrix $P(E|C)$ which is built from a limited amount of Monte Carlo simulation. That is:

$$V = V^{\text{data}} + V^{\text{response}} \tag{5.12}$$

where V^{data} and V^{response} are the contributions to the full covariance matrix V from the observed distribution $n(E)$ and response matrix $P(E|C)$ uncertainties, respectively.

The unfolding uncertainty estimates, as outlined in [55] and [57], are shown below. Here, we only show the final result; full derivations of these terms are shown in the supplemental materials included in [58]. In addition, for convenience we adopt the shorthand notation $P(E_i|C_\mu) \to P_{\mu i}$, $\phi(C_\mu) \to \phi_\mu$, and $n(E_j) \to n_j$. The error contributions from the observed distribution are given by:

$$V^{\text{data}} = \frac{\partial \phi^{i+1}}{\partial n} \text{cov}(n, n') \frac{\partial \phi^{i+1'}}{\partial n} \tag{5.13}$$

with

$$\text{cov}(n_k, n_j) = \begin{cases} n_j(1 - \frac{n_j}{N_{\text{true}}}) & \text{if } k = j \\ -\frac{n_j n_k}{N_{\text{true}}} & \text{if } k \neq j \end{cases} \tag{5.14}$$

and

$$\frac{\partial \phi_\mu^{i+1}}{\partial n_j} = M_{\mu j} + \frac{\phi_\mu^{i+1}}{\phi_\mu^i} \frac{\partial \phi_\mu^i}{\partial n_j} - \sum_{\sigma,k} \epsilon_\sigma \frac{n_k}{\phi_\sigma^i} M_{\mu k} M_{\sigma k} \frac{\partial \phi_\sigma^i}{\partial n_j} \tag{5.15}$$

While the error contributions from the uncertainties on the response matrix are given by:

$$V^{\text{response}} = \frac{\partial \phi^{i+1}}{\partial P} \text{cov}(P, P') \frac{\partial \phi^{i+1'}}{\partial P} \tag{5.16}$$

with

$$\text{cov}(P_{\rho r}, P_{\lambda s}) = \sigma_{\rho r} \sigma_{\lambda s} \delta_{\rho \lambda} \sigma_{rs} \tag{5.17}$$

and

$$\begin{aligned} \frac{\partial \phi_\mu^{i+1}}{\partial P_{\lambda k}} = & \frac{\delta_{\lambda \mu}}{\epsilon_\mu} \left(\frac{n_k \phi_\mu^i}{f_k} - \phi_\mu^{i+1} \right) - \frac{n_k \phi_\lambda^i}{f_k} M_{\mu k} \\ & + \frac{\phi_\mu^{i+1}}{\phi_\mu^i} \frac{\partial \phi_\mu^i}{\partial P_{\lambda k}} - \sum_{\rho,j} n_j \frac{\epsilon_\rho}{\phi_\rho^i} M_{\rho j} M_{\mu j} \frac{\partial \phi_\rho^i}{\partial P_{\lambda k}} \end{aligned}$$

Note that $N_{\text{true}} = \sum_{\mu} \phi(C_\mu)$, $\delta_{\lambda\mu}$ is the Kronecker delta, $\sigma_{\lambda s}$ is the uncertainty on $P_{\lambda s}$, and the superscripts i and $i+1$ refer to the unfolding iteration number. After each unfolding iteration, using the above equations, the uncertainty on the unfolded cause distribution can be determined. With both the iterative unfolding and uncertainty estimation methods established, we can now apply this technique to account for our known composition and energy detector response.

5.4 Energy and Composition Response Matrix

The underlying true distribution we wish to measure for this analysis is the energy spectrum for light and heavy induced air showers. Cast in terms of an inverse problem, the causes are air showers with a particular true primary energy and composition that get smeared into reconstructed energy and composition effects bins. Using simulated light and heavy induced air showers along with the composition classification and energy regression models presented in Chapter 4, we can construct a dataset consisting of the true and reconstructed primary energy and composition for each simulated event. Using this simulation dataset, we can build a true and reconstructed energy-composition histogram that serves as the basis for our detector response matrix.

Iterative unfolding is agnostic to the particular definition of cause and effect bins. However, it is important that once a binning convention is chosen, it remains consistent throughout the unfolding. Here, we choose the energy-composition binning scheme for both the true and observed variables to be arranged in an alternating fashion as illustrated in Figure 5.1. Energy bins are equally spaced in log-energy from $\log_{10}(E/\text{GeV}) = 6.1$ to $\log_{10}(E/\text{GeV}) = 8.0$ with a bin width of 0.1 in $\log_{10}(E/\text{GeV})$.

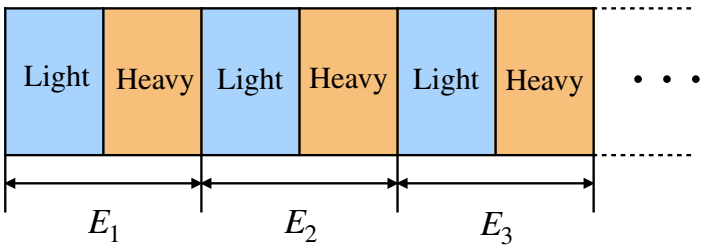


Figure 5.1: The energy-composition bins for both the true and observed variables are arranged in an alternating fashion. Specifically, the first two cause bins contain light and heavy induced air showers, respectively, for the first true energy bin. The third and fourth cause bins contain light and heavy air showers, respectively, for the second true energy bin, and so on. A similar binning scheme is also used for the observed energy and composition.

In order for our energy-composition histogram to be used as a response matrix, it must be normalized to the detection efficiencies as in Equation 5.5. The detection efficiency ϵ is defined to

be:

$$\epsilon(E_i) = \frac{N_{\text{pass}}(E_i)}{N_{\text{generated}}(E_i)} \tag{5.18}$$

where $N_{\text{generated}}(E_i)$ is the number of simulated air showers in the i-th true energy bin, while $N_{\text{pass}}(E_i)$ is the number of showers which pass the data quality cuts (Section 3.4) in the i-th true energy bin. Figure 5.2 shows the detection efficiency as a function of true primary energy for simulated light and heavy induced air showers. To smooth out statistical fluctuations in the

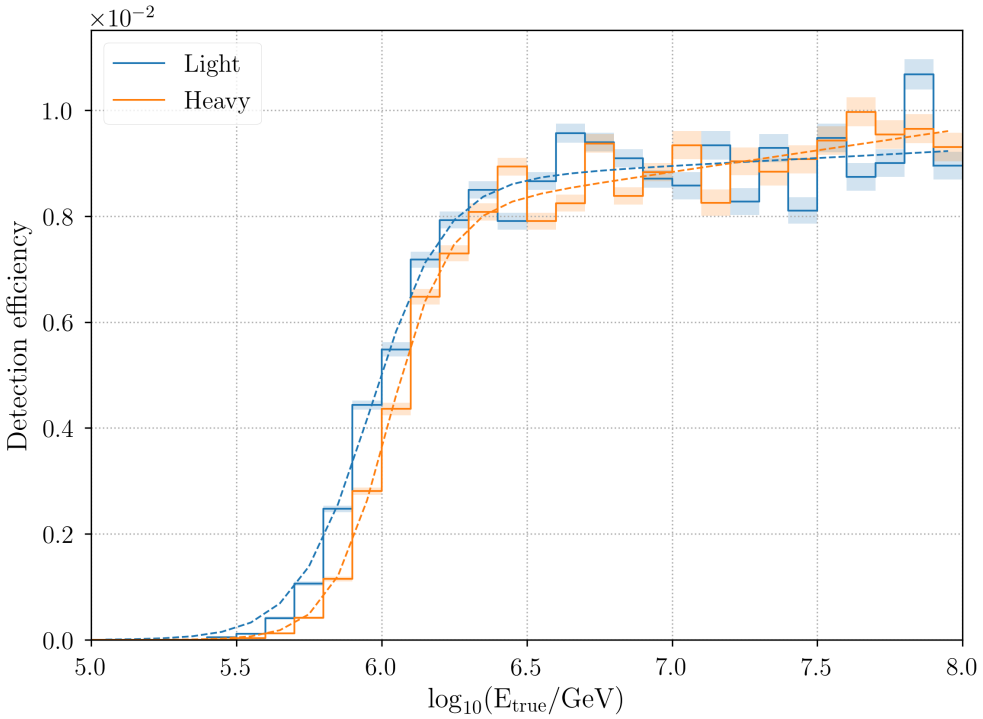


Figure 5.2: Detection efficiency as a function of primary energy for simulated light and heavy induced air showers. A slanted sigmoid function was fit to the efficiencies for each composition group (dashed line). Note that the overall 10^{-2} scale of the efficiencies is due to the relatively large simulation generation area (a circle of radius 1700 m).

efficiencies, a so-called "slanted sigmoid" was fit to the efficiencies for both composition groups independently [59], which has the functional form:

$$\epsilon^{\text{fit}}(E) = \frac{p_0 + p_3 \cdot \log_{10}(E)}{1 + \exp(-p_1 \cdot \log_{10}(E) + p_2)} \tag{5.19}$$

where p_0, p_1, p_2, and p_3 are fit parameters which were determined using the method of least squares. The best fit values for these parameters are listed in Table 5.1. These smoothed detec-

Composition	p_0	p_1	p_2	p_3
Light	7.87	46.91	2.96×10^{-4}	6.88×10^{-3}
Heavy	10.04	60.42	8.20×10^{-4}	3.09×10^{-3}

Table 5.1: Detection efficiency fit parameters for light and heavy induced air showers

tion efficiencies are used to normalize the energy-composition response matrix, which is shown in Figure 5.3. Figure 5.4 displays the same response matrix decomposed into separate panels for each true-reconstruction composition pair. Using the energy-composition response matrix, we can now proceed with unfolding the observed counts distribution.

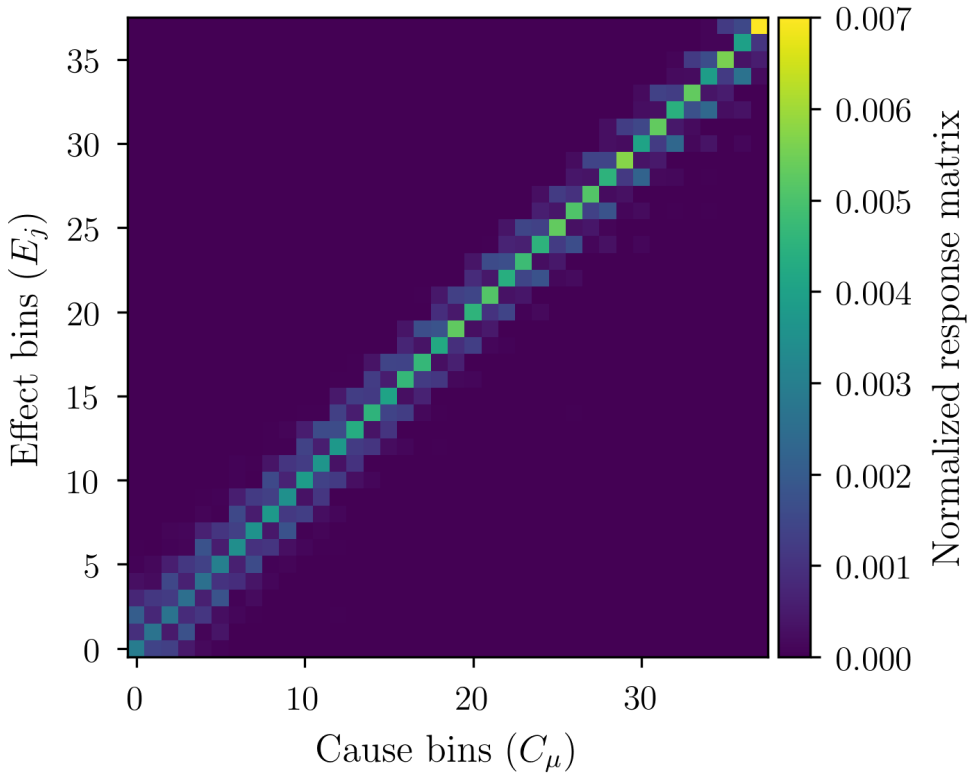


Figure 5.3: Energy and composition detector response matrix. The binning scheme used is illustrated in Figure 5.1. The columns of the response matrix are normalized to the detection efficiencies ϵ shown in Figure 5.2.

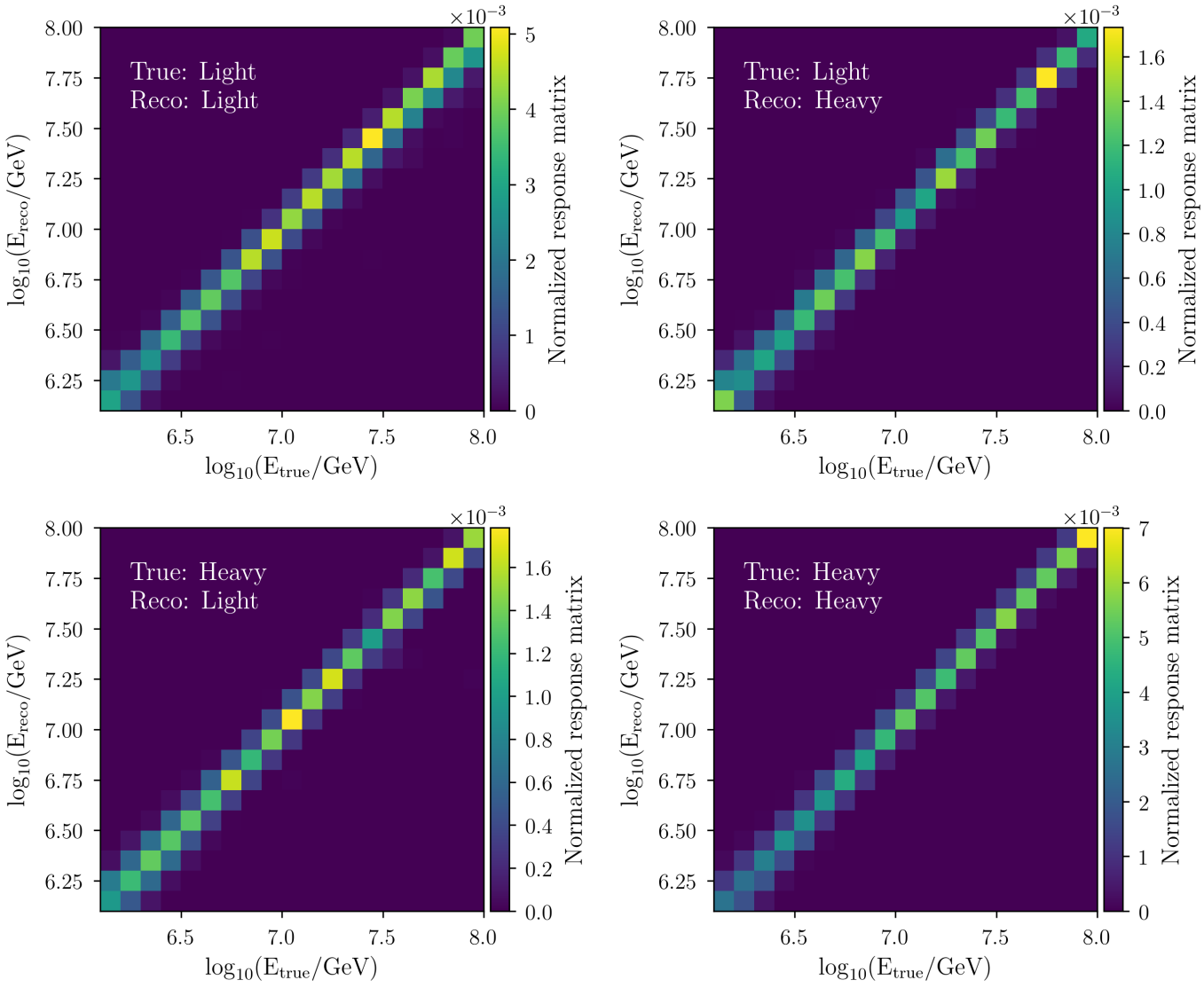


Figure 5.4: Detector response matrix decomposed into separate panels for each true-reconstruction composition pair. The true and classified compositions are indicated in each panel.

It's worth noting that as part of the work for this analysis, and the work presented in [60], a Python package to implement the iterative unfolding algorithm was developed. The package, PyUnfold[1], was published in the Journal of Open Source Software [58] and is used to perform the unfolding in this analysis. Further details regarding PyUnfold, including motivating factors, features, and a toy example, are presented in Appendix C.

[1] https://github.com/jrbourbeau/pyunfold

Chapter 6

Composition-Resolved Cosmic-Ray Spectra

In this chapter, we present the results for the measured light, heavy, and all-particle cosmic-ray energy spectra at PeV energies. We introduce the details of the dataset used and then discuss the process for converting the unfolded counts distribution to a flux. In addition, a detailed study of how various sources of systematic uncertainty impact the results of this measurement is also presented.

6.1 Dataset

The experimental data used in this analysis was collected over a single year by the IceCube Observatory from April 2012 through May 2013. A total of 6,853,826 air shower events passed the applied data quality cuts (Section 3.4) and are used for this work. The number of events in each reconstructed energy bin is, as expected, steeply falling, with the highest energy bin containing 794 events. In addition, while there is no explicit zenith cut applied, the requirement of coincident events that pass through both the IceTop and in-ice arrays place an effective geometric limit on the observable zenith range, as indicated in Figure 6.1. The maximum zenith angle observed is 37.65°.

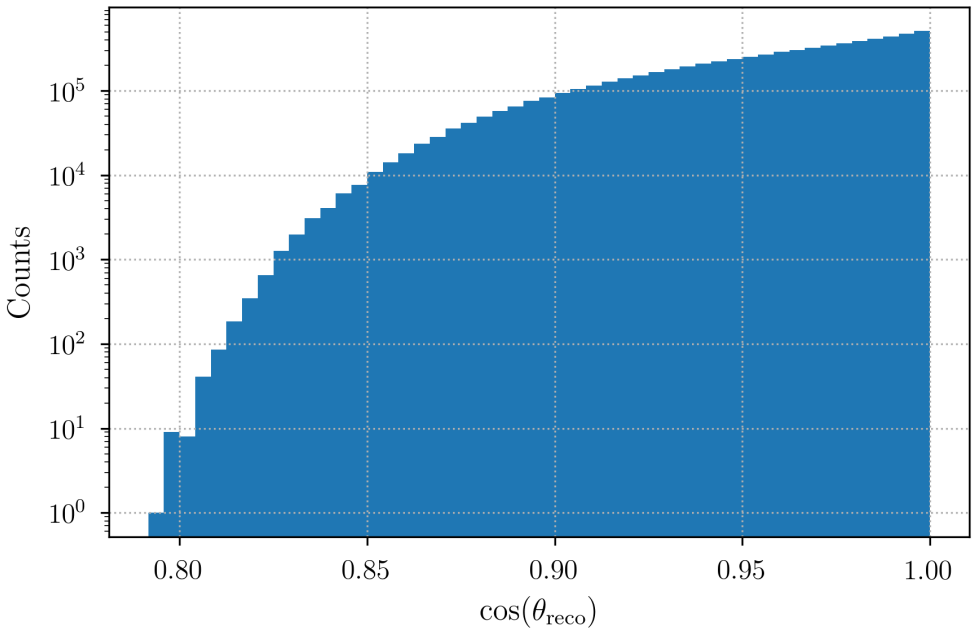


Figure 6.1: Distribution of the reconstructed $\cos\theta$ for dataset. The maximum zenith angle observed is 37.65°.

An important characteristic of the data collection process is the amount of time in which the detector is actively recording data, often called the detector livetime. Ideally, the livetime for a single data taking run would simply be the difference between the end and start times for the run. However, due to periods of instability within a run, a simple time difference is not sufficient to measure the detector livetime. Instead, the distribution of time differences between events is fit using an exponential function $N(\Delta t) = N_0 \exp(-\Delta t/\tau)$, where Δt is the time difference between

events, N_0 is a normalization, and τ is the mean time difference between events. The detector livetime is then calculated via multiplying the total number of events by the mean time difference between events τ. The livetime is determined independently for each month of data. The event time difference distribution and exponential fit for July 2012 is shown in Figure 6.2. The monthly livetimes are then summed together to obtain the livetime for the entire dataset, 329.192 days.

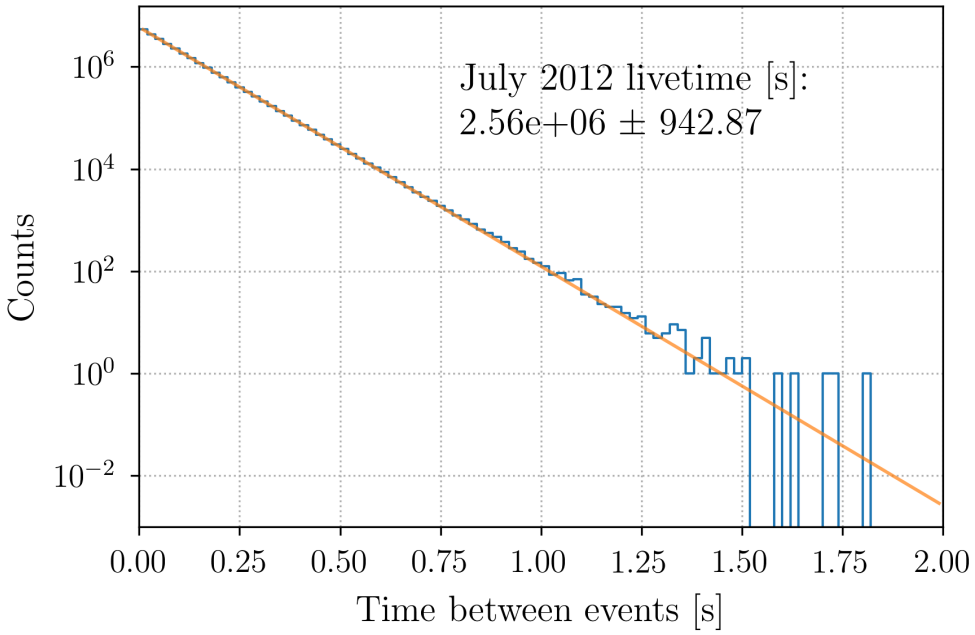


Figure 6.2: Time difference between events distribution for data collected during July 2012. The orange curve shows the exponential fit $N(\Delta t) = N_0 \exp(-\Delta t/\tau)$.

6.2 Composition-Separated Cosmic-Ray Spectra

After running air shower event reconstructions and applying the composition and energy regression models to the detected events in our dataset, we can obtain a measured counts distribution of both light and heavy events for each energy bin. The unfolding procedure, using the detector response matrix from Chapter 5, takes this counts distribution and accounts for energy and composition bin migration and detection efficiencies, providing an unfolded distribution of expected counts for each mass group binned in the true energy.

We can now convert this unfolded counts distribution to a flux. From Equation 1.1, the flux for the i-th energy bin is given by:

$$J(E_i) = \frac{N_i}{\Delta E_i \; A_{\text{eff}} \; \Delta\Omega \; \Delta t} \tag{6.1}$$

where N_i is the number of counts in the i-th energy bin, ΔE_i is the width of the bin, A_{eff} is the effective area, $\Delta\Omega$ is the solid angle, and Δt is the detector livetime. Note that the unfolding method already takes into account detection efficiencies and thus only the simulated thrown area needs be included here.

In order to avoid any a priori assumptions about the spectra, a flat Jefferys prior was chosen as the seed for the unfolding procedure in this analysis. In addition, a Kolmogorov-Smirnov (KS) test is used to compare unfolded distributions between subsequent iterations. Once the corresponding KS test statistic falls below a threshold of 0.005, the unfolding process is said to have converged and the final unfolded distribution is returned. This procedure was shown to provide a final unfolded distribution that is independent of the initial chosen prior within the associated uncertainties. The impact of choosing different priors, as well as a series of data challenges to validate the unfolding method, are presented in Appendix B.

The unfolded cosmic-ray energy spectrum for the light, heavy, and all-particle spectra are shown in Figure 6.3, while Figure 6.4 shows the same spectra scaled by a factor of $E^{2.7}$ to highlight the underlying structure of the distributions. Note that, from the binning scheme outlined in Section 5.4, only the energy distributions for the light and heavy mass groups are obtained directly. The all-particle counts distribution is determined by summing both the light and heavy unfolded counts

in each energy bin. The values for the unfolded flux in each energy bin are listed in Tables A.1, A.2, and A.3.

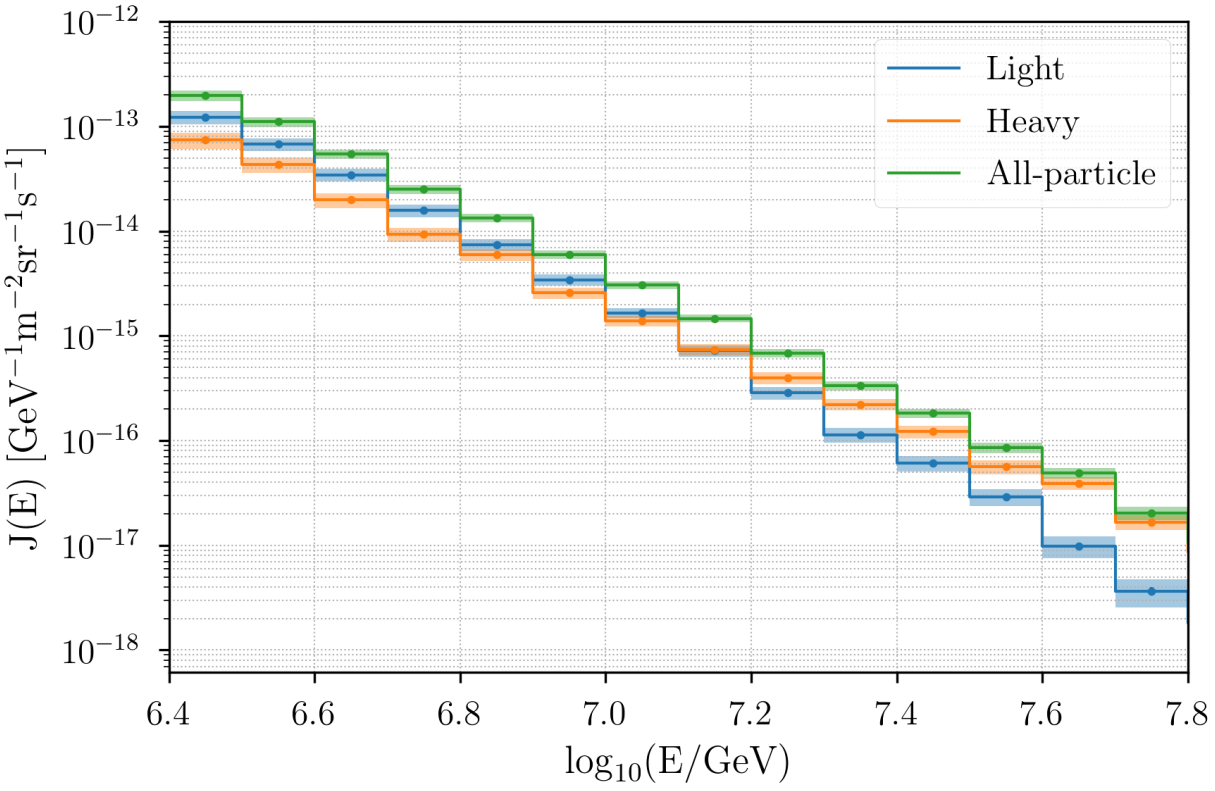


Figure 6.3: Unfolded cosmic-ray energy spectra for the light, heavy, and all-particle mass groups. The shaded bands indicate the uncertainty due to the limited Monte Carlo simulation available to build the unfolding response matrix (V^{response} from Equation 5.16). The vertical error bars show the statistical uncertainties from the experimental data.

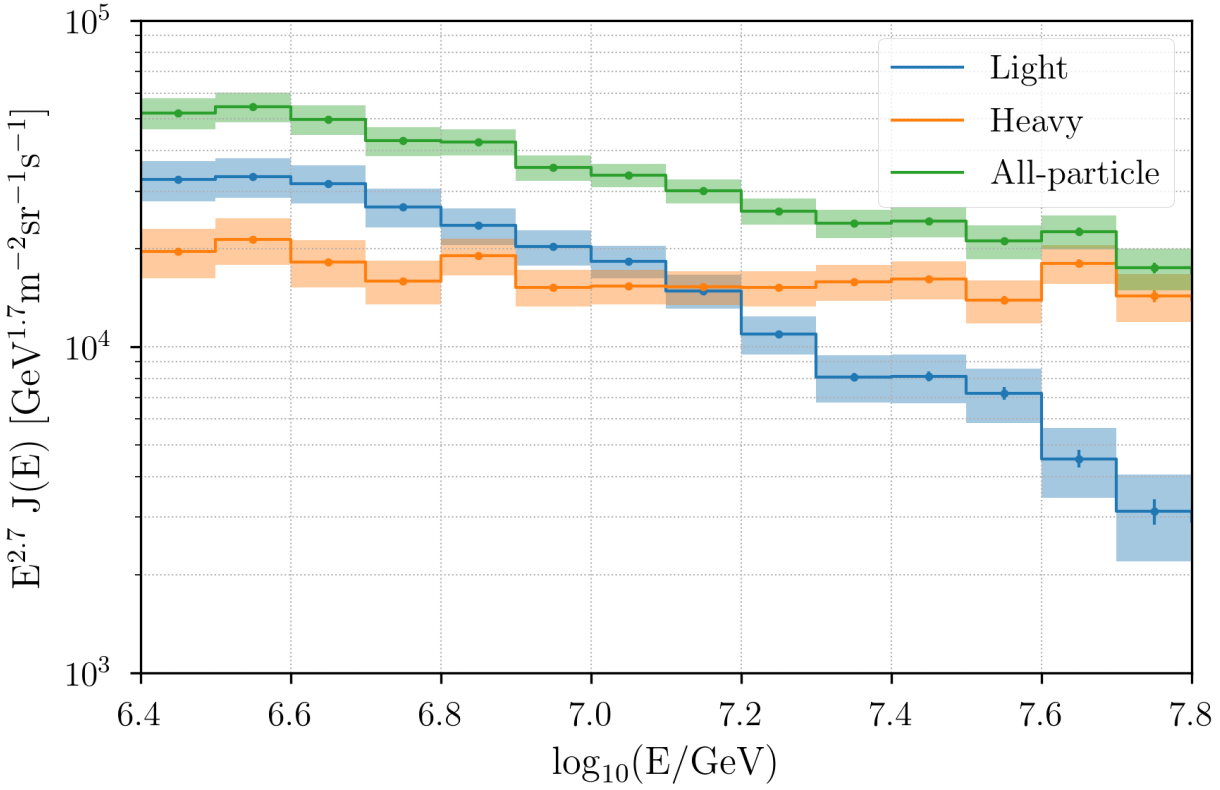


Figure 6.4: Scaled by $E^{2.7}$. Unfolded cosmic ray energy spectra for the light, heavy, and all-particle mass groups. The shaded bands indicate the uncertainty due to the limited Monte Carlo simulation available to build the unfolding response matrix (V^{response} from Equation 5.16). The vertical error bars show the statistical uncertainties from the experimental data.

6.3 Systematic Uncertainties

A detailed study regarding several sources of systematic uncertainties in this analysis has been conducted. Generally, the sources of systematic uncertainty fall into three separate categories:

1. Uncertainties due to limited Monte Carlo statistics for constructing the detector response matrix

2. Uncertainties due to our understanding and calibration of the surface and in-ice arrays

3. Uncertainties due to our modeling of hadronic interactions

For each of these sources of uncertainty, with the exception of the analysis method systematic, a systematic Monte Carlo simulation dataset is constructed. Using this systematic simulation set, a new detector response matrix is constructed and used to unfold the observed counts distribution to produce an unfolded flux. It is important to note that this is the *only* modification made when determining the impact of systematic uncertainties. With this procedure we are estimating how sensitive our spectrum measurement is to changes in our nominal understanding of the detector and interaction models used. A summary of the systematic uncertainties for the light, heavy, and all-particle cosmic-ray spectra is shown in Figure 6.5. Likewise, Tables 6.1, 6.2, and 6.3 provide the percent uncertainty of these systematics for each spectra in the first, middle, and last energy bin considered. A detailed discussion of each systematic is presented in what follows.

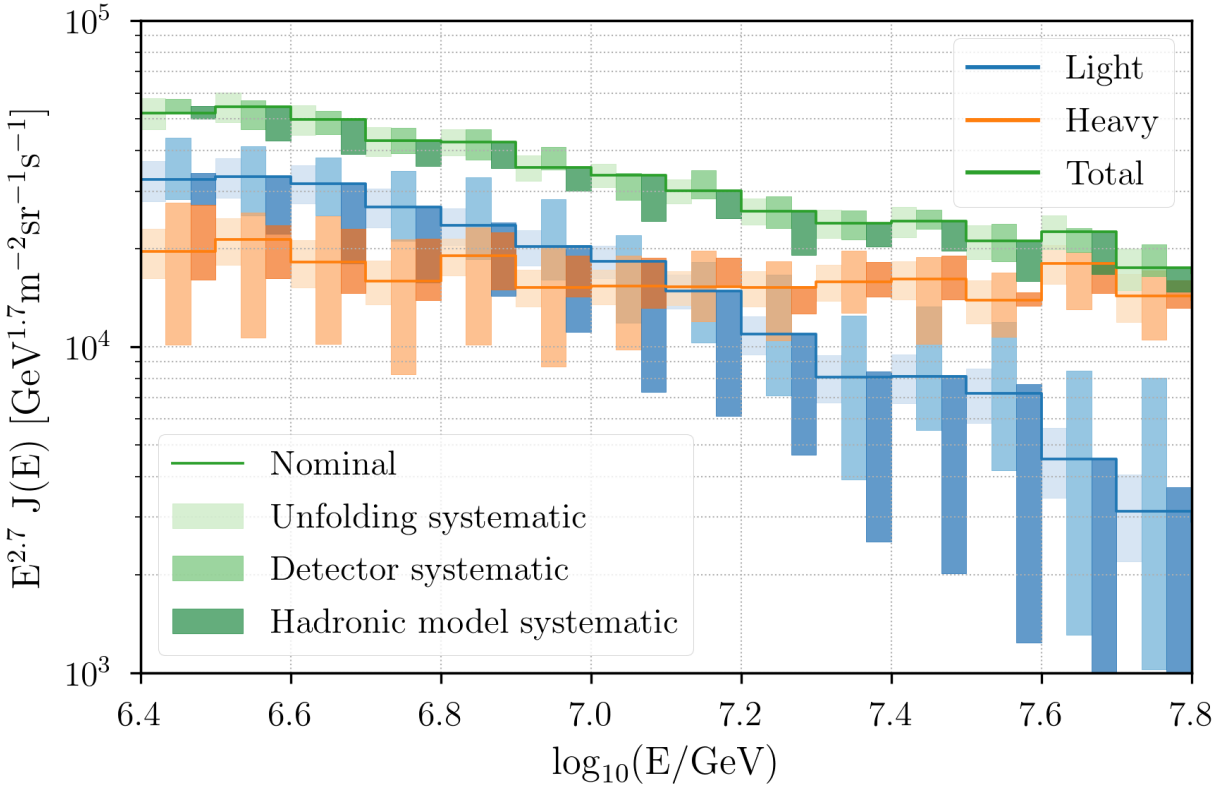


Figure 6.5: Comparison of the unfolding, detector, and hadronic interaction model systematic uncertainties. The nominal value for the light, heavy, and all-particle unfolded flux are indicated by the blue, orange, and green solid lines in each energy bin, respectively. The variation in the flux arising from each source of systematic uncertainty are shown by the corresponding shaded regions. The lower legend indicates which systematic corresponds to which shade.

Systematic	$10^{6.45}$ GeV	$10^{7.15}$ GeV	$10^{7.75}$ GeV
Unfolding	±11.23%	±12.18%	±18.23%
Detector	-11.84% / +35.37%	-30.55% / +22.26%	-73.87% / +155.07%
Hadronic interaction model	-16.33% / +4.63%	-58.62%	-98.30% / +19.01%
Total	-23.09% / +37.40%	-67.22% / +25.37%	-124.31% / +157.29%

Table 6.1: Systematic uncertainty summary for the light mass group cosmic-ray energy spectrum. Estimated effects due to the various sources of uncertainty are shown for the first, middle, and last energy bin.

Systematic	$10^{6.45}$ GeV	$10^{7.15}$ GeV	$10^{7.75}$ GeV
Unfolding	±13.0%	±11.59%	±10.30%
Detector	-48.21% / +42.62%	-21.57% / +30.10%	-26.39% / +21.15%
Hadronic interaction model	-17.79% / +39.44%	+23.08%	-8.33% / +11.39%
Total	-53.01% / +59.51%	-24.49% / +39.66%	-29.53% / +26.14%

Table 6.2: Systematic uncertainty summary for the heavy mass group cosmic-ray energy spectrum. Estimated effects due to the various sources of uncertainty are shown for the first, middle, and last energy bin.

Systematic	$10^{6.45}$ GeV	$10^{7.15}$ GeV	$10^{7.75}$ GeV
Unfolding	±8.55%	±8.50%	±9.09%
Detector	+11.43%	-4.94% / +16.07%	-5.45% / +17.85%
Hadronic interaction model	-3.67% / +4.50%	-17.51%	-15.17% / +0.51%
Total	-9.30% / +14.97%	-20.08% / +18.18%	-18.51% / +20.04%

Table 6.3: Systematic uncertainty summary for the all-particle cosmic-ray energy spectrum. Estimated effects due to the various sources of uncertainty are shown for the first, middle, and last energy bin.

6.3.1 Unfolding Systematic Uncertainty

There is an uncertainty on the unfolded counts distribution due to the finite amount of Monte Carlo simulation available to construct the detector response matrix. The determination of this uncertainty, the V^{response} term in Equation 5.12, is discussed at length in Section 5.3. This term is treated as a systematic uncertainty in the sense that it does not arise from the data themselves and instead is reduced upon an increase of more Monte Carlo simulation statistics. The magnitude of this uncertainty is shown in Figures 6.3 and 6.4 by the shaded error bands.

6.3.2 Detector Systematic Uncertainties

Detector-related uncertainty from the IceTop energy calibration, snow correction, and in-ice light yield also impact this work. Estimates for how these effects manifest themselves are summarized in Table 6.4. Note that the IceTop energy calibration and snow correction systematics influence the shower energy proxy S_{125}, while the in-ice light yield uncertainty alters the muon number proxy dE/dX.

Uncertainty source	Estimated effect
IceTop energy calibration	$\pm 3\%$ in S_{125}
IceTop snow correction	± 0.2m in λ_{eff}
In-ice light yield	+9.6%, -12.5% in in-ice DOM efficiency

Table 6.4: Summary of detector systematic uncertainties and how each effect manifests as it related to this work.

From Section 3.2.1, snow coverage atop each IceTop tank leads to a reduction in the deposited signal measured by the tanks. In order to account for this effect, the measured signal for each tank is corrected via an exponential absorption model (Equation 3.7) with effective absorption length λ_{eff}. The optimal value of $\lambda_{\text{eff}} = 2.25$ m was determined to yield a consistent S_{125} distribution, on average, across the entirety of IceTop with different tank snow heights. However, snow attenuation effects are known to vary with primary particle energy, zenith angle, and mass. An uncertainty of ± 0.2 m in the effective absorption length of λ_{eff} is used to capture these additional variational factors [36].

In addition, the IceTop tank calibration discussed in Section 2.4, in which the conversion factor from number of observed photoelectrons to the VEM charge unit is determined, also introduces some uncertainty to our results. Based on simulation studies [61], tank calibration has been shown to be done on the 2-3% level. This resolution on the photoelectron-to-VEM conversion translates to this analysis as a $\pm 3\%$ uncertainty in the reconstructed S_{125} energy proxy.

Lastly, our knowledge of the in-ice light yield also affects the measured spectra. The light yield in ice is an important aspect of the in-ice reconstruction outlined in Section 3.3, which engineers the composition-sensitive muon energy loss dE/dX variable. Several factors enter into the in-ice light yield uncertainty which are summarized in Table 6.5. First, there is an uncertainty on the photon detection efficiency of the in-ice DOMs of $\pm 3\%$ which linearly translates to a $\pm 3\%$ uncertainty in the measured charge. In addition, our modeling of photon scattering and absorption properties in the ice influence the light yield. Values used to constitute these uncertainties are a $\pm 10\%$ uncertainty on the scattering and -7.1% uncertainty on the scattering and absorption parameters [62]. Work is currently underway to reduce the size of the scattering and absorption errors, and it is known that the estimates used in this work are, in fact, over estimates. Likewise, the effective scattering length of photons in hole ice, the ice which was melted and then allowed to refreeze during detector construction, is also considered. The nominal value for the hole ice scattering length of 50 cm is used, while lengths of 30 cm and 100 cm are used to determine the effect of variations. We assume all in-ice uncertainties are uncorrelated and sum them in quadrature to determine the net uncertainty. This is an oversimplification that results in an over estimate of the combined uncertainty. The net result of in-ice light yield uncertainties is a +9.6% and -12.5% shift in the detection efficiency for the in-ice DOMs.

The resulting impacts of these detector-related uncertainties on the unfolded energy spectra are shown in Figure 6.6. It's clear that the in-ice uncertainties have the largest influence on the final unfolded spectrum measurement. Note that the detector uncertainties for the all-particle spectrum are generally much smaller than for the individual mass groups. This is because the shifts in the light and heavy spectra are anti-correlated in each energy bin and of roughly a similar magnitude, resulting in a much smaller uncertainty on the all-particle spectrum.

Effect	Light yield shift
10% scattering	+3.6%
10% absorption	-11.8%
-7.1% scattering and absorption	+7%
30 cm hole ice scattering	+4.5%
100 cm hole ice scattering	-2.9%
DOM efficiency	±3%
Total light yield effect	+9.6%, -12.5%

Table 6.5: Summary of in-ice light yield sources of uncertainty.

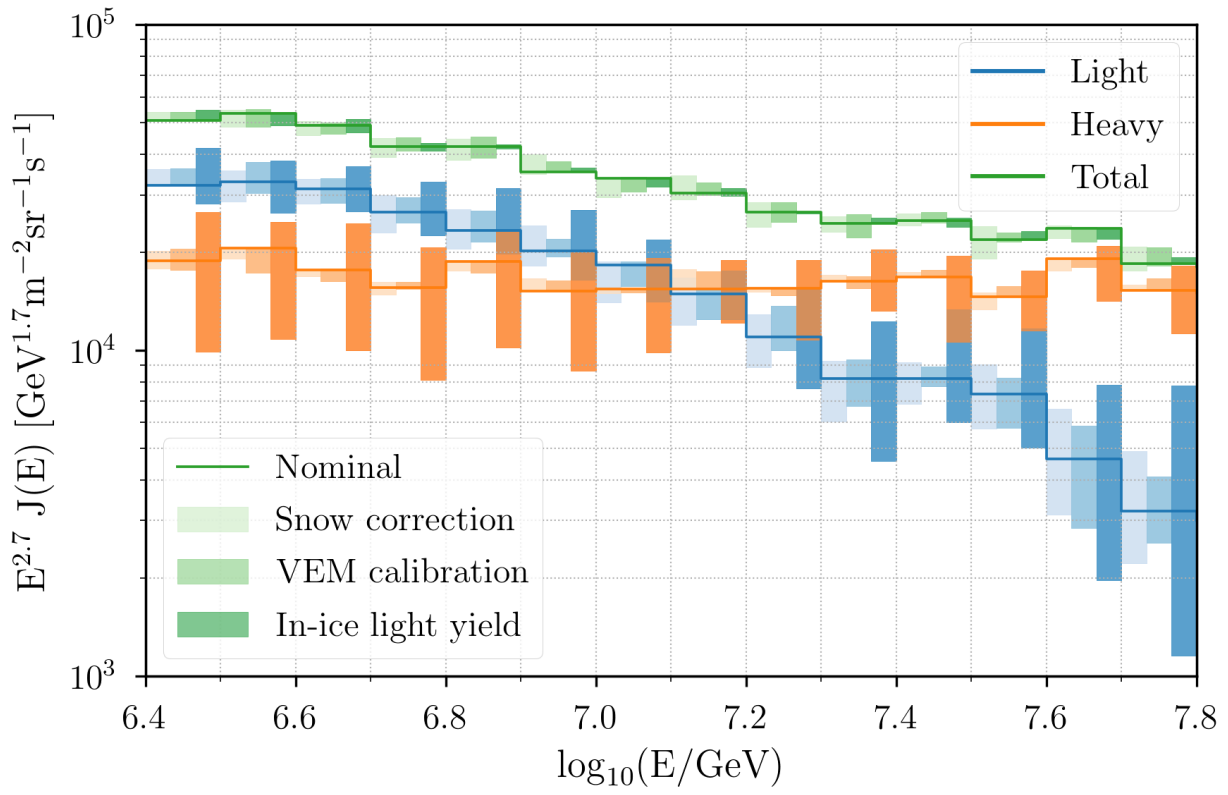


Figure 6.6: Comparison of detector systematic uncertainties. The nominal value for the light, heavy, and all-particle unfolded flux are indicated by the blue, orange, and green solid lines in each energy bin, respectively. The variations in the flux arising from each source of systematic uncertainty are shown by the corresponding shaded regions. The lower legend indicates which systematic corresponds to which shade.

6.3.3 Hadronic Interaction Model Uncertainties

Interpreting detected air shower signals with ground-based experiments relies heavily on modeling the hadronic interactions which occur during shower development. Unfortunately, the interactions needed to model these processes are not well constrained by experimental data. As of Run II at the LHC, protons are accelerated to an energy of 6.5 TeV and positioned to collide, producing a center-of-mass energy of 13 TeV [63]. This corresponds to a roughly 90 PeV fixed-target energy, which is comparable to the cosmic ray case. However, IceCube is sensitive to cosmic ray primaries up to $\sim 10^3$ PeV, well above the maximum energy attainable at the LHC. In addition, accelerator experiments, like ATLAS and CMS at the LHC, have their detectors configured to measure so-called hard processes, in which there are large amounts of momentum transfer. However, cosmic-ray experiments detect particles from soft interactions, where there are relatively low amounts of momentum transfer. Thus, the application of hadronic interaction models to cosmic-ray air shower development involves extrapolation in both energy and phase space.

The three leading interaction models used for air shower simulation are the Sibyll [39], QGSJet [40], and EPOS [41] models. Each of these models is based on the theoretical Gribov-Regge multiple scattering model, in which interactions are represented by an exchange of one or more pomerons [64]. However, the models differ in the assumptions, parameterizations, and experimental data they choose to use, which has important consequences for the final air shower observables.

The QGSJet and Sibyll models are both optimized for air shower simulations, however they do so in different ways. QGSJet has a small set of parameters to minimize the uncertainties introduced by extrapolation to higher energies. However, this comes at the cost of having a less detailed description of final stage hadronic interactions, which limit the data sets to which it can be compared. In contrast, Sibyll is based on the dual parton and minijet model. EPOS, on the other hand, is used by both the cosmic-ray and accelerator physics communities. As opposed to QGSJet and Sibyll, EPOS is optimized to describe accelerator data from heavy ion interactions and has many parameters that must be tuned.

For this analysis, the Sibyll 2.1 hadronic interaction model was used to generate the nominal air shower simulation. However, Sibyll 2.3, QGSJet-II-04, and EPOS-LHC were used to determine

how sensitive the final unfolded spectrum measurements are to changes in the interaction model used. Note that Sibyll 2.3, QGSJet-II-04, and EPOS-LHC are all newer, post-LHC versions of their respective models that have been updated to include recent experimental LHC results. The nominal Sibyll 2.1 model is a pre-LHC version of the Sibyll model.

The primary effect the choice of interaction model has on this analysis is in the number of muons produced in air showers. Figure 6.7 shows the ratio of number of air shower muons for various hadronic interaction models as compared to Sibyll 2.1 as a function of primary particle energy. We can see the post-LHC interaction models generally produce more muons as compared to Sibyll 2.1. This increase in number of muons manifests itself as a shift in the reconstructed in-ice muon number proxy dE/dX. Figure 6.8 illustrates how dE/dX varies for the different interaction models as a function of S_{125}. We can see that the EPOS-LHC model, which is the least optimized for air showers, gives the largest difference in dE/dX. The corresponding influence on the measured light, heavy, and all-particle energy spectrum for different choices of hadronic interaction model is shown in Figure 6.9.

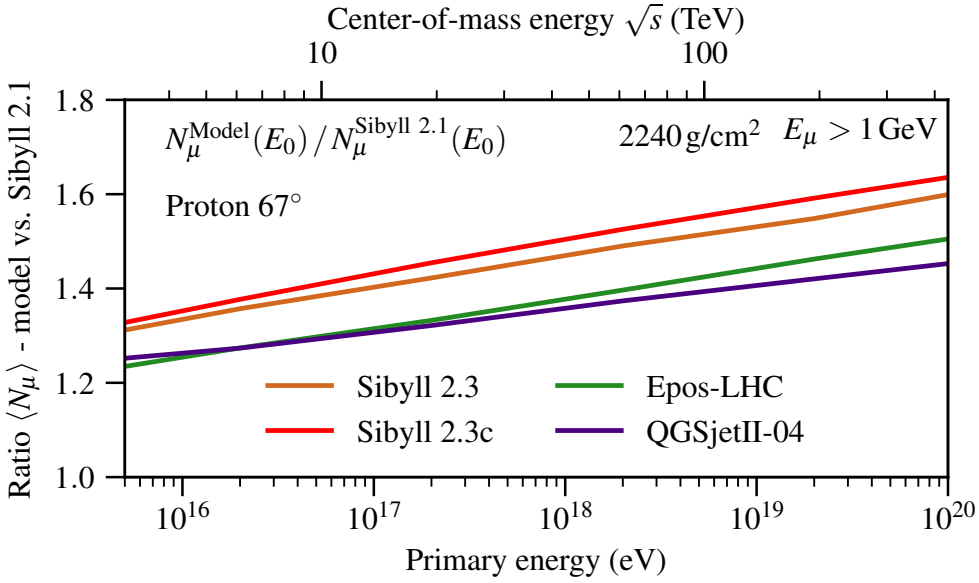


Figure 6.7: Comparison of the average number of air shower muons between Sibyll 2.1 and post-LHC models. Figure from [65].

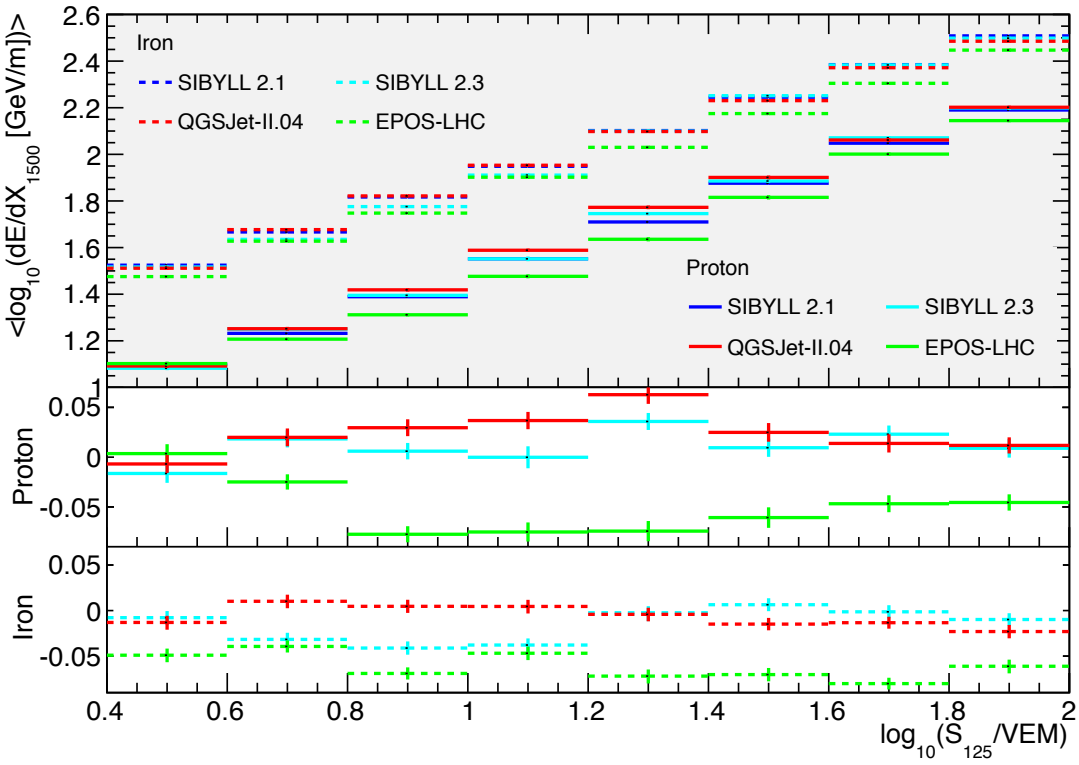


Figure 6.8: Comparison of the reconstructed in-ice muon energy loss dE/dX for various hadronic interaction models. The bottom panels indicate the shift of the new models with respect to Sibyll 2.1. Figure from [66].

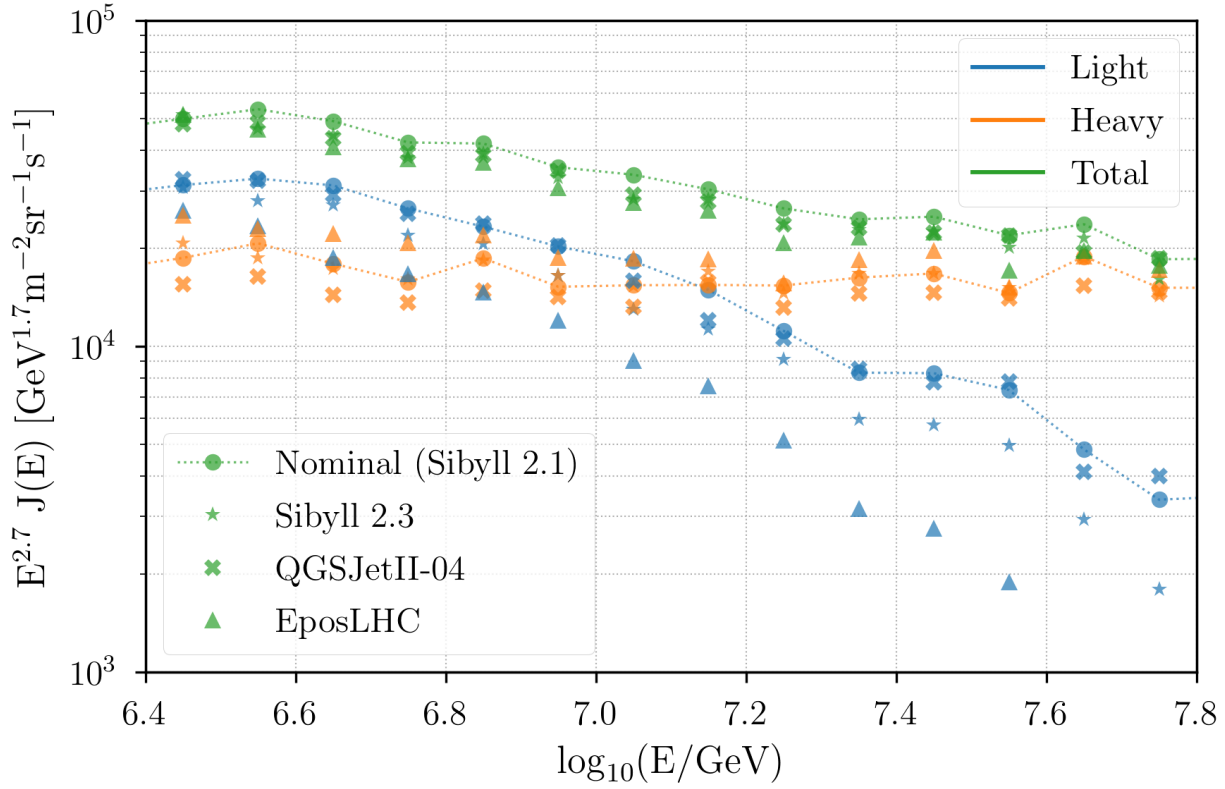


Figure 6.9: Comparison of the unfolded spectra between the Sibyll 2.1 (nominal), Sibyll 2.3, QGSJet-II-04, and EPOS-LHC hadronic interaction models.

6.4 Discussion & Future Prospects

We observe the light mass group spectrum is softer than that of the heavier mass group, which follows a power-law like structure with a spectral index of ~ 2.7 throughout the entire energy range considered. The transition from a primarily light to heavy-dominant spectrum takes place near $\log_{10}(E/\mathrm{GeV}) = 7.1$. This is qualitatively consistent with the findings of a separate IceCube four mass group spectrum measurement [67], recent results from KASCADE-Grande [68], and the H4a spectrum model [69] which are shown in Figures 6.10, 6.11, and 6.12, respectively.

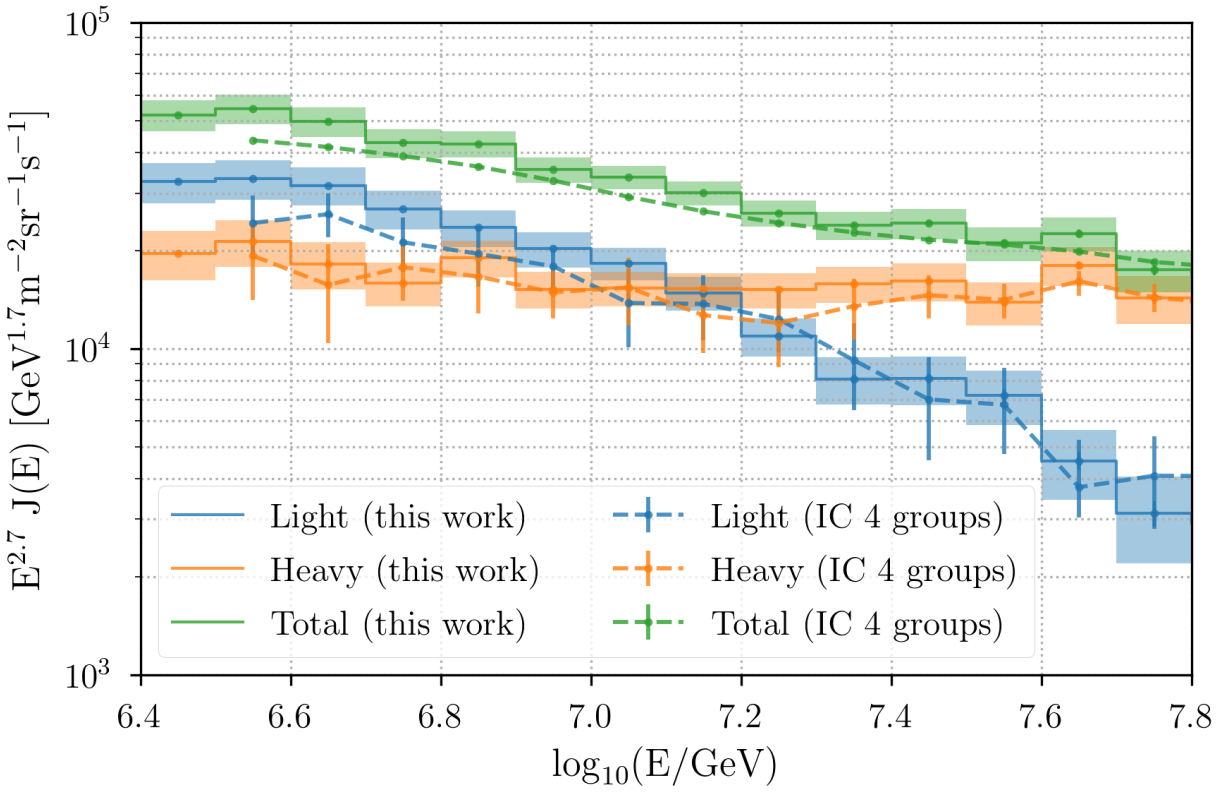


Figure 6.10: Comparison with a four mass group IceCube analysis. The four mass groups are modeled by proton, helium, oxygen, and iron air showers, respectively. For comparison with this work, the proton and helium mass groups have been combined into an effective "light" group, while the oxygen and iron mass groups have been combined to form a "heavy" mass group.

There is ongoing work within the IceCube collaboration to help further improve the mass-separated spectrum measurement presented in this work. One of the primary paths to improve the composition classification performance is to incorporate an IceTop composition-sensitive feature.

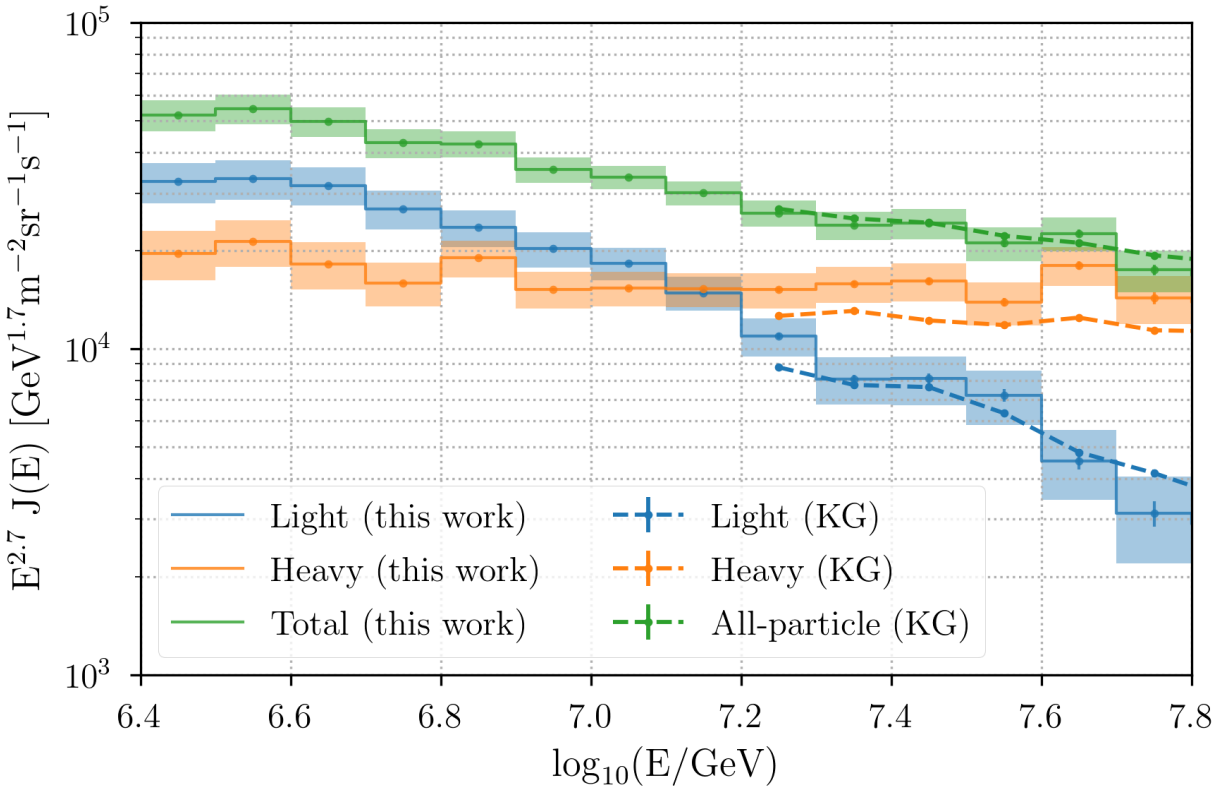


Figure 6.11: Comparison with a two mass group cosmic ray spectrum measurement by the KASCADE-Grande experiment [68]. Note that statistical uncertainties for the KASCADE-Grande measurement are shown in the figure, but are smaller than the marker size used.

One such feature, the IceTop air shower reconstruction β parameter (see Equation 3.6), correlates with the stage of shower development and depends on both primary particle energy and mass. However, this feature was not included in this work as there is currently poor data-MC agreement for this parameter and efforts are currently being devoted to resolve this issue. Additionally, a log-likelihood ratio which depends on the muon content at ground level has been developed and successfully used for discriminating gamma ray showers from cosmic ray showers using IceTop [70]. Such a feature could, in principle, be used to help discriminate between cosmic rays of different masses. Lastly, with the availability of high-energy simulation datasets the work presented here could be extended to higher energies. This would allow for further comparison with the KASCADE-Grande result in which a softening in the energy spectrum for heavier cosmic rays was observed.

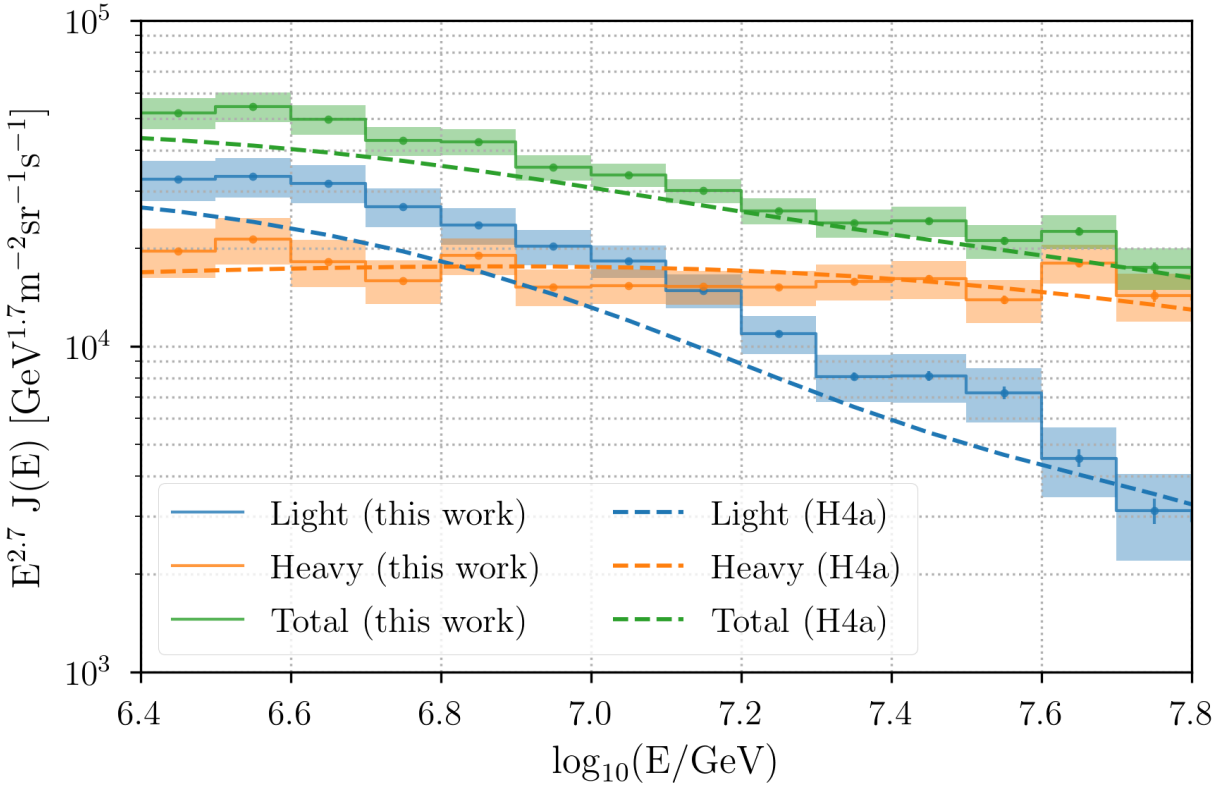


Figure 6.12: Comparison with the H4a cosmic-ray spectrum model [69].

Chapter 7

Energy Spectrum Anisotropy

Equipped with the ability to make light, heavy, and all-particle energy distributions, we can now being to examine the spectrum at different parts of the sky. This chapter presents a search for spectral variations in the light, heavy, and all-particle energy spectrum as a function of arrival direction. We begin by discussing the time and coordinate systems used to define astronomical arrival directions for detected events. Then the analysis method is outlined and corresponding results shown. The chapter ends with an outline of future work and extensions for this analysis.

7.1 Time and Coordinate systems

This work uses the Modified Julian Date (MJD) convention for event times. Based on the historically used Julian Date (JD), the number of days since noon on January 1, 4713 BC, the MJD convention is defined to be:

$$\text{MJD} = \text{JD} - 2400000.5 \tag{7.1}$$

This definition has the effect of shifting the zero point time for the MJD system to be midnight on November 17, 1858. As an example, 9:00 AM CDT on April 26, 2019 has an MJD time of 58599.375.

Astronomical directions for each event detected by IceCube are described using the equatorial coordinate system, which is depicted in Figure 7.1. The equatorial coordinate system, which is a geocentric, fixed-sky spherical coordinate system, is described in terms of two angular coordinates: right ascension α and declination δ. The right ascension coordinate, sometimes abbreviated as RA, for a point on the celestial sphere is defined to be the angular distance, measured counterclockwise along the celestial equator, from the vernal equinox to the hour circle that passes through the point. Right ascension values range from $0^\circ \leq \alpha < 360^\circ$. Declination is the angular distance from the celestial equator to the point on the celestial sphere as measured along the hour circle containing the point. Declination angles cover $-90^\circ \leq \delta \leq 90^\circ$.

An event with local zenith and azimuth coordinates (θ, ϕ) detected at MJD time t can be described in terms of right ascension and declination coordinates (α, δ). Due to IceCube's location at the South Pole, the transformation between local and equatorial coordinates is relatively straightforward. The declination for an event is time-independent and strictly a function of the zenith angle:

$$\delta = \theta - 90^\circ \tag{7.2}$$

In contrast, the right ascension coordinate for an event is time-dependent and determined from the azimuth angle ϕ

$$\alpha = t_s - \phi - 90^\circ \tag{7.3}$$

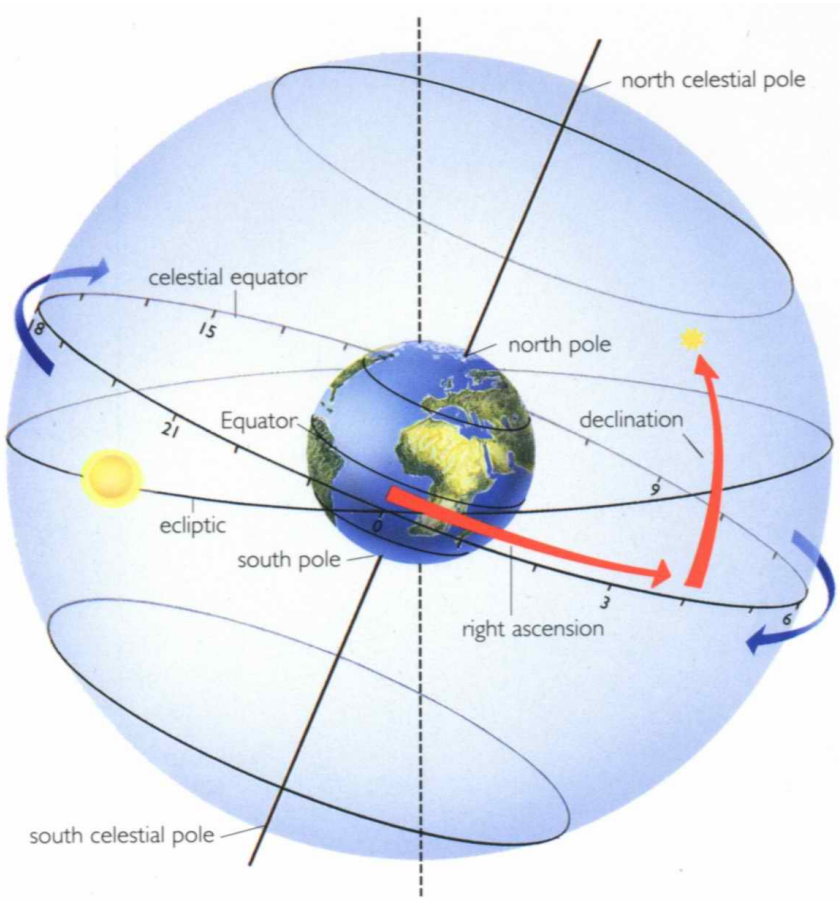


Figure 7.1: The equatorial coordinate system is a geocentric, fixed-sky spherical coordinate system described by two angular coordinates: right ascension α and declination δ. The right ascension α and declination δ for the yellow marker in the illustration are indicated by the angular distance spanned by the red arrows. Figure from [71].

where t_s is given by

$$t_s = 280.46061837 + 1.31850007701 \times 10^7 \; T + 3.87933 \times 10^{-4} \; T^2 - 2.58331181 \times 10^{-8} \; T^3 \quad (7.4)$$

and T is defined as

$$T = \frac{\text{MJD} - 51544.5}{36525} \quad (7.5)$$

7.2 Analysis Method

The analysis method presented here was adapted from the method presented in [28]. The Hierarchical Equal-Area isoLatitude Pixelization (HEALPix) pixelation scheme [72] is used to partition the equatorial coordinate system into equal area pixels on the sky. In this scheme, the number of pixels in the map, n_{pix}, is expressed as $n_{\text{pix}} = 12n_{\text{side}}^2$, where n_{side} is constrained to $n_{\text{side}} = 2^r$ for an integer r between 0 and 10. For this work, we use $n_{\text{side}} = 128$, which corresponds to pixel area of 6.39×10^{-5} sr or 0.2098 square degrees. The arrival direction distribution of all detected events in equatorial coordinates is shown in Figure 7.2.

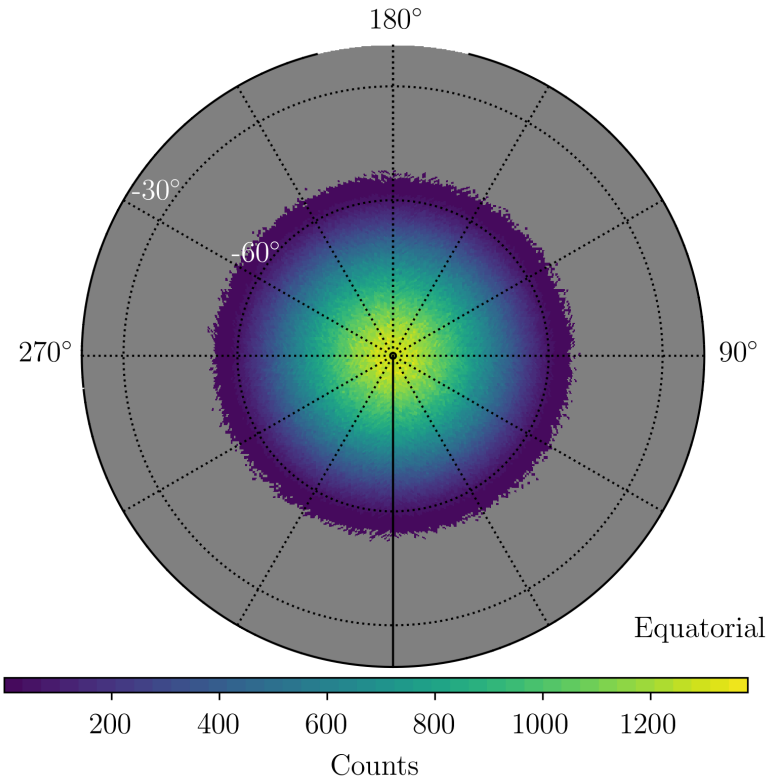


Figure 7.2: Event arrival direction distribution in equatorial coordinates.

To calculate the local deviation in the energy spectrum for a given position on the sky, the spectrum is compared between events in two regions: the on region and the off region. The on region is defined to be centered about the chosen point on the sky and bounded by lines of constant right ascension and declination. The bounding region is the portion of the sky that lies $\pm 5°$ in

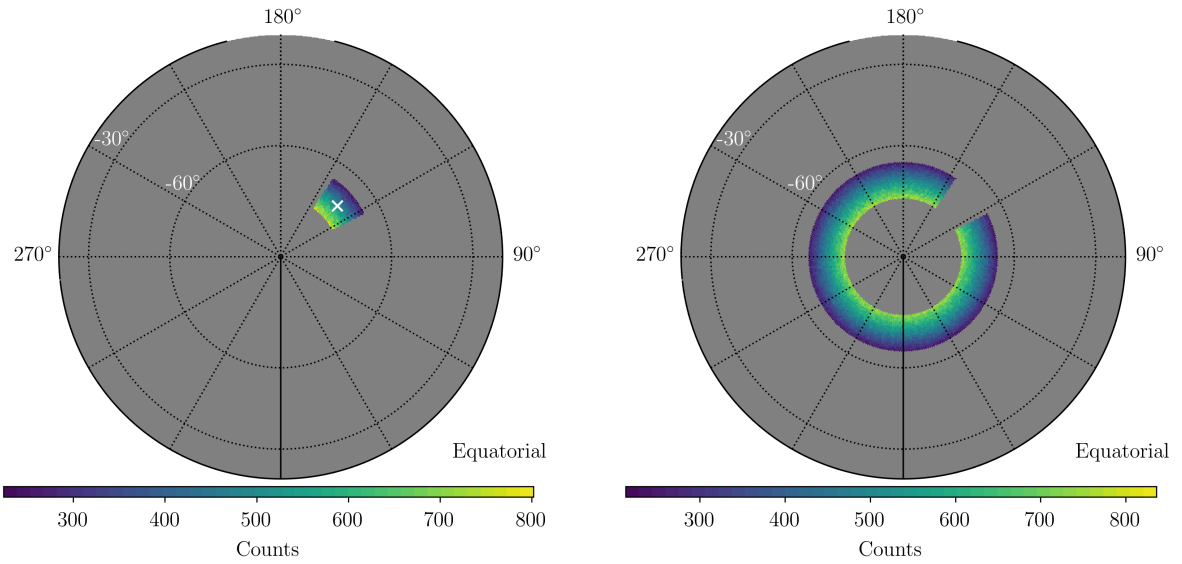


Figure 7.3: Example on and off regions, both shown in equatorial coordinates. The on region is centered at $(\alpha,\ \delta) = (132°, -70°)$ and has a width of 10° in declination. The off region consists of points in the sky that are not contained in the on region but are within the same declination range as the on region.

declination and $\pm 5°/\cos\delta$ in right ascension about the central point. The factor of $\cos\delta$, where δ is the declination of the center of the on region, ensures an equal area on region throughout different locations on the sky. The size of the on region was chosen to provide sufficient data statistics to measure spectrum differences with precision, however there are future plans to further optimize this decision. The associated off region is defined to be the surrounding declination band which contains the on region. Figure 7.3 shows the sky map for events that lie in an example on region centered about $(\alpha,\ \delta) = (132°, -70°)$ and the corresponding declination band off region.

Two energy histograms can then be constructed, one using events that lie within the on region and another for the off region events. In general, the off region will cover a larger portion of the sky and thus contain more events than the corresponding on region. In order to perform a comparison between the shape of the two distributions on equal footing, the energy distribution for the off region is scaled, or normalized, such that it matches the same total number of events as the on region. That is, the number of events in each bin for the off region energy histogram is scaled by the factor $n_{\text{on}}/n_{\text{off}}$, where n_{on} and n_{off} are the total number of events contained in the on and off regions, respectively. Figure 7.4 shows the normalized energy histograms for the example on and

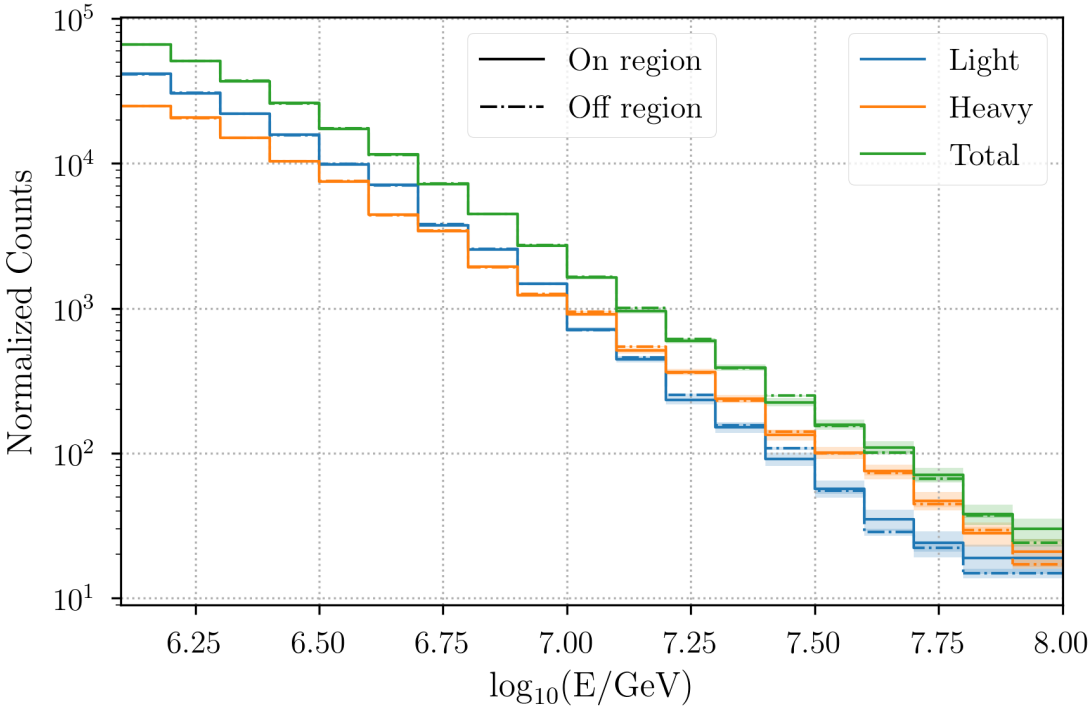


Figure 7.4: Energy distributions for the example on and off regions shown in Figure 7.3. The off region's energy distributions have been normalized such that they contain the same total number of events as the corresponding on region. The shaded bands for the on and off region distribution indicate the $\pm 1\sigma$ statistical uncertainties for the respective distribution.

off regions shown in Figure 7.3.

To quantify how much the on and off region energy distribution deviate from one another, a χ^2 is calculated between the two distributions:

$$\chi^2 = \sum_{i=1} \frac{(N_{\text{on},i} - N_{\text{off},i})^2}{\sigma^2_{\text{on},i} + \sigma^2_{\text{off},i}} \tag{7.6}$$

where $N_{\text{on},i}$ and $N_{\text{off},i}$ are the number of counts in the i-th energy bin for the on and off region histograms, respectively, after normalizing the off region counts distribution. Standard Poisson counting error are assumed for $\sigma^2_{\text{on},i} = N_{\text{on},i}$ and $\sigma^2_{\text{off},i} = N_{\text{off},i}$, where statistical uncertainties are propagated through the off region normalization. Note that the power of this method is that it is concerned with only relative difference between energy distributions. This makes it robust to systematic variations which impact the on and off regions in the same manner.

7.3 Results

The analysis procedure outlined in the previous section can be repeated many times, each time centering an on region about a different pixel in the sky map. In this manner a map of observed spectrum deviation χ^2 values as a function of arrival direction can be built. Here, we restrict to possible declination values between $\delta_{\min} = -85^\circ$ and $\delta_{\max} = -65^\circ$. This restriction is to avoid edge effects at both the equatorial pole, where rectangular on regions extend over the pole, and near the edge of our detector acceptance where there are not sufficient event statistics to produce energy spectra at the highest energies. The resulting χ^2 maps for the light, heavy, and all-particle energy distribution deviations are shown in Figure 7.6.

The observed chi-squared values follow a χ^2 distribution with the expected number of degrees of freedom, one less the number of energy bins due to the off region normalization constraint, as shown in Figure 7.5. This allows us to assign a p-value for each observed χ^2 value by using the cumulative distribution function for the underlying χ^2 distribution. Likewise, it is a common practice to further convert p-values to a significance. The significance, in units of number of standard deviations, is given by $s = \sqrt{2}\ \mathrm{erf}^{-1}(1-p)$ where $\mathrm{erf}(x)$ is the error function $\mathrm{erf}(x) = 2/\sqrt{\pi}\int_0^x \exp(-t^2)dt$ [73].

Significance sky maps are displayed in Figure 7.7. Table 7.1 shows the maximum pre-trial p-value and significance values for all the light, heavy, all-particle spectra. The most significant is the light spectra with a pre-trial significance of 3.98σ, but in all cases we need to determine a post-trial significance.

Spectrum	Light	Heavy	Total
Significance	3.98σ	3.32σ	3.29σ
p-value	6.9×10^{-5}	9.0×10^{-4}	1.0×10^{-3}

Table 7.1: Pre-trial spectrum deviation significance and p-values.

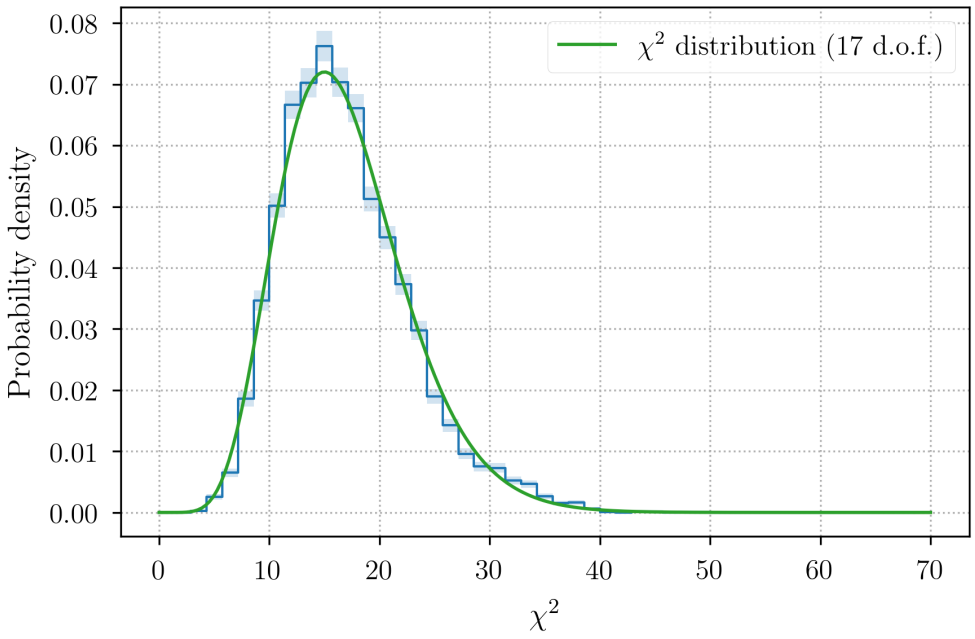
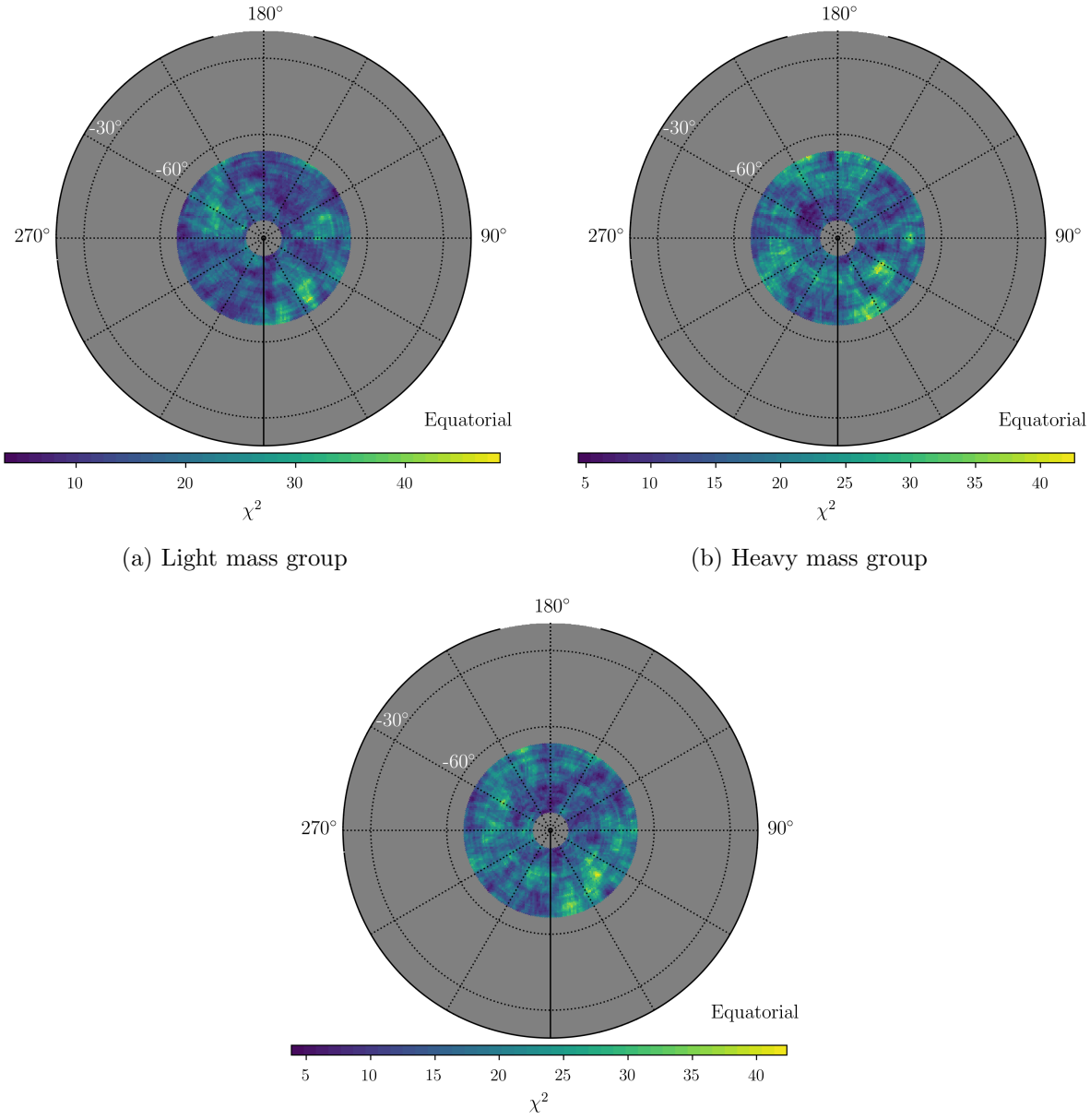


Figure 7.5: Observed chi-squared distribution for the light energy spectrum which has been normalized such that the area is one. A χ^2 probability density function (PDF) with 17 degrees of freedom is shown by the green curve. The observed chi-squared distribution is well described by the χ^2 PDF. There is a similarly good level of agreement between the observed heavy and all-particle chi-squared distributions and the corresponding χ^2 PDFs.

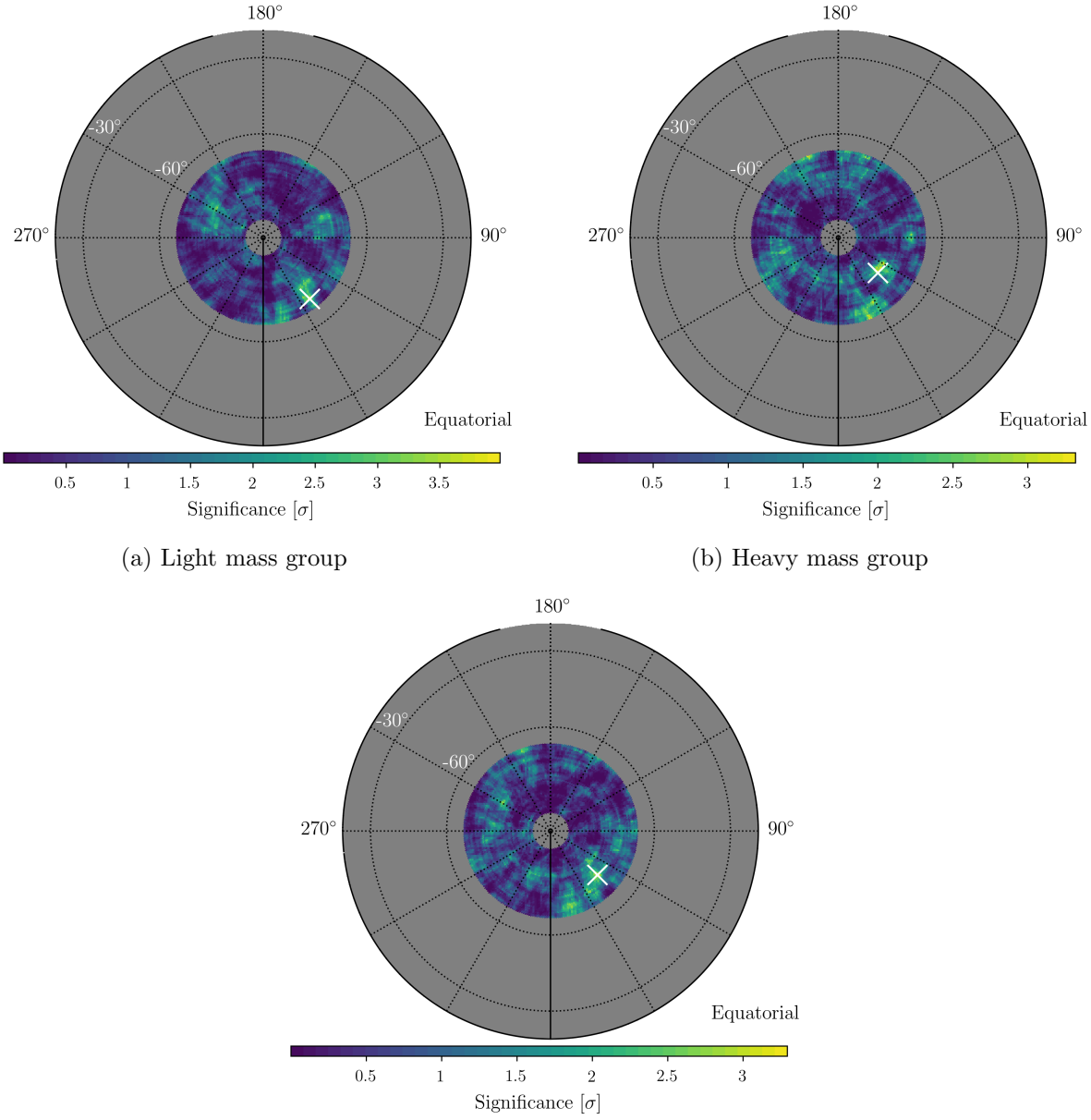


(a) Light mass group

(b) Heavy mass group

(c) All-particle

Figure 7.6: Sky map of observed χ^2 deviations for the light, heavy, and all-particle energy spectra.



(a) Light mass group

(b) Heavy mass group

(c) All-particle

Figure 7.7: Significance distributions for light, heavy, and all-particle energy spectrum deviations. The maximum significance deviation is indicated with the white "X" marker.

7.3.1 Post-trial Significance

To study anisotropies in our dataset, we must first have a detailed understanding of how our detector responds to an isotropic flux of cosmic rays. Anisotropies in the arrival direction distribution can originate from either variations in detector acceptance or from astrophysical origins. Due to both the flat geometry and hexagonal design of the IceCube Observatory, the detected rate of cosmic rays is a strong function of event arrival direction. Figure 7.8 shows the distribution of zenith θ and azimuth ϕ coordinates for the events in our dataset.

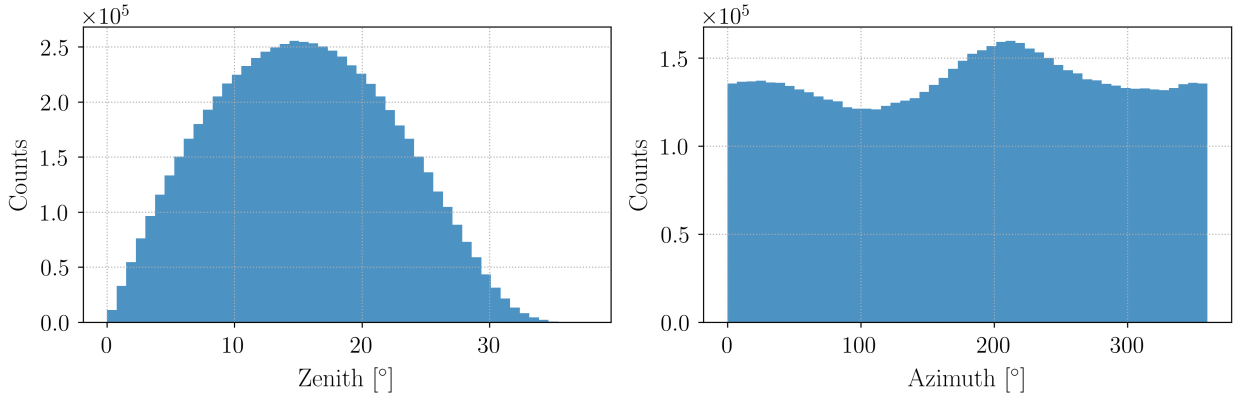


Figure 7.8: Distribution of zenith θ and azimuth ϕ coordinates for detected events. The flat geometry and hexagonal design of IceCube give rise to a strong event arrival direction dependence for the detector acceptance. In addition, a large scale azimuth anisotropy originates from variable snow heights atop the IceTop tanks.

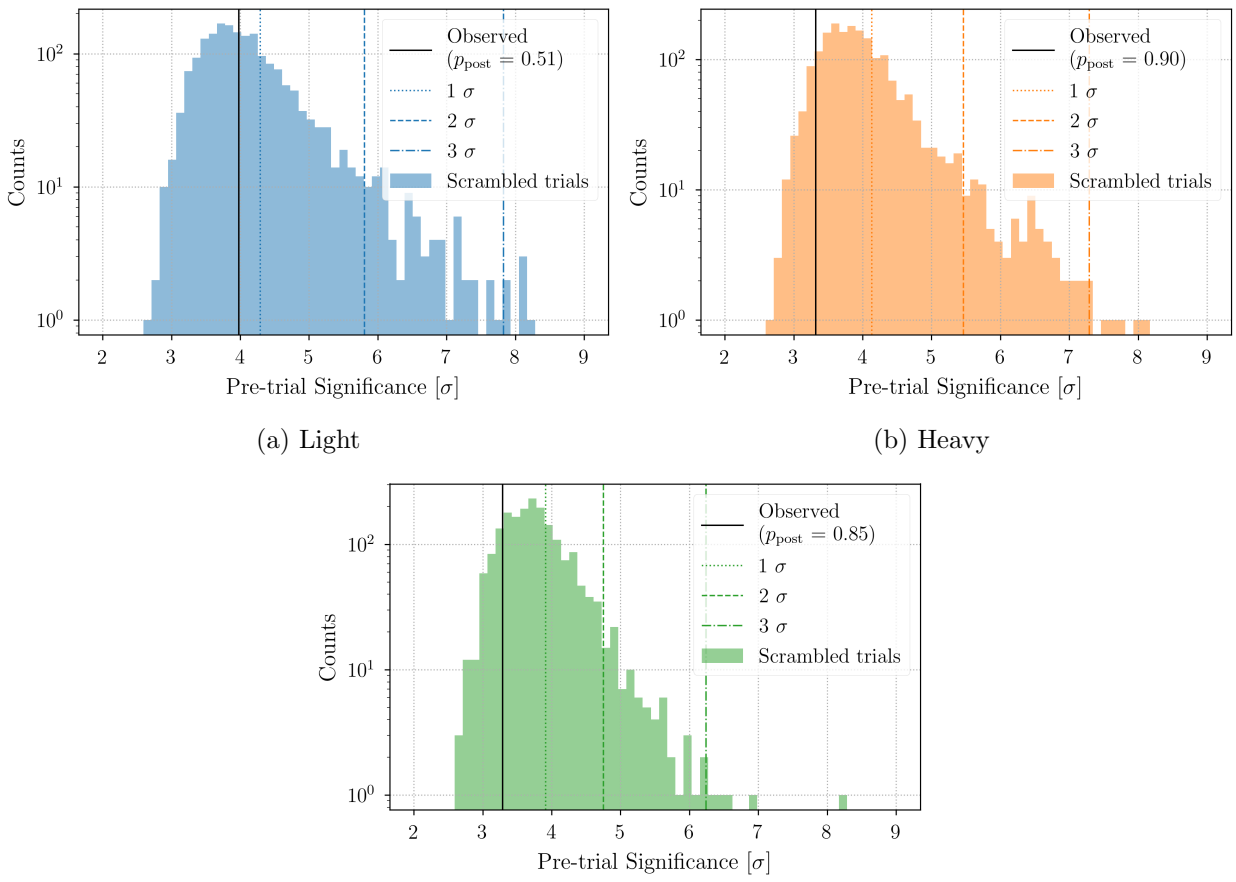
To study anisotropies involving cosmic-ray arrival directions, it's crucial to estimate the detector response to an isotropic incident flux of cosmic rays. This is accomplished by using the well-known time scrambling method presented in [74]. This data-driven algorithm takes as input a dataset of local arrival coordinates and times $\{\theta, \phi, t\}$ and produces an output dataset of modified local coordinates that are an estimate of the exposure-weighted detector response to an isotropic flux. It does this by keeping the detector (θ, ϕ) coordinates for each event fixed, while randomly rearranging the detected times (hence the "time scrambling" name). By randomly sampling from the measured times without replacement, we are able to keep track of gaps in data acquisition. From the discussion in Section 7.1, time scrambling has the effect of randomizing the arrival direction distribution in right ascension, while maintaining the characteristic non-uniformities in detector

coordinates. This decouples the anisotropies in the local and equatorial coordinate systems.

To compute a post-trial significance for the observed spectral deviations shown in Figure 7.7, the time scrambling algorithm was used to produce 2,000 randomized datasets. The same analysis method was then applied to produce a maximum spectrum deviation significance for each randomized trial. The distribution of pre-trial significance values, along with the observed pre-trial significance, for the light, heavy, and all-particle deviations, is shown in Figure 7.9. Post-trial p-values are calculated by comparing the observed pre-trial significance to the corresponding distribution for the scrambled trials. The resulting post-trial p-values are 0.51, 0.90, and 0.85 for the light, heavy, and all-particle spectra, respectively, which are consistent with the background expectation.

Spectrum	Light	Heavy	Total
Significance	0.66σ	0.13σ	0.19σ
p-value	0.51	0.90	0.85

Table 7.2: Post-trial spectrum deviation significance and p-values.



(a) Light

(b) Heavy

(c) All-particle

Figure 7.9: Significance distributions for time scrambled light, heavy, and all-particle energy spectrum deviations.

7.4 Future Work

While we have shown the post-trial significance of the spectrum anisotropy analyses in this work, future studies can be done to interpret the results of our analysis as a limit on spectral variations. To do so we must choose a particular hypothesis to test. Generally speaking, any model which parameterizes the observed energy distributions could be used. However, for simplicity of the discussion here we will consider the case of a single power law.

The steps for setting a limit on the spectral index range this analysis method is sensitive to are outlined in what follows. First, a power law model can be fit to the unfolded off region energy spectrum for the maximum pre-trial significance location on the sky map (Figure 7.7). This fitted index γ_{ref}, serves as the reference expected spectral index for the off region. Next, we inject the reference spectrum plus a small deviation ϵ, i.e a power law spectrum with $\gamma_{\text{inj}} = \gamma_{\text{ref}} + \epsilon$ where ϵ is a deviation parameter that can be varied. The injection is done by forward-folding the injected spectrum through the response matrix to produce an observed counts distribution in the on region. This counts distribution represents the average observed detector response to the true injected spectrum with index γ_{inj}. We can then generate many randomized realizations of this response by adding Poisson fluctuations about the average distribution. For each of these randomized trials, we can determine the deviation relative to the background off region and compute a pre-trial significance. This allows us to build up a distribution of pre-trial significance values similar to those seen in Figure 7.9. The value for the index deviation parameter ϵ can be adjusted, resulting in a shift of the significance distribution for the trials, until a specified confidence level for the post-trial significance is reached. In this manner we can set a limit on the range of possible spectral deviations allowed by our measurement.

In addition, the same method used to generate our current limits can also be used to study how the sensitivity of this analysis changes as a function of the number of energy bins, size of on region, etc. This allows us to tune these parameters such that the analysis sensitivity is optimized. In the future we plan to do these studies and apply the optimizations to additional years of data.

Lastly, a variation of this analysis which uses non-coincident events detected only by IceTop could prove fruitful. The IceTop air shower reconstruction will still provide the energy proxy S_{125}

and local arrival direction θ and ϕ with a comparable performance to what is achieved in this work. Mass information, however, will be lost due to the lack of a composition-sensitive IceTop parameter. Thus this type of analysis would be restricted to studying variation in only the all-particle spectrum. The added benefit in this case is the ability to extend the available field of view to declinations just under -30° while also increasing the detected event rate.

Chapter 8

Conclusions

The results of a composition-resolved cosmic-ray energy spectrum analysis at PeV energies using one year of data collected by the IceCube Observatory has been reported. The analysis focused on making a precise two mass group and all-particle energy spectrum measurement in the range from $6.4 \leq \log_{10}(E/\text{GeV}) \leq 7.8$. The resultant unfolded energy spectra are shown in Figure 6.4, 6.6, and 6.9. We observed the light mass group spectrum is softer than that of the heavier mass group, which follows a power-law like structure with a spectral index of ~ 2.7 throughout the entire energy range considered. The transition from a primarily light to heavy-dominant spectrum takes place near $\log_{10}(E/\text{GeV}) = 7.1$. This feature is characteristic of a potential rigidity-dependent cutoff, or Peters cycle. The change in relative mass abundances could also indicate a possible transition in the source population of cosmic rays.

The impact of systematic uncertainties introduced by the unfolding method, detector-related uncertainties, and choice of hadronic interaction model were also considered, with their effects summarized in Tables 6.1, 6.2, and 6.3. There is work in progress to reduce the uncertainty introduced by the modeling of photon scattering and absorption properties in ice. Furthermore, improvements to the modeling of hadronic interactions is an ongoing effort in both the particle and cosmic-ray physics communities. Both these sources of systematic uncertainty are expected to decrease over time.

Additionally, a study of how the light, heavy, and all-particle energy spectra deviate with arrival

direction was presented. This marks the first time an analysis of this kind has been conducted using the IceCube Observatory. No statistically significant spectrum deviations above the background expectation were observed. The results from this analysis can be used to set a limit on the range of possible spectral deviations. Moreover, a potential future version of the analysis presented here, in which non-coincident events detected by IceTop are used to study variations in the all-particle energy spectrum as a function of arrival direction over a larger field-of-view, could prove valuable.

Appendix A

Energy Spectrum Results

$\log_{10}(E/\mathrm{GeV})$	J(E) [$\mathrm{GeV}^{-1}\mathrm{m}^{-2}\mathrm{sr}^{-1}\mathrm{s}^{-1}$]	$\mathrm{sys}_{\mathrm{MC}}$	stat
6.4-6.5	2.004174e-13	2.207437e-14	3.586689e-16
6.5-6.6	1.127214e-13	1.187095e-14	2.453493e-16
6.6-6.7	5.535148e-14	5.781579e-15	1.533415e-16
6.7-6.8	2.549952e-14	2.593787e-15	9.254202e-17
6.8-6.9	1.361629e-14	1.241593e-15	6.371099e-17
6.9-7.0	6.097723e-15	5.437453e-16	3.801573e-17
7.0-7.1	3.103432e-15	2.553750e-16	2.425460e-17
7.1-7.2	1.491378e-15	1.274732e-16	1.493259e-17
7.2-7.3	6.945824e-16	6.409421e-17	8.896504e-18
7.3-7.4	3.417182e-16	3.436946e-17	5.548250e-18
7.4-7.5	1.861383e-16	1.976928e-17	3.605669e-18
7.5-7.6	8.687299e-17	1.024688e-17	2.180406e-18
7.6-7.7	4.994470e-17	5.932316e-18	1.384380e-18
7.7-7.8	2.072837e-17	3.036319e-18	8.117396e-19

Table A.1: Values of the all-particle cosmic-ray energy spectrum for $6.4 \leq \log_{10}(E/\mathrm{GeV}) \leq 7.8$ including uncertainties. The label "stat" represents the statistical uncertainties, "$\mathrm{sys}_{\mathrm{MC}}$" is for the uncertainties from the Monte Carlo.

$\log_{10}(E/\mathrm{GeV})$	J(E) [$\mathrm{GeV}^{-1}\mathrm{m}^{-2}\mathrm{sr}^{-1}\mathrm{s}^{-1}$]	$\mathrm{sys}_{\mathrm{MC}}$	stat
6.4-6.5	1.250859e-13	1.789549e-14	2.823013e-16
6.5-6.6	6.871850e-14	9.455609e-15	1.915450e-16
6.6-6.7	3.517916e-14	4.728827e-15	1.212248e-16
6.7-6.8	1.601216e-14	2.148505e-15	7.361451e-17
6.8-6.9	7.533558e-15	9.588045e-16	4.683935e-17
6.9-7.0	3.480464e-15	4.274985e-16	2.865391e-17
7.0-7.1	1.688323e-15	1.897800e-16	1.788149e-17
7.1-7.2	7.359461e-16	8.892008e-17	1.045739e-17
7.2-7.3	2.909240e-16	3.932831e-17	5.600456e-18
7.3-7.4	1.155397e-16	1.897893e-17	3.309873e-18
7.4-7.5	6.214693e-17	1.064254e-17	2.153398e-18
7.5-7.6	2.967825e-17	5.594519e-18	1.315054e-18
7.6-7.7	1.004167e-17	2.409821e-18	6.287926e-19
7.7-7.8	3.716937e-18	1.101231e-18	3.335798e-19

Table A.2: Values of the light cosmic-ray mass group energy spectrum for $6.4 \leq \log_{10}(E/\mathrm{GeV}) \leq 7.8$ including uncertainties. The label "stat" represents the statistical uncertainties, "$\mathrm{sys}_{\mathrm{MC}}$" is for the uncertainties from the Monte Carlo.

$\log_{10}(E/\mathrm{GeV})$	J(E) [$\mathrm{GeV}^{-1}\mathrm{m}^{-2}\mathrm{sr}^{-1}\mathrm{s}^{-1}$]	$\mathrm{sys}_{\mathrm{MC}}$	stat
6.4-6.5	7.533151e-14	1.292396e-14	2.208611e-16
6.5-6.6	4.400294e-14	7.177102e-15	1.531416e-16
6.6-6.7	2.017232e-14	3.326385e-15	9.383709e-17
6.7-6.8	9.487359e-15	1.453154e-15	5.605526e-17
6.8-6.9	6.082731e-15	7.888254e-16	4.317800e-17
6.9-7.0	2.617258e-15	3.360117e-16	2.497970e-17
7.0-7.1	1.415109e-15	1.708797e-16	1.638581e-17
7.1-7.2	7.554321e-16	9.133800e-17	1.065905e-17
7.2-7.3	4.036584e-16	5.060978e-17	6.912348e-18
7.3-7.4	2.261785e-16	2.865413e-17	4.452791e-18
7.4-7.5	1.239914e-16	1.666014e-17	2.891987e-18
7.5-7.6	5.719474e-17	8.584862e-18	1.739187e-18
7.6-7.7	3.990303e-17	5.420805e-18	1.233337e-18
7.7-7.8	1.701143e-17	2.829580e-18	7.400301e-19

Table A.3: Values of the heavy cosmic-ray mass group energy spectrum for $6.4 \leq \log_{10}(E/\mathrm{GeV}) \leq 7.8$ including uncertainties. The label "stat" represents the statistical uncertainties, "$\mathrm{sys}_{\mathrm{MC}}$" is for the uncertainties from the Monte Carlo.

Appendix B

Unfolding Validation

The iterative unfolding algorithm presented in Chapter 5, being based on Bayes' theorem, requires the specification of a prior distribution. The choice of prior is a free parameter in the algorithm and acts as the seed, or starting point, for unfolding iterations.

Ideally, the choice of a prior distribution should affect the final unfolded distribution minimally. For example, in the case where two priors are tested, one which is close to the true underlying distribution and one which differs substantially from the true distribution, the desired behavior is for both to converge to the same final unfolded distribution. However, the case in which the prior distribution varies from the true distribution may take more unfolding iterations to converge and result in larger uncertainties in the unfolded distribution.

This appendix presents the results from investigating what effect the choice of prior has on this analysis in two ways: by unfolding the observed data distribution using several different choices of prior and by injecting known simulated spectra and unfolding with different priors.

B.1 Unfolding Data

As discussed in Section 6.2, we have chosen to use the Jeffreys prior in this work to avoid any a priori assumptions about the spectra. However, one should always ensure that other choices of prior do not drastically influence the final unfolded distribution for a measurement. We compare the results from using several priors to unfold our observed composition-energy counts distribution, including uniform, Jeffreys, simple power law, and H3a/H4a [69] composition model priors. The corresponding unfolded spectra are shown in Figure B.1. For ease of visual comparison, we only show the systematic uncertainties due to finite response matrix Monte Carlo for the nominal Jeffreys prior, as indicated by the shaded bands. However, each of the other choices of prior has uncertainty of a comparable magnitude. We see that the final unfolded distributions for each of the choices of prior is consistent with the nominal result using the Jeffreys prior, i.e. the unfolding procedure is performing as expected.

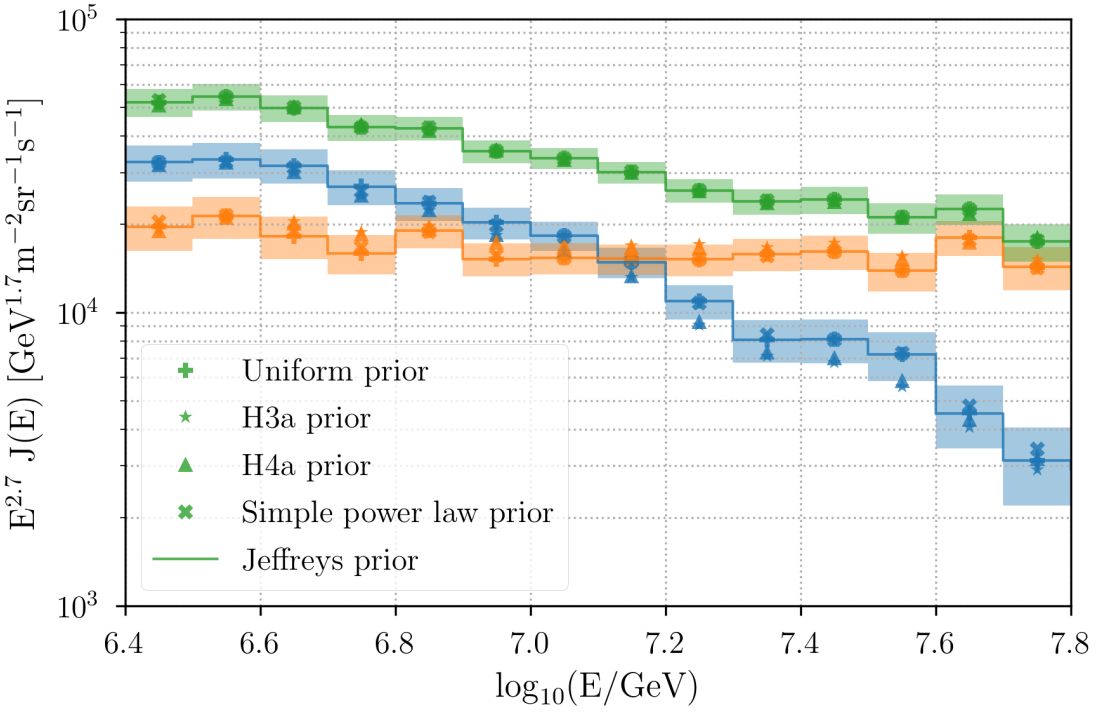


Figure B.1: Unfolded light, heavy, and all-particle energy spectra for various choices of the unfolding prior.

B.2 Unfolding Injected Spectra

The previous validation check ensures that the unfolding process converges to the same final unfolded distribution, irrespective of the prior. However, it does not check whether or not the unfolded solution is consistent with the underlying true cause distribution. To test that this is in fact the case, we weight Monte Carlo simulation to an assumed spectrum, pass this through the analysis procedure, and compare the resulting unfolded spectrum with the original, known injected spectrum. Figures B.2 and B.3 show both the injected and corresponding unfolded flux for a composition-dependent broken power law and the H4a spectrum model, respectively. In both these cases, as well as with all other injected spectra considered, the final unfolded flux was consistent with the injected flux within the unfolding uncertainties.

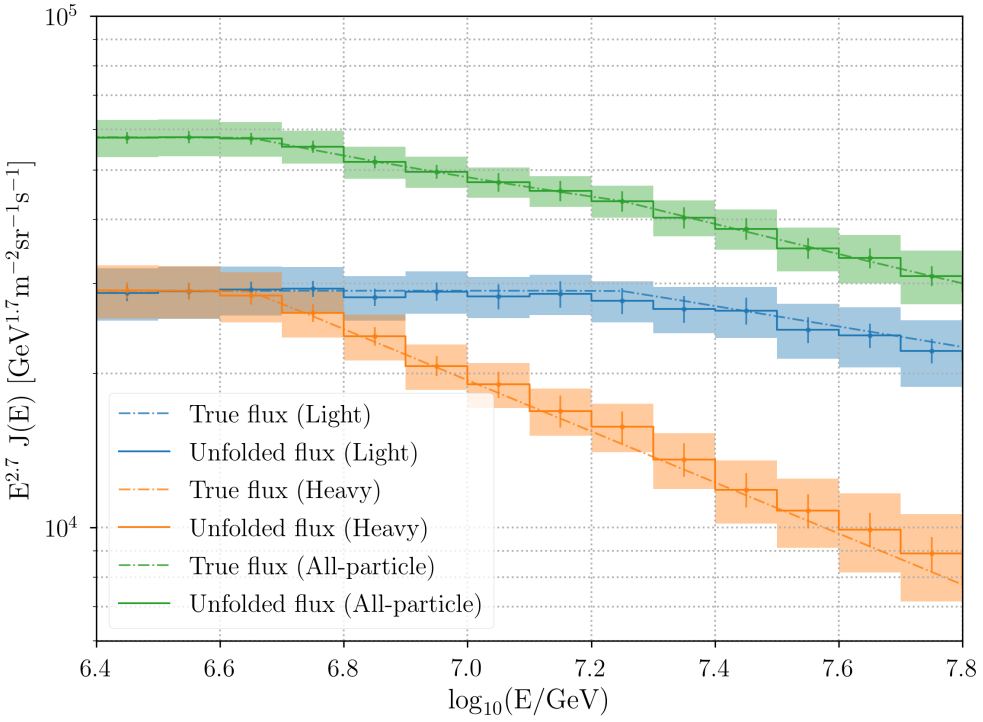


Figure B.2: Unfolded flux comparison for an injected broken power law composition assumption. The injected flux is shown with dashed lines while the corresponding unfolded spectrum is represented by the solid histogram. Shaded bands indicate the uncertainty due to finite Monte Carlo available to build the detector response matrix and the error bars show statistical errors.

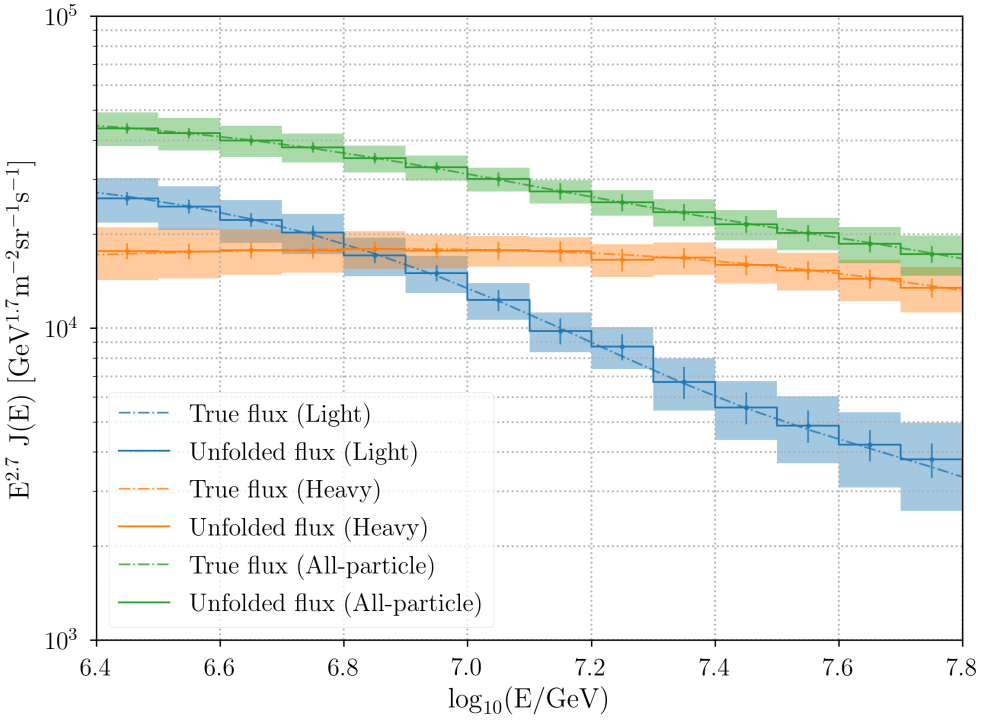


Figure B.3: Unfolded flux comparison for an injected H4a composition assumption. The injected flux is shown with dashed lines while the corresponding unfolded spectrum is represented by the solid histogram. Shaded bands indicate the uncertainty due to finite Monte Carlo available to build the detector response matrix and the error bars show statistical errors.

Appendix C

PyUnfold

As part of the work for this thesis, Zigfried Hampel-Arias and I developed PyUnfold, a Python package for implementing iterative unfolding. The source code for PyUnfold is available on GitHub[1] and the package has been published in the Journal of Open Source Software (JOSS) [58].

C.1 Why Another Package?

The primary motivating factor for creating PyUnfold was to provide an unfolding toolkit that's accessible to members of all scientific disciplines in an easy-to-use package. While unfolding methods are commonly used in scientific analysis in the high-energy physics (HEP) community, the deconvolution packages used in HEP maintain a strong dependence on the ROOT data analysis framework [75]. ROOT, while a very useful piece of software, is rarely used outside the HEP community.

Instead, PyUnfold is built on top of the Python scientific computing stack (i.e. NumPy, SciPy, and pandas [76]), thus broadening its scope to a general scientific audience. PyUnfold has been designed to be both easy to use for first-time users as well as flexible enough for fine-tuning an analysis and testing the robustness of results.

[1] PyUnfold repository is available at https://github.com/jrbourbeau/pyunfold

C.2 Features

PyUnfold provides support for the following features:

- Custom, user defined initial prior probability distributions, the default being the uniform prior. The non-informative Jeffreys prior [56] is accessible as a utility function.
- Unfolding stopping criteria based on test statistic calculations comparing unfolded distributions from one iteration to the next. These include Kolmogorov-Smirnov [77][78], χ^2, relative difference, and Bayes factor [79] tests.
- Tunable spline regularization as a means of ensuring that unfolded distributions do not suffer from growing fluctuations potentially arising from the finite binning of the response matrix.
- Option to choose between Poisson or multinomial forms of the covariance matrices for both the data and response contributions to the uncertainty calculation.
- Multivariate unfolding via definitions of subsets of causes, which are regularized in their respective blocks or groups.

Further mathematical details regarding the iterative unfolding procedure, including complete derivations of the statistical and systematic uncertainty propagation, can be found in the online documentation.

C.3 Example Usage

```
from pyunfold import iterative_unfold

# Observed distributions

data = [100, 150]

data_err = [10, 12.2]

# Response matrix

response = [[0.9, 0.1],

            [0.1, 0.9]]

response_err = [[0.01, 0.01],

                [0.01, 0.01]]

# Detection efficiencies

efficiencies = [1, 1]

efficiencies_err = [0.01, 0.01]

# Perform iterative unfolding

unfolded_result = iterative_unfold(data=data,

                                   data_err=data_err,

                                   response=response,

                                   response_err=response_err,

                                   efficiencies=efficiencies,

                                   efficiencies_err=efficiencies_err)
```

Appendix D

Distributed Electronic Cosmic-Ray Observatory

This appendix provides an overview of the work presented in *Particle Identification In Camera Image Sensors Using Computer Vision* [80]. Please see the publication for full details.

D.1 The DECO Application

The Distributed Electronic Cosmic-ray Observatory (DECO) is a global network of smartphones that use an Android application developed to detect ionizing radiation using smartphone camera image sensors. The DECO application was released publicly in September 2014 and, as of December 2017, has been deployed on every continent in 80 different countries. Figure D.1 shows a world map of DECO data taking locations.

As energetic charged particles pass through a phone's image sensor, electron-hole pairs are created and detected by the sensor. The DECO app takes long-duration exposures (∼50 ms) and then applies onboard filtering to tag potentially interesting images as "events". Tagged event images, along with additional metadata, are then uploaded to a centralized database for further offline analysis.

Figure D.1: World map showing the global network of DECO users. Dots indicate data taking locations and span 80 different countries. Every continent including Antarctica is represented. Lines of data points, such as those in Antarctica and west of the Americas, indicate users running DECO on plane flights. The map is plotted with a Kavrayskiy VII projection and with data collected as of December 2017. Figure from [80].

D.2 Event Types

There are three different types of charged particle events that DECO observes: tracks, worms, and spots. These different categories of events are characterized by their morphology, as described in Table D.1.

Event type	Morphology Description
Track	Long, straight clusters of pixels in an image created by high-energy (GeV) minimum-ionizing cosmic rays
Worm	Curved clusters of pixels caused by the meandering paths of electrons that have undergone multiple Coulomb scattering interactions. Most likely created by photons from radioactive decays which undergo Compton scattering and create energetic electrons.
Spot	Smaller, approximately circular clusters of pixels that can be created by various interactions

Table D.1: DECO event type descriptions.

Figure D.2 shows a few representative examples for each event type. In addition to the three

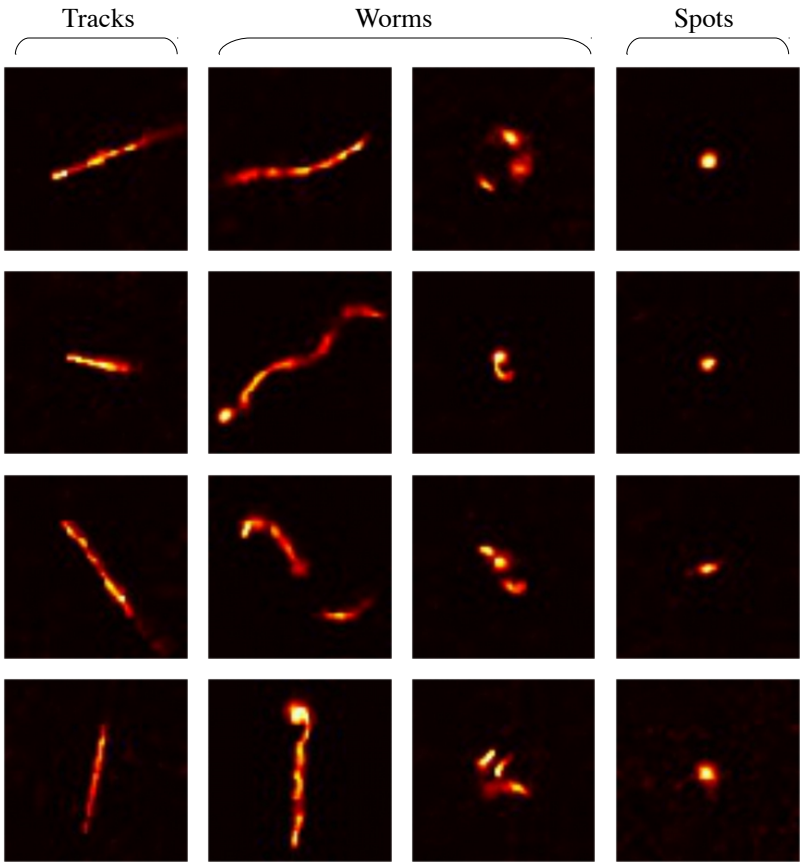


Figure D.2: Representative sample of the three distinct types of charged particle events that require classification. Tracks and spots, left and right columns, respectively, are generally observed to have consistent and predictable features. Worms, middle two columns, are observed to have a much wider variety of features, many of which present potential classification confusion when compared to track-like and spot-like features. Each image above has been converted to grayscale and cropped to 64×64 pixels. Figure from [80].

particle interaction-related event types, there are non-particle induced events in the DECO dataset as well. These events, called noise events, arise from camera sensor hot spots, thermal noise fluctuations, and large-scale sensor artifacts such as entire rows of illuminated pixels. While noise events aren't particularly interesting from an analysis point of view, they do act as a potential source of misidentification for the particle events and need to be accounted for when developing an event classification model.

In order to train an image classification model, sufficient training data is required. Due to the current lack of Monte Carlo event simulation, we chose to construct a human labeled dataset

for training. Labeling the entire DECO dataset (∼45,000 event images) by hand would be a very laborious, and unnecessary, task. Instead, we opted to label 5119 images in total, of which there are 2520 (49%) noise, 1094 (21%) spot, 1063 (21%) worm, and 442 (9%) track images (see Figure D.3).

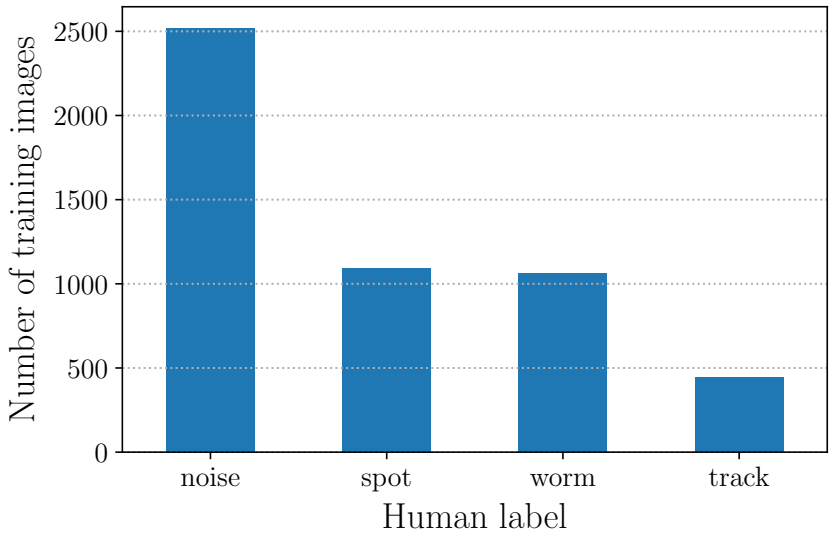


Figure D.3: Number of training images for each event type that was used to train the best performing model. Out of the 5119 total images, there are 2520 (49%) noise, 1094 (21%) spot, 1063 (2%) worm, and 442 (9%) track images. Figure from [80].

D.3 Particle Identification Model

A convolutional neural network (CNN) was constructed and trained with our labeled dataset to identify event images by their event category (track, worm, spot, or noise). The model architecture, specified in Table D.2, is loosely based on the VGG-16 network [81] and consists of several sequential convolution/max pooling units for feature extraction which feeds into a fully connected network for classification. To build and train our CNN model, we used the Keras Python neural network framework [82].

In order to both avoid overfitting and build scale and rotational invariance into our model predictions, we used several data augmentation techniques during the training procedure. Specifically,

	Layer	Features	Size	Activation	Dropout
1	Convolution	64	3×3	Leaky ReLU	-
2	Convolution	64	3×3	Leaky ReLU	-
3	Max Pooling	-	2×2	-	0.2
4	Convolution	128	3×3	Leaky ReLU	-
5	Convolution	128	3×3	Leaky ReLU	-
6	Max Pooling	-	2×2	-	0.2
7	Convolution	256	3×3	Leaky ReLU	-
8	Convolution	256	3×3	Leaky ReLU	-
9	Max Pooling	-	2×2	-	0.2
10	Convolution	512	3×3	Leaky ReLU	-
11	Convolution	512	3×3	Leaky ReLU	-
12	Max Pooling	-	2×2	-	0.2
9	Dense	2048	-	Leaky ReLU	0.4
10	Dense	2048	-	Leaky ReLU	0.4
11	Dense	4	-	softmax	-

Table D.2: Layer-by-layer summary of the best performing network. Each layer name is given followed by the number of feature maps (convolutional layers) or neurons (dense layers), the size of the convolutional filter or pooling region, the activation function used, and, lastly, the dropout probability applied. For the leaky ReLU activation function, the value of α was set to 0.3 in all cases. A max-norm constraint of 3 was used for both 2048 dense (fully connected) layers. Dropout with a probability $P = 0.2$ was also applied to the input layer (not listed in the table). Table from [80].

during each epoch of training, images were subjected to each of the following augmentations:

- Translation: Random left/right and up/down shifts, each by an integer number of pixels uniformly sampled between -8 and +8 with respect to the image center.
- Rescale: Random zoom in/out uniformly sampled between 90% and 110% of the original image size, used for learning scale invariance.
- Reflection: Random horizontal and vertical reflections, each with a probability of 50%.
- Rotation: Random rotation uniformly sampled between 0° and 360°; used for learning rotation invariance. After the rotation, any remaining pixels outside the boundaries of the original input were assigned a value of 0.
- Crop: Crop from 100×100 pixels to 64×64 pixels; used to reduce the amount of empty space created on the boundaries of the image as a result of rotation, translation, and rescaling.

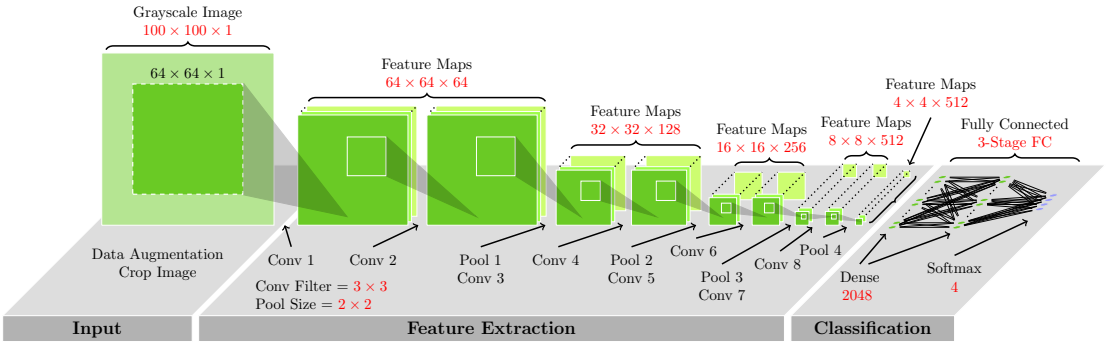


Figure D.4: Block diagram of the best performing network trained in this study. The input and output dimensions for each operation are shown to the left and right of the arrows, respectively. All convolutional filters are 3×3 and all pooling operations are 2×2 max pooling. Following the fourth pooling layer, the feature maps are flattened to a single 1-dimensional vector of length 8196, which is then used as input for the first dense layer. Figure from [80].

D.4 Model Performance

In order to assess the CNN model performance, a stratified 10-fold cross-validation method [83] was used to estimate the quality of the model predictions on unseen data. A summary of the CNN categorization accuracy is shown in Figure D.5. One distinctive feature of this confusion matrix is a bias towards the relative occurrence of each category in the training set. To account for this imbalance, Figure D.6 shows a normalized confusion matrix where each row is normalized by the total number of human-labeled events for each category. From Figure D.6 we can see, for example, the model correctly identifies human-labeled tracks 92% of the time, while incorrectly identifying them as worms 9% of the time. This confusion in the classifier is both expected and comparable to human performance, given that, out of the four categories in the model, track and worm event morphologies are among the most similar.

D.5 Conclusions and Future Work

This classification algorithm has been integrated into the standard DECO processing pipeline and the resulting classification of each event is available along with the event's image and metadata on the public web site[1] within several hours of detection. The CNN classification can be used in

[1] https://wipac.wisc.edu/deco/data

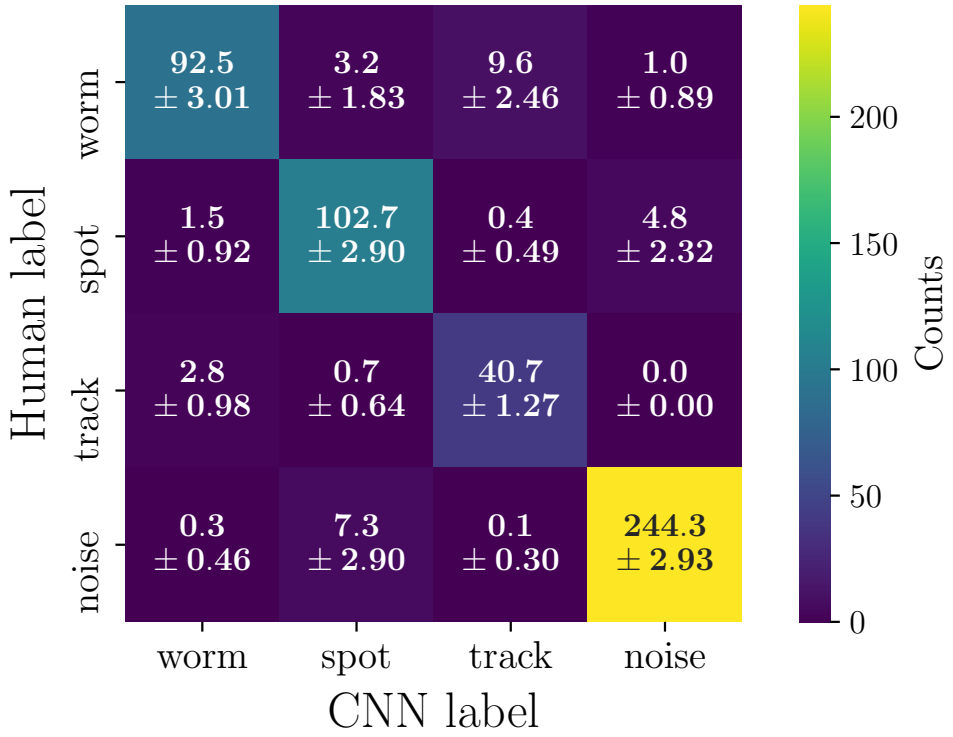


Figure D.5: Confusion matrix summarizing the CNN categorization accuracy. The vertical axis shows the ground truth (human-determined) classification and the horizontal axis shows the classification from the CNN. The values shown in the confusion matrix are the average and standard deviation of the testing set results from the 10-fold cross validation. Figure from [80].

queries, allowing users to select a sample of images of any particle identity, or multiple identities, for analysis and outreach purposes.

In addition to improving the overall experience of DECO users, the new model opens the door for new and improved analyses. For example, the model provides efficient rejection of the radioactive background (i.e., worms), which is necessary to detect extensive air showers using DECO or a similar application.

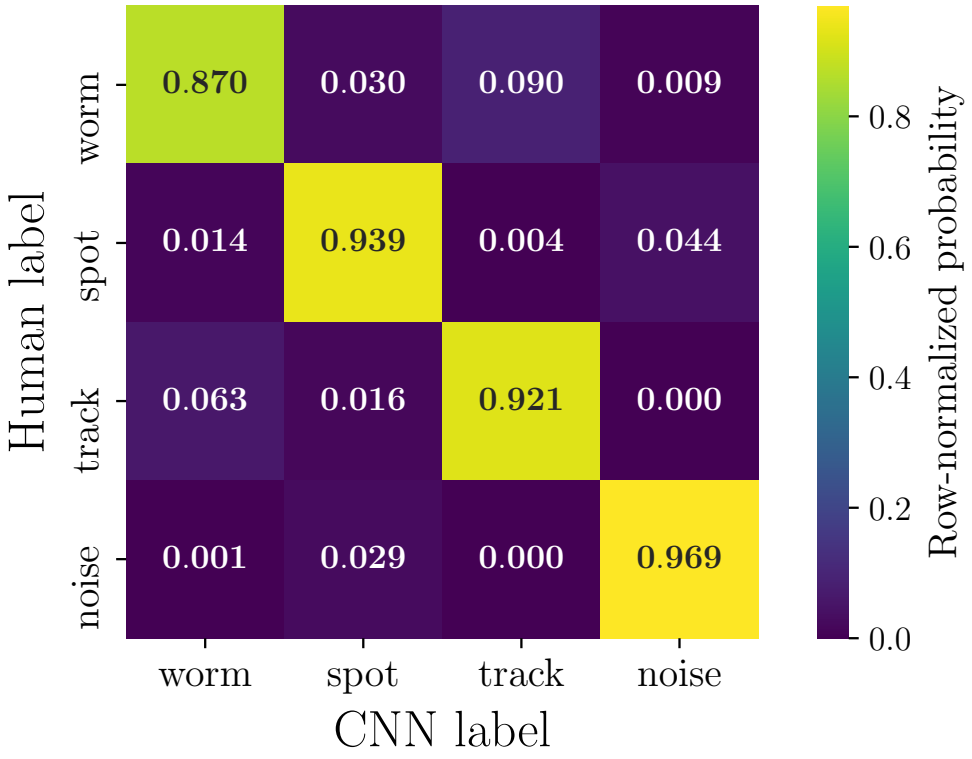


Figure D.6: Row-normalized confusion matrix that accounts for the relative imbalance in the number of testing examples for each category in the training set. Normalization is performed independently for each row and is calculated by dividing each row of the un-normalized confusion matrix by the total number of events in that row. Figure from [80].

Appendix E

Dark Matter Searches with a Mono-Z′ Jet

This appendix provides a brief highlight of the work presented in *Dark Matter Searches with a Mono-Z′ Jet* [84]. Please see the publication for full details.

E.1 Introduction

After the discovery of the Higgs boson in 2012, searching for physics beyond the Standard Model (SM) has become the highest priority at the Large Hadron Collider (LHC). One of the most important new BSM particles is dark matter, whose existence has long been established from astrophysical observations. In spite of a long history of searching for dark matter particles from direct detection, indirect detection and accelerator-based experiments, there is still no clear evidence for the particle nature of dark matter.

While many existing studies [85] [86] [87] have concentrated on identifying signals using the initial state radiation (ISR) of partons inside an accelerated proton, less attention has been paid to potential dark matter final state radiation (FSR). The basic process is that dark matter is pair produced, after which one of the particles can radiate a dark Z′, as illustrated in Figure E.1. The Z′ from FSR can decay back to SM particles and result in visible signals in a collider, while there can still be substantial missing transverse momentum from the dark matter particles.

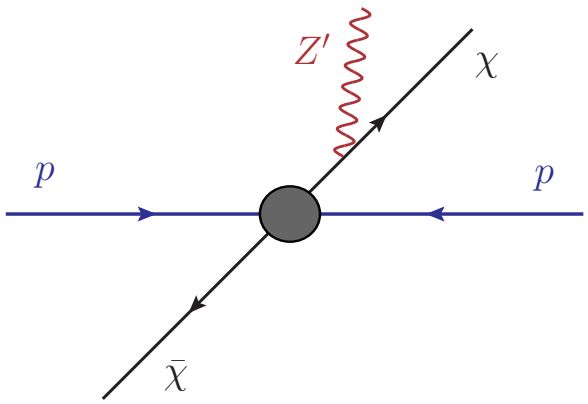


Figure E.1: An illustrative diagram for the mono-Z′ signature at hadron colliders. The Z′ is mainly produced from dark matter final state radiation. For a GeV-scale Z′ decaying to hadrons, this gives a unique mono-Z′ jet signature.

We focus on the possibility that the dominant decay of the Z′ is into quarks. When the Z′ mass scale is light (GeV-scale) there are two important effects to consider. First, the hadronic decay of a boosted Z′ gives a new collider signature: the Z′ appears as a jet with a very narrow cone of radiation and a small multiplicity of charged particles which is referred to as a Z′-jet. Second, the rate for dark matter FSR of Z′ jets can be larger than the rate for ISR jets. Taking advantage of both these effects, we demonstrate that a dedicated collider analysis based on the mono-Z′ signature will dramatically improve our understanding of the dark matter interactions with visible particles.

E.2 Results

How dark matter is produced at colliders is model-dependent, but generically there are two possibilities. The first possibility, which we call the "secluded dark Z′ model", is that the SM particles are charge-neutral under the dark U(1)′, but have additional interactions with dark matter particles. The second possibility, which we call the "public dark Z′ model", is to have some SM particles also charged under the dark U(1)′ gauge symmetry.

In the secluded model schenerio we were able to set constraints on the dark matter-proton cross section, as shown in Figure E.2. While in the public dark model case we set constraints on the Z′ couplings to quarks (g_q) and dark matter (g_χ) as shown in Figure E.3.

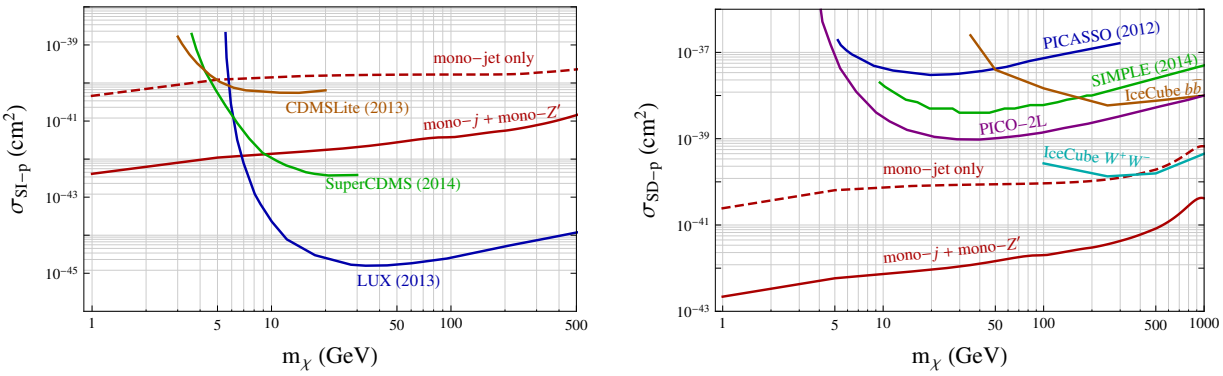


Figure E.2: Left panel: Projected constraints on dark matter-proton spin-independent scattering cross sections from the standard mono-jet analysis at the 14 TeV LHC with 100 fb^1 and the mono-Z′ jet-substructure based analysis. The model parameters are $M_{Z'} = 1$ GeV and $g_\chi = 1.0$. Also shown are the current constraints from direct detection experiments: LUX [88], SuperCDMS [89], and CDMSLite [90]. Right panel: similar to the left panel but for dark matter-proton spin-dependent scattering cross sections. The current experimental bounds are from: PICASSO [91], SIMPLE [92], PICO-2L [93], and IceCube [94].

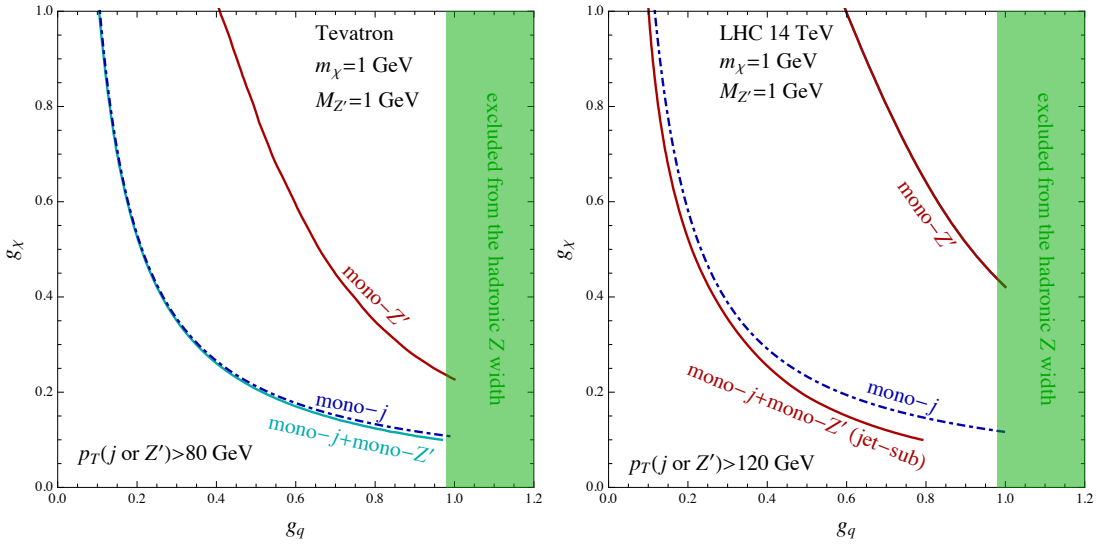


Figure E.3: Left panel: the 90% C.L. constraints on the light Z′ couplings from the Tevatron with 1.96 TeV and 1.0 fb^{-1}. The shaded region is excluded by the Z boson hadronic width at 90% C.L. Right panel: the projected sensitivity at the 14 TeV LHC with 100 fb^1. The systematic error is assumed to be 5%.

Bibliography

[1] **IceCube** Collaboration, K. Rawlins, "Cosmic ray spectrum and composition from three years of IceTop and IceCube," *J. Phys. Conf. Ser.* **718** (2016), no. 5, 052033.

[2] V. F. Hess, "ber Beobachtungen der durchdringenden Strahlung bei sieben Freiballonfahrten," *Phys. Z.* **13** (1912) 1084–1091.

[3] W. Kolhörster, "Messungen der Durchdringenden Strahlung im Freiballon in Grösseren Höhen," *Physikalische Zeitschrift* **14** (1913) 1153.

[4] W. Bothe and W. Kolhörster, "The nature of the penetrating radiation," *Nature* (January, 1929).

[5] B. Rossi, "Directional Measurements on the Cosmic Rays Near the Geomagnetic Equator," *Phys. Rev.* **45** (Feb, 1934) 212–214.

[6] P. Auger, P. Ehrenfest, R. Maze, and et al., "Extensive Cosmic-Ray Showers," *Rev. Mod. Phys.* **11** (Jul, 1939) 288–291.

[7] D. J. Bird *et al.*, "Detection of a cosmic ray with measured energy well beyond the expected spectral cutoff due to cosmic microwave radiation," *Astrophys. J.* **441** (1995) 144–150, astro-ph/9410067.

[8] **IceCube** Collaboration, R. Abbasi *et al.*, "All-particle cosmic ray energy spectrum measured with 26 IceTop stations," *Astropart. Phys.* **44** (2013) 40–58, 1202.3039.

[9] **KASCADE Grande** Collaboration, W. D. Apel *et al.*, "Kneelike structure in the spectrum of the heavy component of cosmic rays observed with KASCADE-Grande," *Phys. Rev. Lett.* **107** (2011) 171104, 1107.5885.

[10] **IceCube** Collaboration, M. G. Aartsen *et al.*, "Measurement of the cosmic ray energy spectrum with IceTop-73," *Phys. Rev.* **D88** (2013), no. 4, 042004, 1307.3795.

[11] J. J. Beatty and S. Westerhoff, "The Highest-Energy Cosmic Rays," *Annual Review of Nuclear and Particle Science* **59** (2009), no. 1, 319–345, https://doi.org/10.1146/annurev.nucl.58.110707.171154.

[12] **Particle Data Group** Collaboration, K. A. Olive *et al.*, "Review of Particle Physics," *Chin. Phys.* **C38** (2014) 090001.

[13] X. H. Sun, W. Reich, A. Waelkens, and et al., "Radio observational constraints on Galactic 3D-emission models," *Astron. Astrophys.* **477** (2008) 573, 0711.1572.

[14] **Pierre Auger** Collaboration, A. Aab *et al.*, "Large-scale cosmic-ray anisotropies above 4 EeV measured by the Pierre Auger Observatory," *Astrophys. J.* **868** (2018), no. 1, 4, 1808.03579.

[15] E. Fermi, "On the Origin of the Cosmic Radiation," *Phys. Rev.* **75** (1949) 1169–1174.

[16] B. Peters, "Primary cosmic radiation and extensive air showers," *Il Nuovo Cimento* **22** (Nov., 1961) 800–819.

[17] T. K. Gaisser, T. Stanev, and S. Tilav, "Cosmic Ray Energy Spectrum from Measurements of Air Showers," *Front. Phys.(Beijing)* **8** (2013) 748–758, 1303.3565.

[18] S. Thoudam, J. P. Rachen, A. van Vliet, and et al., "Cosmic-ray energy spectrum and composition up to the ankle: the case for a second Galactic component," *Astron. Astrophys.* **595** (2016) A33, 1605.03111.

[19] L. Merten, C. Bustard, E. G. Zweibel, and J. Becker Tjus, "The Propagation of Cosmic Rays from the Galactic Wind Termination Shock: Back to the Galaxy?," *Astrophys. J.* **859** (2018), no. 1, 63, 1803.08376.

[20] C. Bustard, E. G. Zweibel, and C. Cotter, "Cosmic Ray Acceleration by a Versatile Family of Galactic Wind Termination Shocks," *Astrophys. J.* **835** (2017), no. 1, 72, 1610.06565.

[21] K. M. Ferriere, "The interstellar environment of our galaxy," *Rev. Mod. Phys.* **73** (2001) 1031–1066, astro-ph/0106359.

[22] J. S. George, K. A. Lave, M. E. Wiedenbeck, and et al., "ELEMENTAL COMPOSITION AND ENERGY SPECTRA OF GALACTIC COSMIC RAYS DURING SOLAR CYCLE 23," *The Astrophysical Journal* **698** (jun, 2009) 1666–1681.

[23] H. Ahn, P. Allison, M. Bagliesi, and et al., "Discrepant hardening observed in cosmic-ray elemental spectra," *The Astrophysical Journal Letters* **714** (2010), no. 1, L89.

[24] D. Ikeda, "Recent results of the energy spectrum and mass composition from Telescope Array Fluorescence Detector," *Journal of Physics: Conference Series* **409** (feb, 2013) 012097.

[25] K.-H. Kampert and M. Unger, "Measurements of the cosmic ray composition with air shower experiments," *Astroparticle Physics* **35** (May, 2012) 660–678, 1201.0018.

[26] **KASCADE** Collaboration, T. Antoni *et al.*, "KASCADE measurements of energy spectra for elemental groups of cosmic rays: Results and open problems," *Astropart. Phys.* **24** (2005) 1–25, astro-ph/0505413.

[27] W. D. Apel *et al.*, "KASCADE-Grande measurements of energy spectra for elemental groups of cosmic rays," *Astropart. Phys.* **47** (2013) 54–66, 1306.6283.

[28] **Telescope Array** Collaboration, R. U. Abbasi *et al.*, "Evidence of Intermediate-Scale Energy Spectrum Anisotropy of Cosmic Rays E$\geq 10^{19.2}$ eV with the Telescope Array Surface Detector," *Astrophys. J.* **862** (2018), no. 2, 91, 1802.05003.

[29] **Telescope Array** Collaboration, R. U. Abbasi *et al.*, "Indications of Intermediate-Scale Anisotropy of Cosmic Rays with Energy Greater Than 57 EeV in the Northern Sky Measured with the Surface Detector of the Telescope Array Experiment," *Astrophys. J.* **790** (2014) L21, 1404.5890.

[30] P. Grieder, *Extensive Air Showers: High Energy Phenomena and Astrophysical Aspects - A Tutorial, Reference Manual and Data Book*. Astrophysics and space science library. Springer Berlin Heidelberg, 2010.

[31] K.-H. Kampert and A. A. Watson, "Extensive Air Showers and Ultra High-Energy Cosmic Rays: A Historical Review," *Eur. Phys. J.* **H37** (2012) 359–412, 1207.4827.

[32] J. Matthews, "A Heitler model of extensive air showers," *Astropart. Phys.* **22** (2005) 387–397.

[33] **Particle Data Group** Collaboration, C. Patrignani *et al.*, "Review of Particle Physics," *Chin. Phys.* **C40** (2016), no. 10, 100001.

[34] D. Chirkin and W. Rhode, "Muon Monte Carlo: A High-precision tool for muon propagation through matter," hep-ph/0407075.

[35] X. Bai, D. Chirkin, T. Gaisser, and et al., "Muon bundle energy loss in deep underground detector," in *ICRC Proceedings*. 2009.

[36] R. Abbasi, Y. Abdou, M. Ackermann, and et al., "IceTop: The surface component of IceCube. The IceCube Collaboration," *Nuclear Instruments and Methods in Physics Research A* **700** (February, 2013) 188–220, 1207.6326.

[37] D. Heck, J. Knapp, J. N. Capdevielle, and et al., "CORSIKA: A Monte Carlo code to simulate extensive air showers,".

[38] G. Battistoni, F. Cerutti, A. Fassò, and et al., "The FLUKA code: description and benchmarking,".

[39] R. Fletcher, T. Gaisser, P. Lipari, and et al., "s i b y l l: An event generator for simulation of high energy cosmic ray cascades," *Physical Review D* **50** (1994), no. 9, 5710.

[40] S. Ostapchenko, "QGSJET-II: towards reliable description of very high energy hadronic interactions," *Nuclear Physics B-Proceedings Supplements* **151** (2006), no. 1, 143–146.

[41] T. Pierog and K. Werner, "EPOS model and ultra high energy cosmic rays," *Nuclear Physics B-Proceedings Supplements* **196** (2009) 102–105.

[42] P. A. Čerenkov, "Visible Radiation Produced by Electrons Moving in a Medium with Velocities Exceeding that of Light," *Phys. Rev.* **52** (Aug, 1937) 378–379.

[43] T. Benson, J. Cherwinka, M. Duvernois, and et al., "IceCube Enhanced Hot Water Drill functional description," *Annals of Glaciology* **55** (2014), no. 68, 105114.

[44] **IceCube** Collaboration, M. G. Aartsen *et al.*, "The IceCube Neutrino Observatory: Instrumentation and Online Systems," *JINST* **12** (2017), no. 03, P03012, 1612.05093.

[45] **IceCube** Collaboration, R. Abbasi *et al.*, "Calibration and Characterization of the IceCube Photomultiplier Tube," *Nucl. Instrum. Meth.* **A618** (2010) 139–152, 1002.2442.

[46] **IceCube** Collaboration, R. Abbasi *et al.*, "The IceCube Data Acquisition System: Signal Capture, Digitization, and Timestamping," *Nucl. Instrum. Meth.* **A601** (2009) 294–316, 0810.4930.

[47] IceCube Collaboration, F. Kislat, S. Klepser, and A. V. Overloop, "A Lateral Distribution Function and Fluctuation Parametrisation for IceTop," *IceCube Internal Report* (2007).

[48] F. Kislat, "Study of charge and time fluctuations of signals in the IceTop detector,".

[49] N. Whitehorn, J. van Santen, and S. Lafebre, "Penalized Splines for Smooth Representation of High-dimensional Monte Carlo Datasets," *Comput. Phys. Commun.* **184** (2013) 2214–2220, 1301.2184.

[50] S. Raschka and V. Mirjalili, *Python Machine Learning, 2nd Ed.* Packt Publishing, Birmingham, UK, 2 ed., 2017.

[51] J. VanderPlas, *Python Data Science Handbook*. O'Reilly Media, Incorporated, 2017.

[52] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *Journal of Machine Learning Research* **13** (2012), no. Feb, 281–305.

[53] J. Friedman, T. Hastie, and R. Tibshirani, *The elements of statistical learning*, vol. 1. Springer series in statistics New York, 2001.

[54] F. Pedregosa, G. Varoquaux, A. Gramfort, and et al., "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research* **12** (2011) 2825–2830.

[55] G. D'Agostini, "A multidimensional unfolding method based on Bayes' theorem," *Nucl. Instrum. Meth. A* **362** (1995) 487–498.

[56] H. Jeffreys, "An Invariant Form for the Prior Probability in Estimation Problems," *Proc. of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **186** (1946), no. 1007, 453–461.

[57] T. Adye, "Corrected Error Calculation for Iterative Bayesian Unfolding," 2011.

[58] J. Bourbeau and Z. Hampel-Arias, "PyUnfold: A Python package for iterative unfolding," *The Journal of Open Source Software* **3** (June, 2018) 741, 1806.03350.

[59] T. Feusels, *Measurement of cosmic ray composition and energy spectrum between 1 PeV and 1 EeV with IceTop and IceCube*. PhD thesis, Ghent University, 2014.

[60] **HAWC** Collaboration, R. Alfaro *et al.*, "All-particle cosmic ray energy spectrum measured by the HAWC experiment from 10 to 500 TeV," *Phys. Rev.* **D96** (2017), no. 12, 122001, 1710.00890.

[61] **IceCube** Collaboration, A. Van Overloop, "Simulation of IceTop VEM calibration and the dependency on the snow layer," in *Proceedings, 32nd International Cosmic Ray Conference (ICRC 2011): Beijing, China, August 11-18, 2011*, vol. 1, pp. 97–101. 2011.

[62] **IceCube** Collaboration, M. G. Aartsen *et al.*, "Measurement of South Pole ice transparency with the IceCube LED calibration system," *Nucl. Instrum. Meth.* **A711** (2013) 73–89, 1301.5361.

[63] **CMS** Collaboration, M. A. Shah *et al.*, "The CMS RPC Detector Performance and Stability during LHC RUN-2," in *14th Workshop on Resistive Plate Chambers and Related Detectors (RCP2018) Puerto Vallarta, Jalisco State, Mexico, February 19-23, 2018*. 2018. 1808.10488.

[64] V. N. Gribov, "A REGGEON DIAGRAM TECHNIQUE," *Sov. Phys. JETP* **26** (1968) 414–422. [Zh. Eksp. Teor. Fiz.53,654(1967)].

[65] F. Riehn, H. P. Dembinski, R. Engel, and et al., "The hadronic interaction model SIBYLL 2.3c and Feynman scaling," *PoS* **ICRC2017** (2018) 301, 1709.07227. [35,301(2017)].

[66] S. De Ridder, *Sensitivity of IceCube Cosmic Ray measurements to the hadronic interaction models*. PhD thesis, Ghent University, 2019.

[67] M. Plum. personal communication.

[68] **KASCADE-Grande** Collaboration, M. Bertaina *et al.*, "KASCADE-Grande energy spectrum of cosmic rays interpreted with post-LHC hadronic interaction models," *PoS* **ICRC2015** (2016) 359.

[69] T. K. Gaisser, "Spectrum of cosmic-ray nucleons, kaon production, and the atmospheric muon charge ratio," *Astropart. Phys.* **35** (2012) 801–806, 1111.6675.

[70] **IceCube** Collaboration, H. Pandya, "A Composition Sensitive Log-Likelihood Ratio for Cosmic Rays and Gamma Rays," *PoS* **ICRC2017** (2018) 514.

[71] Z. Hampel-Arias, *Cosmic Ray Observations at the TeV Scale with the HAWC Observatory*. PhD thesis, University of Wisconsin, Madison, 2017.

[72] K. M. Gorski, E. Hivon, and A. J. e. a. Banday, "HEALPix - A Framework for high resolution discretization, and fast analysis of data distributed on the sphere," *Astrophys. J.* **622** (2005) 759–771, astro-ph/0409513.

[73] L. Demortier, "P Values: What They Are and How to Use Them,".

[74] D. E. Alexandreas *et al.*, "Point source search techniques in ultrahigh-energy gamma-ray astronomy," *Nucl. Instrum. Meth.* **A328** (1993) 570–577.

[75] ROOT, "An Object Oriented Data Analysis Framework," *Nucl. Instr. & Meth. in Phys. Rev. A* **389** (1997) 81–86.

[76] W. McKinney, "Data Structures for Statistical Computing in Python," in *Proceedings of the 9th Python in Science Conference*, S. van der Walt and J. Millman, eds., pp. 51 – 56. 2010.

[77] A. N. Kolmogorov, "Sulla Determinazione Empirica di Una Legge di Distribuzione," *Giornale dell'Istituto Italiano degli Attuari* **4** (1933) 83–91.

[78] N. Smirnov, "Table for Estimating the Goodness of Fit of Empirical Distributions," *Ann. Math. Statist.* **19** (1948), no. 2, 279–281.

[79] S. Y. BenZvi, B. M. Connolly, S. Westerhoff, and et al., "A Bayesian Approach to Comparing Cosmic Ray Energy Spectra," *The Astrophysical Journal* **738** (2011), no. 1, 82.

[80] M. Winter, J. Bourbeau, S. Bravo, and et al., "Particle identification in camera image sensors using computer vision," *Astroparticle Physics* **104** (2019) 42–53.

[81] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556* (2014).

[82] F. Chollet *et al.*, "Keras." `https://keras.io`, 2015.

[83] R. Kohavi *et al.*, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Ijcai*, vol. 14, pp. 1137–1145, Montreal, Canada. 1995.

[84] Y. Bai, J. Bourbeau, and T. Lin, "Dark matter searches with a mono-Z jet," *JHEP* **06** (2015) 205, `1504.01395`.

[85] J. Goodman, M. Ibe, A. Rajaraman, and et al., "Constraints on Dark Matter from Colliders," *Phys. Rev.* **D82** (2010) 116010, `1008.1783`.

[86] A. Birkedal, K. Matchev, and M. Perelstein, "Dark matter at colliders: A Model independent approach," *Phys. Rev.* **D70** (2004) 077701, `hep-ph/0403004`.

[87] Y. Bai and T. M. P. Tait, "Searches with Mono-Leptons," *Phys. Lett.* **B723** (2013) 384–387, `1208.4361`.

[88] **LUX** Collaboration, D. S. Akerib *et al.*, "First results from the LUX dark matter experiment at the Sanford Underground Research Facility," *Phys. Rev. Lett.* **112** (2014) 091303, `1310.8214`.

[89] **SuperCDMS** Collaboration, R. Agnese *et al.*, "Search for Low-Mass Weakly Interacting Massive Particles with SuperCDMS," *Phys. Rev. Lett.* **112** (2014), no. 24, 241302, 1402.7137.

[90] **SuperCDMS** Collaboration, R. Agnese *et al.*, "Search for Low-Mass Weakly Interacting Massive Particles Using Voltage-Assisted Calorimetric Ionization Detection in the SuperCDMS Experiment," *Phys. Rev. Lett.* **112** (2014), no. 4, 041302, 1309.3259.

[91] **PICASSO** Collaboration, S. Archambault *et al.*, "Constraints on Low-Mass WIMP Interactions on ^{19}F from PICASSO," *Phys. Lett.* **B711** (2012) 153–161, 1202.1240.

[92] **SIMPLE** Collaboration, M. Felizardo *et al.*, "The SIMPLE Phase II Dark Matter Search," *Phys. Rev.* **D89** (2014), no. 7, 072013, 1404.4309.

[93] **PICO** Collaboration, C. Amole *et al.*, "Dark Matter Search Results from the PICO-2L C_3F_8 Bubble Chamber," *Phys. Rev. Lett.* **114** (2015), no. 23, 231302, 1503.00008.

[94] **IceCube** Collaboration, M. G. Aartsen *et al.*, "Search for dark matter annihilations in the Sun with the 79-string IceCube detector," *Phys. Rev. Lett.* **110** (2013), no. 13, 131302, 1212.4097.