THE UNIVERSITY OF CHICAGO

SEARCHING FOR DARK MATTER WITH THE SUDBURY NEUTRINO OBSERVATORY

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF PHYSICS

BY
ANTHONY LATORRE

CHICAGO, ILLINOIS
MARCH 2021

To my family and Marie.

Well-chosen, non-frivolous epigraphs can enhance a thesis. – Dave Clarke

Table of Contents

List of Figures

List of Tables

ACKNOWLEDGMENTS

This thesis would not have been possible without the help of a huge number of people. First and foremost, I want to thank my advisor, Ed Blucher, for being a constant source of support and guidance throughout my whole graduate career.

Thank you to Josh Klein, who inspired me to join the field of physics as an undergrad at the University of Texas and then taught me almost everything I know about being an experimental physicist at the University of Pennsylvania.

Thank you to Sean Keel at the University of Texas for making me realize that it was possible to understand the foundations of everything in the math and sciences.

Thank you to Gabriel Orebi Gann for giving me a chance to work at Berkeley before I started graduate school in Chicago, and for organizing everything which allowed me to analyze the data from the SNO experiment.

Thank you to all of my SNO and SNO+ colleagues, especially Andy Mastbaum and Javi Caravaca for their help on understanding the intricacies of the SNO data. Thank you to Andy Mastbaum and Richie Bonventre since their work provided the foundations for this whole analysis. Thank you to Freija Descamps for leading the DAQ group during some of the best times I had on the experiment. Also, thank you to Neil Mccauley and Chris Kyba for their help on all things related to SNOMAN.

Thank you to everybody I worked with at Chicago over the years, especially Kevin Labe and Matt Strait. I still remember the great times we had travelling to Sudbury and going to collaboration meetings. Thanks also to Kevin for the numerous interesting discussions on quantum mechanics, artificial intelligence, and philosophy.

Thank you to Mengfei He, Johnny Ho, Pavel Motloch, and Albert Ryu for making life in Chicago so great. I will always remember our Friday night boardgames, trips to Strings, and videogame nights.

Thank you to Marie Vidal for sticking with me through this whole process and for all the help reviewing my analysis and this thesis.

Thank you to my family, especially my parents for being supportive and more generally awesome people.

Thanks to all of the University of Chicago administrators and staff who made my whole career at Chicago so great.

This research was done using resources provided by the Open Science Grid[1, 2], which is supported by the National Science Foundation and the U.S. Department of Energy's Office of Science. The final analysis would not have been remotely possible without the huge amount of CPU time I was able to utilize via OSG Connect.

Also, thanks to Don Groom for rerunning his old Fortran code to produce new stopping power tables for heavy water!

ABSTRACT

Dark matter currently makes up approximately 84% of the matter in our universe, but has yet to be observed. A recent model by Grossman, Harnik, Telem, and Zhang proposes a new form of dark matter called self-destructing dark matter which decays to standard model leptons after an interaction in Earth. Motivated by this model, we perform two distinct analyses looking at high energy events in the Sudbury Neutrino Observatory data between 1999 and 2003. In the first, we perform a null hypothesis test on the data between 20 MeV and 10 GeV to look for *any* data which is not consistent with atmospheric neutrinos and find no evidence for new physics. In the second analysis, we perform a dedicated search for back to back lepton pairs from self-destructing dark matter. We find no evidence for the self-destructing dark matter and place new limits on the rate of these events.

CHAPTER 1

INTRODUCTION

According to current best estimates[1], approximately 84% of the matter in our universe is still unobserved[3]. Evidence for dark matter first appeared as early as the 1930's in observations of galaxy clusters. Later in the 1970's, detailed measurements of galactic rotation curves provided strong evidence supporting the hypothesis that the majority of the mass in these galaxies was invisible. Over the past few decades, searches for this invisible matter have been primarily focused on looking for weakly interacting massive particles, or WIMPs for short. These experiments typically look for WIMP nuclear recoils in large liquid noble gas detectors or in high purity bolometric crystals. However, despite all this effort no definitive evidence for WIMP nuclear recoils has been found[2].

Given the importance of discovering what makes up such a large fraction of the matter in our universe, new models have recently been proposed which are more complex than the single WIMP model. In 2017 Grossmann, Harnik, Telem, and Zhang published a new class of models for dark matter called self-destructing dark matter[4]. In these models, some fraction of dark matter is made up of a cosmologically stable bound state which can undergo a transition to a short lived state through an interaction with normal matter in Earth. This short lived state then decays into two or more dark photons which then decay into standard model leptons such as electrons or muons[4]. Such a model predicts a visible signal on the order of the mass of the dark matter particle instead of the kinetic energy. This puts these models in reach of large neutrino detectors like the Sudbury Neutrino Observatory and Super Kamiokande.

---

1. Specifically, according to the benchmark ΛCDM cosmology[3].

2. With the exception of the DAMA/LIBRA experiment which has seen a strong annual modulation in sodium iodide scintillation detectors for many years. However, interpreting the energy of the excess events and size of the annual modulation under a single WIMP hypothesis leads to a prediction which has been ruled out by other direct detection dark matter experiments.

Motivated by this model, in this analysis I look for anomalous events in the energy range from 20 MeV to 10 GeV in data taken by the Sudbury Neutrino Observatory between 1999 and 2003.

1.1 Self-Destructing Dark Matter

In their recent paper[4], Grossman et. al. discuss three possible models for self-destructing dark matter. In the first model, dark matter consists of a particle χ and its antiparticle $\bar{\chi}$ which form a positronium-like bound state which is prevented from decaying by the fact that they orbit in a high angular momentum state. Through a scattering process with normal matter, this bound state can transition to a low angular momentum state which promptly decays into two or three dark photons V which couple to the standard model photon. Each of the dark photons can then decay into an electron-positron or muon-antimuon pair. The other two models also involve a dark matter particle and antiparticle pair which can annihilate through an interaction with normal matter in Earth[4]. Figure 1.1 shows an example of a dark matter decaying into an electron-positron pair visible in the detector.

Each of these models has a distinctive characteristics one could search for. In their paper they discuss several characteristics such as event kinematics, opening angle, and directionality to distinguish different models. However, as suggested by the authors themselves, in an attempt to be as model independent as possible I will focus on only those characteristics common to all of the models. In all of the self-destructing dark matter models the common components are a dark matter bound state Ψ which can transition to a short lived state Ψ' after an interaction with normal matter in Earth. This short lived state then decays to two or more dark photons V which can decay into standard model leptons. The most striking experimental signature for all of these models is one or more high energy pairs of leptons with a fixed invariant mass. In the case where the dark matter pair decays to 2 Vs these lepton pairs will also have a fixed energy.

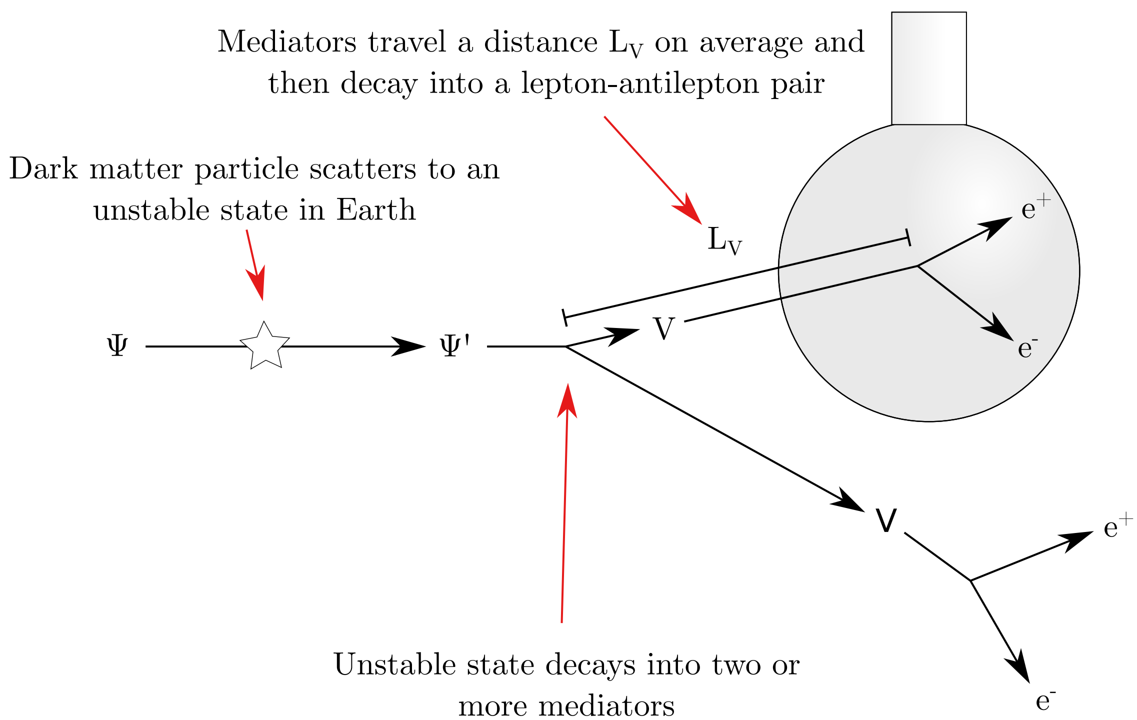

Figure 1.1: Cartoon showing a dark matter decay.

In this analysis I will focus on the case where the dark matter pair decays to 2 Vs and where the decay length of the mediator, L_V is much greater than the size of the SNO detector. In this case I expect to see a single lepton pair with a fixed energy and mass.

1.2 Analysis

In this thesis, I will perform two distinct analyses:

1. A null hypothesis test to test if the events in the energy range 20 MeV - 10 GeV match what is expected from atmospheric neutrino events.

2. A detailed search for an electron-positron or muon-antimuon pair coming from a particle with a fixed mass.

The first test aims to be as broad as possible and will be sensitive to *any* departure from known physics, not just a potential dark matter signal, whereas the second analysis will look for specific signatures of self-destructing dark matter. The challenge in both cases is correctly accounting for the expected atmospheric neutrino background and the uncertainties associated with the flux and neutrino cross sections. The process for modeling this background is discussed in Chapter 4.

In the first analysis, I will perform a 1D fit in energy to all multi-particle prompt events which pass a series of cuts[3]. We then perform a multinomial test on the fit results and report a p-value.

In the second analysis, I will add a term for a lepton-antilepton pair with a fixed invariant mass and energy to the likelihood. I will focus on the simplest case of a slow moving mediator which decays to a back to back lepton-antilepton pair. I will present a limit on the event rate for electron-positron and muon-antimuon pairs per unit volume in the detector as a function of these two parameters.

3. The run and event selection criteria are discussed in Chapter 7.

CHAPTER 2

THE SNO DETECTOR

The SNO detector is a large water Čerenkov detector located approximately 2 km underground in an active nickel mine in Sudbury, Ontario. The detector consists of approximately 10,000 photomultiplier tubes (PMTs) attached to an approximately spherical 16 meter diameter PMT support structure (PSUP). These PMTs surround a 12 m diameter acrylic vessel (AV) containing 1 metric kton of heavy water (D^2O)[5]. Figures 2.1 and 2.2 show a schematic drawing of the SNO detector and a picture taken from a camera mounted to the PSUP.

Each PMT is connected via a long RG59-like cable to electronic racks on the deck above the detector. The detector is triggered when a certain number of PMT hits[1] occur within a 100 ns time window. When this happens the charge and time for each PMT hit are recorded in a 400 ns window around the time of the trigger. All of the PMT hits recorded during this window are then assembled together and called an event.

In addition to the approximately 10,000 regular PMTs the SNO detector also has 91 outward-looking PMTs (OWLs) mounted to the PSUP[5]. These OWL PMTs are useful in tagging external muons.

The SNO detector observes charged particles traveling through the detector from the Čerenkov light produced when they travel above the local speed of light

$$v = \frac{c}{n}, \tag{2.1}$$

where n is equal to the index of refraction. This threshold corresponds to a kinetic energy of approximately 0.8 MeV for electrons and 53 MeV for muons. The Čerenkov light, produced in a cone with an opening angle of approximately 42 degrees, travels to the PMTs where it is detected. As a rough rule of thumb, the detector will see approximately 7 PMT hits per

1. The threshold was typically somewhere around 20 PMT hits.

Figure 2.1: An artist's conception of the SNO detector.

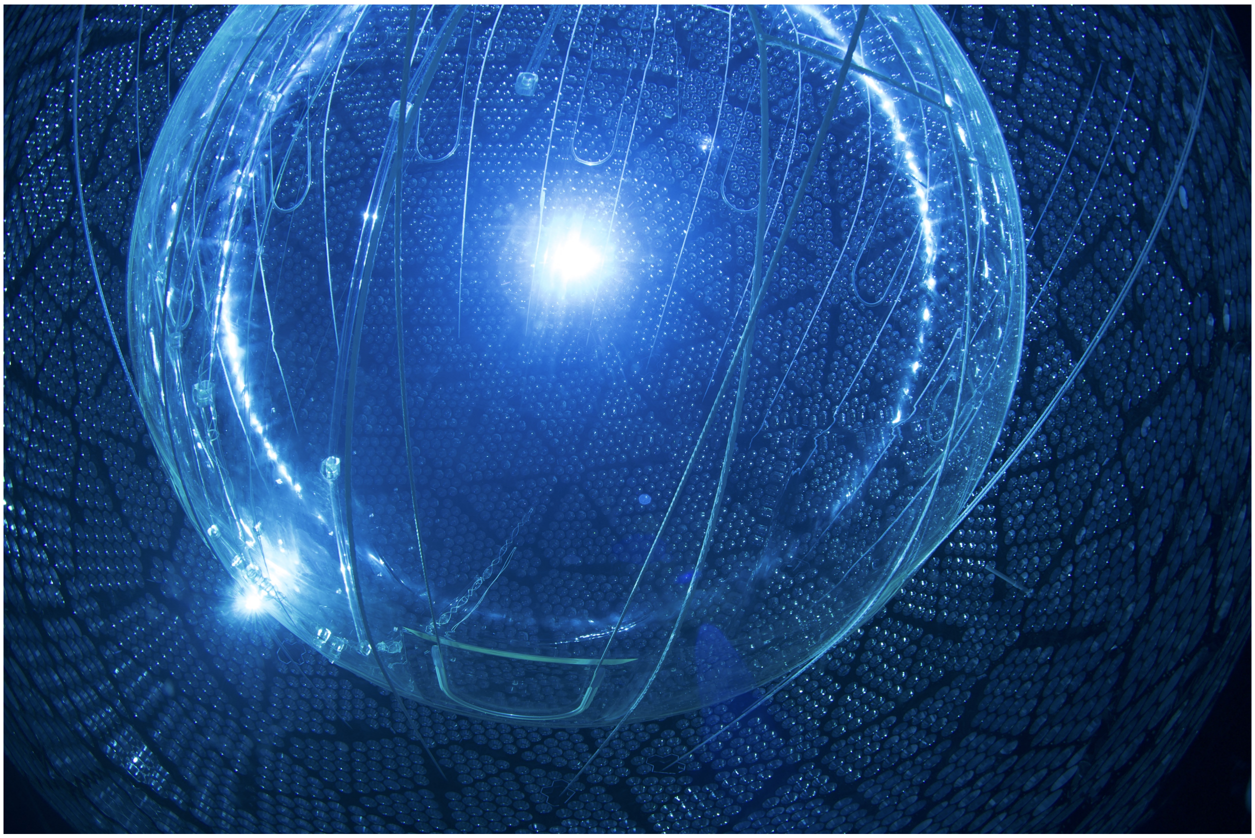

Figure 2.2: A picture of the detector taken from a camera mounted to the PSUP. This picture was taken after the SNO detector was upgraded to work for the SNO+ experiment. The major difference between the original SNO detector and this picture is the addition of hold-down ropes to prevent the acrylic vessel from floating when scintillator is added.

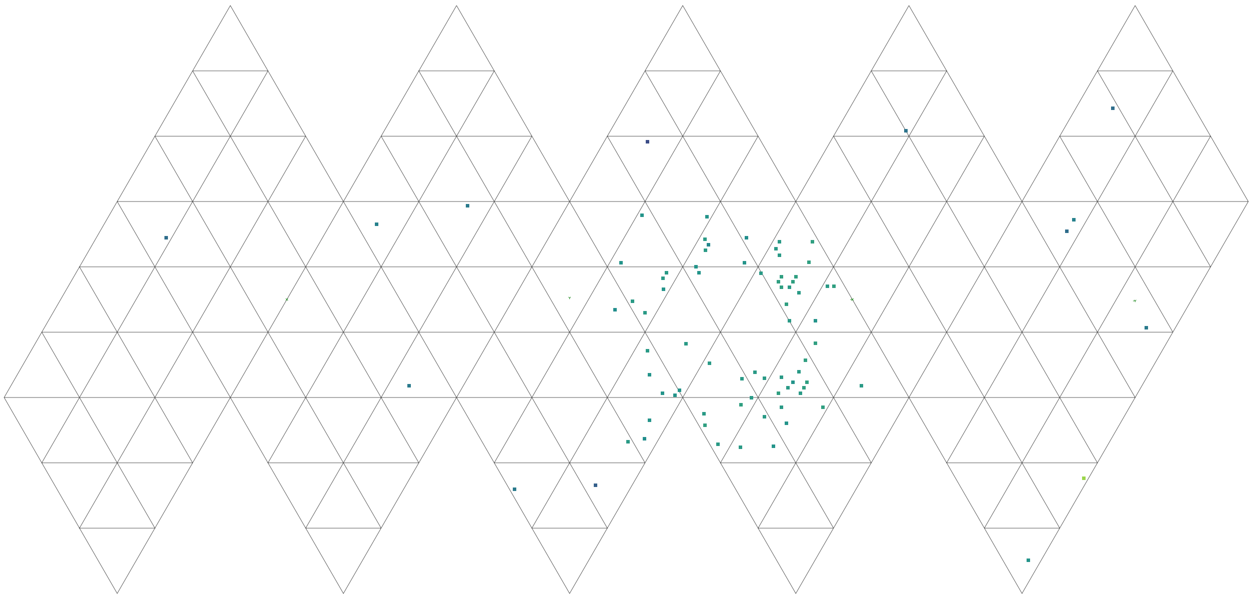

Figure 2.3: A possible neutrino event in the SNO detector shown using the XSnoed event display program. Each dot represents a PMT hit in the detector and the color represents the time of the hit.

MeV of energy for electrons. Figure 2.3 shows an example of a neutrino event in the SNO detector.

The primary goal of the SNO experiment was to measure the ratio of the number of charged current solar neutrino reactions, to the number of neutral current reactions in order to resolve the solar neutrino problem. The charged current reaction,

$$\nu_e + d \rightarrow p + p + e^- \tag{2.2}$$

is only sensitive to electron type neutrinos, while the neutral current reaction

$$\nu_x + d \rightarrow p + n + \nu_x \tag{2.3}$$

is sensitive to all neutrino flavors equally. The SNO detector operated between 1999 and 2006 in three distinct phases: the D2O phase, the salt phase, and the NCD phase[5]. The primary difference between the phases was how the neutron from the neutral current reaction was

	Livetime (days)
D2O	196.2
Salt	485.4

Table 2.1: Livetime for the D2O and Salt phases.

measured. During the D2O phase, which took place between November 1999 and May 2001, the acrylic vessel was filled with pure heavy water or D2O[5]. In this phase the neutrons captured on the heavy water emitting a 6.25 MeV gamma ray. During the salt phase, which took place between July 2001 and August 2003, sodium chloride was added to the heavy water in the acrylic vessel[5]. Due to the higher capture cross section, the neutrons would primarily capture on chlorine and emit multiple gamma rays with a combined energy of 8.6 MeV. Finally, in the NCD phase which took place between November 2004 and November 2006[5], an array of 2 m long ^{3}He filled proportional counters called neutral current detectors (NCD) was added to the acrylic vessel to detect the neutrons via the reaction

$$^3\mathrm{He} + n \rightarrow ^3\mathrm{H} + p. \tag{2.4}$$

In this analysis I will only be looking at the data from the D2O and salt phases since the proportional counters added during the NCD phase make modeling the optics of the detector difficult.

CHAPTER 3

EVENT RECONSTRUCTION

In this chapter I will discuss the process of reconstructing the position, time, energy, direction, and particle ID of each event from the individual PMT hits. This process is necessary for several reasons; first, the reconstructed energy will be the primary observable in the final analysis. Second, the reconstructed position is used to cut events near the PSUP since the vast majority of the instrumental backgrounds originate from there. Finally, the particle ID is used to cut single lepton events, reducing the atmospheric neutrino background by more than a factor of two.

Events are reconstructed using a maximum likelihood method. The likelihood function calculates the probability of observing the data in an event given a proposed particle, vertex position, energy, time, and track direction. In Section 3.1 I discuss the overall formulation of the likelihood function. Then, in Sections 3.2 and 3.3 I discuss the two most important inputs to the likelihood function: the expected charge and time distribution from direct Čerenkov light expected at each PMT. In Section 3.4, I discuss how the photons from electromagnetic showers and delta rays are added to the likelihood function. In Sections 3.5 and 3.6, I discuss the implementation of the algorithms used to seed the position and directions of the likelihood fit. In Section 3.7 I discuss the calculation of a number ψ which is used as a goodness of fit parameter and is used to cut instrumental backgrounds. In Section 3.8 I discuss the process by which I select the total number of particles and particle type for each ring. Finally, in Sections 3.9 and 3.10 I show several results used to benchmark the performance of the fitter on single leptons and lepton pairs.

3.1 Likelihood

For a given position, energy, direction, and time, the likelihood of an event is equal to

$$\mathscr{L}(E, \vec{x}, \vec{v}, t_0) = P(\vec{q}, \vec{t} \mid E, \vec{x}, \vec{v}, t_0), \tag{3.1}$$

where E, $\vec{x}$, $\vec{v}$ represent the initial particle's kinetic energy, position, and direction respectively, t_0 represents the initial time of the event, $\vec{q}$ is a vector representing the charge seen by each PMT, and $\vec{t}$ is the time recorded by each PMT.

The right hand side of Equation (3.1) is not factorizable in general since for particle tracks that scatter there will be correlations between the PMT hits. However, to make the problem analytically tractable, we assume that the probability of each PMT being hit is approximately independent of the others. With this assumption we can factor the right hand side of the likelihood as:

$$\mathscr{L}(E, \vec{x}, \vec{v}, t_0) = \left(\prod_i P(\text{not hit} \mid E, \vec{x}, \vec{v}, t_0) \right) \left(\prod_j P(\text{hit}, q_j, t_j \mid E, \vec{x}, \vec{v}, t_0) \right) \tag{3.2}$$

where the first product is over all PMTs that weren't hit and the second product is over all of the hit PMTs.

If we introduce the variable n, which represents the number of photoelectrons detected, we can write the likelihood as:

$$\mathscr{L}(E, \vec{x}, \vec{v}, t_0) = \left(\prod_i \sum_{n=0}^{\infty} P(\text{not hit}, n \mid E, \vec{x}, \vec{v}, t_0) \right) \left(\prod_j \sum_{n=1}^{\infty} P(n, q_j, t_j \mid E, \vec{x}, \vec{v}, t_0) \right). \tag{3.3}$$

We can then condition on n and write the likelihood as:

$$\mathscr{L}(E,\vec{x},\vec{v},t_0) = \left(\prod_i \sum_{n=0}^{\infty} P(\text{not hit} \mid n, E, \vec{x}, \vec{v}, t_0) P(n \mid E, \vec{x}, \vec{v}, t_0)\right)$$
$$\left(\prod_j \sum_{n=1}^{\infty} P(q_j, t_j \mid n, E, \vec{x}, \vec{v}, t_0) P(n \mid E, \vec{x}, \vec{v}, t_0)\right). \quad (3.4)$$

We now assume that the charge and time observed at a given PMT are independent and write the likelihood as:

$$\mathscr{L}(E,\vec{x},\vec{v},t_0) = \left(\prod_i \sum_{n=0}^{\infty} P(\text{not hit} \mid n) P(n \mid E, \vec{x}, \vec{v}, t_0)\right)$$
$$\left(\prod_j \sum_{n=1}^{\infty} P(q_j \mid n) P(t_j \mid n, E, \vec{x}, \vec{v}, t_0) P(n \mid E, \vec{x}, \vec{v}, t_0)\right). \quad (3.5)$$

Since there are many photons produced in each event each of which has a small probability to hit a given PMT, we assume that the probability of detecting n photons at a given PMT is Poisson distributed, i.e.

$$P(n \mid E, \vec{x}, \vec{v}, t_0) = e^{-\mu} \frac{\mu^n}{n!} \quad (3.6)$$

We can therefore write the likelihood as:

$$\mathscr{L}(E,\vec{x},\vec{v},t_0) = \left(\prod_i \sum_{n=0}^{\infty} P(\text{not hit} \mid n) e^{-\mu_i} \frac{\mu_i^n}{n!}\right)$$
$$\left(\prod_j \sum_{n=1}^{\infty} P(q_j \mid n) P(t_j \mid n, E, \vec{x}, \vec{v}, t_0) e^{-\mu_j} \frac{\mu_j^n}{n!}\right) \quad (3.7)$$

where μ_i is the expected number of photoelectrons detected at the i^{th} PMT (given an initial particle's energy, position, and direction).

Finally, we also add the probability that a channel is miscalibrated by calculating

$$\mathscr{L}(E,\vec{x},\vec{v},t_0) = \left(\prod_i P_{\text{miscal}} + (1-P_{\text{miscal}})\sum_{n=0}^{\infty} P(\text{not hit} \mid n)e^{-\mu_i}\frac{\mu_i^n}{n!}\right)$$
$$\left(\prod_j P_{\text{miscal}}\left(\frac{1}{4096}\right)^2 + (1-P_{\text{miscal}})\sum_{n=1}^{\infty} P(q_j \mid n)P(t_j \mid n,E,\vec{x},\vec{v},t_0)e^{-\mu_j}\frac{\mu_j^n}{n!}\right) \quad (3.8)$$

3.2 Expected Charge

First, we'll calculate the expected number of photoelectrons for a single non-showering track that undergoes multiple scattering through small angles. In this case, we can calculate the expected number of photoelectrons as

$$\mu_i = \int_x \mathrm{d}x \int_\lambda \mathrm{d}\lambda \frac{\mathrm{d}^2N}{\mathrm{d}x\mathrm{d}\lambda} P(\text{detected} \mid E,x,v), \quad (3.9)$$

where $\frac{\mathrm{d}^2N}{\mathrm{d}x\mathrm{d}\lambda}$ is the number of photons produced per unit length and wavelength, x is the position along the track and λ is the wavelength of the light.

If the particle undergoes many small angle Coulomb scatters, the net angular displacement of the particle after a distance x will be a Gaussian distribution by the central limit theorem[3]. The distribution of the net angular displacement at a distance x along the track is then given by

$$f(\theta,\phi) = \frac{\theta}{2\pi\theta_0^2}e^{-\frac{\theta^2}{2\theta_0^2}}, \quad (3.10)$$

where

$$\theta_0 = \frac{13.6\ \text{MeV}}{\beta c p} z\sqrt{\frac{x}{X_0}}\left[1 + 0.038\ln\left(\frac{xz^2}{X_0\beta^2}\right)\right], \quad (3.11)$$

and p, βc, and z are the momentum, velocity, and charge of the particle, and X_0 is the radiation length of the particle[3].

Now, we integrate over the angular displacement of the track around the original velocity:

$$\mu_i = \int_x \mathrm{d}x \int_\lambda \mathrm{d}\lambda \frac{\mathrm{d}^2 N}{\mathrm{d}x \mathrm{d}\lambda} \int_\theta \mathrm{d}\theta \int_\phi \mathrm{d}\phi P(\text{detected} \mid \theta, \phi, E, x, v) f(\theta, \phi) \tag{3.12}$$

The probability of being detected can be factored into several different components:

$$\mu_i = \int_x \mathrm{d}x \int_\lambda \mathrm{d}\lambda \frac{\mathrm{d}^2 N}{\mathrm{d}x \mathrm{d}\lambda} P(\text{not scattered or absorbed} \mid \lambda, E, x, v) \epsilon(\eta) \mathrm{QE}(\lambda)$$
$$\int_\theta \mathrm{d}\theta \int_\phi \mathrm{d}\phi P(\text{emitted towards PMT i} \mid \theta, \phi, E, x, v) f(\theta, \phi) \tag{3.13}$$

where η is the angle between the vector connecting the track position x to the PMT position and the normal vector to the PMT, $\epsilon(\eta)$ is the collection efficiency, and $\mathrm{QE}(\lambda)$ is the quantum efficiency of the PMT.

We now make the assumption that the probability is uniform across the face of the PMT[1] and write the probability that a photon is emitted directly towards a PMT as a delta function:

$$P(\text{emitted towards PMT i} \mid \theta, \phi, E, x, v) = \frac{1}{2\pi} \delta \left(\frac{1}{n(\lambda)\beta} - \cos\theta'(\theta, \phi, x) \right) \Omega(x) \tag{3.14}$$

where θ' is the angle between the track and the PMT and $\Omega(x)$ is the solid angle subtended by the PMT. In a coordinate system with the z axis aligned along the original particle velocity and with the PMT in the x-z plane, the angle θ' is defined by:

$$\cos\theta' = \sin\theta \cos\phi \sin\theta_1 + \cos\theta \cos\theta_1 \tag{3.15}$$

where θ_1 is the angle between the PMT and the original particle velocity. We can now solve

1. This approximation is valid as long as the track is far away from the PMT. This assumption can cause problems later on when this is not valid by creating discontinuities in the likelihood based on exactly what point along the track is sampled when numerically integrating along the track. To partially mitigate this issue when computing the likelihood I calculate a minimum value for the RMS scattering angle θ_0 based on the angle subtended by the PMT concentrator.

the integral on the right hand side of Equation (3.13) as:

$$P(\text{emitted towards PMT i}) = \frac{\Omega(x)}{2\pi}\frac{1}{2\pi\theta_0^2}\int_\theta \mathrm{d}\theta \int_\phi \mathrm{d}\phi \delta\left(\frac{1}{n(\lambda)\beta} - \sin\theta\cos\phi\sin\theta_1 - \cos\theta\cos\theta_1\right)\theta e^{-\frac{\theta^2}{2\theta_0^2}}. \quad (3.16)$$

We now assume θ is small (which should be valid for small angle scatters), so that we can rewrite the delta function as:

$$P(\text{emitted towards PMT i}) = \frac{\Omega(x)}{2\pi}\frac{1}{2\pi\theta_0^2}\int_\theta \mathrm{d}\theta \int_\phi \mathrm{d}\phi \delta\left(\frac{1}{n(\lambda)\beta} - \theta\cos\phi\sin\theta_1 - \cos\theta_1\right)\theta e^{-\frac{\theta^2}{2\theta_0^2}}. \quad (3.17)$$

We can rewrite the argument of the delta function

$$P(\text{emitted towards PMT i}) = \frac{\Omega(x)}{2\pi}\frac{1}{2\pi\theta_0^2}\int_\theta \mathrm{d}\theta \int_\phi \mathrm{d}\phi \frac{1}{|\cos\phi\sin\theta_1|}\delta\left(\theta - \frac{\frac{1}{n(\lambda)\beta} - \cos\theta_1}{\cos\phi\sin\theta_1}\right)\theta e^{-\frac{\theta^2}{2\theta_0^2}}, \quad (3.18)$$

and solve the integral as

$$P(\text{emitted towards PMT i}) = \frac{\Omega(x)}{2\pi}\frac{1}{\sqrt{2\pi}\theta_0}\frac{1}{|\sin\theta_1|}e^{-\frac{1}{2\theta_0^2}\left(\frac{\frac{1}{n(\lambda)\beta}-\cos\theta_1}{\sin\theta_1}\right)^2}. \quad (3.19)$$

To simplify this expression, we can write

$$P(\text{emitted towards PMT i}) = \frac{\Omega(x)}{2\pi}\frac{1}{\sqrt{2\pi}\theta_0}\frac{1}{|\sin\theta_1|}e^{-\frac{\Delta^2(\lambda)}{2\theta_0^2}}, \quad (3.20)$$

where

$$\Delta(\lambda) = \frac{\frac{1}{n(\lambda)\beta} - \cos\theta_1}{\sin\theta_1}. \tag{3.21}$$

Plugging this back into Equation (3.13) we get

$$\mu_i = \frac{1}{\sqrt{2\pi}\theta_0}\int_x \mathrm{d}x \frac{\Omega(x)}{2\pi}\frac{1}{|\sin\theta_1|}\epsilon(\eta) \\ \int_\lambda \mathrm{d}\lambda \frac{\mathrm{d}^2 N}{\mathrm{d}x\mathrm{d}\lambda} P(\text{not scattered or absorbed} \mid \lambda, E, x, v)\mathrm{QE}(\lambda)e^{-\frac{\Delta^2(\lambda)}{2\theta_0^2}}. \tag{3.22}$$

Ideally we would just evaluate this double integral for each likelihood call, however the double integral is too computationally expensive to perform for every likelihood call. We therefore make some assumptions in order to make it more computationally tractable. First, since the scattering and absorption lengths to do not change drastically over the wavelength range that the PMTs are sensitive to we pull that factor out of the second integral:

$$\mu_i = \frac{1}{\sqrt{2\pi}\theta_0}\int_x \mathrm{d}x \frac{\Omega(x)}{2\pi}\frac{1}{|\sin\theta_1|}\epsilon(\eta) P_{\mathrm{eff}}(\text{not scattered or absorbed} \mid E, x, v) \\ \int_\lambda \mathrm{d}\lambda \frac{\mathrm{d}^2 N}{\mathrm{d}x\mathrm{d}\lambda}\mathrm{QE}(\lambda)e^{-\frac{\Delta^2(\lambda)}{2\theta_0^2}}. \tag{3.23}$$

The number of Čerenkov photons produced per unit length and per unit wavelength is given by[3]

$$\frac{\mathrm{d}^2 N}{\mathrm{d}x\mathrm{d}\lambda} = \frac{2\pi\alpha z^2}{\lambda^2}\left(1 - \frac{1}{\beta^2 n^2(\lambda)}\right), \tag{3.24}$$

where α is the fine-structure constant and z is the charge of the particle in units of the electron charge. We can therefore write the second integral in Equation (3.22) as

$$N(\beta, \cos\theta) = \int_\lambda \mathrm{d}\lambda \frac{2\pi\alpha z^2}{\lambda^2}\left(1 - \frac{1}{\beta^2 n^2(\lambda)}\right)\mathrm{QE}(\lambda)e^{-\frac{\Delta^2(\lambda)}{2\theta_0^2}}. \tag{3.25}$$

This integral can be parameterized in terms of only two parameters: $\beta\cos(\theta)$ and $\beta\sin(\theta)\theta_0$.

Therefore we simply pre-calculate this integral for values of $\beta \cos(\theta)$ between -1 and 1 and for values of $\beta \sin(\theta)\theta_0$ between 0 and 1.

The effective scattering and absorption probabilities in Equation (3.22) are calculated as:

$$P_{\text{eff}}(\text{scattered} \mid x) = \frac{\int_\lambda \frac{1}{\lambda^2} QE(\lambda) P(\text{scatter} \mid \lambda, x)}{\int_\lambda \frac{1}{\lambda^2} QE(\lambda)}. \tag{3.26}$$

We can therefore write the expected charge as

$$\mu_i = \frac{1}{\sqrt{2\pi}\theta_0} \int_x \mathrm{d}x \frac{\Omega(x)}{2\pi} \frac{1}{|\sin\theta_1|} \epsilon(\eta) P_{\text{eff}}(\text{not scattered or absorbed} \mid E, x, v) N(\beta, \cos\theta). \tag{3.27}$$

This integral over the particle track is calculated numerically each time the likelihood is evaluated.

3.3 Time Distribution

In this section I discuss the calculation of the probability of observing a given PMT hit time. Since the majority of direct Čerenkov light will hit each PMT in a time window much smaller than the transit time of each PMT, we assume the final time distribution is Gaussian. Therefore, in order to calculate the probability of getting a hit at a certain time, we only need to calculate the mean time of the photon arrival at the PMT and then take the n^{th} first order statistic[2] of a Gaussian distribution with that mean and a standard deviation equal to the PMT transit time.

The mean time of a photon arrival at a PMT is given by an integral along the particle track of the time of flight from each point along the track to the PMT weighted by the expected charge at each point.

2. The first order statistic of a distribution is the distribution of values obtained by sampling n random variables from the distribution and selecting the smallest one.

$$\mathbb{E}(t_i) = \frac{1}{\sqrt{2\pi}\theta_0} \int_x \mathrm{d}x \frac{l(x)}{c} \frac{\Omega(x)}{2\pi} \frac{1}{|\sin\theta_1|} \epsilon(\eta) P_{\text{eff}}(\text{not scattered or absorbed} \mid E, x, v) N(\beta, \cos\theta) \tag{3.28}$$

where $l(x)$ is the effective path length from the track to the PMT:

$$l(x) = l_{\text{d2o}}(x) n_{\text{d2o}} + l_{\text{h2o}}(x) n_{\text{h2o}}. \tag{3.29}$$

The time distribution for direct light is then given by:

$$p_{\text{direct}}(t_i) = \frac{1}{\sqrt{2\pi}\sigma_{\text{TTS}}} e^{-\frac{(t_i - \mathbb{E}(t_i))^2}{2\sigma_{\text{TTS}}^2}} \tag{3.30}$$

where σ_{TTS} is the single PE transit time spread of the PMTs.

Reflected and scattered light is treated in an approximate way since it is less important for the fit and it is difficult to calculate analytically. We assume that all reflected and scattered light has a flat time distribution starting at the mean time of direct light and ending 160 ns later (two times the time it would take light to travel across the PSUP), i.e.

$$p_{\text{indirect}}(t) = \begin{cases} \frac{1}{2\Delta t_{\text{PSUP}}} & \mathbb{E}(t) < t < \mathbb{E}(t) + 2\Delta t_{\text{PSUP}} \\ 0 & \text{otherwise} \end{cases}, \tag{3.31}$$

where Δt_{PSUP} is approximately the time it takes light to cross the PSUP (80 ns).

The total time distribution is given by a charge weighted sum of the two distributions

$$\begin{aligned} p(t) = {} & \frac{\mu_i}{\mu_i + \mu_{\text{indirect}}} \frac{1}{\sqrt{2\pi}\sigma_{\text{TTS}}} e^{-\frac{(t_i - \mathbb{E}(t_i))^2}{2\sigma_{\text{TTS}}^2}} \\ & + \frac{\mu_{\text{indirect}}}{\mu_i + \mu_{\text{indirect}}} \begin{cases} \frac{1}{2\Delta t_{\text{PSUP}}} & \mathbb{E}(t) < t < \mathbb{E}(t) + 2\Delta t_{\text{PSUP}} \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{3.32}$$

where μ_{indirect} is the expected number of PE from scattering and reflections. This quantity is currently calculated by keeping track of the reflected and scattered light when integrating Equation (3.27) for each PMT, adding up all these contributions, and then assuming that a certain fraction[3] of it is equally distributed over all PMTs.

Finally, to evaluate the probability of observing a hit at a given time, t, given that n PE were detected, we compute the first order statistic of Equation (3.32)

$$P(t \mid n, E, \vec{x}, \vec{v}, t_0) = np(t)(1 - P(t))^{n-1}, \tag{3.33}$$

where $P(t)$ is the cumulative distribution function of Equation (3.32)[4].

3.4 Electromagnetic Showers and Delta Rays

In addition to the direct Čerenkov light from the primary particle, there is also a significant amount of light created from electromagnetic showers for high energy electrons and muons and from delta rays created by muons. Since both of these processes are very complex, we model them in a *very* approximate way. For both processes we assume that the number of photons emitted along the track is independent from the angular distribution of the light along the track, i.e. that we can approximate

$$N(x, \theta) \approx N(x)f(\theta)$$

where x is the distance along the track and θ is the angle between a PMT and the particle track[5]. We describe both $N(x)$ and $f(\theta)$ using simple functional forms.

3. This fraction is currently set to 40% which was determined by floating that parameter in several fits of electrons and muons. This value is reasonable since the approximate coverage of the SNO detector is 50% and we expect to lose slightly more than that due to absorption (when doing the initial sum of all reflected and scattered light the quantum efficiency of the PMT is already taken into account).

4. When computing the likelihood function I use the pt1 variable in the SNO data structure for the PMT hit times. I use this instead of the multi-photon PCA (PMT Calibration) time since that time was designed to correct for the first order statistic effect. I also don't use the regular pt time since that was only calibrated to work with single photons.

5. It should be noted that the angular distribution of the light *definitely* changes as a function of the longitudinal position along the particle track. By assuming the angular distribution is constant we are not even self-consistent since we might use a particular form for $f(\theta)$ at 100 MeV, but a different form at 200 MeV even though at some distance along the latter's track its kinetic energy will be 100 MeV. Nevertheless we stick with a simple form because it makes the problem much more computationally tractable.

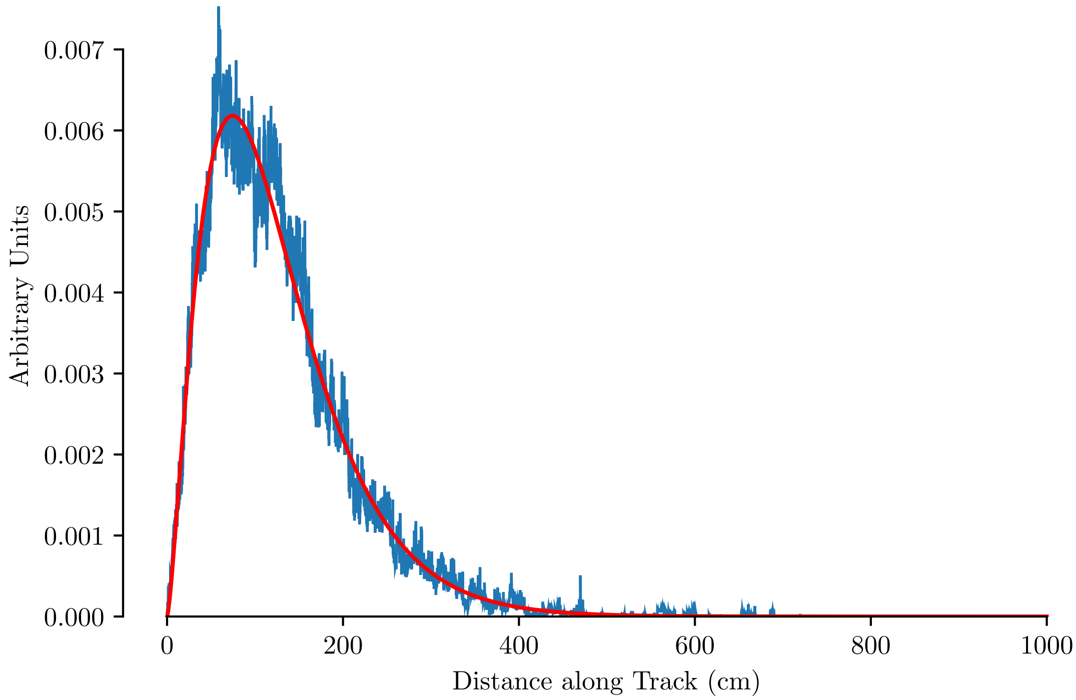

Figure 3.1: Position distribution of photons from the electromagnetic shower of a 1 GeV electron.

For electromagnetic showers, we model the longitudinal profile of the emitted photons with a gamma distribution,

$$N(x) = \frac{1}{\Gamma(k)\theta^k} x^{k-1} e^{-\frac{x}{\theta}}. \tag{3.34}$$

This model comes from the "Passage of Particles Through Matter" review in the PDG[3], which states that the *energy* distribution in an electromagnetic shower is well described by a gamma distribution. An example fit is shown in Figure 3.1.

The angular distribution for electromagnetic showers is described by the function:

$$f(\cos\theta) \propto e^{-\frac{|\cos\theta - \mu|^\alpha}{\beta}}. \tag{3.35}$$

This functional form does not have any theoretical motivation that I'm aware of but fits

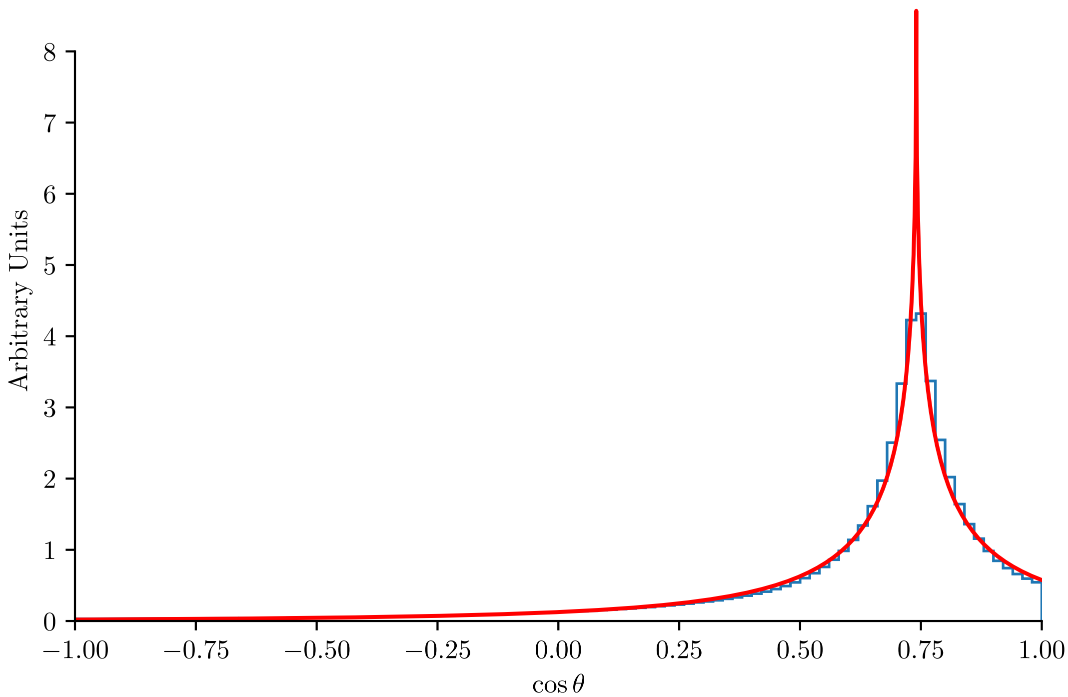

Figure 3.2: Angular distribution of photons from the electromagnetic shower of a 1 GeV electron. The red line shows a fit to the functional form shown in Equation (3.35), where we fit for α, β, and μ.

the angular distribution of Čerenkov light very accurately. Figure 3.2 shows the angular distribution of light from the electromagnetic shower of a 1 GeV electron along with a fit to Equation (3.35).

Therefore, to describe the light from an electromagnetic shower we need five parameters: k, θ, μ, α, and β. Instead of creating lookup tables, we parameterize each of these variables as a function of energy for electrons and muons. This approach has the advantage that the results consist of only a handful of numbers instead of large tables and the parameterization means that the parameters all change smoothly as a function of energy which prevents the likelihood from having discontinuous derivatives. The parameterization of these values for electrons is discussed in Section 3.4.1 and for muons in Section 3.4.2.

3.4.1 Electrons

Electromagnetic Showers

To determine the longitudinal distribution of Čerenkov photons for electrons, I first simulated electrons with energies between 10 MeV and 1 GeV using RAT-PAC (RAT, Plus Additional Codes)[6] in a volume composed of pure heavy water[6]. I then fit Equation (3.34) to the starting vertices of all Čerenkov photons in the wavelength range 200 nm to 800 nm. The θ parameter is almost constant above 100 MeV where electromagnetic showers start to become important, and so I use a constant value of 43.51 for this parameter. The k parameter is then calculated using Equation 33.36 in the PDG "Passage of Particles Through Matter" review[7]

$$t_{\max} = X_0 \frac{k-1}{\theta} = \ln y + C_e, \tag{3.36}$$

where $t_{\max}$ is the longitudinal peak of the shower distribution in units of the radiation length, X_0 is the radiation length in water, y is the energy of the electron in units of the critical energy, C_e = -0.5, and k and θ are the parameters in Equation (3.34). Using the right hand side of Equation (3.36) we are able to calculate $t_{\max}$, and then we can use the fitted value of k to determine θ. The bottom two plots in Figure (3.3) show the fitted values of k and θ as a function of the kinetic energy of the electron.

For the angular distribution, μ is assumed to be equal to the Čerenkov angle. Based on the simulations, this is a decent approximation at low energies and a very good approximation at higher energies. We then fit the energy dependence of α and β to the following form:

---

6. RAT-PAC is an open source spin-off of the RAT package which is used to simulate events in the SNO+ experiment which uses an upgraded version of the SNO detector.

7. The actual equation in the PDG is given as $t_{\max} = \frac{a-1}{b}$. The form given here is obtained by switching into the units we use here where $a \rightarrow k$ and $b \rightarrow \frac{X_0}{b}$.

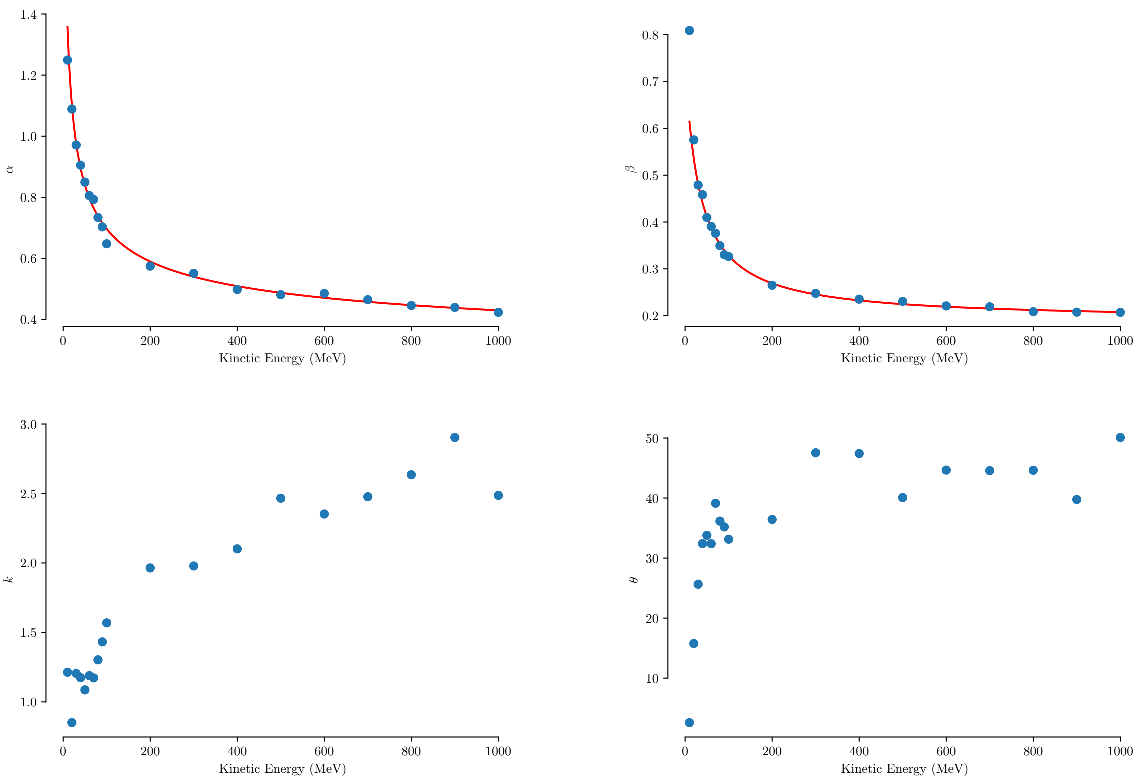

Figure 3.3: Electromagnetic shower parameters as a function of kinetic energy for electrons. The top two plots show α and β for the angular distribution of the electromagnetic shower photons. The bottom two plots show k and θ which describe the longitudinal position of the electromagnetic shower photons.

	Electrons		Muons	
Parameter	α	β	α	β
c_0	3.14×10^{-1}	1.35×10^{-1}	8.24×10^{-1}	2.24×10^{-1}
c_1	2.08×10^{-1}	2.22×10^{-1}	3.90×10^{-3}	5.38×10^{-3}
c_2	6.33×10^{-3}	1.96×10^{-2}	1.58×10^{-5}	1.20×10^{-5}
c_3	1.19	1.24	9.99×10^{-1}	1.00

Table 3.1: Parameters describing the angular distribution of Čerenkov light from electromagnetic showers for electrons and muons as a function of initial kinetic energy.

$$\alpha(T) = c_0 + \frac{c_1}{\log(c_2 T + c_3)} \tag{3.37}$$

$$\beta(T) = c_0 + \frac{c_1}{\log(c_2 T + c_3)} \tag{3.38}$$

where T is the initial kinetic energy of the electron.

Table 3.1 shows the fit results for α and β as a function of energy. The top two plots in Figure 3.3 show α and β as a function of the kinetic energy of the electron along with the fits.

The total number of shower photons as a function of energy is fit to the following form:

$$n(T) = T_{\text{rad}} \left(c_0 + \frac{c_1}{\log(c_2 T + c_3)} \right) \tag{3.39}$$

$$\tag{3.40}$$

where T_{rad} is the total energy loss to radiation. The energy lost to radiation is computed by numerically integrating lookup tables downloaded from the ESTAR website run by NIST[7]. Table 3.2 shows the constants used in Equation (3.40).

Parameter	Value (n)
c_0	406.75
c_1	0.272
c_2	5.31×10^{-5}
c_3	1.00

Table 3.2: Parameters describing the number of electromagnetic shower photons for electrons as a function of initial kinetic energy.

3.4.2 Muons

Electromagnetic Showers

To describe the light from electromagnetic showers and delta rays for muons, I simulated muons with total energies between 300 MeV and 20 GeV using RAT-PAC in a volume composed of pure heavy water[6]. I then fit Equation (3.34) to the starting vertices of all Čerenkov photons in the wavelength range 200 nm to 800 nm. To describe the θ parameter as a function of energy, I used a single degree polynomial fit,

$$\theta(T) = -7.8 + 0.118928T. \tag{3.41}$$

The k parameter was fit to a constant value of 1.5. The values of θ and k as a function of energy are shown in the middle two plots in Figure 3.4. The linear fit and constant value we use here are *not* a very good fit to the data but since the number of photons from electromagnetic showers is dwarfed by the number of photons from delta rays it is not expected to affect the reconstruction very much.

The angular distribution is described using the same functional form as for electrons and, similarly, μ is assumed to be equal to the Čerenkov angle. The constants used to describe the variation of α and β with energy are shown in Table 3.1 and the data is shown in the top two plots of Figure 3.4.

The total number of shower photons as a function of energy was fit to the following form:

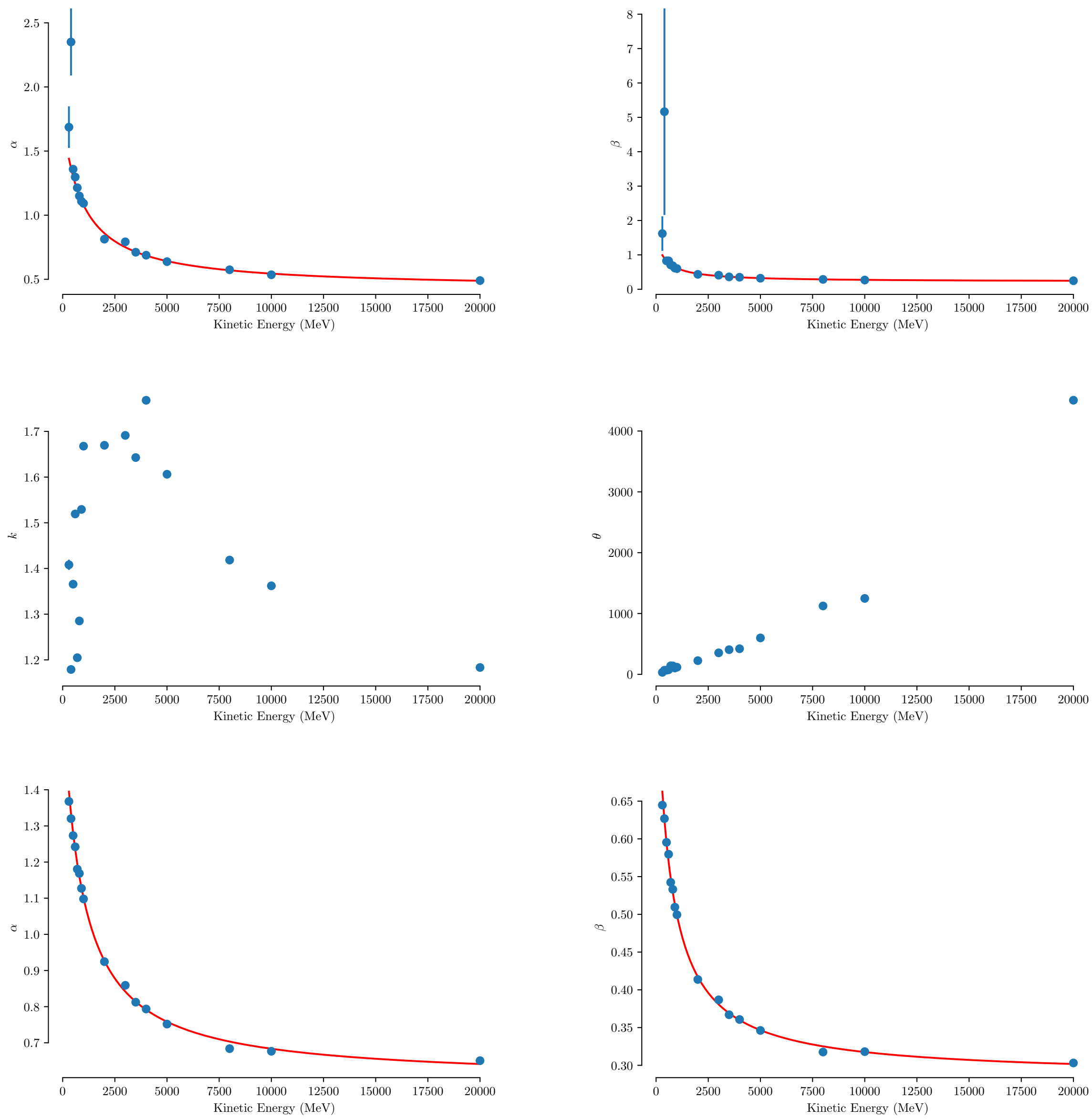

Figure 3.4: Electromagnetic shower and delta ray parameters as a function of kinetic energy for muons. The top two plots show α and β for the angular distribution of the electromagnetic shower photons. The middle two plots show k and θ which describe the longitudinal distribution of the electromagnetic shower photons. The bottom two plots show α and β for the angular distribution of the photons from delta rays.

Parameter	Value (n)
c_0	9.289×10^3
c_1	8.40×10^2

Table 3.3: Parameters describing the number of electromagnetic shower photons for muons as a function of initial kinetic energy.

Parameter	α	β
c_0	3.46×10^{-1}	2.30×10^{-1}
c_1	1.11×10^{-2}	4.09×10^{-3}
c_2	5.66×10^{-6}	8.22×10^{-6}
c_3	1.01	1.01

Table 3.4: Parameters describing the angular distribution of Čerenkov light from delta rays as a function of initial muon kinetic energy.

$$n(T) = T_{\text{rad}} \left(c_0 \left(1 - e^{-\frac{T0}{c_1}} \right) \right), \tag{3.42}$$

where T_{rad} is the total amount of energy loss to radiation. These values are calculated by numerically integrating the muon energy loss to radiation from lookup tables downloaded from the PDG website[8]. Table 3.3 shows the constants used in Equation (3.42).

Delta Rays

The angular distribution of the light from delta rays is described using the same functional form as the shower photons. The parameters describing the angular distribution of the delta ray light are shown in Table 3.4 and the data is shown in the bottom two plots of Figure 3.4.

The total number of Čerenkov photons in the wavelength range from 200 nm to 800 nm from delta rays as a function of energy is parameterized as

$$n(T) = \max\left(0.0, -7532.39 + 39.4548T\right). \tag{3.43}$$

3.5 Vertex Seed

When running the log likelihood minimization, the minimization procedure is started at a point called the "seed" of the fit. Starting the fit at a seed position close to the true position is important for a couple of different reasons: first, many complex likelihood functions have several local minima away from the true global minimum. Since the fitting procedure is not guaranteed to find the global minimum and instead only looks for local minima, by starting the fit further away from the global minimum we increase the probability that the fitter fails to converge to the true global minimum. Second, the further the algorithm starts from a minimum, the longer it will take to reach the minimum. It therefore pays to invest a small amount of time up front to use approximate techniques to start off near the minimum since it will result in a large net time savings.

The seed position for the fitter is determined using a slightly modified version of the QUAD fitter developed for SNO[9]. The original QUAD fitter was implemented in SNOMAN and works by sampling 4 PMT hits at a time and computing the unique position and time consistent with producing those four PMT hits. This process of selecting 4 random hits and computing a time and position is repeated over and over, producing a "cloud" of quad points. The final best fit time and position was then computed by searching for a local maximum in the density of the cloud points using the AMOEBA algorithm from Numerical Recipes[9].

In my version of the QUAD fitter there are two significant differences:

1. Instead of using the AMOEBA algorithm to find the maximum density in the cloud points, I simply take the median of the cloud points[8]

2. Instead of randomly sampling 4 PMT hits at each iteration, points are sampled proportionally to the probability that they are produced from more than one photon. For the high energy events I'm interested in, PMT hits with single photon hits are more

8. This procedure is the same as the one used in the SNO+ version of QUAD.

likely to be due to scattered light, which will produce a bad cloud point.

3. Since the particle tracks are often tens of centimeters or meters long for high energy particles, we added the option to filter the final quad points taking only those with a time less than the 10% quantile of all the times. The best fit point is then returned from the median of the remaining points. By selecting the points less than the 10% quantile we get a much more accurate starting position for muons that travel several meters in the detector.

This second option is useful for particles like external muons but can make the initial guess worse for smaller energy events. Therefore, we perform two fits for each event and particle combination; one with no filtering of the quad points and one where we only select the quad points with a time less than the 10% quantile of all the times[9].

3.6 Direction Seed

In addition to seeding the position of the fit, it is also necessary to seed the initial direction of the fit. For a single particle, this is simple and one can use a simple method like finding the centroid of the hit PMTs, however for multi-track fits it is necessary to have a more complicated ring finding algorithm. Therefore, I designed a ring finding algorithm based on a custom Hough transform inspired by the one employed in a previous atmospheric SNO analysis[10].

The algorithm starts by constructing a 2 dimensional array representing bins in a pro-

9. This is very inefficient since it effectively doubles the amount of time needed to fit an event. I started looking into a smarter way to try and select the beginning of a track from a set of quad points, but didn't have enough time to test it sufficiently. The idea was to compute a Mahalanobis distance from the quad points and then select a spot with the earliest time but within some specified Mahalanobis distance like 1 or 2. For particles without an extended track the requirement that the Mahalanobis distance be relatively small means we shouldn't bias ourselves too much. However, for particles like external muons with a large track the hope was that the Mahalanobis distance would be small out to the edge of the cloud of quad points. Another idea which I wasn't able to test was to fit a linear regression to the quad cloud and then select the earliest point along this line which still has a high density of nearby points.

jection of the surface of the detector. The columns represent evenly spaced bins between 0 and π for the polar angle θ, and the rows represent evenly spaced bins between 0 and 2π for the azimuthal angle ϕ. I then construct a weight, w for each PMT hit which represents the probability that the PMT hit was caused by multiple photons:

$$w = P(\text{multiple photons} \mid q) = 1 - P(1 \text{ PE} \mid q) = 1 - \frac{P(q \mid 1 \text{ PE})P(1 \text{ PE})}{P(q)}, \tag{3.44}$$

where $P(1 \text{ PE})$ and $P(q)$ are calculated assuming the number of photons striking each PMT is Poisson distributed with a mean equal to the total charge in the event divided by the number of PMTs hit[10]. This is intended to be a rough proxy for the probability that the light is direct and not scattered or reflected. I then loop over every single PMT hit, discarding hits whose time residual (based on the position found by QUAD) is greater than 10 ns. For each PMT hit we add a value to each bin in the 2D array proportional to

$$we^{-(\cos\theta - \frac{1}{n})^2/0.01}, \tag{3.45}$$

where $\cos\theta$ is the angle between the PMT hit and the current bin and $\frac{1}{n}$ is the cosine of the Čerenkov angle in heavy water[11]. Finally, the peak in the 2D array is then found and that is the first ring.

The whole process is repeated a specified number of times with the only difference being that on subsequent iterations PMTs are skipped which are within the Čerenkov cone of previously found rings. A PMT hit is assumed to be a part of a previous ring if

10. This is a *very* rough approximation and we don't expect all the PMTs to have the same mean number of photons, but the term which dominates the equation is $P(q \mid 1 \text{ PE})$ and so this is a good enough approximation for the other terms.

11. The form of the weighting function, a Gaussian like weight as a function of the angle between the PMT hit and the bin, was chosen since it is the same form as the expected angular distribution from direct Čerenkov light calculated in Chapter 3. The value of 0.01 in the denominator was chosen by trial and error to give good results based on the atmospheric Monte Carlo.

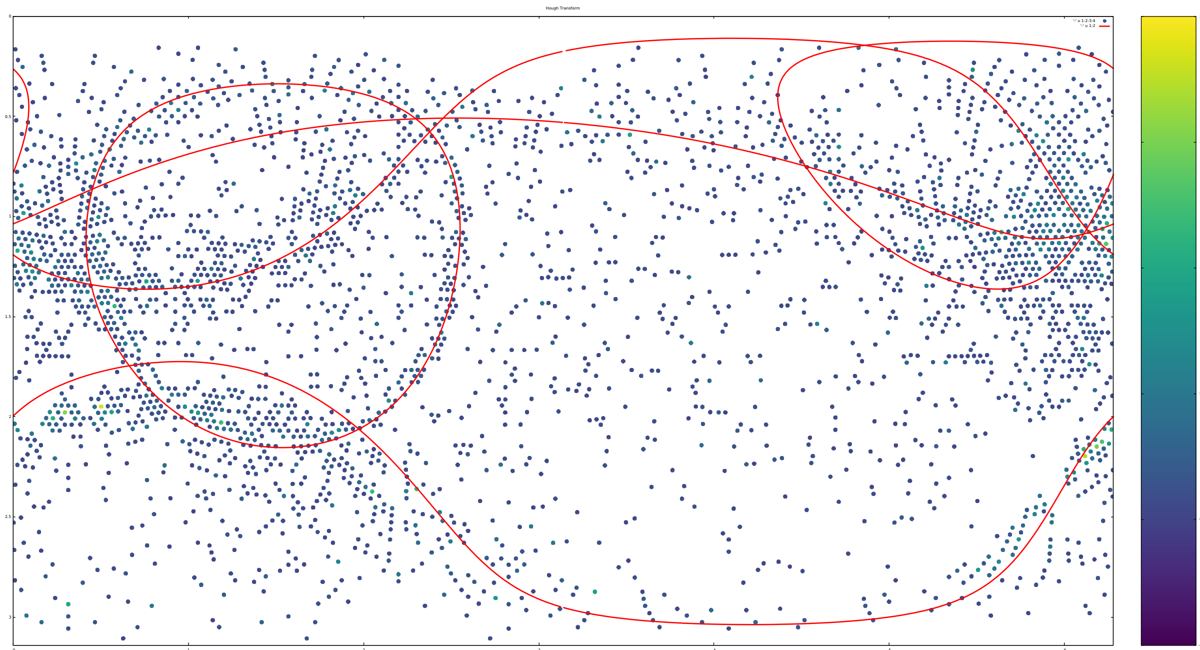

Figure 3.5: Plot showing the 5 peaks detected in an atmospheric neutrino event.

$$\left|\cos\theta - \frac{1}{n}\right| < 0.1, \tag{3.46}$$

where $\cos\theta$ is the cosine of the angle between the previous direction and the current PMT. Finally, we ignore rings if they are within 0.1 radians of any previously found ring. Figure 3.5 shows the results of the algorithm applied to an atmospheric neutrino event with multiple rings.

This algorithm has the advantage that it works well and is very simple to implement. The one disadvantage is that it is unable to actually predict the number of rings and instead will always return a specified number of rings. Although it may be possible to add something to determine when a new ring falls below some threshold, I chose to keep the algorithm as is and instead let the likelihood fits determine the total number of particles.

When performing the likelihood minimization, before running the “full” minimization we do several shorter quick minimizations (with a maximum of 1000 steps each) using every possible combination of particle type and direction seed. The best of these minimizations is

Electron	Muon
1	1
1	2
1	3
2	1
2	2
2	3
3	1
3	2
3	3

Table 3.5: Example showing the 9 different possible "quick" fits performed when fitting for an electron and muon with 3 seed directions. The numbers in each row represent which one of the 3 possible directions is the seed direction for the electron or muon.

then selected and we continue the minimization process. For example, if we were fitting for the particle combination of an electron and a muon and we have 3 direction seeds, labeled 1 to 3, then we would perform the fits shown in Table 3.5 during the quick minimization phase.

3.7 Goodness of Fit Parameter ψ

The goodness of fit parameter ψ is designed to measure how well the event is reconstructed (similar to a χ^2 value for a least squares fit). The ψ parameter is defined as the log of the likelihood ratio between the maximum likelihood of the fit and the likelihood of the best hypothesis in a restricted class of models Ω, i.e.

$$\psi \equiv \log \mathscr{L}_{\text{fit}} - \log \mathscr{L}_{\Omega}, \tag{3.47}$$

where $\mathscr{L}_{\text{fit}}$ represents the maximum likelihood of the fit and $\mathscr{L}_{\Omega}$ represents the likelihood of the best hypothesis in Ω. In our case, Ω represents the class of all models where we specify a mean number of expected PE and the mean time that we expect each PMT to be hit. In other words, in Ω we still assume that the number of photons hitting each PMT is Poisson distributed, that the charge distribution for single PEs is given by the measured distribution in SNO PMTs, and that the time resolution of the PMT is 1.6 ns. But in this class of models we are free to vary the expected number of PE and the hit time for each PMT independently. In some sense, this model is intended to capture the *best* possible likelihood value one could hope for.

To actually calculate $\mathscr{L}_\Omega$, we first determine the mean number of PE hitting each PMT, μ, such that the probability of observing the known charge is maximized, i.e. we maximize

$$P(q \mid \mu) = \sum_i P(q \mid n) P(n \mid \mu)$$

as a function of μ. We then choose the mean hit time t to be equal to the value such that the first order statistic of μ samples is equal to the actual hit time. For PMTs which aren't hit we assume μ is equal to the noise rate of the PMTs. With μ and t calculated for each PMT we then evaluate Equation (3.8) to calculate the likelihood.

As an example of the discriminatory power of ψ, the distributions of ψ/Nhit for atmospheric neutrino events and tagged flashers are shown in Figure 3.6 .

3.8 Particle ID

To determine the particle ID and multiplicity of events, we fit a single event under multiple hypotheses and then perform a likelihood ratio test. By default we fit each event with up to two tracks and consider each track as being from an electron or muon. Therefore, each event is fit under the single electron, single muon, double electron, electron plus muon, and double muon hypotheses[12].

To pick the most likely hypothesis, the best fit is selected according to the fit with the highest likelihood multiplied by an "Ockham factor" and a correction term. The Ockham

12. Although we are actually looking at finding electron-positron pairs, electrons and positrons produce a very similar signal in the detector. Therefore as far as the reconstruction is concerned, it is looking for a signal of 2 electrons. The same argument applies for muons and antimuons. Therefore, when speaking of doing particle ID I will refer to these events as 2 electrons or 2 muons even though we are really talking about a particle anti-particle pair.

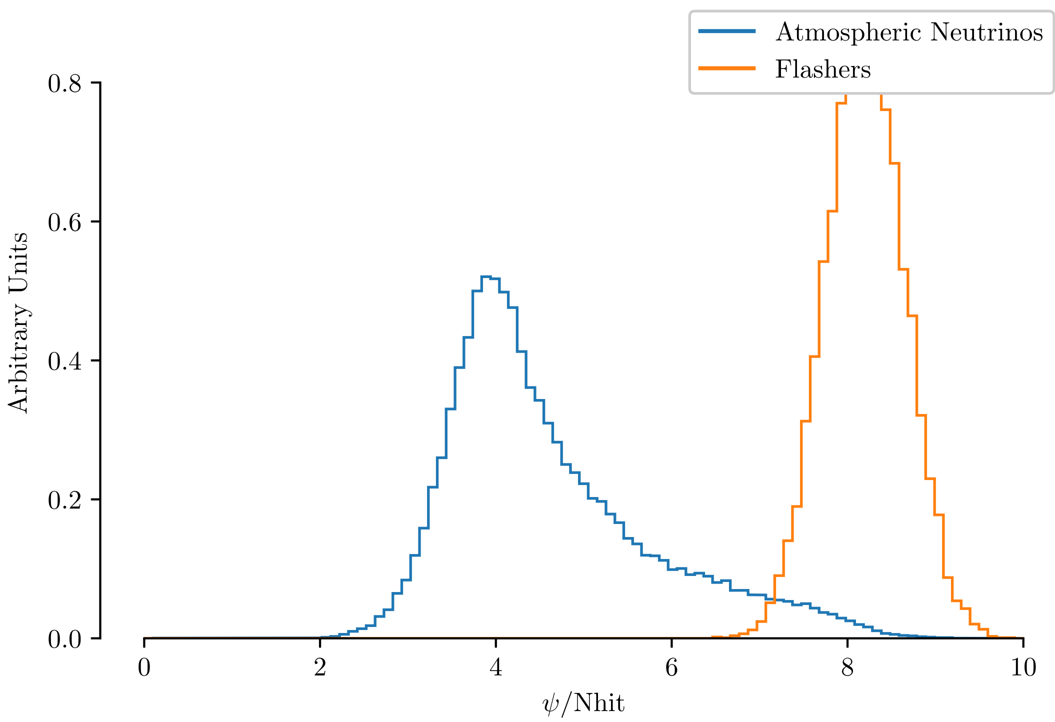

Figure 3.6: Distribution of ψ/Nhit for atmospheric neutrinos and flashers.

factor accounts for the fact that we require the likelihood to be significantly better before accepting a significantly more complex hypothesis (more particles) and the correction term is to account for the fact that the likelihood is not able to model muon scattering and electromagnetic showers correctly. The Ockham factor in its most general form is given by:

$$W = \int \frac{\mathscr{L}(\vec{\theta})}{\mathscr{L}_{\text{max}}} P(\vec{\theta} \mid \text{I}) \mathrm{d}\vec{\theta}, \tag{3.48}$$

where $\vec{\theta}$ represents all the variables being fitted for (position, energy, direction, etc.), $\mathscr{L}_{\text{max}}$ represents the maximum value of the likelihood, and $P(\vec{\theta} \mid \text{I})$ represent priors on the variables being fitted for[11].

Although it is possible to calculate this quantity approximately using something like a Markov Chain Monte Carlo, instead we assume that the prior density is very broad (we will use flat priors) and we can approximate the integral as:

$$W = V(\Omega')P(\vec{\theta} \mid \text{I}), \tag{3.49}$$

where $V(\Omega')$ is the volume of some high-likelihood region Ω'. For example, this volume could be computed as the volume of likelihood space where the negative log likelihood is less than 0.5 from the minimum. Although calculating something like this is possible, we instead make another very simple assumption that the width of the likelihood space is the same as the average uncertainty on an ensemble of fits[13]. For example, we assume that the width of the likelihood space in each of the position directions is approximately 10 cm since that is the average position resolution. Furthermore we assume the various parameters are uncorrelated and so we can calculate W as:

13. This assumption is not likely to be true in practice for a couple of reasons. First, some of the uncertainties have a very strong radial dependence and so events near the edge of the detector may have a much larger uncertainty on the position. Second, for multi-particle events the situation is a lot more complex and depends on the exact direction and energy of the two particles. Ideally here we would numerically estimate the Ockham factor by running a Markov Chain Monte Carlo at the final fit position.

$$W = V(x)V(y)V(z)P(x, y, z \mid \mathrm{I}) \prod_i V(T_i)V(\omega_i)P(T_i, \omega_i \mid \mathrm{I}) \tag{3.50}$$

where i sums over the total number of particles and ω represents the volume of solid angle for the direction. Since the position resolution is roughly 10 cm for both electrons and muons we can approximate $V(x)V(y)V(z)$ as $(10 \text{ cm})^3$. Similarly the energy resolution is approximately 5% for both electrons and muons so we can write:

$$V(T) \approx 0.05T.$$

Finally, we take the volume of the direction to be equal to the approximate solid angle of a cone with opening angle equal to the direction resolution which is approximately 1 degree for muons and goes from 4 degrees to 1 degree for electrons depending on the energy[14].

For the priors we simply assume flat priors for all the variables. Therefore,

$$\log(W) \approx \log\left(\frac{(10 \text{ cm})^3}{\frac{3}{4}\pi R_{\text{PSUP}}^3}\right) + \sum_i \log\left(\frac{0.05T_i}{1 \text{ GeV}}\right) + \log\left(\frac{\pi\Delta\theta_i^2}{4\pi}\right). \tag{3.51}$$

In addition to the terms shown in Equation (3.51) we also add a correction factor which is currently a constant of +100 to any events with two particles. The reason for this is that there are several aspects of the likelihood function which aren't modeled perfectly. This includes Rayleigh scattered light, photons reflected from the concentrators, and hard scattering from muons. Because of these factors, the fit often wants to "fix" some of this mismodeling by adding a second particle. This factor can be thought of as a prior on the likelihood actually containing two particles given the results of the fit. For example, this factor could in principle look at the results of the two particle fit and if the two directions are close enough and one of the particles has significantly less energy it may be trying to "correct" for a hard shower. In practice, I've found that a constant value of +100 works well

14. In the final analysis I actually used a constant factor of $\frac{1\times10^{-4}}{4\pi}$ for the solid angle term. In practice, the correction factor of +100 for each additional particle dominates the Ockham factor, and so the solid angle term has little effect.

enough for this analysis and so I have not attempted to optimize it further.

3.9 Single Track Performance

To benchmark the performance of the reconstruction algorithm, I simulated electrons and muons isotropically in the acrylic vessel. The electrons were simulated with a total energy from 100 MeV to 1 GeV in 100 MeV steps and the muons from 300 MeV to 1 GeV in 100 MeV steps.

3.9.1 Particle ID

In Figure 3.7 the log of the likelihood ratio for electrons over muons is shown as a function of reconstructed electron kinetic energy. Combining all events, the fraction of electrons and muons which are correctly identified in this energy range is 99%.

3.9.2 Energy Resolution and Bias

The energy resolution and bias are shown in Figures 3.8 and 3.9 respectively. The bias is calculated as the median of the difference between the reconstructed and true kinetic energy. The resolution is calculated as the interquartile range divided by 1.35 which gives a good approximation to the standard deviation of the central part of the distribution. This metric was chosen instead of directly computing the standard deviation of the residuals because a small number of events occasionally misreconstruct in a local minimum far away from the global minimum. A small number of these events can have an arbitrarily large impact on the standard deviation causing it to not be representative of the underlying distribution. The "tails" of these distributions are naturally taken into account in the final analysis since the expected signal is derived solely from reconstructed Monte Carlo simulations.

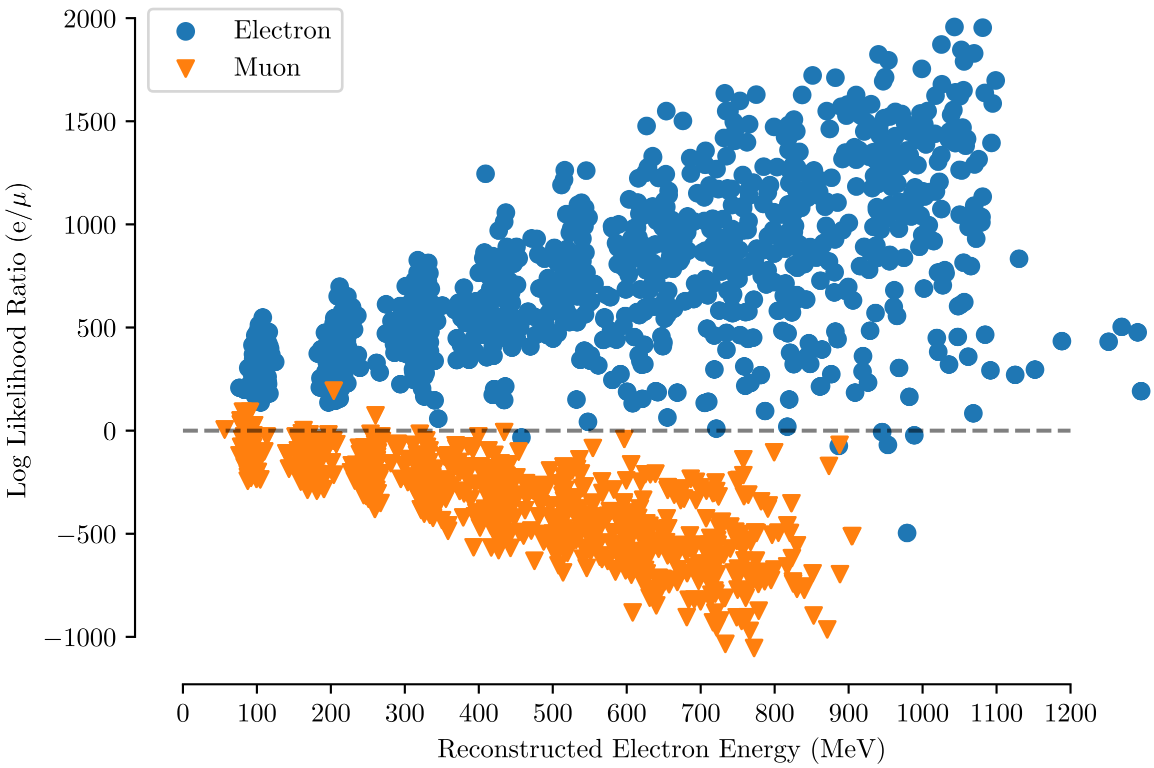

Figure 3.7: Log likelihood ratio vs reconstructed electron energy for single electrons and muons with kinetic energy between approximately 100 MeV and 1 GeV. The dashed line represents a log likelihood ratio value of 0 which is where we cut to decide whether an event is an electron or muon.

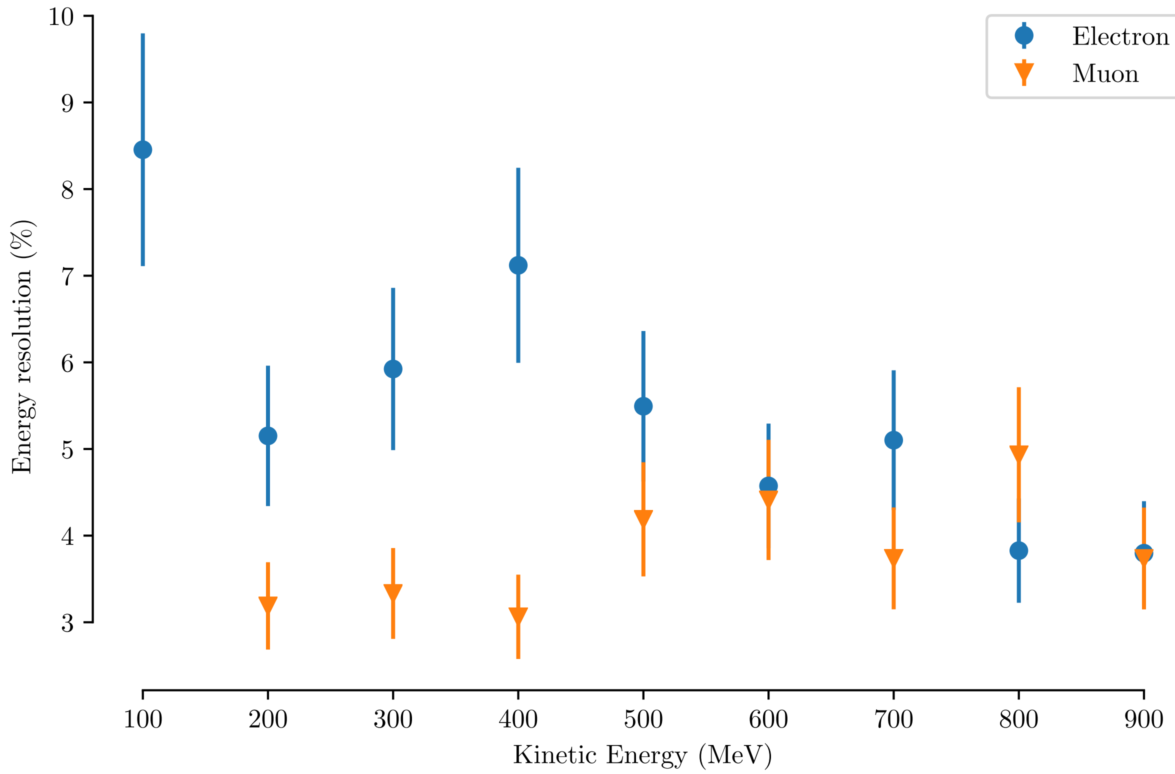

Figure 3.8: Energy resolution as a function of kinetic energy for single electrons and muons. The resolution is plotted as a fraction of the kinetic energy. The energy resolution is defined here as 1.35 times the interquartile range of the difference between the reconstructed and true kinetic energy.

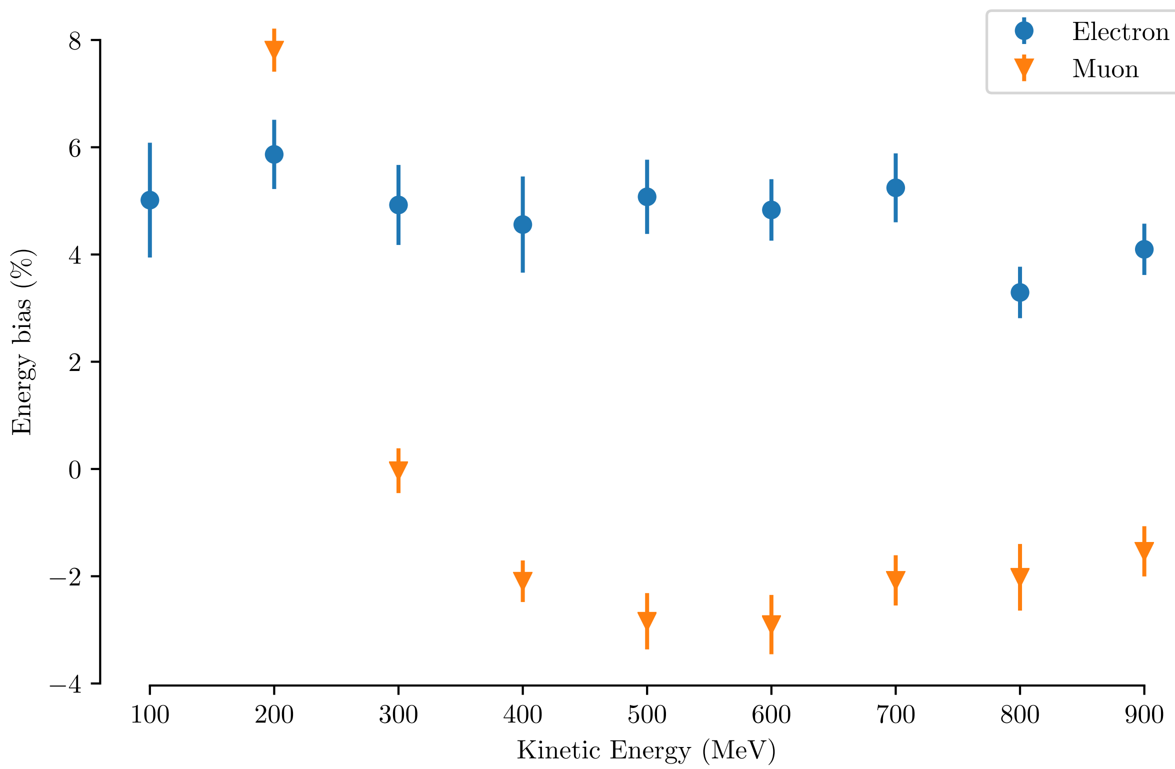

Figure 3.9: Energy bias as a function of kinetic energy for single electrons and muons. The bias is plotted as a fraction of the kinetic energy. The bias is calculated by finding the median of the difference between the reconstructed and true kinetic energy.

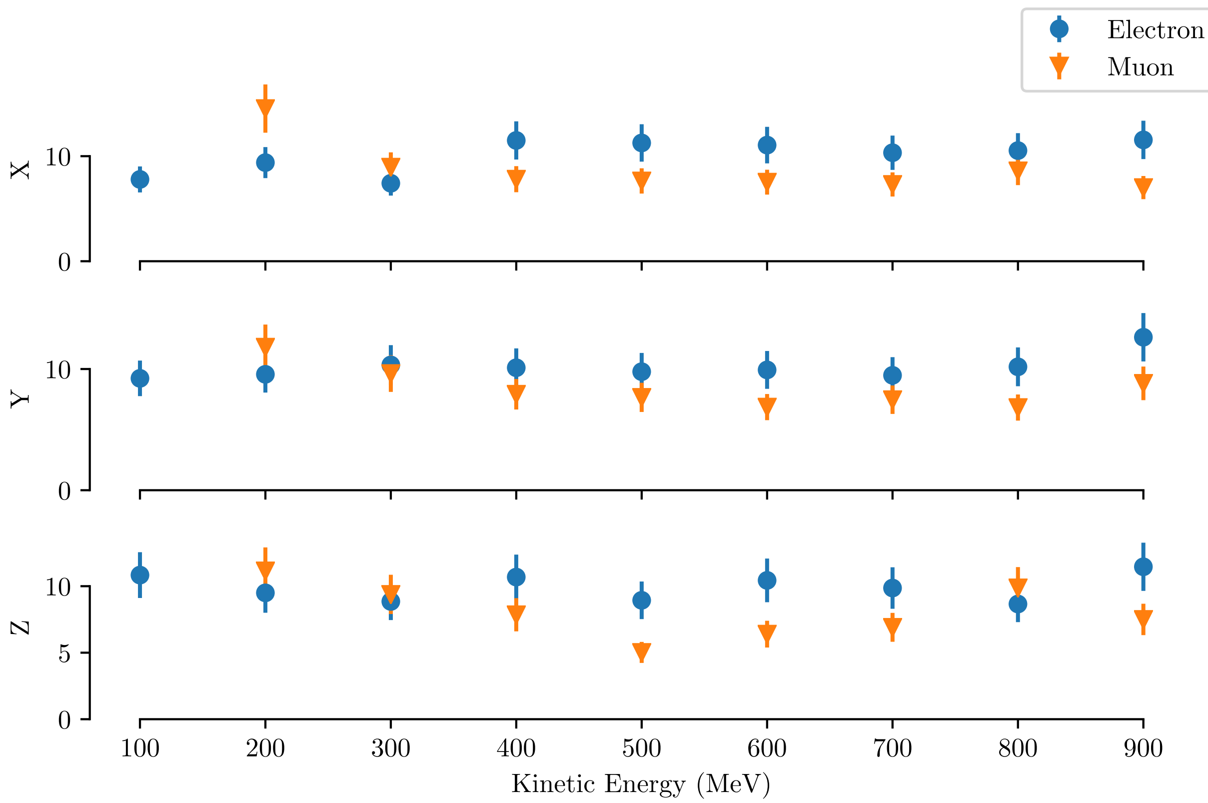

Figure 3.10: Position resolution in cm as a function of kinetic energy for single electrons and muons. The position resolution is defined here as 1.35 times the interquartile range of the difference between the reconstructed and true positions.

3.9.3 Position Resolution and Bias

Figures 3.10 and 3.11 show the position resolution and bias respectively. The bias and resolutions are characterized in the same way as for energy by the median and interquartile range divided by 1.35. The position resolution is approximately 10 cm across the entire energy range for both electrons and muons and there is no significant bias for electrons or muons.

3.9.4 Angular Resolution

The angular resolution is shown in Figure 3.12. The resolution here is defined as the standard deviation of the angle between the true initial track direction and the reconstructed track

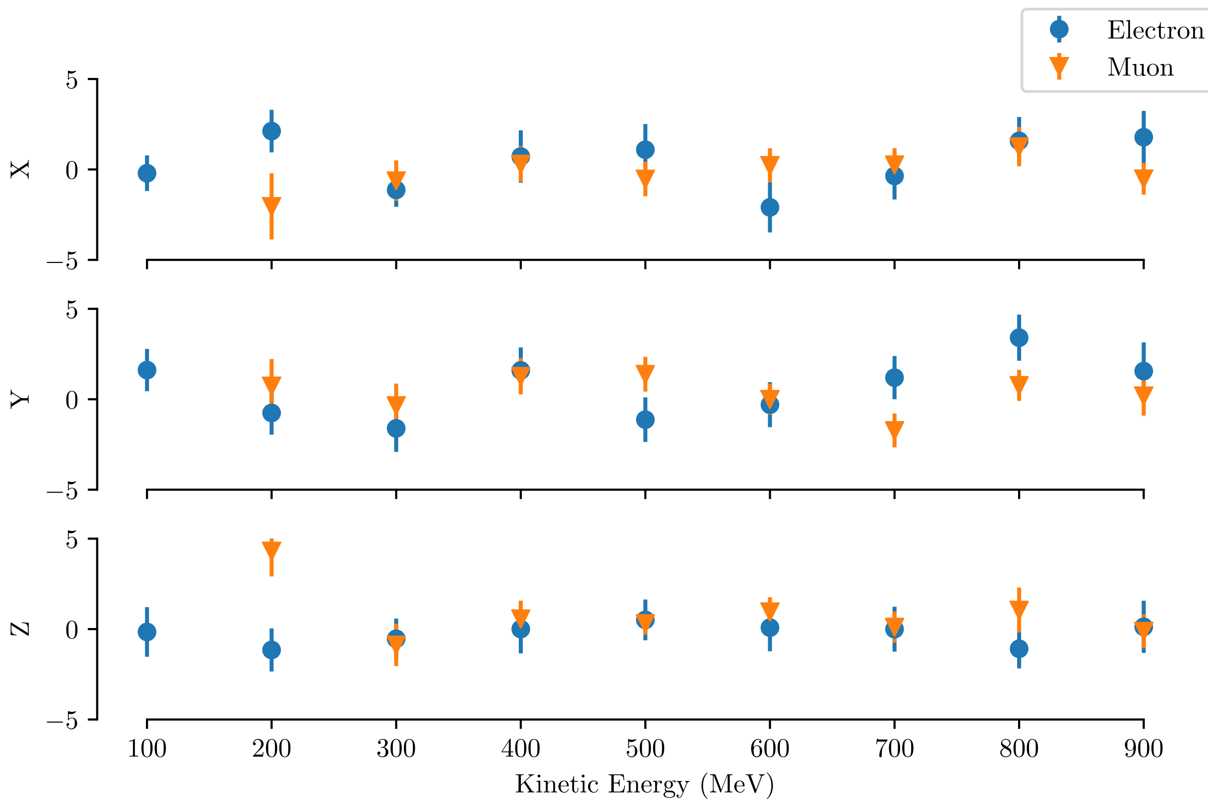

Figure 3.11: Position bias in cm as a function of kinetic energy for single electrons and muons. The bias is calculated by finding the median of the difference between the reconstructed and true position.

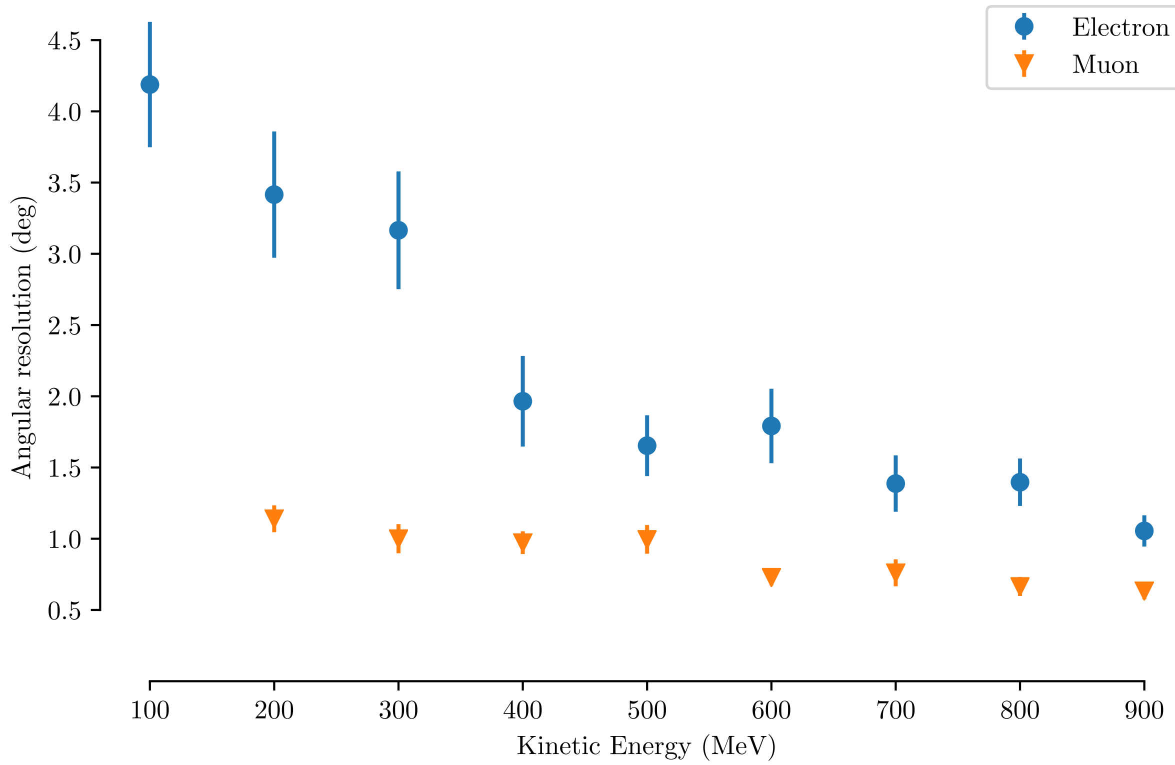

Figure 3.12: Angular resolution as a function of kinetic energy for single electrons and muons. The angular resolution is defined here as the standard deviation of the angle between the reconstructed and true directions.

direction. The angular resolution for muons is constant at approximately 1 degree while the resolution for electrons varies between 4 degrees and 1 degree, getting better at higher energies.

3.10 Multi-Track Performance

In this section I discuss the performance of the reconstruction algorithm when fitting for multiple tracks from the same position. A full characterization of the performance is difficult since, when considering 2 particles, it is necessary to consider the probability of reconstructing various quantities as a function of the energy, particle id, opening angle, and position of both particles. Therefore, in this section I will simply choose a few key quantities to look at

	Electrons					Muons				
T (MeV)	e	μ	ee	eμ	$\mu\mu$	e	μ	ee	eμ	$\mu\mu$
100	92		8			-	-	-	-	-
200	89		9	2		-	-	-	-	-
300	93		7			8	92			
400	91		8	1			98	2		
500	87		10	3		2	98			
600	87		13				96		3	1
700	86		10	4			89		11	
800	91		4	5			94		4	1
900	84		8	8			96		4	
1000	84	1	6	8			94		6	

Table 3.6: The probability of reconstructing a given particle combination from single electrons and muons in the AV for various kinetic energies.

to give an overview of the performance of the fitter.

3.10.1 Particle ID

We first consider the probability of reconstructing a given particle combination (i.e. single electron, single muon, electron + muon, etc.) from a known single particle type. Since the majority of the atmospheric background events involve a single charged lepton and our signal is expected to show up as two or more leptons, this gives us a handle on the extra contamination from single lepton atmospheric events we can expect. A table showing these probabilities based on fitting single electrons and muons with energies from 100 MeV to 1 GeV and 300 MeV to 1 GeV for electrons and muons respectively is shown in Table 3.6.

Second, we consider the probability of correctly tagging lepton pairs. Table 3.7 shows the probability that an electron-positron and muon-antimuon pair reconstruct as various particle combinations. For both the electron-positron and muon-antimuon pair, the fraction correctly identified is close to 100% at lower energies and decreases to approximately 90% at higher energies. At higher energies the pair is boosted and the two rings from each particle have a higher chance of overlapping, making it more difficult to distinguish the pair from a

	$\psi \rightarrow e^- + e^+$					$\psi \rightarrow \mu^- + \mu^+$				
T (MeV)	e	μ	ee	eμ	$\mu\mu$	e	μ	ee	eμ	$\mu\mu$
0			92	8						100
100			100							100
200			100							100
300			93	7						100
400			97	3						100
500			93	7						100
600	3		90	6						100
700			94	6					8	92
800			92	8			2		2	96
900			88	12			7		5	88
1000			91	9			7			93

Table 3.7: The probability of reconstructing a given particle combination from a 1 GeV dark matter mediator with various energies decaying to a positron and electron or a muon and antimuon for events which reconstruct inside the acrylic vessel. For each row, 100 events were simulated. The probabilities shown represent the fraction of the events reconstructing as each particle type after these cuts.

single higher energy particle. An example of a highly boosted event is shown in Figure 3.13.

3.11 Energy Resolution

To compute the energy resolution we take the best fit (i.e. we don't assume we know the particle ids), and add together the kinetic energies. This is identical to what will happen in the final analysis. We use the sum of the kinetic energies as opposed to the total energy because in this case a particle misidentification will have less of an impact.

The energy resolution and bias for a 1 GeV dark matter particle decaying to an electron-positron and muon-antimuon pair are shown in Figures 3.14 and 3.16. The energy resolution for both the electron-positron and muon-antimuon pair is approximately 10% for all values of the kinetic energy. Figure 3.15 shows the energy resolution as a function of radius for these same events.

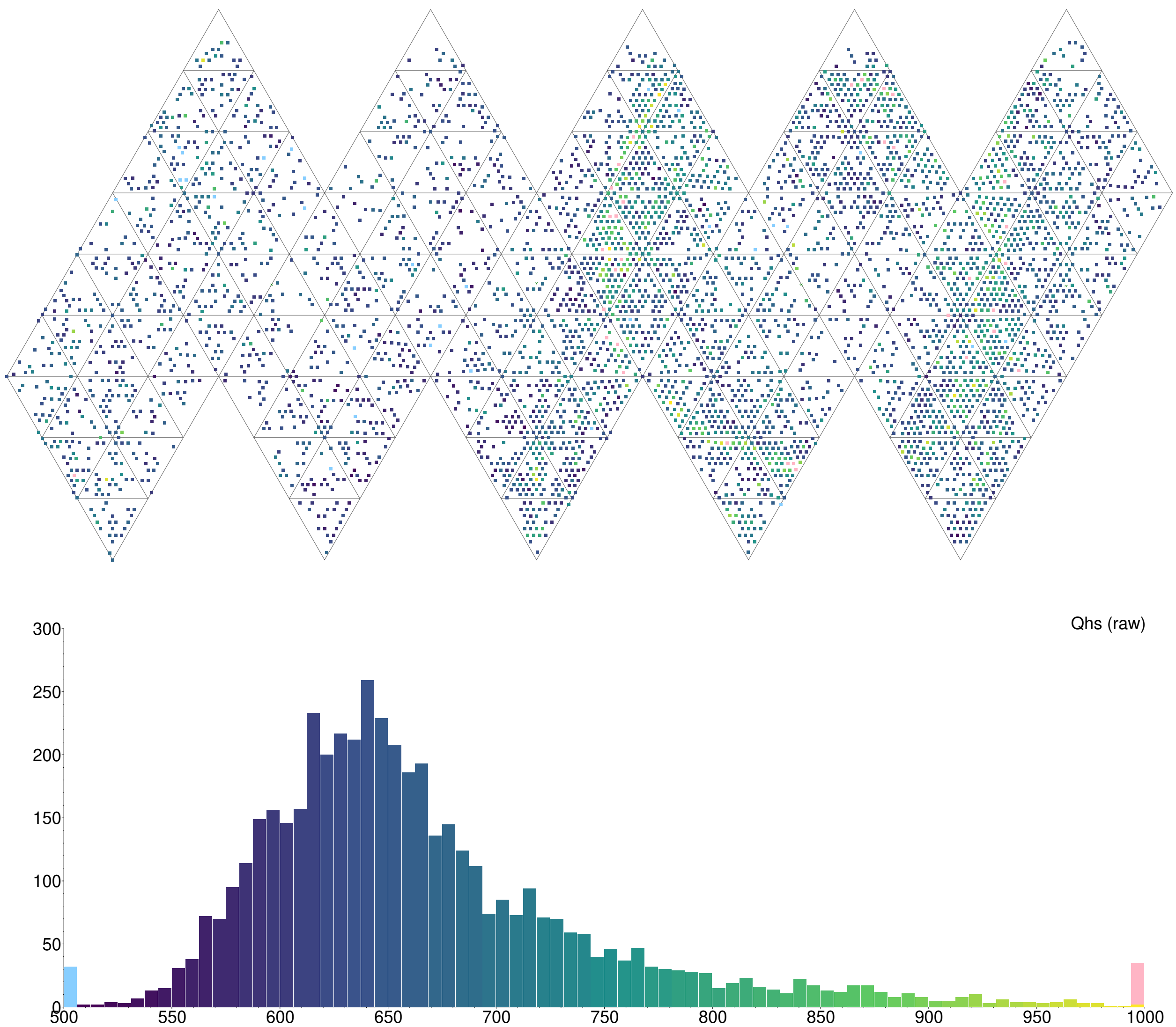

Figure 3.13: XSnoed event display showing a 100 MeV dark matter mediator with a total energy of 1 GeV decaying into an electron positron pair. This particular event was correctly reconstructed as an event with 2 electron like rings.

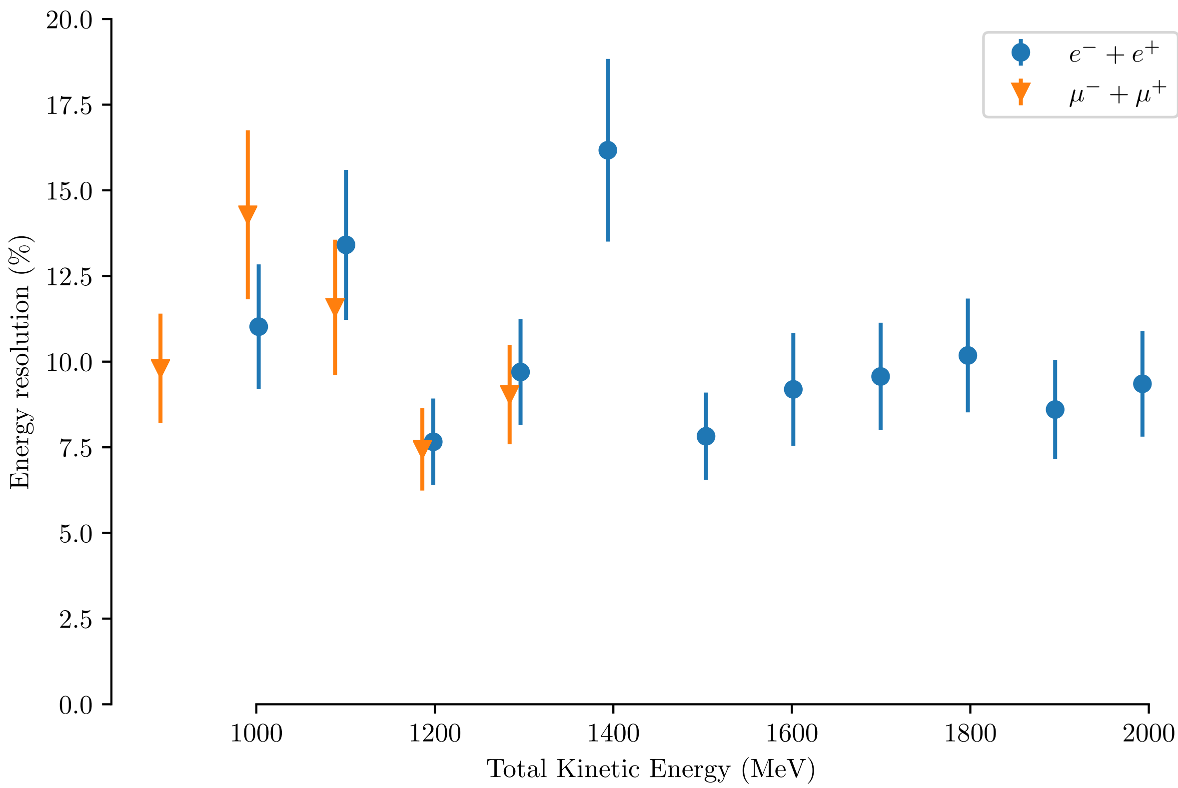

Figure 3.14: Energy resolution as a function of total kinetic energy for a 1 GeV dark matter particle decaying to an electron-positron and muon-antimuon pair. The total kinetic energy here is calculated as the sum of the kinetic energy of the lepton pair. The resolution is plotted as a fraction of the total kinetic energy. The energy resolution is defined here as 1.35 times the interquartile range of the difference between the reconstructed and true kinetic energy. Above approximately 1.4 GeV, the resolution is not shown for the muon-antimuon pair because it is off the scale. Above this energy a large fraction of the muons start to escape the detector at which point it is not possible to accurately reconstruct the energy.

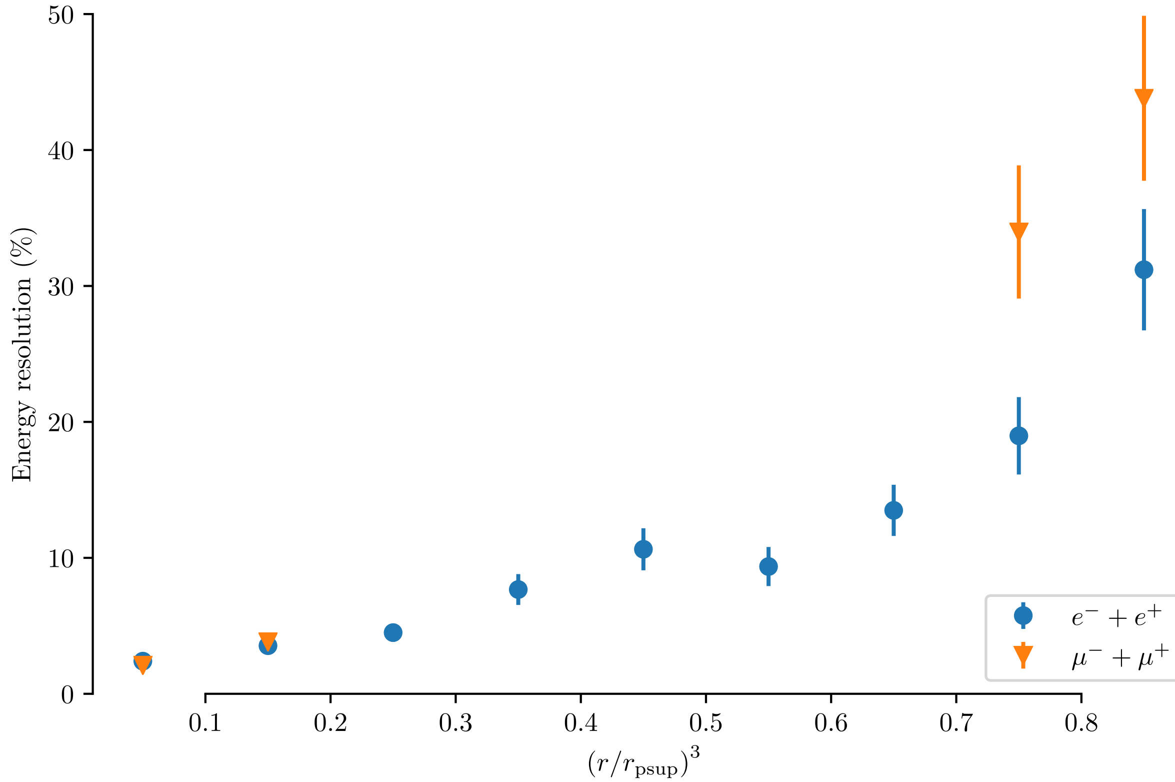

Figure 3.15: Energy resolution as a function of reconstructed radius for a 1 GeV dark matter particle decaying to an electron-positron and muon-antimuon pair. The total kinetic energy here is calculated as the sum of the kinetic energy of the lepton pair. The resolution is plotted as a fraction of the total kinetic energy. The energy resolution is defined here as 1.35 times the interquartile range of the difference between the reconstructed and true kinetic energy.

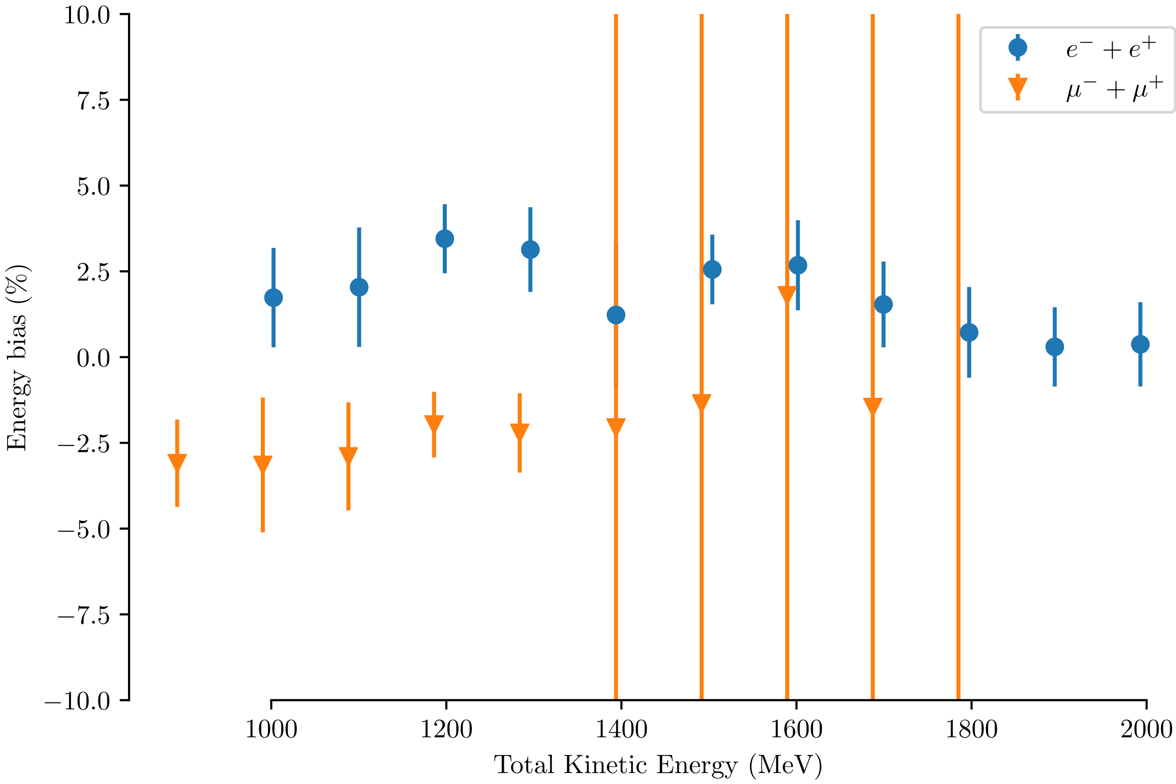

Figure 3.16: Energy bias as a function of the total kinetic energy for a 1 GeV dark matter particle decaying to an electron-positron and muon-antimuon pair. The total kinetic energy here is calculated as the sum of the kinetic energy of the lepton pair. The bias is plotted as a fraction of the total kinetic energy. The bias is calculated by finding the median of the difference between the reconstructed and true kinetic energy. The error on the bias is off the scale for the muon-antimuon pair starting at approximately 1.4 GeV. Above this energy a large fraction of the muons start to escape the detector at which point it is not possible to accurately reconstruct the energy.

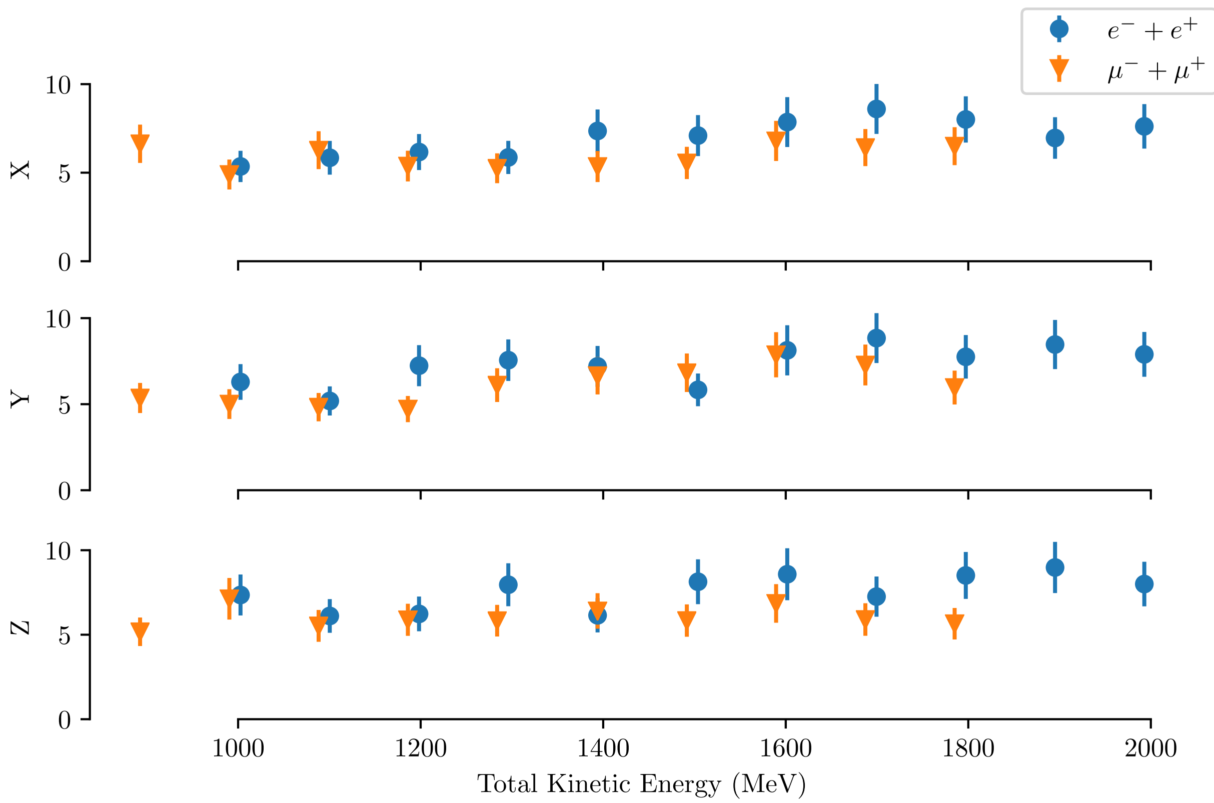

Figure 3.17: Position resolution in cm as a function of the total kinetic energy for a 1 GeV dark matter particle decaying to an electron-positron and muon-antimuon pair. The position resolution is defined here as 1.35 times the interquartile range of the difference between the reconstructed and true positions.

3.12 Position Resolution

Figures 3.17 and 3.18 show the position resolution and bias for a 1 GeV dark matter particle decaying to an electron-positron and muon-antimuon pair. The bias and resolutions are characterized in the same way as for energy by the median and interquartile range divided by 1.35. The position resolution is approximately 10 cm across the entire energy range for both pairs which is similar to the results in the single particle case.

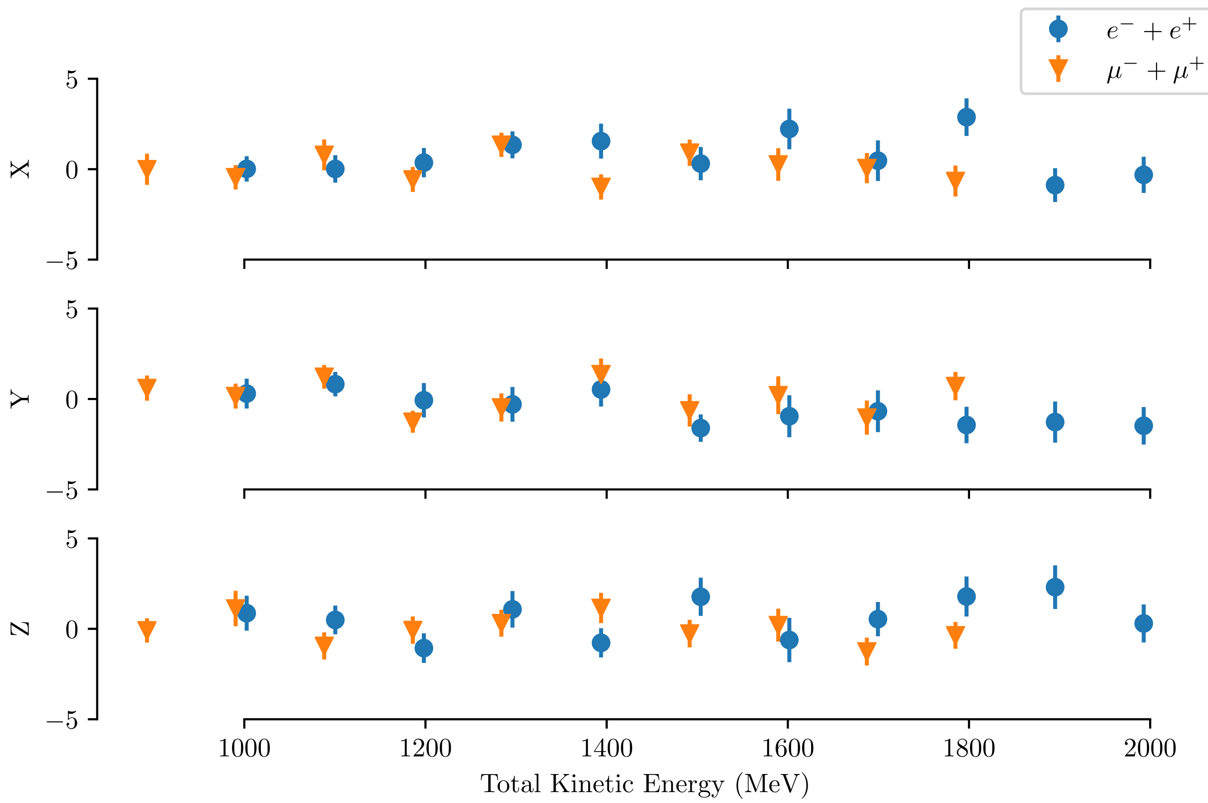

Figure 3.18: Position bias in cm as a function of the total kinetic energy for a 1 GeV dark matter particle decaying to an electron-positron and muon-antimuon pair.The bias is calculated by finding the median of the difference between the reconstructed and true position.

CHAPTER 4

ATMOSPHERIC NEUTRINOS

Atmospheric neutrino interactions represent the primary irreducible background in this analysis. Atmospheric neutrinos are produced from high energy interactions between cosmic particles and nuclei in the upper atmosphere. These interactions produce lots of pions that subsequently decay to a muon neutrino and a muon. Many of these muons then decay to an electron and a muon neutrino and an electron neutrino. The energy of these neutrinos can be anywhere from 10 MeV to 100s of GeV, and so they cover the full energy range considered in this analysis.

Atmospheric neutrinos interact with the water, heavy water, and acrylic in the SNO detector in a variety of different ways. For energies less than approximately 2 GeV, the primary mode of interaction is charged current quasi-elastic scattering[12]:

$$\overline{\nu}_e + p \rightarrow n + e^+ \tag{4.1}$$
$$\overline{\nu}_\mu + p \rightarrow n + \mu^+ \tag{4.2}$$
$$\nu_e + n \rightarrow p + e^- \tag{4.3}$$
$$\nu_\mu + n \rightarrow p + \mu^-. \tag{4.4}$$

In addition to these relatively simple processes, it is also possible to produce pions in an interaction called resonant single pion production. With sufficient energy, the neutrino interaction can leave the nucleon in an excited state which then decays back to a proton or neutron plus a pion[12]. For example:

$$\nu_\mu + p \rightarrow \mu^- + p + \pi^+ + \gamma. \tag{4.5}$$

Atmospheric neutrinos can also interact via a neutral current process:

Fraction of Events (%)	Interaction
5.0	$\nu_\mu + n \rightarrow \mu^- + p$
4.7	$\nu_e + n \rightarrow e^- + p$
3.3	$\nu_\mu + n \rightarrow \mu^- + \gamma + p$
3.1	$\nu_e + n \rightarrow e^- + \gamma + p$
1.9	$\overline{\nu}_\mu + p \rightarrow \mu^+ + n$
1.7	$\overline{\nu}_e + p \rightarrow e^+ + n$
1.3	$\nu_\tau + p \rightarrow \nu_\tau + p$
1.3	$\nu_\mu + p \rightarrow \nu_\mu + p$
1.3	$\overline{\nu}_\mu + p \rightarrow \mu^+ + \gamma + n$
1.2	$\nu_\mu + p \rightarrow \mu^- + p + \pi^+$

Table 4.1: List of the 10 most common atmospheric neutrino interactions in the SNO D2O and salt phases. Note that this is not a list of the 10 most common *visible* interactions, and many of these will not produce a visible signal in the detector.

$$\overline{\nu}_x + p \rightarrow \overline{\nu}_x + p \tag{4.6}$$
$$\overline{\nu}_x + n \rightarrow \overline{\nu}_x + n \tag{4.7}$$
$$\nu_x + p \rightarrow \nu_x + p \tag{4.8}$$
$$\nu_x + n \rightarrow \nu_x + n. \tag{4.9}$$

Neutrinos can also interact coherently with the entire nucleus via either the charged current or neutral current interaction to produce a final pion. These interactions produce pions scattered in the forward direction with no nuclear recoil[12].

In addition to these interactions there are a handful of more exotic interactions producing multiple pions and kaons. For more information see Reference [12]. Table 4.1 shows the 10 most common atmospheric neutrino interactions expected in the SNO D2O and salt phases above 100 MeV as predicted by the simulation discussed in the following sections.

4.1 Simulating Atmospheric Neutrino Events

The atmospheric neutrino flux prediction comes from two different sources: from 10 MeV to 100 MeV we use the fluxes provided by Battistoni et al., and from 100 MeV to 10 GeV

Name	Value
$\sin^2\theta_{23}$	0.512
$\sin^2\theta_{13}$	0.0218
$\sin^2\theta_{12}$	0.307
Δm^2_{21}	$7.53 \times 10^{-5}\,\mathrm{eV}^2$
Δm^2_{32}	$2.444 \times 10^{-3}\,\mathrm{eV}^2$
δ_{CP}	0

Table 4.2: Neutrino oscillation parameters used to oscillate the atmospheric neutrino flux.

we use the fluxes provided by Giles Barr on his website, which were created using a three dimensional calculation using the TARGET2.1 generator[13, 14]. The Barr fluxes are given as a function of both energy and cosine zenith angle whereas the Battistoni fluxes are only given as a function of energy. Therefore, I assumed that the angular distribution of the flux below 100 MeV was given by that of the lowest energy bin in the Barr fluxes.

The atmospheric fluxes are then oscillated using the output from the nuCraft package[15**?**] with the oscillation parameters shown in Table 4.2. The oscillation probabilities for an atmospheric muon neutrino to oscillate to an electron, muon, or tau neutrino are shown in Figures 4.2, 4.3, and 4.4. Figures 4.1 and 4.5 show the initial and oscillated atmospheric fluxes.

We then simulate the initial products of the atmospheric neutrino interaction by passing these fluxes and a simplified SNO geometry created by Andy Mastbaum to the GENIE generator package[17]. The expected event rate from the output of GENIE is shown in Tables 4.3 and 4.4 for the D2O and salt phases respectively. The output from GENIE is then converted to the MCPL file format and the interaction products are simulated using the SNO detector simulation program SNOMAN. An example event is shown using the XSnoed event display in Figure 4.6[18]. These events are then reconstructed on the Open Science Grid using the reconstruction algorithm described in Chapter 3[1, 2].

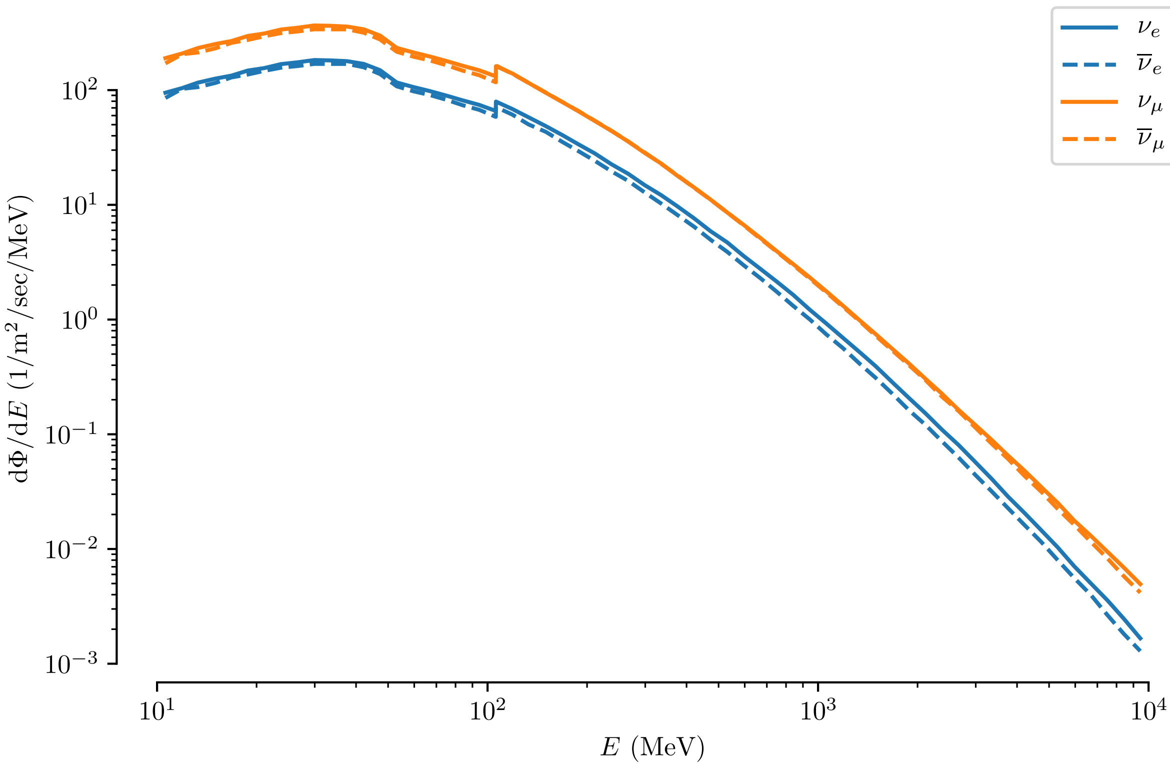

Figure 4.1: Unoscillated atmospheric neutrino fluxes for electron and muon neutrinos as a function of energy. The discontinuity at 100 MeV comes from the fact that below 100 MeV we use the FLUKA calculations and above 100 MeV we use 3D atmospheric neutrino flux calculations done using the IRC01 primary flux. We currently take the fluxes for the solar maximum since the majority of the data was taken between 1999 and 2003 which is close to the solar maximum in 2001[16], and treat the difference as a systematic error.

	AV			PSUP		
	CC	NC	Total	CC	NC	Total
ν_e	56.1	20.2	76.3	137.1	51.0	188.0
$\overline{\nu}_e$	13.9	8.1	22.0	39.6	20.1	59.6
ν_μ	65.1	24.7	89.9	159.8	65.8	225.5
$\overline{\nu}_\mu$	19.7	10.9	30.6	53.6	29.0	82.6
ν_τ	0.2	13.7	13.9	0.6	35.5	36.1
$\overline{\nu}_\tau$	0.1	5.8	5.9	0.3	15.2	15.5

Table 4.3: Expected event rates per year in the D2O phase.

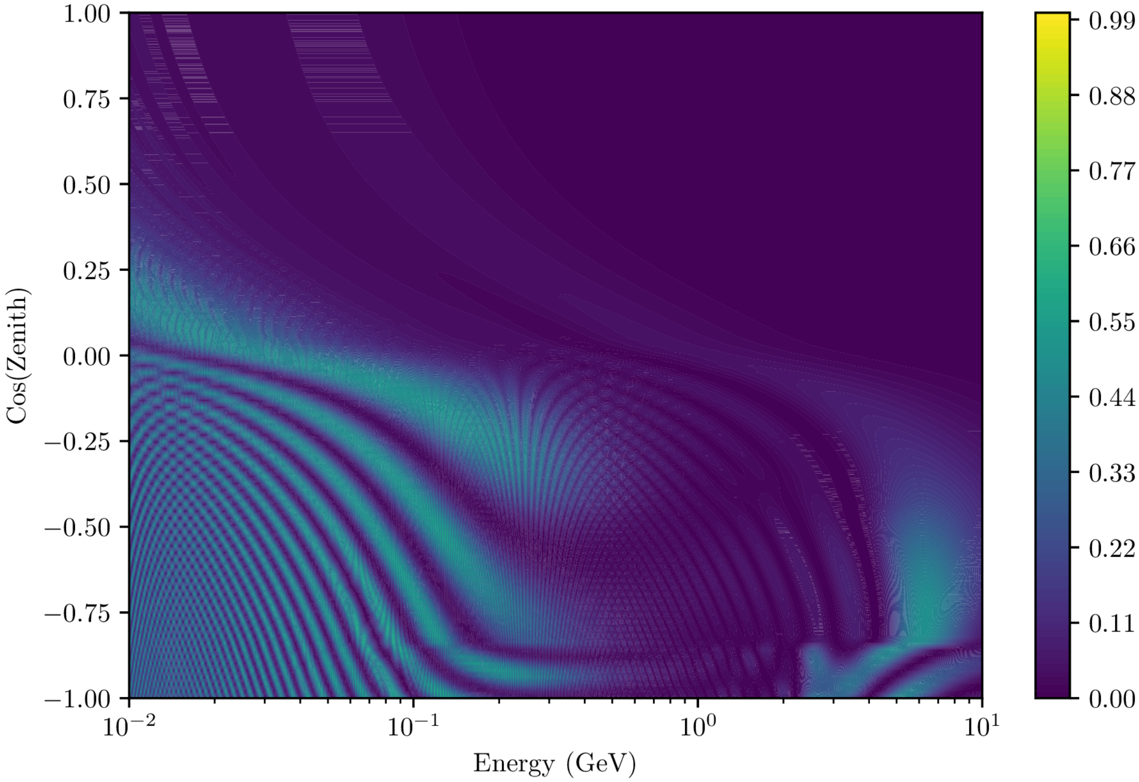

Figure 4.2: Probability for a ν_μ to oscillate into a ν_e as a function of energy and zenith angle.

	AV			PSUP		
	CC	NC	Total	CC	NC	Total
ν_e	54.5	19.8	74.3	136.3	50.2	186.6
$\overline{\nu}_e$	14.2	7.4	21.6	39.8	19.6	59.3
ν_μ	63.1	25.2	88.3	157.7	66.9	224.5
$\overline{\nu}_\mu$	18.4	10.9	29.3	52.6	28.8	81.3
ν_τ	0.2	13.7	13.9	0.6	35.3	35.9
$\overline{\nu}_\tau$	0.1	6.2	6.3	0.3	16.1	16.4

Table 4.4: Expected event rates per year in the salt phase.

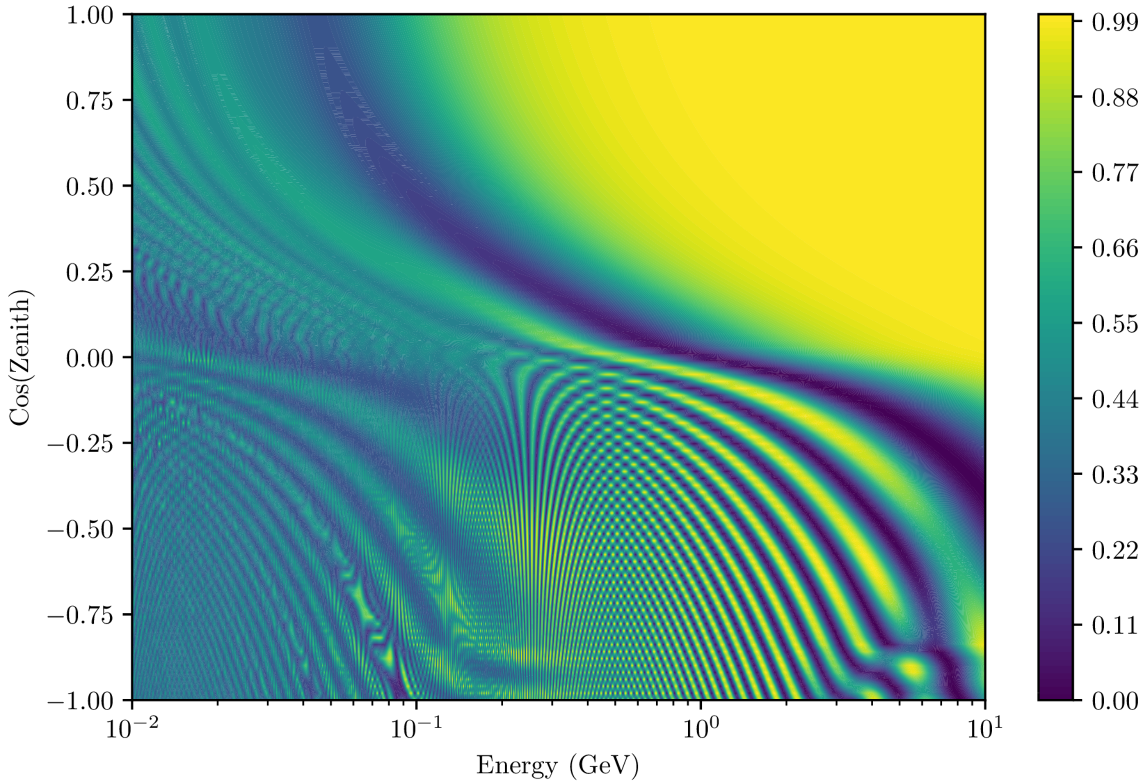

Figure 4.3: Probability for a ν_μ to stay as a ν_μ as a function of energy and zenith angle.

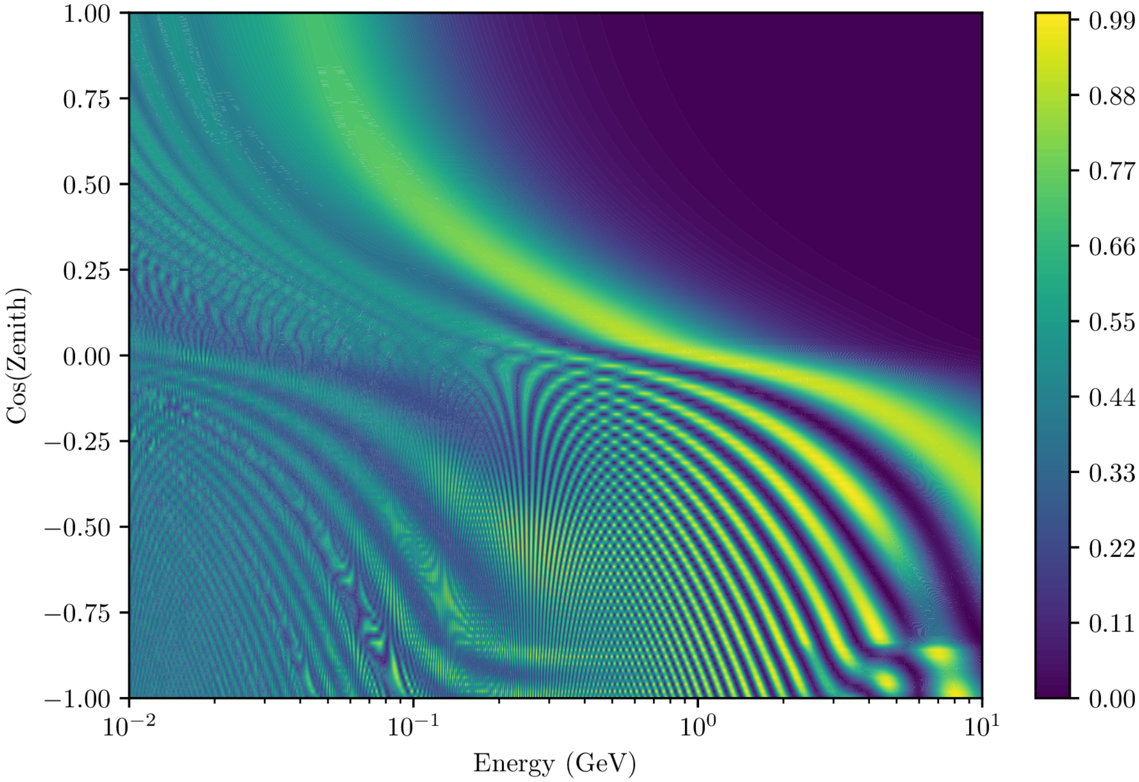

Figure 4.4: Probability for a ν_μ to oscillate into a ν_τ as a function of energy and zenith angle.

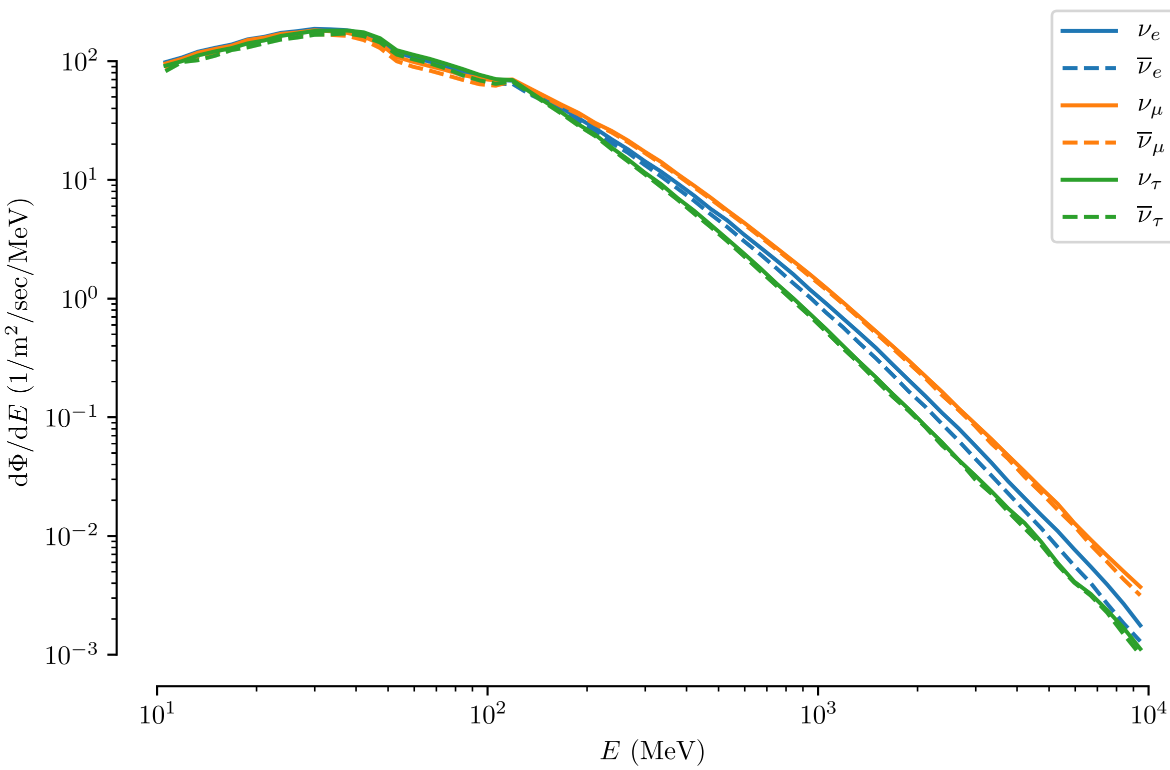

Figure 4.5: Oscillated atmospheric neutrino fluxes as a function of energy. The discontinuity at 100 MeV comes from the fact that below 100 MeV we use the FLUKA calculations and above 100 MeV we use 3D atmospheric neutrino flux calculations done using the IRC01 primary flux.

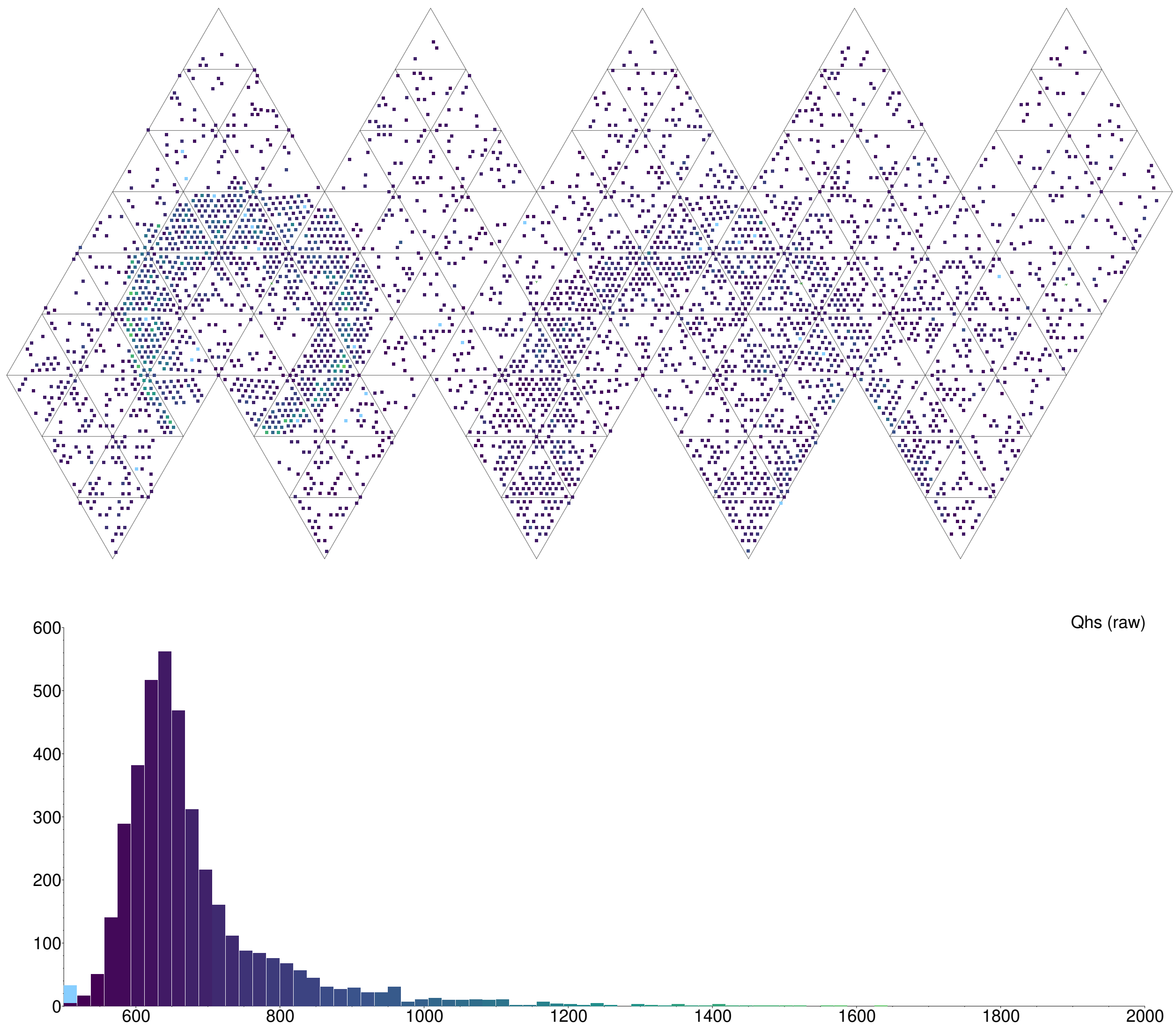

Figure 4.6: XSnoed event display showing an atmospheric neutrino event. This particular interaction is $\nu_\mu + n \rightarrow \mu^- + \gamma + p + p + n + \pi^0$. The ring on the left is the muon and the ring on the right is the π^0.

4.2 Cuts to Reduce the Atmospheric Background

We have two main handles to reduce the atmospheric neutrino background: the number of particles and neutron followers. In the dark matter candidate events of interest, all events have at least 2 leptons, whereas a sizable fraction of the atmospheric events have either a single outgoing particle or only a single charged particle above the Čerenkov threshold. Therefore by cutting events that reconstruct as a single particle, we reduce the atmospheric background by around 70%. The second handle is that many atmospheric neutrino interactions result in a free neutron. These neutrons will eventually capture on either deuterium or chlorine in the salt phase. The nucleus that captured the neutron will then emit gammas as it de-excites to the ground state, thus producing a measurable signal in the detector. By tagging events with following events that look like neutrons we can reduce the background further by a factor of approximately 33% in the D2O phase and 47% in the salt phase. Events tagged with a neutron follower also provide an important side channel to double check the atmospheric neutrino Monte Carlo. For example, if the null hypothesis test shows that the data are not consistent with the expected atmospheric neutrino background one may naturally suspect that the atmospheric neutrino Monte Carlo is wrong in some way. With this sideband we are able to test this hypothesis and if the events tagged with a neutron follower match the Monte Carlo, then this suggests that the Monte Carlo is correct. Figure 4.7 shows the expected number of atmospheric background events per year both before and after these cuts.

4.3 Systematic Uncertainties

The systematic uncertainties associated with the initial neutrino cross section, hadronization and resonance decay, and final state interactions in GENIE are handled using an event reweighting framework within GENIE. To use the reweighting framework, we use a script

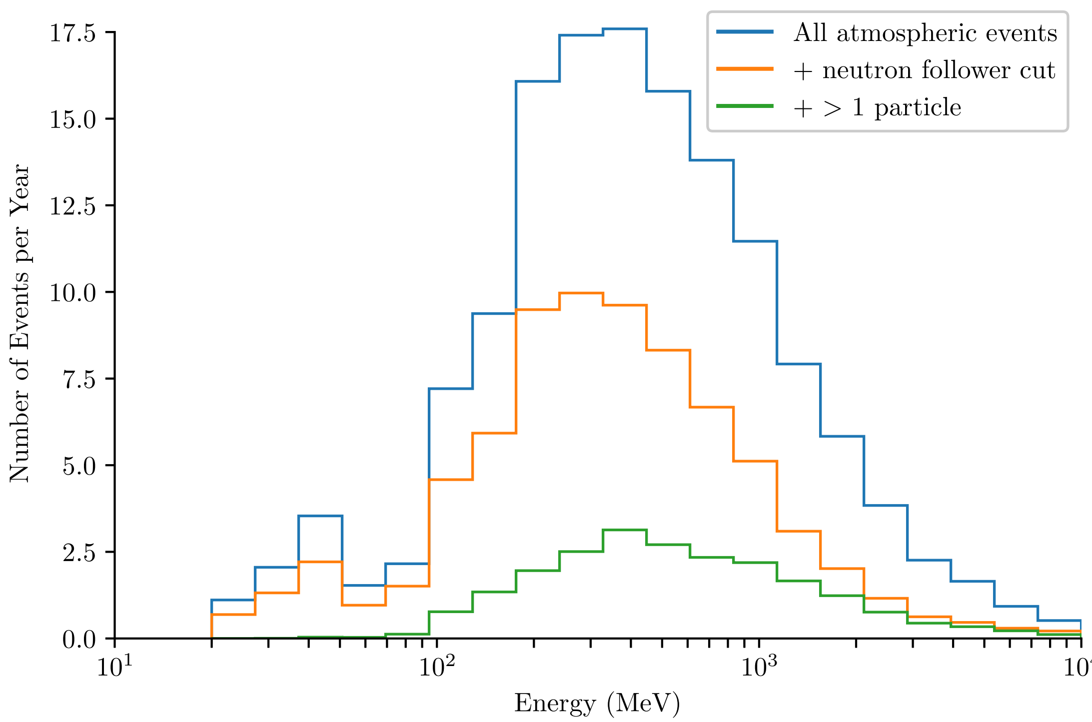

Figure 4.7: Expected number of atmospheric events in the AV per year as a function of the total reconstructed kinetic energy based on the GENIE Monte Carlo. The blue line shows the total number of atmospheric events before any cuts, the orange line after the neutron follower cut, and the green line after the cut requiring at least two particles.

which samples values for the GENIE parameters shown in Table 4.5 according to the known uncertainties on each parameter. With these parameters selected, the GENIE framework then returns a weight associated with each event which is used to weight that event when constructing histograms. The procedure for incorporating these weights into the final analysis is discussed in Section 9.1.3.

Table 4.5 contains nearly all the available GENIE parameters except for the following:

1. The CCQE vector form factor and AGKY x_F and p_T for performance reasons.

2. The reweighting of the CCQE p distribution to a spectral function model since it is not implemented for our target nuclei.

GENIE Parameter	1σ Uncertainty	Description
MaNCEL	25%	NC Elastic axial mass
EtaNCEL	30%	NC Elastic strange FF eta
MaCCQE	+25% -15%	QE axial mass
MaCCRES	20%	CC resonance axial mass
MvCCRES	10%	CC resonance vector mass
MaNCRES	20%	NC resonance axial mass
MvNCRES	10%	NC resonance vector mass
MaCOHpi	40%	Coherent pion prod. axial mass
R0COHpi	10%	Coherent pion prod. nuclear size
NonRESBGvpCC1pi	50%	Non-resonance background in νp CC1π reactions
NonRESBGvpCC2pi	50%	Non-resonance background in νp CC2π reactions
NonRESBGvpNC1pi	50%	Non-resonance background in νp NC1π reactions
NonRESBGvpNC2pi	50%	Non-resonance background in νp NC2π reactions
NonRESBGvnCC1pi	50%	Non-resonance background in νn CC1π reactions
NonRESBGvnCC2pi	50%	Non-resonance background in νn CC2π reactions
NonRESBGvnNC1pi	50%	Non-resonance background in νn NC1π reactions
NonRESBGvnNC2pi	50%	Non-resonance background in νn NC2π reactions
NonRESBGvbarpCC1pi	50%	Non-resonance background in $\overline{\nu}p$ CC1π reactions
NonRESBGvbarpCC2pi	50%	Non-resonance background in $\overline{\nu}p$ CC2π reactions
NonRESBGvbarpNC1pi	50%	Non-resonance background in $\overline{\nu}p$ NC1π reactions
NonRESBGvbarpNC2pi	50%	Non-resonance background in $\overline{\nu}p$ NC2π reactions
NonRESBGvbarnCC1pi	50%	Non-resonance background in $\overline{\nu}n$ CC1π reactions
NonRESBGvbarnCC2pi	50%	Non-resonance background in $\overline{\nu}n$ CC2π reactions
NonRESBGvbarnNC1pi	50%	Non-resonance background in $\overline{\nu}n$ NC1π reactions
NonRESBGvbarnNC2pi	50%	Non-resonance background in $\overline{\nu}n$ NC2π reactions
BR1gamma	50%	Resonance decays, radiative decay BR
BR1eta	50%	Resonance decays, single η BR
Theta_Delta2Npi	-	Resonance decays, angular distribution
AhtBY	25%	DIS form factor, high-twist BY scaling
BhtBY	25%	DIS form factor, high-twist BY scaling
CV1uBY	30%	DIS form factor, GRV98 PDF correction
CV2uBY	40%	DIS form factor, GRV98 PDF correction
DISNuclMod	100%	DIS nuclear model modification
FormZone	50%	Hadron formation zone
CCQEPauliSupViaKF	30%	Fermi Gas Model, Pauli suppression k_F
MFP_N	20%	Intranuke FSI model, nucleon mean free path
FrCEx_N	50%	Intranuke FSI model, nucleon charge exchange probability
FrElas_N	30%	Intranuke FSI model, nucleon elastic reaction probability
FrInel_N	40%	Intranuke FSI model, nucleon inelastic reaction probability
FrAbs_N	20%	Intranuke FSI model, nucleon absorption probability
MFP_pi	20%	Intranuke FSI model, π mean free path
FrCEx_pi	50%	Intranuke FSI model, π charge exchange probability
FrElas_pi	10%	Intranuke FSI model, π elastic reaction probability
FrInel_pi	40%	Intranuke FSI model, π inelastic reaction probability
FrAbs_pi	30%	Intranuke FSI model, π absorption probability
CCQEMomDistroFGtoSF	-	Reweights incoming nucleon momentum distribution from Fermi Gas (Bodek-Ritchie) to a spectral function[19]

Table 4.5: Parameters in the GENIE cross section model which varied to account for systematic uncertainties in the model[20]. The Theta_Delta2Npi and CCQEMomDistroFGtoSF parameters don't have a fractional uncertainty and are instead varied uniformly between 0 and 1.

CHAPTER 5

INSTRUMENTAL AND EXTERNAL BACKGROUNDS

Besides atmospheric neutrinos, instrumental backgrounds and external muons represent the only other source of backgrounds for this analysis. The instrumental backgrounds represent a serious problem since the rate of these events is orders of magnitude larger than events coming from atmospheric neutrinos or a potential dark matter signal. Luckily, these backgrounds were studied extensively during SNO and many data cleaning cuts were designed to reject these events. Based on these cuts, I have developed several data cleaning cuts more suitable for the high energy range used in this analysis.

In the following sections we describe the instrumental and external muon backgrounds and the data cleaning cuts designed to reject them.

5.1 External Muons

Both cosmic ray muons and muons created from atmospheric neutrinos interacting in the surrounding rock present a background for this analysis. In both cases, it is necessary to cut events that start *outside* the PSUP and enter the detector. A typical muon is shown in Figure 5.1.

During SNO, these events were cut using the MUON cut which tagged events with at least 150 hits, 5 or more outward-looking (OWL) PMT hits, and with a time RMS of less than 90 nanoseconds. This cut would have a negligible loss (referred to as sacrifice) for any contained atmospheric or dark matter candidate events, but could potentially cut events which produce an energetic muon which then exits the detector. Therefore, I have slightly modified this cut to *also* require that at least 1 OWL tube is both early and has a high charge relative to the nearby normal PMTs. We define an early and high charge tube by

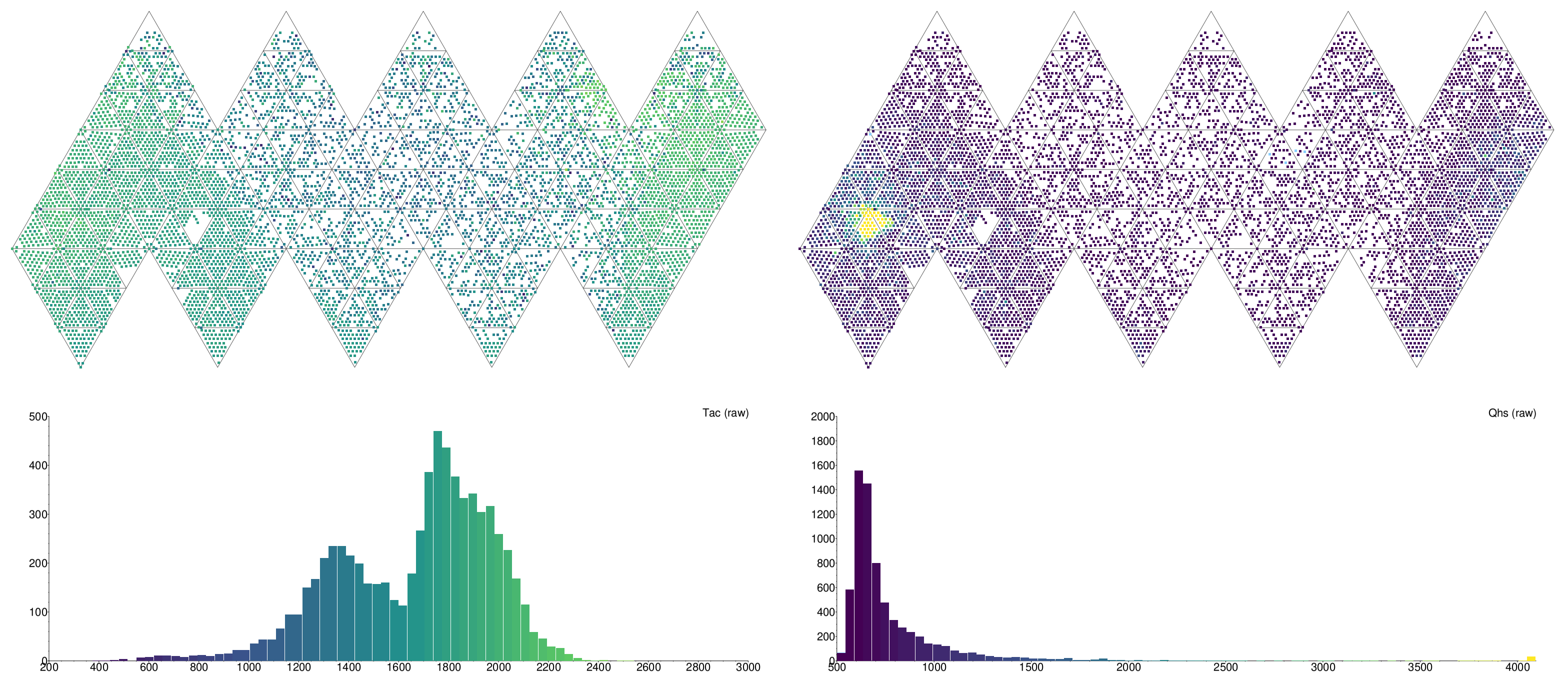

Figure 5.1: XSnoed event display showing a typical muon. The left image shows the PMT hits colored according to the TAC value and the right image according to the QHS value.

creating an array of the ECA[1] calibrated hit times (we can't use PCA calibrated times since the OWL tubes were never calibrated via PCA) and of the best uncalibrated charge (see Appendix J) for all normal PMTs within 3 meters of each hit OWL PMT. We then compute the median charge and time for these normal PMTs. We then compute how many OWL PMT hits are *both* earlier than the median normal PMT time and have a higher charge than the surrounding PMTs. If at least 1 OWL PMT hit satisfies this criteria and all the other criteria from the SNO MUON cut are satisfied (except the time RMS part) then it's tagged as a muon.

5.2 Noise Events

Noise events refer to events triggered by sources that do not actually create light in the detector. The two most common sources are "ringing" after large events and electrical pickup on deck. An example of a typical electrical pickup event is shown in Figure 5.2.

These events are tagged by the QvNHIT and ITC[2] cuts which are identical to their SNO counterparts aside from minor updates[3].

---

1. ECA stands for Electronic Calibration Analysis. The ECA calibrations generate pedestal values for the charge and time slope calibrations. These constants are then used to subtract off the baseline values for the charge and convert the raw time values into a time in nanoseconds.

2. ITC stands for In Time Channel and cuts on the number of PMT hits in a 93 ns sliding window.

3. The original ITC cut used during SNO used the fully calibrated hit time for each PMT. In this analysis, the ITC cut uses the pt1 time which is the time without the charge walk calibration. We use this time since otherwise the cut may fail to tag an event which consists of mostly electronics noise which has charge too low to apply PCA (Javi, personal communication, June 12, 2019). Similarly, the original SNO QvNHIT cut only looked at channels which had good calibrations (i.e. the calibration processor was able to apply the charge walk calibration) whereas the QvNHIT cut used in this analysis does not require good calibrations. The reason is the same as for the ITC cut; the cut may fail to tag electronic pickup events in which all the channels have a charge too low to apply PCA. In addition, the SNO version of the QvNHIT cut tagged events in which the charge to Nhit ratio was less than 0.25 whereas I use 0.5 in my cut. The reason for this is that after investigating the SNOMAN code, I discovered that all the charges were accidentally divided by 2.

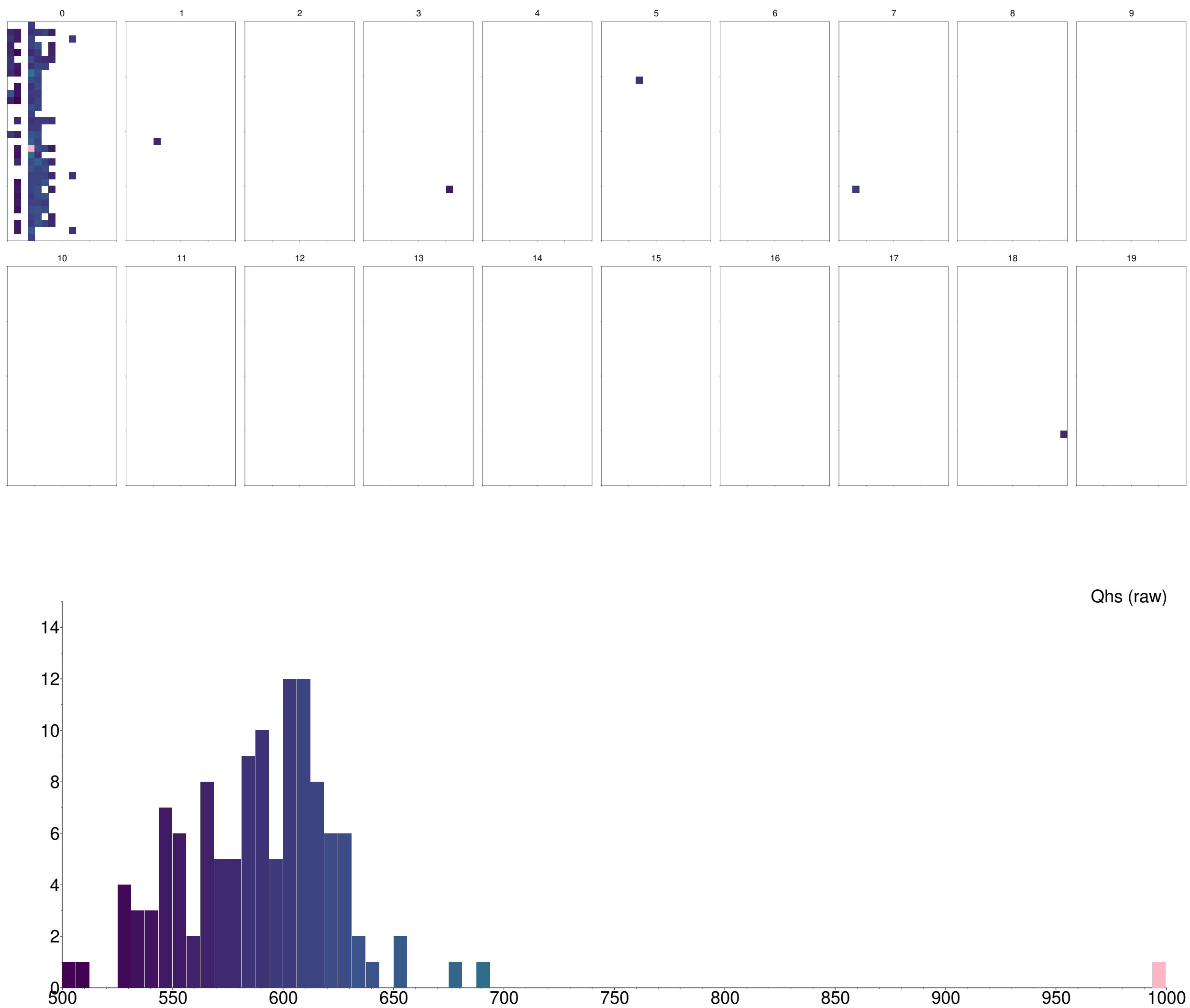

Figure 5.2: XSnoed event display showing a typical noise event. The plot shows a histogram of the QHS values which are a measure of the charge in each PMT before calibrations. This event was likely caused by pickup near crate 0.

5.3 Neck Events

Neck events are caused by light produced in or leaking through the glove box on top of the detector[21]. An example neck event is shown in Figure 5.3. The original SNO neck event cut is defined as:

> This cuts events containing neck tubes. It requires that either both tubes in the neck fire, or that one of those tubes fires and it has a high charge and is early. High charge is defined by a neck tube having a pedestal subtracted charge greater than 70 or less than -110. Early is defined by the neck tube having an ECA time 70 ns or more before the average ECA time of the PSUP PMTS with z less than 0. After the cable changes to the neck tubes this time difference changes to 15 ns.

Similarly to the MUON cut, I've used the SNO criteria but added an additional requirement to avoid tagging high energy upwards going events. The neck cut I use also has a requirement that 50% of the hit PMTs must have a z coordinate of less than 4.25 meters *or* 50% of the ECA calibrated QHS charge must be below 4.25 meters.

5.4 Flashers

Flashers are probably the most difficult and common source of instrumental background for this analysis. They occur at a rate of approximately 100 events per hour which is orders of magnitude greater than events from atmospheric neutrinos. A flasher event occurs when there is an electrical short in the PMT base or dynode stack which causes light to be emitted from the PMT and hit the opposite side of the detector[citation needed]. Because this event is caused by actual light in the detector, it is particularly hard to cut while also maintaining a small signal sacrifice. A typical flasher event is shown in Figure 5.4.

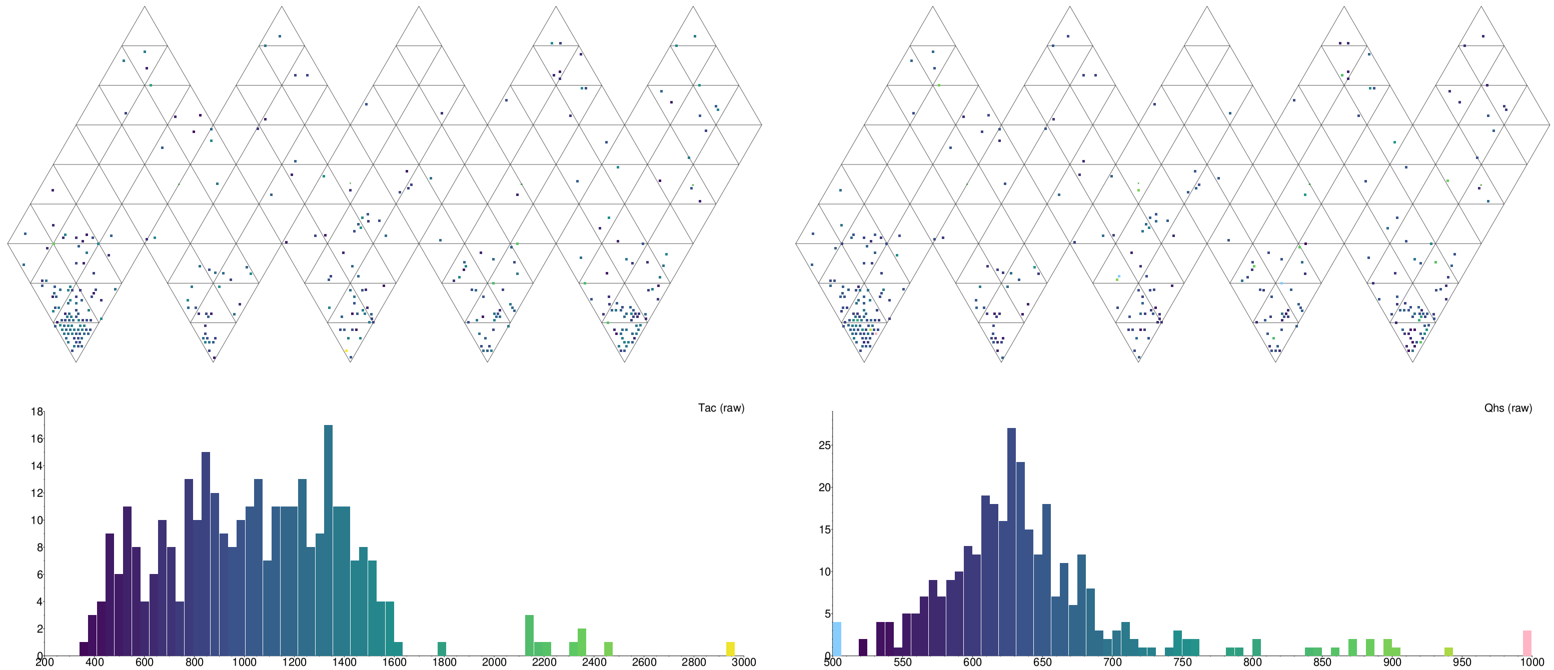

Figure 5.3: XSnoed event display showing a typical neck event. The left image shows the PMT hits colored according to the TAC value and the right image according to the QHS value.

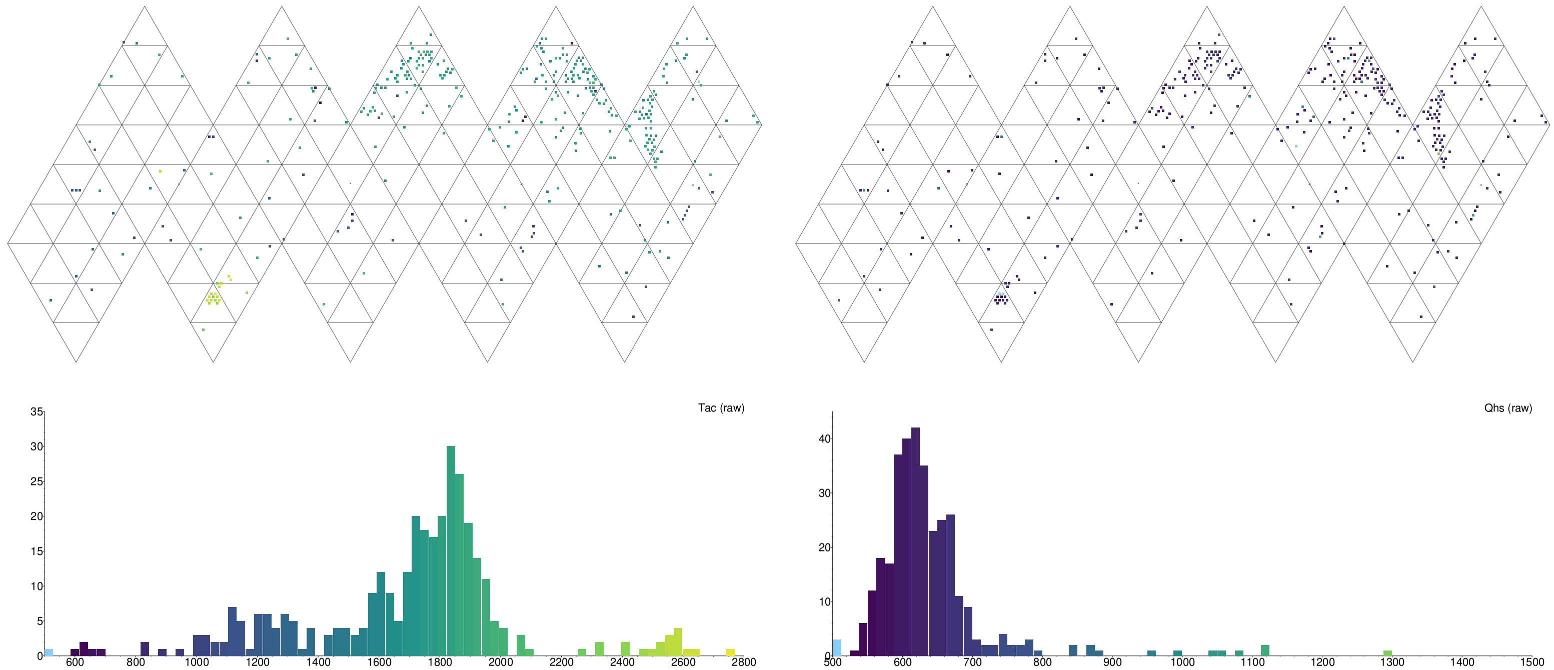

Figure 5.4: XSnoed event display showing a typical flasher. The left image shows the PMT hits colored according to the TAC value and the right image according to the QHS value.

The algorithm used to cut flashers is similar to two different cuts used during SNO: the flasher geometry cut and the QvT[4] cut, however it adds additional criteria in order to make the cut more robust against tagging high energy physics events. The full algorithm is described in pseudocode in Algorithm 1.

5.5 Breakdowns

Wet end breakdowns (WEBs for short) are believed to be caused by an arc somewhere in the PMT base circuitry which causes repeated large flashes of light in the detector. These calamitous events occasionally stop on their own, in which case they are referred to as "friendly WEBs"; otherwise it is the shift operator who is responsible for ending the run and powering down the PMT which is causing the breakdown. Breakdowns are very similar to flashers except that they produce *much* more light[5].

Since breakdowns often cause many of the electronics to saturate, it is *very* difficult to find a single common characteristic on which to cut. However, the one thing that does seem to be common among almost all breakdowns is that the channels in the same crate as the breakdown trigger on electronic pickup from the channel which is breaking down. The channels which trigger on the electronic pickup come much earlier in the event than the rest of the PMT hits.

Therefore, the breakdown cut tags any event which has at least 1000 PMT hits and in which the crate with the highest median TAC has at least 256 hits and is 500 TAC counts away from the next highest crate (with at least 20 hits).

Occasionally a breakdown is so big that it causes issues with the TAC measurement and many of them end up reading outside of the linear TAC region. Therefore we also tag any

4. The QvT cut, or charge vs time cut, cuts events in which a very high charge PMT hit occurs much earlier than the rest of the PMT hits.

5. In fact, I think there is a continuous spectrum between flashers and breakdowns, but the distinction is still helpful since the ways to tag the two are very different.

Algorithm 1 Flasher Cut Algorithm

```
if nhit < 31 then
   return 0
end if
# This condition is similar to the SNO QvT cut except we require that 70% of the normal hit PMTs be
12 meters from the high charge channel and that 70% of the normal hit PMTs be at least 50 ns after the
high charge channel.
if highest QLX > second highest QLX +80 then
   Collect all hit times from the same slot as the high charge channel and compute the median hit time
   if At least 4 hits in the slot and 70% of the normal hit PMTs with good calibration are more than 12
   meters from the high charge channel and 70% of the normal hit PMTs with good calibration are more
   than 50 ns after the median hit time in the slot then
      return 1
   end if
end if
for All PC with at least 4 hits do
   Collect the QHS, QHL, and QLX charges and the ECA calibrated hit times (EPT) for each PMT in
   the PC sending charge values below 300 to 4095
   t ← median(EPT)
   QHS_1 ← max(QHS)
   QHL_1 ← max(QHL)
   QLX_1 ← max(QLX)
   QHS_2 ← second highest(QHS)
   QHL_2 ← second highest(QHL)
   QLX_2 ← second highest(QLX)
   if QHS_1 > QHS_2 + 1000 then
      if 70% of the normal hit PMTs with good calibration are more than 12 meters from the high charge
      channel and 70% of the normal hit PMTs with good calibration are more than 50 ns after t then
         return 1
      end if
   else if QHL_1 > QHL_2 + 1000 then
      if 70% of the normal hit PMTs with good calibration are more than 12 meters from the high charge
      channel and 70% of the normal hit PMTs with good calibration are more than 50 ns after t then
         return 1
      end if
   else if QLX_1 > QLX_2 + 80 then
      if 70% of the normal hit PMTs with good calibration are more than 12 meters from the high charge
      channel and 70% of the normal hit PMTs with good calibration are more than 50 ns after t then
         return 1
      end if
   else
      for All normal PMT channels not hit in PC do
         if more hits in slot than surrounding 4 meters or median hit time in slot is 10 ns earlier than
         PMTs within 4 meters then
            if 70% of the normal hit PMTs with good calibration are more than 12 meters from the high
            charge channel and 70% of the normal hit PMTs with good calibration are more than 50 ns
            after t then
               return 1
            end if
         end if
      end for
   end if
end for
return 0
```

event in which more than 70% of the PMT hits have a TAC value below 400.

5.6 Additional Data Cleaning Cuts

In addition to the data cleaning cuts previously discussed, I also designed two additional data cleaning cuts that target all of the instrumental events more generally.

5.6.1 Calibrated Nhit Fraction

The first cut is called the "Calibrated Nhit Fraction" cut and cuts any event in which less than 80% of the PMT hits are properly calibrated. The value of 80% was chosen by plotting this fraction for external muons which represent a worst case scenario and placing the cut such that virtually none of them were tagged. Figure 5.5 shows the calibrated Nhit fraction for muons, flashers, and neck events.

5.6.2 Burst Cut

The second cut is called the "Burst Cut" and is intended to cut any burst of high Nhit events. These bursts often come from a burst of neck events or from a breakdown. To define a burst, we first label "Prompt 50" events, which are defined as:

1. Any event with at least 100 PMT hits

2. Where the last PMT event with at least 100 PMT hits was at least 50 ms ago

The idea here is to tag the "prompt" high nhit events without tagging any events which might retrigger based on ringing or late pulsing in the PMTs.

Next, we cut any events in a 1 second sliding window in which there are 3 or more "Prompt 50" events. The sliding window is allowed to be longer than 1 second as long as there is no gap between events with at least 100 PMT hits longer than 1 second.

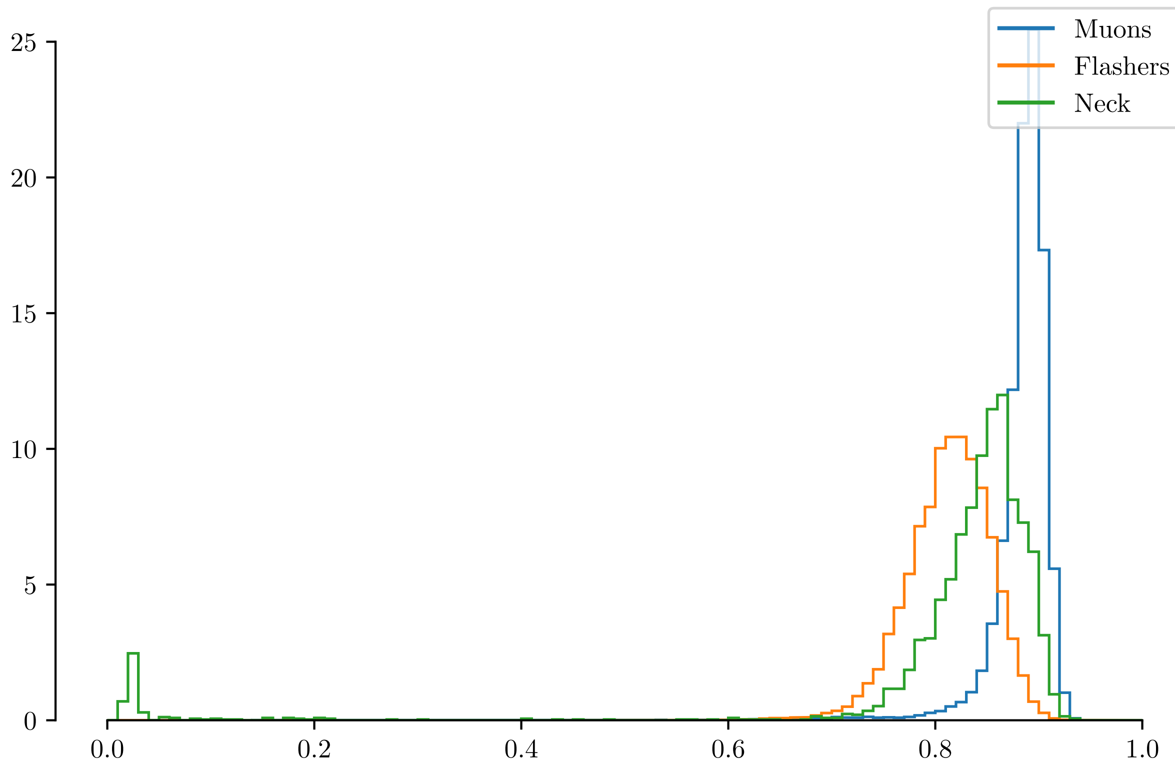

Figure 5.5: Calibrated Nhit fraction for muons, flashers, and neck events.

CHAPTER 6

CUT EFFICIENCIES

There are two important quantities necessary to characterize the performance of the data cleaning cuts described in Chapter 5 and the high-level cuts discussed in Chapter 7: the signal sacrifice and the background contamination. The signal sacrifice refers to the events of interest which are "accidentally" cut, while the background contamination refers to background events which are not cut. Since all of the data cleaning and high level cuts are primarily designed to cut instrumentals and not atmospheric events, for the rest of this chapter I will consider atmospheric neutrino events to be "signal" events even though they ultimately represent a background for the dark matter search.

Section 6.1 describes the sacrifice for the instrumental and muon cuts described in Chapter 5. Finally, in Section 6.2 I describe the new method used to estimate the overall rates and contamination for the backgrounds.

6.1 Sacrifice

We can group the data cleaning cuts into two distinct groups. For the first group, the data that the cut uses can be accurately simulated and so the sacrifice can easily be estimated from Monte Carlo. The second group contains cuts which rely on data which is not simulated, and so we have to try and estimate the sacrifice independent of the Monte Carlo.

The majority of the cuts fall into the first group and we can estimate the sacrifice from Monte Carlo[1]. For these cuts, the sacrifice will automatically be taken into account in the

1. The Monte Carlo does not capture many low level hardware issues that we do expect to appear in real data like cross talk and shark fins (a shark fin is a type of event pathology characterized by a single channel with a high charge where the trigger signal looks like a shark fin). In all SNO analyses therefore the sacrifice was always measured using real data from calibration sources. Unfortunately there are no calibration sources that mimic the events we are interested in for this analysis. We do use Michel electrons and stopping muons to estimate systematic uncertainties on the energy reconstruction, but these events do not represent the broad range of types of events we expect to see from atmospheric neutrino or dark matter events. Therefore, I believe the best estimate of the sacrifice is to be obtained by looking at Monte Carlo.

Cut	# of Events	Fraction (%)	Δ (%)
Total	103892	100.00	100.00
Junk	103892	100.00	100.00
Crate Isotropy	103892	100.00	100.00
QvNHIT	103892	100.00	100.00
Flasher	103861	99.97	99.97
ITC	103561	99.68	99.71
Breakdown	103561	99.68	100.00
Radius cut	37505	36.10	36.22
ψ cut	35753	34.41	95.33

Table 6.1: Atmospheric neutrino sacrifice for the data cleaning and high-level cuts used in the final analysis. The first column shows the cut that is applied while the second and third columns show the number of events and fraction remaining after each cut is successively applied. The fourth column shows the fraction of events cut by each cut individually.

final analysis since the final analysis uses Monte Carlo and applies the same cuts. To get a rough idea for the size of the sacrifice, Table 6.1 shows the number of events remaining after each of the data cleaning and high-level cuts for the atmospheric neutrino Monte Carlo sample. The combined fraction of atmospheric events cut by the data cleaning cuts alone is only 0.8%. The radial cut has the largest sacrifice and cuts approximately 64% of the atmospheric events, which is expected for events uniformly distributed throughout the PSUP. The sacrifice from the goodness-of-fit ψ cut is only 5%, although it would be larger had we extended the fiducial volume outside the AV. This is primarily because outside the AV some of the assumptions made in the likelihood function do not hold. This can be seen in the second row and first column of Figure 6.1.

The cuts which fall into the second group which can't be estimated from Monte Carlo are the muon, neck, calibrated nhit, and burst cut. The muon and neck cut sacrifice cannot be estimated from Monte Carlo because SNOMAN does not simulate either the OWL or neck PMTs (which are the primary input for these cuts). The calibrated nhit cannot be estimated from Monte Carlo since the Monte Carlo does not include all the electronic effects which lead to channels being miscalibrated like cross talk. The burst cut sacrifice cannot be estimated from Monte Carlo because the most likely reason for a physics event to be cut

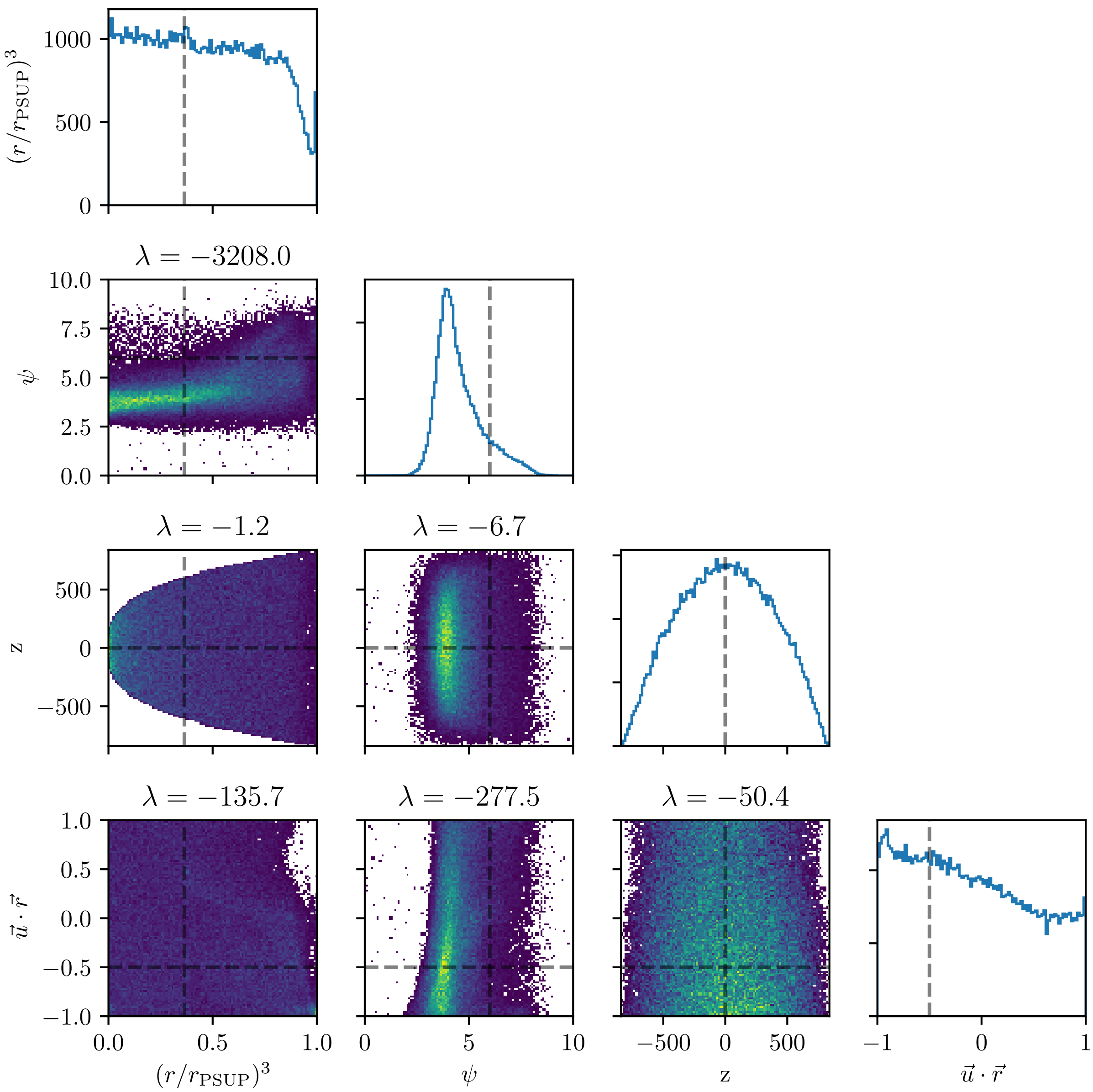

Figure 6.1: Distribution of high-level variables for the atmospheric neutrino Monte Carlo sample. The x axes on the plots represent the reconstructed radius, reconstructed z position, $\vec{u} \cdot \vec{r}$, and the goodness-of-fit parameter ψ. The values of λ above each plot represent a likelihood ratio test to show if the two variables are correlated and is discussed in Appendix F.

by this is due to another instrumental event, like a flasher, preceding it. Since we do not simulate instrumental events, we cannot estimate it from Monte Carlo.

6.1.1 Muon Cut

For the Muon cut, we can try to estimate the sacrifice by writing the sacrifice as:

$$
\begin{aligned}
P(\mathrm{M} \mid \mathrm{signal}) = P(\mathrm{M} \mid \text{fully contained}, \mathrm{signal}) P(\text{fully contained} \mid \mathrm{signal}) \\
+ P(\mathrm{M} \mid \text{not contained}, \mathrm{signal}) P(\text{not contained} \mid \mathrm{signal})
\end{aligned}
\tag{6.1}
$$

which, if we assume the probability of tagging a fully contained event is negligible since the Muon tag requires 5 OWL hits, becomes

$$
P(\mathrm{M} \mid \mathrm{signal}) \simeq P(\mathrm{M} \mid \text{not contained}, \mathrm{signal}) P(\text{not contained} \mid \mathrm{signal}).
$$

It may be possible to estimate the first term from through-going muons and ignoring the OWL PMTs near the entry point and the second term from Monte Carlo. However, in the end we chose to apply the fiducial volume and energy cuts such that we do not expect any signal events in the range of interest to exit the PSUP. Therefore, the second term is zero and this cut should have a negligible sacrifice.

6.1.2 Neck Cut

We can estimate an upper bound for the neck event cut sacrifice by using a slightly modified version of the cut. Instead of requiring that at least 2 neck PMTs get hit or 1 neck PMT gets hit with a high charge and early, we can instead require the same using the three highest normal PMTs instead of the neck PMTs.

The most likely type of physics event we are interested in which might accidentally get tagged by the neck PMT cut would be an event where two charged particles come off back to back near the top of the detector. The light from the particle going upwards could cross the water air interface and cause the neck PMTs to fire. Since a lot of light is required to make it past the water air interface and reflect up the acrylic neck, we would also expect that many of the PMTs in the upper hemisphere would also get hit by light from the particle traveling upward. Therefore, instead of looking at the neck PMT hits (which we can't simulate), we can instead look for hits in the three highest normal PMTs. The number of events which fail the neck cut using these PMTs instead of the neck PMTs should give us an upper bound on the sacrifice (since it is easier to hit the normal PMTs than for light to travel up the neck to hit the neck PMTs). The one other assumption we have to make is that these PMTs act similarly to the neck PMTs. I performed an analysis on the atmospheric neutrino Monte Carlo using this technique, and found that 309 events out of 42390 simulated prompt events failed the cut, or 0.7% of events. This amount is negligible compared to the other uncertainties in the atmospheric analysis and because it is an upper bound it will be ignored for the rest of this analysis.

6.1.3 Burst Cut

For the burst cut, I looked at all events tagged as an external muon by the burst cut as a proxy for signal events. Looking over all the runs in the D2O phase, I found that approximately 30 events were cut. I hand scanned all 30 of these events and found that only 16 were plausible muons (the rest were neck events or breakdown events which accidentally got tagged as a muon). The most likely reason for the muon events to get tagged was that they were preceded or followed by neck or flasher events. The total number of muon events was 19674 which gives a sacrifice of 0.08% which is negligible compared to the other uncertainties in this analysis.

6.1.4 Calibrated Nhit Fraction

I estimated the sacrifice for the Calibrated Nhit Fraction cut by looking at tagged muons which pass the ψ cut. Since these events have a very high density of charge, they represent an extreme case of what we expect for atmospheric neutrino events in the energy range of interest. Looking over a large fraction of the final data, I found that 6 out of 3,560 events tagged as a muon also failed the calibrated nhit fraction cut, or a 0.17% sacrifice. This sacrifice is negligible compared to other uncertainties in this analysis.

6.2 Contamination

The contamination of the instrumental and muon cuts is much more difficult to measure since we don't have any way of accurately simulating the instrumental backgrounds. We can't measure the contamination of the muon data cleaning cut either because although we can simulate external muons in SNOMAN, SNOMAN does not simulate the OWL PMTs.

To properly estimate the contamination from instrumentals and external muons I have developed a new method inspired by the bifurcated analysis used in SNO (the bifurcated analysis method used in SNO and its problems are discussed in Appendix K). This method works by first assuming that we can tag the instrumental events with the low-level cuts to obtain pure samples of each background, i.e. we have some way of tagging noise, neck, flasher, breakdown, and muon events to obtain pure samples of each[2]. Using the pure samples, we can determine what the distribution of reconstructed variables like radius and goodness of fit look like for these backgrounds, which will allow us to infer how many of them are in the untagged sample. To measure the number of events in the untagged sample, we construct a likelihood function which looks at the distribution of these high-level variables in the events

2. This step isn't strictly necessary for the method to work, however, if you do not assume something like this the number of terms in the likelihood function quickly becomes unmanageable since you have to consider every single possible combination of data cleaning cut and background.

that were *not* tagged by a low-level cut and is able to measure the residual contamination.

6.2.1 Partitioning the Backgrounds

The first step in implementing this analysis method is to define a new set of tags which hopefully captures only a single type of background for each tag. Although the existing data cleaning cuts are already mostly geared towards a single background, they are not written to be mutually exclusive and there is some mixing. Therefore, I will define a new set of low-level tags[3] tailored to each of the different sources of background and define them such that they are mutually exclusive, i.e.

Noise Tag Any event which fails the Junk, Crate Isotropy, QvNhit, ESUM, or ITC cuts

Neck Event Tag Any event which fails the Neck cut and is not already tagged as a noise event

Flasher Tag Any event which fails the Flasher cut and is not already tagged as a noise or neck event and has < 1000 nhit

Breakdown Tag Any event which fails the Flasher or Breakdown cut and is not already tagged as a noise or neck event and has $>= 1000$ nhit

Muon Tag Any event which fails the Muon cut and is not already tagged as a noise, neck, flasher, or breakdown event

6.2.2 Setting up the Observables

For the high-level cuts we use four different variables: reconstructed radius, reconstructed z position, $\vec{u} \cdot \vec{r}$, and the goodness-of-fit parameter ψ.

3. We will refer to these as tags instead of cuts to distinguish them from the data cleaning cuts.

$\vec{\theta}_j$	Radius cut (cm)	Z cut (cm)
$\vec{\theta}_1$	$r > 800$	$z > 0$
$\vec{\theta}_2$	$r > 800$	$z < 0$
$\vec{\theta}_3$	$r < 800$	$z > 0$
$\vec{\theta}_4$	$r < 800$	$z < 0$

Table 6.2: An example showing how you might bin the high-level variables if only looking at reconstructed radius and z position.

The observables in our likelihood function are the number of events which get tagged with each low-level cut (or no cut) and all possible combinations of the high-level cuts. We can calculate the expected number of each of these by taking the product between a matrix representing the probability of tagging an event with a given low-level cut and all possibilities of the high-level cuts for each of the different background sources multiplied by a vector representing the expected number of background and signal events:

$$\begin{pmatrix} \mu_{i,\theta_1} \\ \mu_{i,\theta_2} \\ \vdots \\ \mu_{i,\theta_n} \end{pmatrix} = \begin{bmatrix} P(i,\vec{\theta}_1 \mid \text{signal}) & P(i,\theta_1 \mid \text{noise}) & \cdots & P(i,\theta_1 \mid \text{muon}) \\ P(i,\vec{\theta}_2 \mid \text{signal}) & P(i,\theta_2 \mid \text{noise}) & \cdots & P(i,\theta_2 \mid \text{muon}) \\ \vdots & & \ddots & \\ P(i,\theta_n \mid \text{signal}) & & & P(i,\theta_n \mid \text{muon}) \end{bmatrix} \vec{\mu} \tag{6.2}$$

where

$$\vec{\mu} = \begin{pmatrix} \mu_{\text{signal}} \\ \mu_{\text{noise}} \\ \mu_{\text{neck}} \\ \mu_{\text{flasher}} \\ \mu_{\text{breakdown}} \\ \mu_{\text{muon}} \end{pmatrix}$$

represents the number of signal and background events, i stands for one of the low-level tags (and one for no low-level tag), and θ_j represents some binning of the high-level variables. Table 6.2 shows an example binning of θ_j if we were only considering high-level cuts on radius and z position. Note that there is a separate Equation (6.2) for each of the low-level tags, i.

To simplify Equation (6.2) we condition on the low-level cuts, i.e.

$$P(i, \theta \mid \text{background}) = P(\theta \mid i, \text{background}) P(i \mid \text{background})$$

which then becomes

$$\begin{pmatrix} \mu_{i,\theta_1} \\ \mu_{i,\theta_2} \\ \vdots \\ \mu_{i,\theta_n} \end{pmatrix} = \begin{bmatrix} P(\theta_1 \mid i, \text{signal}) & P(\theta_1 \mid i, \text{noise}) & \cdots & P(\theta_1 \mid i, \text{muon}) \\ P(\theta_2 \mid i, \text{signal}) & P(\theta_2 \mid i, \text{noise}) & \cdots & P(\theta_2 \mid i, \text{muon}) \\ \vdots & & \ddots & \\ P(\theta_n \mid i, \text{signal}) & & & P(\theta_n \mid i, \text{muon}) \end{bmatrix} \vec{\epsilon_i} \odot \vec{\mu} \qquad (6.3)$$

where

$$\vec{\epsilon_i} = \begin{pmatrix} P(i \mid \text{signal}) \\ P(i \mid \text{noise}) \\ P(i \mid \text{neck}) \\ P(i \mid \text{flasher}) \\ P(i \mid \text{breakdown}) \\ P(i \mid \text{muon}) \end{pmatrix}$$

and $\odot$ represents component-wise multiplication.

Next, we assume no mixing between the backgrounds, i.e.

$$P(i \mid \mu_j) = 0 \quad \text{when} \quad i \neq j \quad \forall \quad i, j \in \{\text{muon}, \text{noise}, \text{neck}, \text{flasher}, \text{breakdown}\}$$

except for the neck tag and muons.

In the following subsections I will describe Equation (6.3) for each of the background tags separately. First, I will discuss which high-level variables we will assume are independent of each other for each source. For example, if we assume that all four of the high-level variables are independent for a given tag we can write:

$$P(\theta_i) = P(r) P(\psi) P(z) P(\vec{u} \cdot \vec{r})$$

This assumption is not strictly necessary for the method to work, however, it greatly reduces the final number of variables we need to fit for and so is practically necessary to avoid fitting for hundreds of variables.

Second and more importantly I will discuss the independence between the low and high-level cuts. In order to make any measurement of the sacrifice it is necessary to assume independence between at least some of the high-level cuts and the low-level tags (otherwise there would be no way to ever measure the contamination). Therefore, I will discuss in each subsection what assumptions about the independence are made and the reasoning for them.

To determine if two high-level variables are independent, I calculate a likelihood ratio test with an Ockham factor. The calculation of this ratio is discussed in Appendix F. If the likelihood ratio is greater than zero, the ratio favors the independent hypothesis and if it's less than zero it favors the hypothesis that they are correlated. However, since assuming they are correlated makes this analysis more difficult we will require that the likelihood ratio be less than -1 to assume that the variables are correlated, i.e. that the correlated hypothesis is more than 2.7 times more likely to require it.

Throughout the rest of the chapter I will use the shorthand

$$\begin{aligned}
\mathrm{M} &= \text{Muon Tag} \\
\mathrm{N} &= \text{Noise Tag} \\
\mathrm{Ne} &= \text{Neck Event Tag} \\
\mathrm{F} &= \text{Flasher Tag} \\
\mathrm{B} &= \text{Breakdown Tag} \\
\mathrm{S} &= \text{No Tag (Signal-like event).}
\end{aligned}$$

to refer to the background tags defined at the beginning of the section.

6.2.3 Muon Tag

For the muon tag, we assume a priori that the reconstructed quantities are independent of the low-level cut. The reason for this is that the low-level cut almost exclusively relies on detecting OWL PMT hits which aren't even used in the reconstruction. Therefore, we will assume that

$$P(\theta_i \mid \mathrm{M}, \mathrm{muon}) = P(\theta_i \mid \mathrm{muon})$$

and we can rewrite Equation (6.3) for the muon tag as:

$$\begin{pmatrix} \mu_{\mathrm{M},\theta_1} \\ \mu_{\mathrm{M},\theta_2} \\ \vdots \\ \mu_{\mathrm{M},\theta_n} \end{pmatrix} = \begin{bmatrix} P(\theta_1 \mid \mathrm{M}, \mathrm{signal}) & P(\theta_1 \mid \mathrm{M}, \mathrm{noise}) & \cdots & P(\theta_1 \mid \mathrm{muon}) \\ P(\theta_2 \mid \mathrm{M}, \mathrm{signal}) & P(\theta_2 \mid \mathrm{M}, \mathrm{noise}) & \cdots & P(\theta_2 \mid \mathrm{muon}) \\ \vdots & & \ddots & \\ P(\theta_n \mid \mathrm{M}, \mathrm{signal}) & & & P(\theta_n \mid \mathrm{muon}) \end{bmatrix} \begin{pmatrix} P(\mathrm{M} \mid \mathrm{signal}) \\ 0 \\ 0 \\ 0 \\ 0 \\ P(\mathrm{M} \mid \mathrm{muon}) \end{pmatrix} \odot \vec{\mu}$$

For the high-level variables, we will assume that none of them are independent based on the plots and Ψ values in Figure B.1.

Therefore,

$$P(\theta \mid \mathrm{muon}) = P(r, \psi, z, \vec{u} \cdot \vec{r} \mid \mathrm{muon})$$

6.2.4 Noise Tag

Since the noise tag encompasses a wide variety of low-level cuts (Junk, Crate Isotropy, QvNhit, ESUM, or ITC cuts), it is difficult to reason about the independence between the low-level tag and the reconstructed quantities. In addition, it is difficult to imagine how a noise event could make it past the cuts. The vast majority of the noise events in the data are pickup on one or more crates. These events typically fail both the QvNhit cut and the calibrated nhit cut. In addition, the ψ cut should reject the vast majority of any pickup events since they are clustered in crate space.

To proceed, I will assume that the high-level cuts are independent of the tag for these events, being careful to hand scan events which get tagged but pass the ψ cut in order to guess at possible ways in which an event can sneak past the cuts.

Based on the plots shown in Figure B.2 we will assume that r and $\vec{u} \cdot \vec{r}$ are the only two dependent variables[4] and we can write

$$P(\theta \mid \text{noise}) = P(r, \vec{u} \cdot \vec{r} \mid \text{noise}) P(\psi \mid \text{noise}) P(z \mid \text{noise}).$$

6.2.5 *Neck Event Tag*

For neck events, we will assume that the most likely reason for a neck event to be missed is that the neck PMTs failed to fire. In this case, we expect the reconstructed quantities to be independent of the low-level cuts because the neck PMTs are not used in the reconstruction. In this case,

$$P(\theta_i \mid \text{Ne}, \text{neck}) = P(\theta_i \mid \text{neck}),$$

and we can rewrite Equation (6.3) for the neck tag as:

$$\begin{pmatrix} \mu_{\text{Ne},\theta_1} \\ \mu_{\text{Ne},\theta_2} \\ \vdots \\ \mu_{\text{Ne},\theta_n} \end{pmatrix} = \begin{bmatrix} P(\theta_1 \mid \text{Ne}, \text{signal}) & \cdots & P(\theta_1 \mid \text{neck}) & \cdots & P(\theta_1 \mid \text{Ne}, \text{muon}) \\ P(\theta_2 \mid \text{Ne}, \text{signal}) & \cdots & P(\theta_2 \mid \text{neck}) & \cdots & P(\theta_2 \mid \text{Ne}, \text{muon}) \\ \vdots & & & \ddots & \\ P(\theta_n \mid \text{Ne}, \text{signal}) & & P(\theta_n \mid \text{neck}) & & P(\theta_n \mid \text{Ne}, \text{muon}) \end{bmatrix} \begin{pmatrix} P(\text{Ne} \mid \text{signal}) \\ 0 \\ 0 \\ P(\text{Ne} \mid \text{neck}) \\ 0 \\ P(\text{Ne} \mid \text{muon}) \end{pmatrix} \odot \vec{\mu}.$$

Note that we assume there is a nonzero probability of accidentally tagging an external muon as a neck event. This is because a muon traveling directly through the neck region of the detector is expected to cause PMT hits in the neck region and get tagged by the neck event cut.

Based on the plots and Ψ values in Figure B.3, we will assume that ψ is independent of r, $\vec{u} \cdot \vec{r}$ and z, but that the latter three are all dependent, i.e.

4. Actually we should include all of z, r, and $\vec{u} \cdot \vec{r}$, but the r and $\vec{u} \cdot \vec{r}$ correlation is the strongest.

$$P(\theta_j \mid \mathrm{Ne}, \mathrm{neck}) = P(r, z, \vec{u} \cdot \vec{r} \mid \mathrm{neck}) P(\psi \mid \mathrm{neck}).$$

Finally, we will assume that any muon that gets tagged as a neck event has the same r distribution as muons tagged with the muon tag and has $z > 0$, $\psi < 6$, and $\vec{u} \cdot \vec{r} < -0.5$:

$$P(\theta_j \mid \mathrm{Ne}, \mathrm{muon}) = \begin{cases} P(r \mid \mathrm{muon}) & \text{if } z > 0,\, \psi < 6 \text{ and } \vec{u} \cdot \vec{r} < -0.5 \\ 0 & \text{otherwise} \end{cases}.$$

6.2.6 Flasher Tag

The flasher tag is probably the most complicated cut and looks for a high charge channel in a paddle card with multiple hits and with the majority of the other PMT hits later and on the opposite side of the detector. Thus, it's not obvious whether the high-level variables depend on the low-level cuts. However, we will assume here that the most likely way for a flasher event to not get tagged is by being a "blind flasher". A blind flasher refers to an event that looks like a typical flasher event except the hits from the flashing PMT and the surrounding channels aren't recorded in the event[5]. Therefore, to test whether the high-level and data cleaning cuts are independent, I took events tagged as a flasher and removed all the hits from the crate with the flasher. The reconstructed quantities from these events looked identical to the tagged flashers and so we will therefore assume that they are independent, i.e.

$$P(\theta_i \mid \mathrm{F}, \mathrm{flasher}) = P(\theta_i \mid \mathrm{flasher})$$

and we can rewrite Equation (6.3) for the flasher tag as:

5. This can happen for any number of reasons. Typically when someone refers to a blind flasher they are referring to the case where the sequencers are disabled or the readout is broken but the channel is still at high voltage. However, when testing out the flasher cut, I found several examples where only the flashing PMT hit was gone from the event (i.e. was not due to a whole paddle card's sequencers shut off) or where the flasher appeared in the previous event as what looked like a sharkfin.

$$
\begin{pmatrix} \mu_{\mathrm{F},\theta_1} \\ \mu_{\mathrm{F},\theta_2} \\ \vdots \\ \mu_{\mathrm{F},\theta_n} \end{pmatrix} = \begin{bmatrix} P(\theta_1 \mid \mathrm{F}, \mathrm{signal}) & \cdots & P(\theta_1 \mid \mathrm{flasher}) & P(\theta_1 \mid \mathrm{F}, \mathrm{muon}) \\ P(\theta_2 \mid \mathrm{F}, \mathrm{signal}) & \cdots & P(\theta_2 \mid \mathrm{flasher}) & P(\theta_2 \mid \mathrm{F}, \mathrm{muon}) \\ \vdots & & \vdots & \vdots \\ P(\theta_n \mid \mathrm{F}, \mathrm{signal}) & & P(\theta_n \mid \mathrm{flasher}) & P(\theta_n \mid \mathrm{F}, \mathrm{muon}) \end{bmatrix} \begin{pmatrix} P(\mathrm{F} \mid \mathrm{signal}) \\ 0 \\ 0 \\ 0 \\ P(\mathrm{F} \mid \mathrm{flasher}) \\ 0 \end{pmatrix} \odot \vec{\mu}
$$

Based on the plots shown in Figure B.4 we will assume that z and $\vec{u} \cdot \vec{r}$ are correlated and that the other two high level variables are uncorrelated with the rest, i.e.

$$
P(\theta \mid \mathrm{flasher}) = P(r \mid \mathrm{flasher}) P(\psi \mid \mathrm{flasher}) P(z, \vec{u} \cdot \vec{r} \mid \mathrm{flasher}).
$$

6.2.7 *Breakdown Tag*

For breakdowns we will assume that, similar to flashers, the low-level tags are independent of the high-level quantities, i.e.

$$
P(\theta_i \mid \mathrm{B}, \mathrm{breakdown}) = P(\theta_i \mid \mathrm{breakdown})
$$

and we can rewrite Equation (6.3) for the breakdown tag as:

$$
\begin{pmatrix} \mu_{\mathrm{B},\theta_1} \\ \mu_{\mathrm{B},\theta_2} \\ \vdots \\ \mu_{\mathrm{B},\theta_n} \end{pmatrix} = \begin{bmatrix} P(\theta_1 \mid \mathrm{F}, \mathrm{signal}) & \cdots & P(\theta_1 \mid \mathrm{breakdown}) & P(\theta_1 \mid \mathrm{F}, \mathrm{muon}) \\ P(\theta_2 \mid \mathrm{F}, \mathrm{signal}) & \cdots & P(\theta_2 \mid \mathrm{breakdown}) & P(\theta_2 \mid \mathrm{F}, \mathrm{muon}) \\ \vdots & & \vdots & \vdots \\ P(\theta_n \mid \mathrm{F}, \mathrm{signal}) & & P(\theta_n \mid \mathrm{breakdown}) & P(\theta_n \mid \mathrm{F}, \mathrm{muon}) \end{bmatrix} \begin{pmatrix} P(\mathrm{M} \mid \mathrm{signal}) \\ 0 \\ 0 \\ 0 \\ P(\mathrm{B} \mid \mathrm{breakdown}) \\ 0 \end{pmatrix} \odot \vec{\mu}
$$

Since breakdowns are so rare in the data, it's not possible to have confidence about the independence of the high-level variables for breakdowns based on the data. Although all of the high-level variables appear to be independent based on the Ψ test in Figure B.5, we will assume that r and $\vec{u} \cdot \vec{r}$ are possibly dependent since that is the case for flashers:

$$P(\theta \mid \text{breakdown}) = P(r, \vec{u} \cdot \vec{r} \mid \text{breakdown}) P(\psi \mid \text{breakdown}) P(z \mid \text{breakdown}).$$

6.2.8 No Tag

Now we consider Equation (6.3) for signal-like events, i.e. events with no tag:

$$\begin{pmatrix} \mu_{\mathrm{S},\theta_1} \\ \mu_{\mathrm{S},\theta_2} \\ \vdots \\ \mu_{\mathrm{S},\theta_n} \end{pmatrix} = \begin{bmatrix} P(\theta_1 \mid \mathrm{S}, \text{signal}) & P(\theta_1 \mid \mathrm{S}, \text{noise}) & \cdots & P(\theta_1 \mid \mathrm{S}, \text{muon}) \\ P(\theta_2 \mid \mathrm{S}, \text{signal}) & P(\theta_2 \mid \mathrm{S}, \text{noise}) & \cdots & P(\theta_2 \mid \mathrm{S}, \text{muon}) \\ \vdots & & \ddots & \\ P(\theta_n \mid \mathrm{S}, \text{signal}) & & & P(\theta_n \mid \mathrm{S}, \text{muon}) \end{bmatrix} \vec{\epsilon_{\mathrm{S}}} \odot \vec{\mu}.$$

Since a given event must either have one of the low-level tags or no tag, we can write:

$$\vec{\epsilon_{\mathrm{S}}} = 1 - \sum_i \vec{\epsilon_i}$$

where the sum goes over the low-level tags for muon, noise, neck, flasher, and breakdown. Therefore, we can rewrite the equation as:

$$\begin{pmatrix} \mu_{\mathrm{S},\theta_1} \\ \mu_{\mathrm{S},\theta_2} \\ \vdots \\ \mu_{\mathrm{S},\theta_n} \end{pmatrix} = \begin{bmatrix} P(\theta_1 \mid \mathrm{S}, \text{signal}) & P(\theta_1 \mid \mathrm{S}, \text{noise}) & \cdots & P(\theta_1 \mid \mathrm{S}, \text{muon}) \\ P(\theta_2 \mid \mathrm{S}, \text{signal}) & P(\theta_2 \mid \mathrm{S}, \text{noise}) & \cdots & P(\theta_2 \mid \mathrm{S}, \text{muon}) \\ \vdots & & \ddots & \\ P(\theta_n \mid \mathrm{S}, \text{signal}) & & & P(\theta_n \mid \mathrm{S}, \text{muon}) \end{bmatrix} \left(1 - \sum_i \vec{\epsilon_i}\right) \odot \vec{\mu}. \tag{6.4}$$

Now, I will show for muons that the high-level variables are independent of the fact that they were not tagged by a data cleaning cut (we ignore here the very small probability that a muon gets tagged by the neck tag). First, we expand the probability of the high-level variables conditioned on each of the low level tags including no tag:

$$
\begin{aligned}
P(\theta_i \mid \text{muon}) &= \sum_i P(\theta_i \mid i, \text{muon}) P(i \mid \text{muon}) \\
&= P(\theta_i \mid M, \text{muon}) P(M \mid \text{muon}) + P(\theta_i \mid \text{S}, \text{muon}) P(\text{S} \mid \text{muon}).
\end{aligned}
$$

Therefore,

$$
P(\theta_i \mid \text{S}, \text{muon}) = \frac{P(\theta_i \mid \text{muon}) - P(\theta_i \mid \text{M}, \text{muon}) P(\text{M} \mid \text{muon})}{P(\text{S} \mid \text{muon})}
$$

We assume that the high-level cuts and the low-level cuts are independent for muons tagged with the muon tag. Therefore,

$$
\begin{aligned}
P(\theta_i \mid \text{S}, \text{muon}) &= \frac{P(\theta_i \mid \text{muon}) - P(\theta_i \mid \text{muon}) P(\text{M} \mid \text{muon})}{P(\text{S} \mid \text{muon})} \\
&= P(\theta_i \mid \text{muon}) \frac{1 - P(\text{M} \mid \text{muon})}{P(\text{S} \mid \text{muon})} \\
&= P(\theta_i \mid \text{muon})
\end{aligned}
$$

Since this same argument applies equally well to the other background sources, Equation (6.4) becomes

$$
\begin{pmatrix} \mu_{\text{S},\theta_1} \\ \mu_{\text{S},\theta_2} \\ \vdots \\ \mu_{\text{S},\theta_n} \end{pmatrix} = \begin{bmatrix} P(\theta_1 \mid \text{S}, \text{signal}) & P(\theta_1 \mid \text{noise}) & \cdots & P(\theta_1 \mid \text{muon}) \\ P(\theta_2 \mid \text{S}, \text{signal}) & P(\theta_2 \mid \text{noise}) & \cdots & P(\theta_2 \mid \text{muon}) \\ \vdots & & \ddots & \\ P(\theta_n \mid \text{S}, \text{signal}) & & & P(\theta_n \mid \text{muon}) \end{bmatrix} \left(1 - \sum_i \vec{\epsilon_i}\right) \odot \vec{\mu}
$$

Finally, we fit for all the unknown probabilities using a likelihood:

$$
\mathscr{L}(\vec{\mu}, \vec{p} \mid \text{data}) = \prod_i \prod_j e^{-\mu_{i,\theta_j}} \frac{\mu_{i,\theta_j}^{n_{i,j}}}{n_{i,j}!}, \tag{6.5}
$$

where $\vec{p}$ stands for all the unknown probabilities (for example, $P(M \mid \text{muon})$), θ_j stands for the jth high-level observable, and $n_{i,j}$ stands for the number of events observed with tag i and high-level quantities j.

Because we are fitting for so many variables (50 free parameters!) and most of them have constraints (probabilities must be between 0 and 1), we use a Markov Chain Monte Carlo combined with a traditional minimizer to estimate all the posteriors.

6.3 Results

The likelihood function shown in Equation (6.5) is first minimized using the SBPLX routine from the nlopt python package to find the global minimum[22, 23]. Starting from this minimum, I then perform the Markov Chain Monte Carlo using the emcee python package[24]. The proposal function is a custom function written to deal with the fact that the vast majority of the unknown free parameters are probabilities and must therefore be between 0 and 1. Therefore the proposal function is based on the truncated normal distribution from 0 to 1 with the standard deviations given by 10% of the error for each parameter found by scanning the log likelihood from the minimum along each parameter direction and looking for it to change by 0.5 from the value at the minimum.

The Markov Monte Carlo Chain is run for 100,000 iterations. The mean acceptance fraction is 10%, and the autocorrelation time for each of the parameters is shown in Table 6.3. One of the main concerns when running a Markov Chain Monte Carlo is whether it has sufficiently sampled the space. Although I'm not aware of any definitive methods to prove this, there are a few rules of thumb. First, the marginalized probabilities all look "smooth" and the walker positions appear to have fully sampled the space (see Figures 6.3 and 6.4). Second, the autocorrelation times are all verified to be approximately 10 times less than the full length of the chain.

We use the final chain to determine the marginalized distribution for the expected number of background events in our signal sample. These distributions are shown in Figure 6.2. From

Parameter	Step Size	Autocorrelation Time (samples)
Signal events	21	624
Muon events	330	171
Noise events	21	228
Neck events	99	454
Flasher events	720	184
Breakdown events	9.9	1258
$P(\mathrm{M} \mid \mathrm{muon})$	0.00093	706
$P(\mathrm{N} \mid \mathrm{noise})$	0.019	341
$P(\mathrm{Ne} \mid \mathrm{neck})$	0.0023	996
$P(\mathrm{F} \mid \mathrm{flasher})$	0.00029	310
$P(\mathrm{B} \mid \mathrm{breakdown})$	0.093	1135
$P(r < r_{\mathrm{AV}}, \psi < 6, z < 0, \vec{u} \cdot \vec{r} < -0.5 \mid \mathrm{muon})$	0.00027	2075
$P(r < r_{\mathrm{AV}}, \psi < 6, z < 0, \vec{u} \cdot \vec{r} > -0.5 \mid \mathrm{muon})$	0.00049	3094
$P(r < r_{\mathrm{AV}}, \psi < 6, z > 0, \vec{u} \cdot \vec{r} < -0.5 \mid \mathrm{muon})$	0.00056	6150
$P(r < r_{\mathrm{AV}}, \psi < 6, z > 0, \vec{u} \cdot \vec{r} > -0.5 \mid \mathrm{muon})$	0.00049	3452
$P(r < r_{\mathrm{AV}}, \psi > 6, z < 0, \vec{u} \cdot \vec{r} < -0.5 \mid \mathrm{muon})$	0.00041	2236
$P(r < r_{\mathrm{AV}}, \psi > 6, z < 0, \vec{u} \cdot \vec{r} > -0.5 \mid \mathrm{muon})$	0.00017	1880
$P(r < r_{\mathrm{AV}}, \psi > 6, z > 0, \vec{u} \cdot \vec{r} < -0.5 \mid \mathrm{muon})$	0.00038	1956
$P(r < r_{\mathrm{AV}}, \psi > 6, z > 0, \vec{u} \cdot \vec{r} > -0.5 \mid \mathrm{muon})$	0.00018	3921
$P(r > r_{\mathrm{AV}}, \psi < 6, z < 0, \vec{u} \cdot \vec{r} < -0.5 \mid \mathrm{muon})$	0.00056	5213
$P(r > r_{\mathrm{AV}}, \psi < 6, z < 0, \vec{u} \cdot \vec{r} > -0.5 \mid \mathrm{muon})$	0.00058	8717
$P(r > r_{\mathrm{AV}}, \psi < 6, z > 0, \vec{u} \cdot \vec{r} < -0.5 \mid \mathrm{muon})$	0.00059	10099
$P(r > r_{\mathrm{AV}}, \psi < 6, z > 0, \vec{u} \cdot \vec{r} > -0.5 \mid \mathrm{muon})$	0.00058	7476
$P(r > r_{\mathrm{AV}}, \psi > 6, z < 0, \vec{u} \cdot \vec{r} < -0.5 \mid \mathrm{muon})$	0.00053	4027
$P(r > r_{\mathrm{AV}}, \psi > 6, z < 0, \vec{u} \cdot \vec{r} > -0.5 \mid \mathrm{muon})$	0.00045	3244
$P(r > r_{\mathrm{AV}}, \psi > 6, z > 0, \vec{u} \cdot \vec{r} < -0.5 \mid \mathrm{muon})$	0.00054	4845
$P(z < 0 \mid \mathrm{noise})$	0.022	179
$P(\psi < 6 \mid \mathrm{noise})$	0.015	178
$P(r < r_{\mathrm{AV}}, \vec{u} \cdot \vec{r} < -0.5 \mid \mathrm{noise})$	0.015	173
$P(r < r_{\mathrm{AV}}, \vec{u} \cdot \vec{r} > -0.5 \mid \mathrm{noise})$	0.022	173
$P(r > r_{\mathrm{AV}}, \vec{u} \cdot \vec{r} < -0.5 \mid \mathrm{noise})$	0.012	176
$P(r < r_{\mathrm{AV}}, z < 0, \vec{u} \cdot \vec{r} < -0.5 \mid \mathrm{neck})$	0.0033	956
$P(r < r_{\mathrm{AV}}, z < 0, \vec{u} \cdot \vec{r} > -0.5 \mid \mathrm{neck})$	0.0081	344
$P(r < r_{\mathrm{AV}}, z > 0, \vec{u} \cdot \vec{r} < -0.5 \mid \mathrm{neck})$	0.0099	311
$P(r < r_{\mathrm{AV}}, z > 0, \vec{u} \cdot \vec{r} > -0.5 \mid \mathrm{neck})$	0.0033	796
$P(r > r_{\mathrm{AV}}, z < 0, \vec{u} \cdot \vec{r} < -0.5 \mid \mathrm{neck})$	0.0081	366
$P(r > r_{\mathrm{AV}}, z < 0, \vec{u} \cdot \vec{r} > -0.5 \mid \mathrm{neck})$	0.016	479
$P(r > r_{\mathrm{AV}}, z > 0, \vec{u} \cdot \vec{r} < -0.5 \mid \mathrm{neck})$	0.022	633
$P(\psi < 6 \mid \mathrm{neck})$	0.031	1411
$P(z < 0, \vec{u} \cdot \vec{r} < -0.5 \mid \mathrm{flasher})$	0.0026	1861
$P(z < 0, \vec{u} \cdot \vec{r} > -0.5 \mid \mathrm{flasher})$	0.0018	384
$P(z > 0, \vec{u} \cdot \vec{r} < -0.5 \mid \mathrm{flasher})$	0.0026	1817
$P(r < r_{\mathrm{AV}} \mid \mathrm{flasher})$	0.0019	178
$P(\psi < 6 \mid \mathrm{flasher})$	9.6×10^{-5}	905
$P(r < r_{\mathrm{AV}}, \vec{u} \cdot \vec{r} < -0.5 \mid \mathrm{breakdown})$	0.04	284
$P(r < r_{\mathrm{AV}}, \vec{u} \cdot \vec{r} > -0.5 \mid \mathrm{breakdown})$	0.01	1035
$P(r > r_{\mathrm{AV}}, \vec{u} \cdot \vec{r} < -0.5 \mid \mathrm{breakdown})$	0.057	601
$P(\psi < 6 \mid \mathrm{breakdown})$	0.069	166
$P(z < 0 \mid \mathrm{breakdown})$	0.051	174
$P(\mathrm{Ne} \mid \mathrm{muon})$	0.0012	1777

Table 6.3: Step sizes and autocorrelation times for parameters in the Markov Chain Monte Carlo fit. One important check to make when trying to decide if a Markov chain has properly sampled the space is to check that the autocorrelation time is much smaller than the length of the chain. In our case, the chain is 100,000 steps and the autocorrelation time for all but one of the parameters is less than 10% of the number of steps.

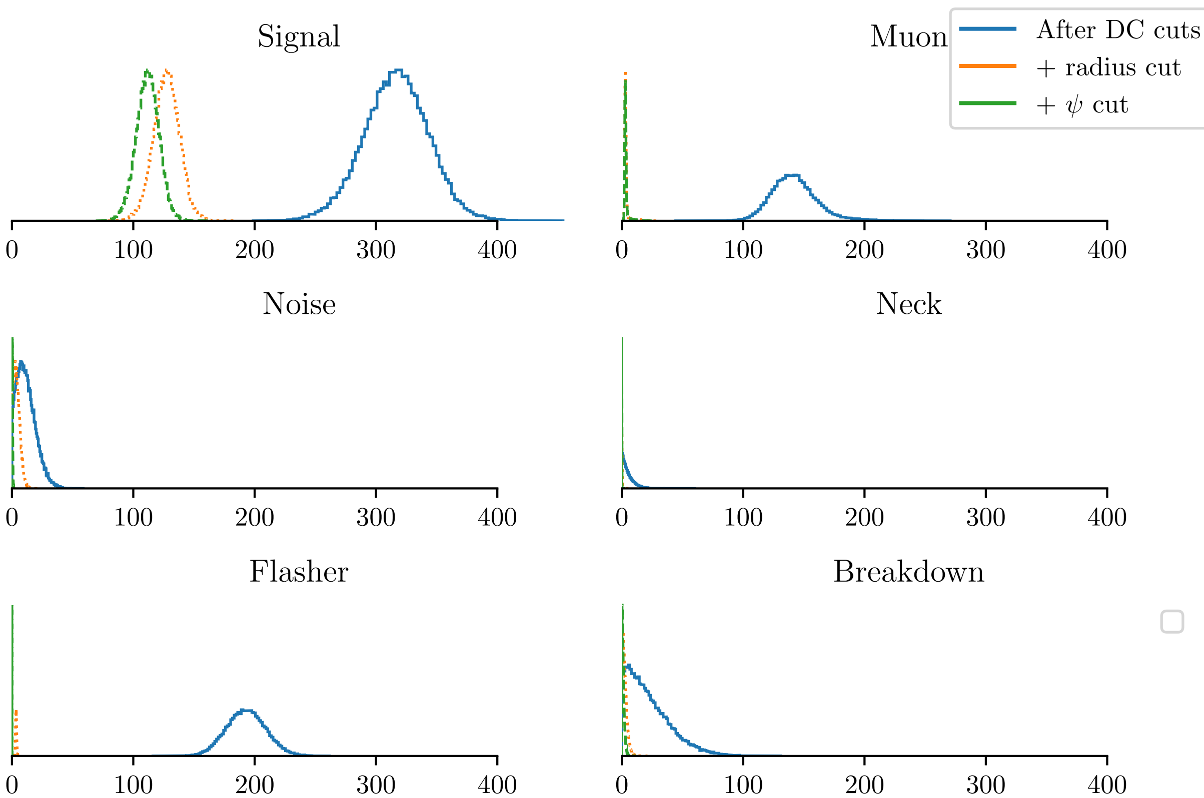

Figure 6.2: Marginalized distributions for the background contamination. The x axis of each plot represents the number of events. The distribution is shown in blue, orange, and green for the distribution after the data cleaning cuts, the data cleaning plus the radius cut, and the data cleaning cuts plus the radius and goodness of fit cut respectively.

this figure you can see that all of the instrumental backgrounds are reduced to negligible levels after the combination of the data cleaning cuts, radius cut, and goodness of fit cut.

6.4 Monte Carlo Closure Test

To double check that the analysis works as intended, I performed a Monte Carlo closure test. To perform the test, I take the untagged data samples from Monte Carlo and then I sample 100 events from each of the *tagged* background samples in data and add them to the untagged data sample. The tagged samples come from data. I then run the fit and extract the mean and standard deviation of the posteriors and create a pull plot.

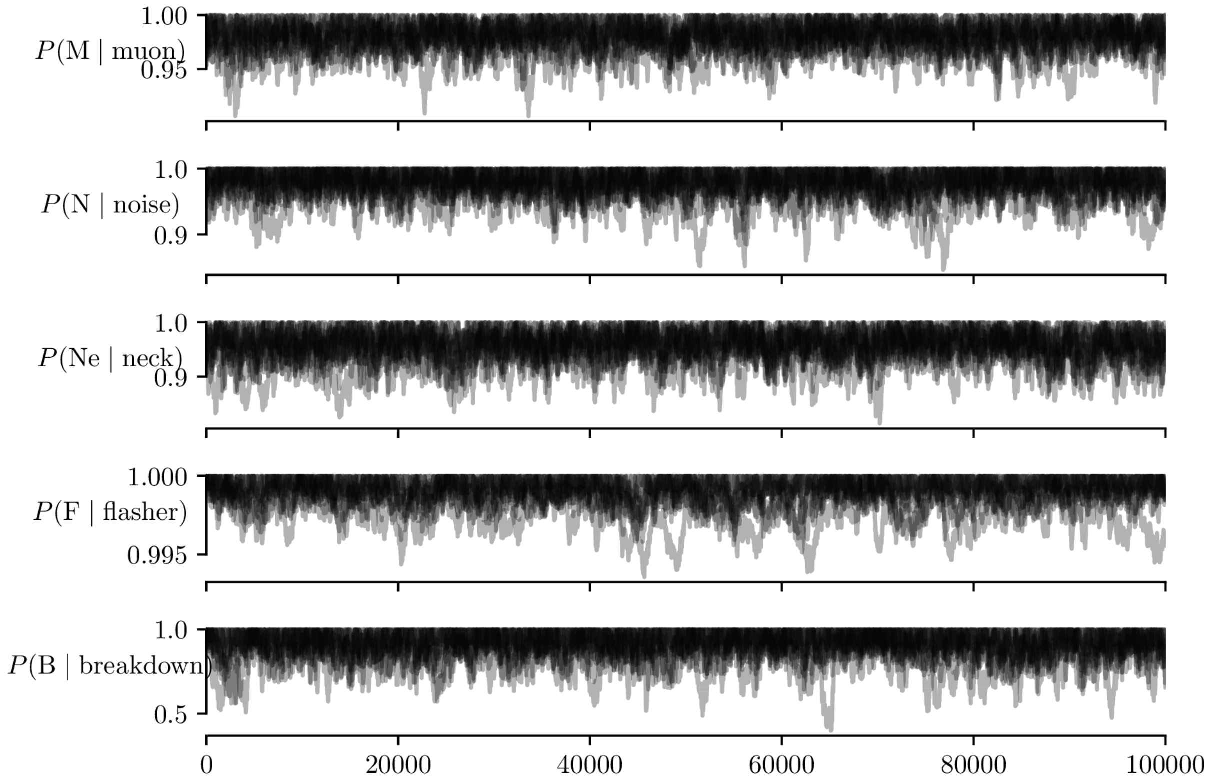

Figure 6.3: Walker positions for the data cleaning tag efficiencies. The walker positions show the value of the Markov chain at each step during the fit. Ideally the walker positions should be "fuzzy", indicating that the chain is sampling the high likelihood area and not going on a random walk.

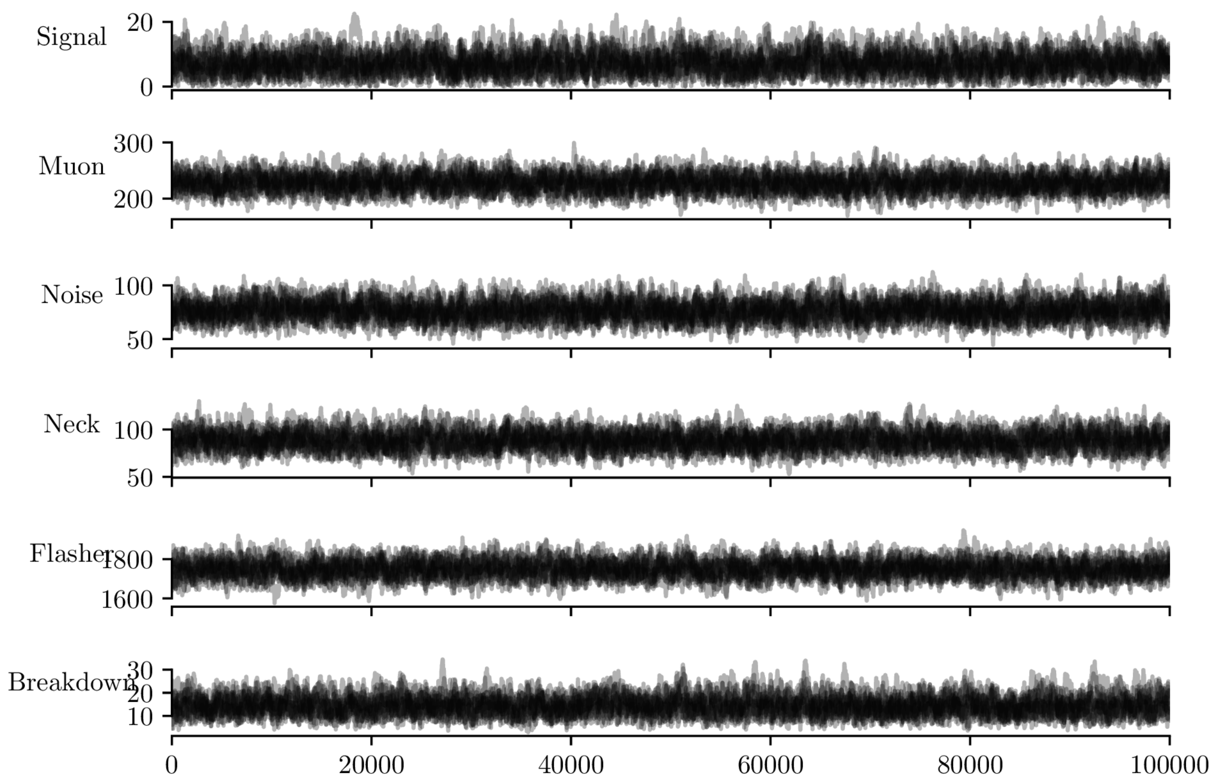

Figure 6.4: Walker positions for the total number of events. The walker positions show the value of the Markov chain at each step during the fit. Ideally the walker positions should be "fuzzy", indicating that the chain is sampling the high likelihood area and not going on a random walk.

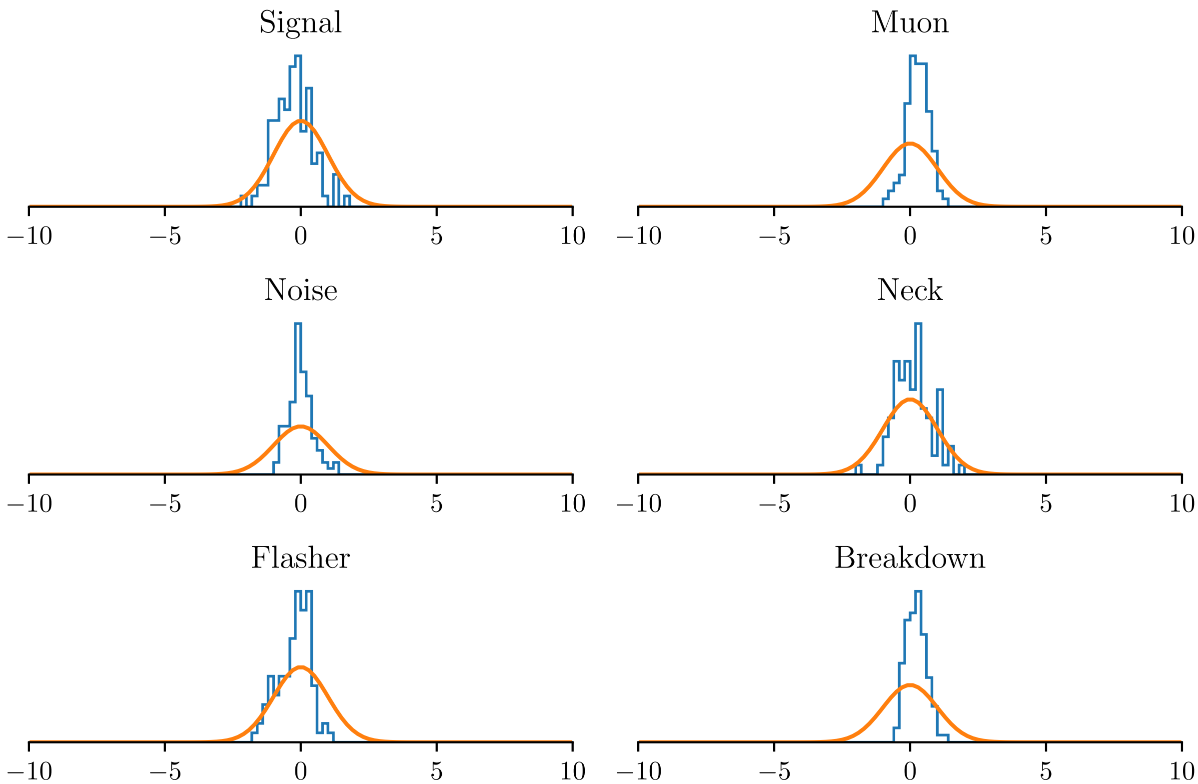

Figure 6.5: Pull plots for the different backgrounds. The x axis on each plot represents the fit result minus the true value divided by the fit uncertainty. For a properly calibrated fit, each distribution should be equivalent to a standard normal distribution.

Figure 6.5 shows the pull plots for each of the backgrounds. Although the pull plots are not completely consistent with a standard normal distribution, none show a significant bias or underestimation of the error. The fact that some of the distributions appear to be narrower than a standard normal suggests the fit is overestimating the error and may be because the likelihood is not perfectly Gaussian.

CHAPTER 7

EVENT SELECTION

7.1 Run Selection

The data taken by the SNO detector is grouped into runs lasting anywhere from an hour to several days. These runs provide a convenient way to organize the data and each run is assigned a run type which marks the run as physics, calibration, etc.

The first step in producing high quality data for analysis is the run selection group which decides which runs meet the requirement to be included in a physics analysis. For the initial SNO analyses a number of "golden" run lists were produced. Later atmospheric neutrino analyses used a set of run lists called the "muon neutrino" run lists, which were produced by adding runs to the previous golden run lists runs which hadn't been included only because they had a high amount of radon in the external water[25]. Although radon is a problem for the low energy solar neutrino analyses, it is not a problem for analyses at higher energies.

For this analysis, I started with the "muon neutrino" run lists and then applied a run level cut on the number of orphans (orphans refer to any PMT hit data which is never associated with a detector event, and occur frequently during PMT breakdowns). This extra cut is discussed in Appendix C. The final run lists for the D2O and salt phases can be found in Appendices D and E respectively.

7.2 Event Selection

The next step in determining the final set of events in the analysis is to apply a series of event level cuts. Section 7.2.1 lists the criteria for the prompt event selection which is the primary cut used for every single event selection in this analysis. This cut is designed to select the first event in a time window with potentially multiple follower events like Michel electrons, neutron capture, or detector ringing. This cut also applies the data cleaning cuts and so reduces the data down to a sample of mostly physics events. Section 7.2.2 describes the atmospheric event selection criteria used to select physics events with a neutron follower; these events are used as a sideband to check the atmospheric Monte Carlo in the final analysis. Section 7.2.3 lists the event selection criteria for the final events in our signal sample. Finally, Sections 7.2.4 and 7.2.5 discuss the event selection criteria for stopping muons and Michel electrons which are used to constrain the reconstruction systematic uncertainties in Chapter 8.

7.2.1 Prompt Event Selection

Prompt events are defined as any event satisfying the following criteria:

- Nhit > 100
- Last > 100 nhit event was more than 250 ms ago

In addition, we then apply the following set of basic data cleaning cuts[1]:

- retrigger cut - skip all events which came less than 500 ns from the last event
- Junk Cut
- Crate Isotropy Cut
- QvNHIT cut
- Flasher cut
- Neck Cut
- ITC Cut
- Breakdown Cut

1. The order here is important since if we were to apply the data cleaning cuts first we might end up selecting a prompt event caused by the "ringing" of an event cut by the data cleaning cuts.

- 00-orphan Cut
- Calibrated nhit > 100
- At least 1 NHIT trigger
- Calibrated Nhit Fraction > 0.8
- Burst Cut

7.2.2 Atmospheric Event Selection

Atmospheric events are defined as prompt events which pass the external muon cut and have a neutron follower. A neutron follower is defined as any event which passes all of the previous data cleaning cuts plus:

- ESUM
- OWL
- OWL Trigger
- FTS Cut
- has a valid FTP and RSP energy,
- FTP radius satisfies $r < r_{\mathrm{AV}}$
- RSP energy is greater than 4 MeV
- event time is $> 20\ \mu$s and < 250 ms after prompt event

Finally, atmospheric events must also satisfy the following cuts:

- $r < r_{\mathrm{AV}}$
- $\psi <= 6$

7.2.3 Signal Selection

Signal events are selected by looking at prompt events, filtering atmospheric events, and then applying the following cuts:

- $r < r_{\mathrm{AV}}$
- $\psi <= 6$

7.2.4 Stopping Muon Event Selection

Stopping muons are selected by looking for prompt events which pass all the data cleaning cuts but **fail** the muon cut with a Michel follower. In addition we also apply the following cuts:

- $\psi <= 6$
- Reconstructed kinetic energy < 10 GeV
- $\cos\theta < -0.5$

The kinetic energy cut is designed to get rid of events which have a Michel but are through-going (like double muons or external atmospheric events). The $\cos\theta$ cut is designed to select only cosmic muons and reduce the contamination from muons produced from atmospheric neutrinos (which could produce extra particles like pions).

7.2.5 Michel Event Selection

Michel events are then selected by looking at all non-prompt events (not just external muons) which satisfy the basic data cleaning cuts given in Section 7.2.1 and also satisfy the following criteria:

- ESUM Trigger Cut

- OWL Cut
- OWL Trigger Cut
- FTS Cut
- event comes more than 800 ns but less than 20 μs after a prompt event
- The associated stopping muon has Calibrated Nhit < 2500

The last nhit requirement is designed to reduce the effect of ringing and after-pulsing in the detector which is not properly modeled in the Monte Carlo.

CHAPTER 8

SYSTEMATIC UNCERTAINTIES

In this chapter I will discuss the systematic uncertainties associated with the two primary observables in this analysis: the energy reconstruction and particle ID[1]. The other major source of systematic uncertainty is the atmospheric neutrino interaction cross section which is discussed in Section 9.1.3. The source of these uncertainties is differences between the SNOMAN simulation and the real data. The physics and detector properties of the SNOMAN simulation were all validated during SNO using deployed sources at energies below 20 MeV, so differences at higher energies are to be expected.

One example that is expected to make a big difference is the model of the single photoelectron charge distribution. This distribution is modeled as a double Polya in both SNOMAN and my reconstruction algorithm. The parameters for the distribution were fit to single photoelectron data during SNO and capture the distribution above the discriminator threshold very well. For analyses below 20 MeV this was all that was necessary since the vast majority of PMT hits were single photons. For multi-photon PMT hits however, even if we assume the distribution is Gaussian by the central limit theorem, we need to know the mean and standard deviation of the single PE charge distribution, *including charges below the discriminator threshold*. Any discrepancy in the way the distribution is modeled below the discriminator threshold will show up as an energy bias at higher energies.

For this analysis, I will attempt to get a handle on any significant differences between Monte Carlo and data by looking at two natural calibration sources: Michel electrons and stopping muons. Stopping muons result from lower energy cosmic muons which enter the detector and decay within the PSUP. These muons will decay via the process

1. There may also be some difference between the other reconstructed quantities like position and direction between data and MC. However, the position will only have an effect insofar as it moves events in and outside of the fiducial volume which will only have the effect of scaling the expected number of atmospheric events. Since we are already floating the total atmospheric flux and the uncertainty is 20%, this effect will be negligible.

$$\mu \rightarrow e + \nu_e + \overline{\nu}_\mu. \tag{8.1}$$

The resulting electron is called a Michel electron. The muons have kinetic energies ranging from 200 MeV to a few GeV thus providing a good cross check at higher energies. The Michel electrons on the other hand provide a good calibration source closer to the lower limit of our analysis with a distribution spanning the energy range of 20 MeV - 60 MeV.

Stopping muons are first identified by looking for a Michel event following an event tagged as a muon. The event selection criteria are described in Sections 7.2.4 and 7.2.5.

8.1 Energy Scale and Resolution

8.1.1 Michel Electrons

Figure 8.1 shows the energy distribution of Michel electrons from data and Monte Carlo. The p-value for obtaining a result at least as extreme as the data is 65% which is consistent with the data being accurately modeled by the Monte Carlo. I also fit the two distributions while floating an energy bias parameter applied to the data and an additional energy resolution parameter applied to the Monte Carlo. The results of the fit were an energy bias of $1.5 \pm 3.4\%$ and an additional energy resolution of $0 \pm 5\%$.

8.1.2 Stopping Muons

Since the stopping muons do not have a well defined energy distribution, we instead look at the difference between the reconstructed energy and the energy as determined by the track length. To determine the energy of stopping muons from the track length, I first take the reconstructed initial position of the muon and project it back to the PSUP[2]. I then take the difference in position between the entry point at the PSUP and the reconstructed position

2. Since the event was tagged as an external muon, it definitely came from outside the PSUP. This step corrects for the fact that events near the PSUP are not always properly reconstructed at the PSUP.

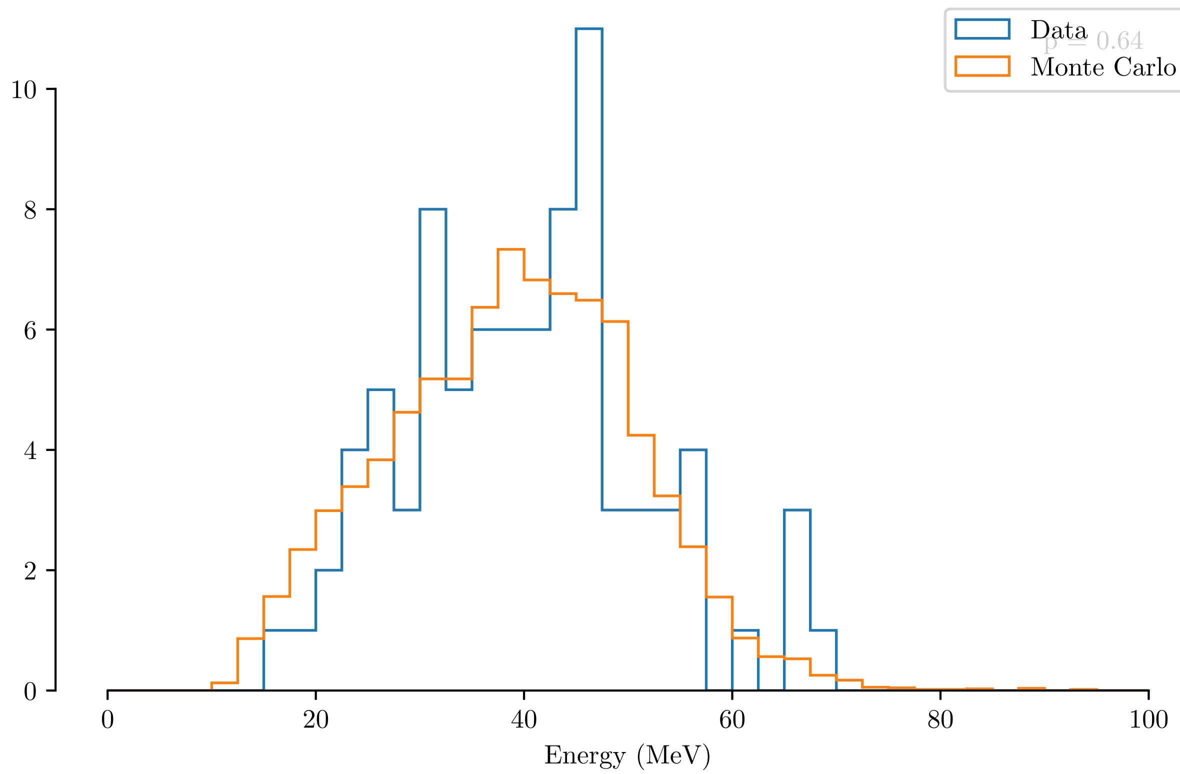

Figure 8.1: Energy distribution of Michel electrons.

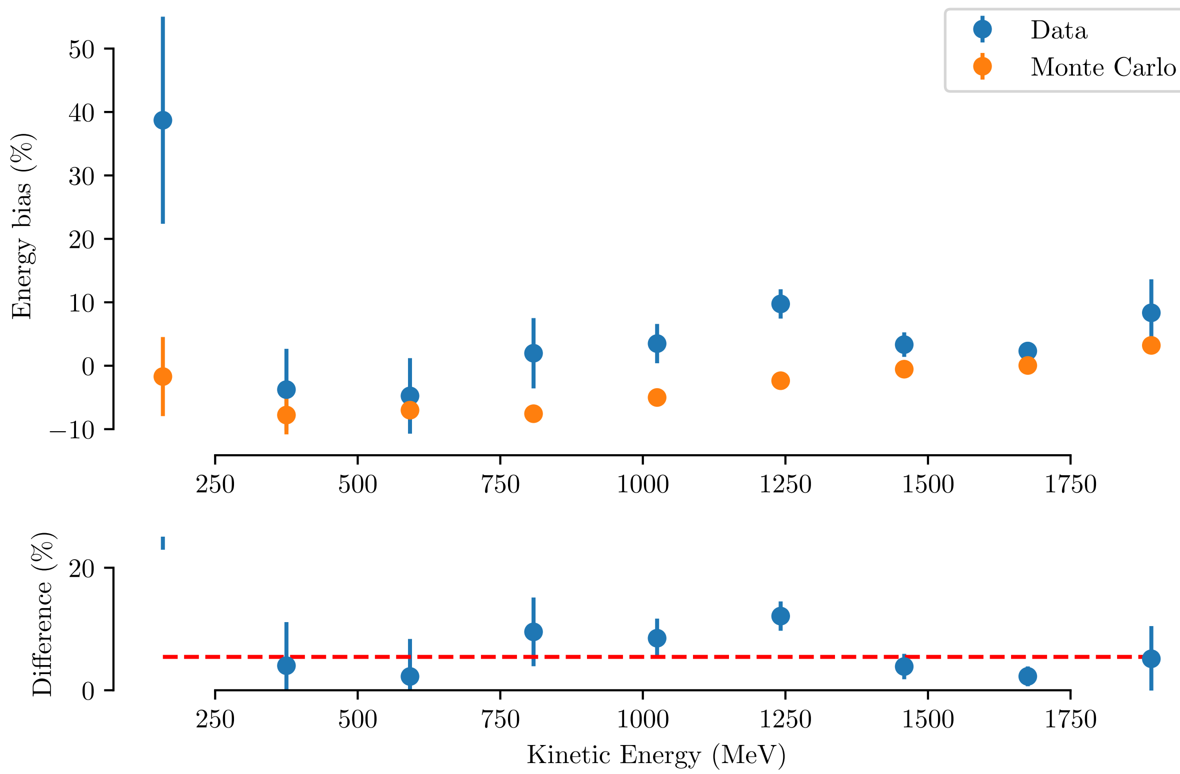

Figure 8.2: Energy bias for stopping muons. The top plot shows the energy bias for stopping muons as a fraction of the kinetic energy for both data and Monte Carlo. In the bottom plot, the data bias minus the Monte Carlo bias is shown along with a dashed red line representing the best fit to a constant difference. The bias is consistent with a constant energy bias of approximately 5%.

of the Michel electron. This distance is then used to determine the muon's initial kinetic energy by interpolating the CSDA range table for muons produced by the PDG[8].

The Monte Carlo for the stopping muons was a simulation of cosmic muons propagated by MUSIC[26], a 3D muon propagation code, and then simulated in the SNO detector.

The bias and resolution for stopping muons is shown in Figures 8.2 and 8.3 respectively. The energy resolution is consistent with no difference between data and Monte Carlo. Fitting a straight line to the difference between data and Monte Carlo gives a result of $1 \pm 1\%$. The bias shows a consistently higher energy bias in data relative to Monte Carlo. Fitting a straight line to the difference in the bias gives a result of $5.4 \pm 1\%$.

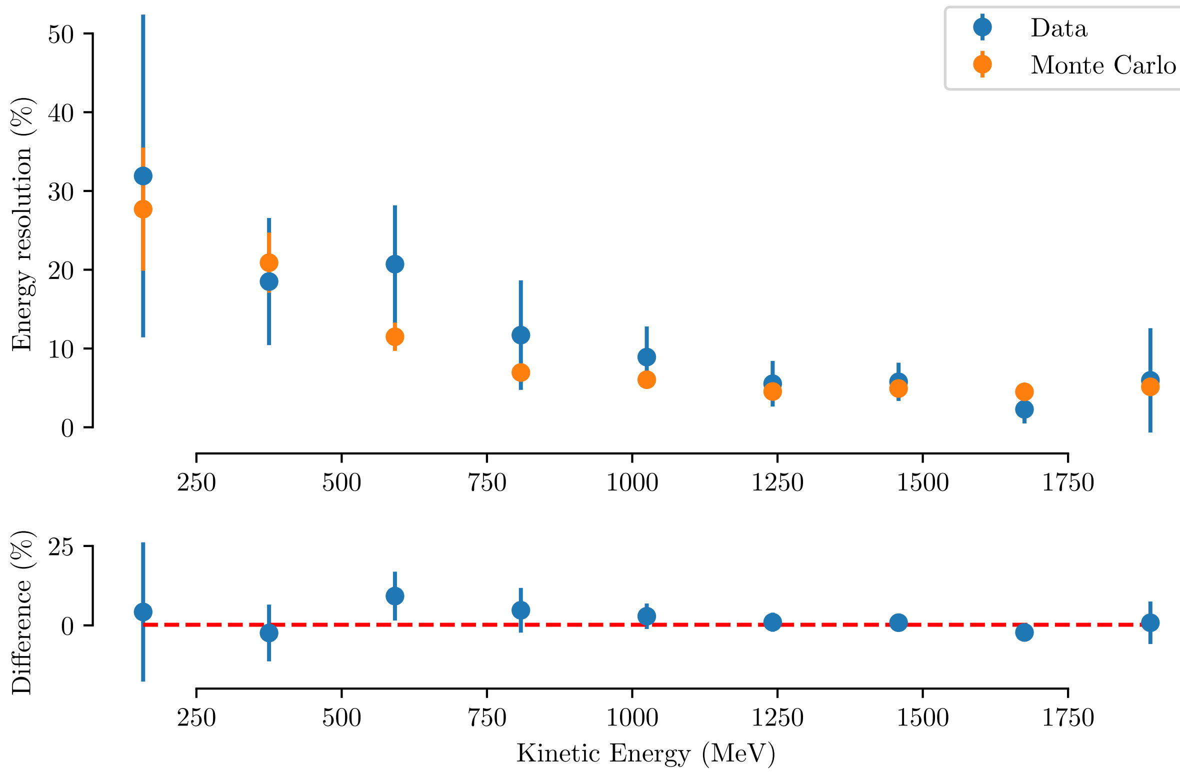

Figure 8.3: Energy resolution for stopping muons. The top plot shows the energy resolution for stopping muons as a fraction of the kinetic energy for both data and Monte Carlo. In the bottom plot, the data resolution minus the Monte Carlo resolution is shown along with a dashed red line representing the best fit to a constant difference. The data is consistent with no difference in resolution between data and Monte Carlo.

Particle ID	Monte Carlo (%)	Data (%)
e	99.2 ± 0.1	100 ± 2
μ	0.05 ± 0.03	0 ± 1
ee	0.75 ± 0.10	0 ± 1
$e\mu$	0.02 ± 0.02	0 ± 1
$\mu\mu$	0.0 ± 0.1	0 ± 1

Table 8.1: Probability of reconstructing a given particle ID for Michel electrons.

Particle ID	Monte Carlo (%)	Data (%)
e	3.2 ± 0.6	5 ± 2
μ	50 ± 2	44 ± 4
ee	0.1 ± 0.2	1.5 ± 1.2
$e\mu$	45 ± 2	48 ± 4
$\mu\mu$	1.4 ± 0.4	0.8 ± 1.0

Table 8.2: Probability of reconstructing a given particle ID for stopping muons.

8.2 Particle ID

8.2.1 Michel Electrons

Table 8.1 shows the probability of reconstructing various particle IDs for Michel electrons. The particle ID probabilities are all consistent with the values from the Monte Carlo.

8.2.2 Stopping Muons

Table 8.2 shows the probability of reconstructing various particle IDs for stopping muons. The particle ID probabilities are all consistent with the values from the Monte Carlo.

CHAPTER 9

RESULTS

9.1 Null Hypothesis Test

9.1.1 The Likelihood Function

To perform the null hypothesis test, I apply the event selection criteria described in Chapter 7 to select the signal events. I then perform a Bayesian fit with the following parameters:

1. The atmospheric neutrino flux scale
2. The energy bias for electrons
3. Additional energy resolution for electrons
4. The energy bias for muons
5. Additional energy resolution for muons
6. The number of external muons

The likelihood function is computed by first applying the energy bias and resolution terms to the Monte Carlo. To do this, all Monte Carlo events are grouped based on the reconstructed particle ID, i.e. we separately histogram events which reconstruct as a single electron, single muon, double electron, electron + muon, and double muon. For each group, I apply the energy bias and resolution parameters to the Monte Carlo and histogram the results by computing

$$h_{\text{MC},i} = \sum_j \Phi\left(\frac{T_j - b_{i+1}}{\sigma_j}\right) - \Phi\left(\frac{T_j - b_i}{\sigma_j}\right) \tag{9.1}$$

where i represents the i^{th} bin, j represents an index running over every single MC event, $\Phi(x)$ represents the normal cumulative distribution function, T_j represents the scaled Monte

Carlo energy, σ_j represents the additional energy resolution, and b_i represents the lower edge of the i^{th} bin. The scaled energy is calculated by multiplying the reconstructed energy for each particle in the fit by:

$$T_j = \sum_k T'_k (1 + \delta_{e/\mu,k}) \tag{9.2}$$

where k loops over the particles in the fit, T' is the original reconstructed kinetic energy, and $\delta_{e/\mu}$ is the bias parameter associated with either electrons or muons depending on the particle ID. The additional energy resolution is similarly calculated as

$$\sigma_j = \sqrt{\sum_k \left(T'_k \times \sigma_{e/\mu,k}\right)^2}. \tag{9.3}$$

This process is also done for the external muon Monte Carlo separately. The total number of expected events is then calculated by multiplying the atmospheric neutrino Monte Carlo histogram by the flux term A, the external muon histogram by the normalization term M, and adding the two together.

Next, the real data is also grouped based on the reconstructed particle ID, and for each group the total kinetic energy is histogrammed. The likelihood is then calculated by computing the product of the Poisson probability of observing n events and the multinomial probability of observing the data given the atmospheric neutrino Monte Carlo for all five distributions:

$$\mathcal{L} = \prod_{\text{ID}} P(n|N) P(\vec{h}|A \times \vec{h}_{\text{MC,atmo}} + M \times \vec{h}_{\text{MC,muon}}) \tag{9.4}$$

$$= \prod_{\text{ID}} e^{-N} \frac{N^n}{n!} \frac{n!}{\prod_i h_i!} \prod_i \left(p_{\text{MC},i}\right)^{h_i} \tag{9.5}$$

where the outer product is over the six different possible particle IDs (single electron, single muon, double electron, etc.), the $p_{\text{MC},i}$ represent the fraction of the atmospheric neutrino Monte Carlo and external muons in bin i, N represents the total expected number of expected

Parameter	Symbol	Central Value	1σ Uncertainty
Atmospheric Flux	A	1.0	0.2
Energy Bias (e)	δ_e	0.015	0.034
Energy Resolution (e)	σ_e	0.0	0.049
Energy Bias (μ)	δ_μ	0.054	0.01
Energy Resolution (μ)	σ_μ	0.01	0.01
External Muon Scale	M	0	10

Table 9.1: Table showing the fit parameters along with the central value and uncertainty for any priors. Parameters with a dash have a flat prior with no constraint.

Monte Carlo events, n is the number of data events, and h_i represents the number of data events.

9.1.2 Priors

The priors for the parameters in the fit are shown in Table 9.1. We use a conservative 20% error on the total atmospheric neutrino flux scale. This uncertainty comes from the uncertainty in the computed total flux which is expected to be approximately 15%[27], as well as additional uncertainty from the livetime, fiducial volume, and the fact that the MC was generated at the solar maximum[1]. The energy bias and resolution terms correct for the difference in the energy reconstruction between Monte Carlo and data (see Chapter 8). The priors for the energy bias and resolution parameters for muons come from the fits shown in Figures 8.2 and 8.3. For electrons, these priors come from a fit to the Michel energy distribution where we allow the energy bias and resolution to float, which is discussed in Section 8.1.1. Finally, the constraint on the external muon scale comes from the data cleaning analysis discussed in Section 6.2.

1. The solar cycle also has an effect on the energy distribution of the atmospheric neutrino flux. However, the solar cycle abruptly jumped to a maximum right at the start of the SNO data and so the data from both the D2O and salt phases is almost perfectly aligned with the solar maximum.

9.1.3 GENIE Systematics

In order to account for systematic uncertainties in neutrino cross sections, the GENIE software package includes a library for reweighting events based on changes in input parameters to the models GENIE uses to calculate neutrino cross sections. Because there are so many parameters and the GENIE reweighting procedure is fairly computationally expensive, it is not feasible to float these terms in the fit. One standard procedure for dealing with systematics which are not floated in the final fit is to vary each one individually to assess the impact on the final result. However, this method ignores correlated effects between the parameters. Therefore, I use a slightly different method.

In order to incorporate these uncertainties into the fit, I use a script, created by Andy Mastbaum, which creates 1000 different "universes". For each "universe" a value for each of the systematics listed in Table 4.5 is randomly selected according to the uncertainty listed in the table. Then, given these values, we use the reweighting package in GENIE to assign a weight to every single event in the atmospheric Monte Carlo.

Finally, in order to incorporate these results into our analysis, we do the following:

1. run a Markov Chain Monte Carlo on our likelihood and priors, and select the set of parameters with the highest likelihood

2. Loop over every single "universe", apply the weights to the atmospheric Monte Carlo, and then choose the universe with the highest likelihood

3. run the Markov Chain Monte Carlo again with the most likely universe to produce the fit posteriors

9.1.4 P-Value

After calculating the final fit posteriors, I calculate a p-value for each particle ID. The p-value represents the probability of obtaining a test statistic $-2\log(\lambda)$ value at least as extreme as the data. The test statistic I use for the p-value is

$$-2\log(\lambda) = 2\left(N - n + \sum_i O_i \log\left(\frac{O_i}{E_i}\right)\right) \tag{9.6}$$

where N is the total number of expected atmospheric events, n is the total number of observed events, and O_i and E_i are the number of observed and expected atmospheric events in bin i. This statistic is based on a likelihood ratio and is discussed in Appendix G.

The p-value is computed by randomly sampling parameters from the posterior of the fit using a Markov Chain Monte Carlo. This procedure is repeated 1000 times to produce a distribution of possible p-values. The 50th percentile of these p-values is the final p-value presented for each distribution.

9.1.5 Monte Carlo Closure Test

A Monte Carlo closure test was run in order to verify that the likelihood fit was unbiased and reported the correct errors. To perform this test, I used simulation-based calibration (SBC) tests, which are described in Appendix H. To produce the SBC histograms, I randomly sample the parameters in the fit from their prior distributions and then run the fit. This process is repeated over and over, and then the rank statistic for the truth values in the distribution of the posterior is computed. Figure 9.1 shows the SBC plots for 1000 different runs. These plots show that the fit is well calibrated and there are no significant biases or problems with the posterior.

9.1.6 P-Value Coverage

Although it is not possible to obtain uniform coverage perfectly for a model with unknown parameters, it is still useful to be able to visualize the coverage for a typical set of parameters. Since I calculate the p-value using the posterior results, the coverage is expected to be conservative.

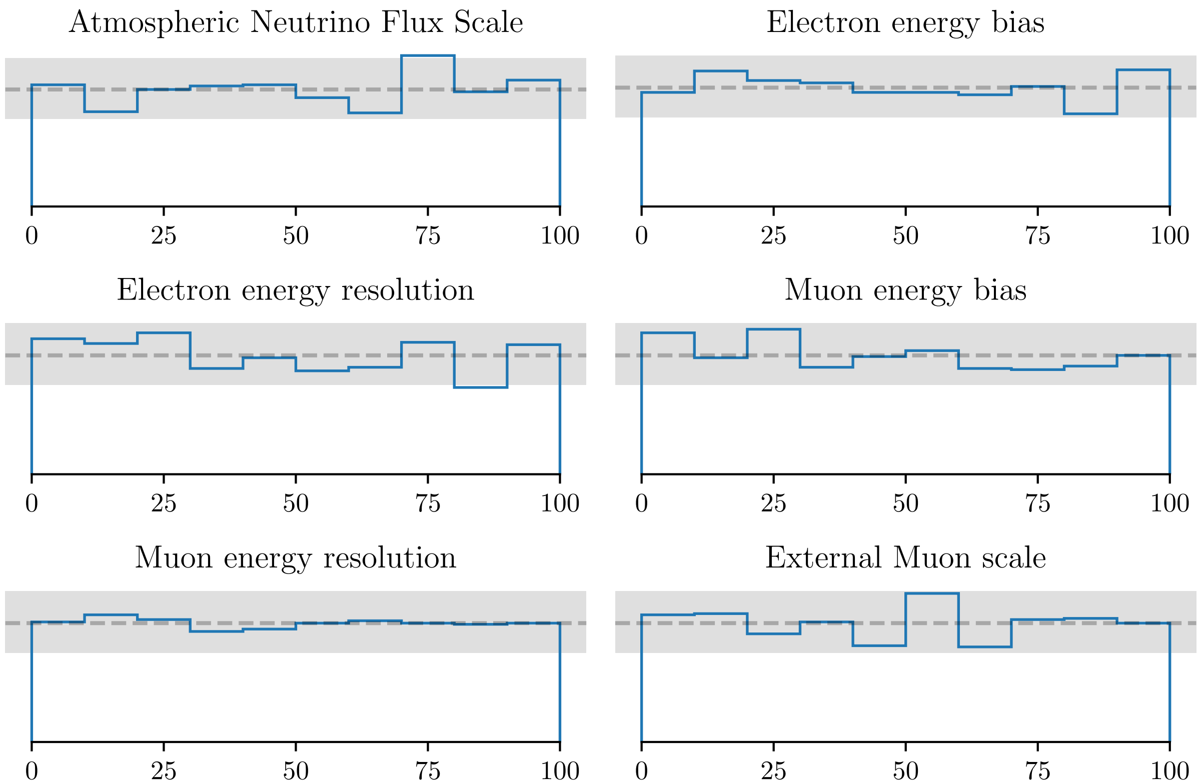

Figure 9.1: Simulation-based calibration (SBC) histograms for parameters in the final fit. The x axes represents the percentile of the simulated quantity in the Markov chain. The grey dashed line shows a uniform distribution and the grey band shows the range in which we expect 99% of the bins to fall assuming the underlying distribution is uniform.

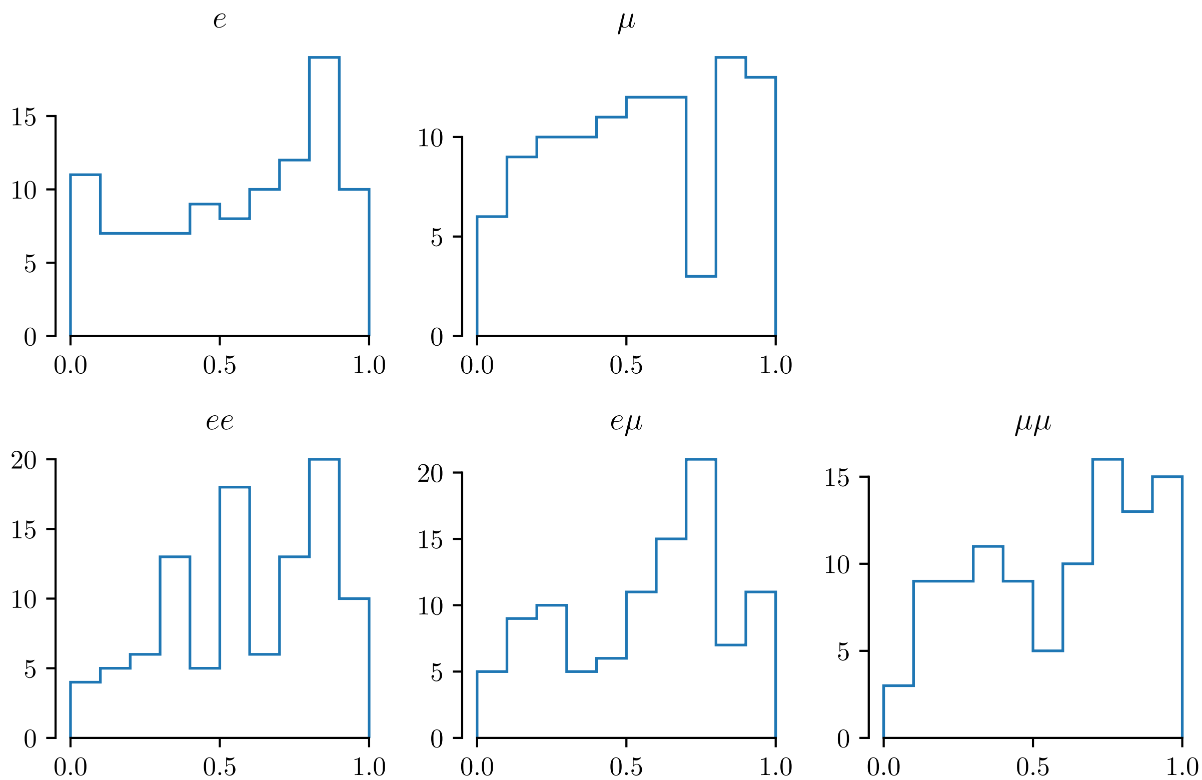

Figure 9.2: This plot shows the p-value coverage for 100 different randomly sampled data from a typical set of parameters. To generate the data, we chose a single set of the GENIE systematic parameters and a single set of the fit parameters and then randomly generated the data 100 times and fit it. The distributions are "conservative" in the sense that they tend to produce distributions which are weighted towards a p-value of 1.0, which is what is expected when calculating a posterior predictive p-value.

To generate a coverage plot for a typical set of parameters, I first draw random values for the fit parameters according to their prior distributions. I then select the first "universe" for the GENIE systematics. Using the randomly drawn values and the GENIE weights, I then sample the Monte Carlo data to produce a data set drawn from the null hypothesis and run the fit to produce a p-value for each particle ID. This process is repeated (using the same values for the parameters as the first time) 100 times to produce a distribution of p-values. The distribution of p-values is shown in Figure 9.2.

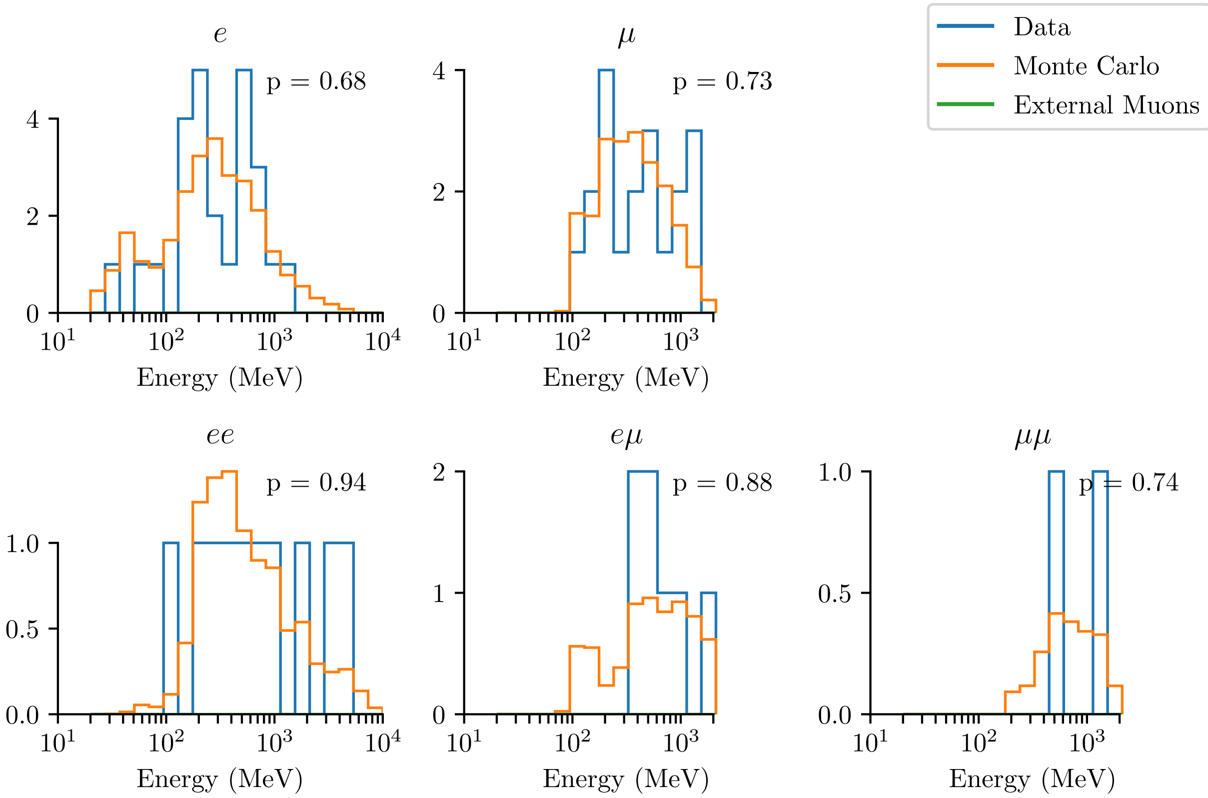

Figure 9.3: Energy distribution of signal events. The p-value shown represents the probability of obtaining a χ^2 value at least as extreme as the data.

9.1.7 Results

For the analysis presented here, we analyzed approximately 50% of the total data with a livetime of 234 days. The energy distribution of the signal events along with the p-value for each particle ID is shown in Figure 9.3 and the posteriors for the fit parameters are shown in Figure 9.4. All the distributions are consistent with the events being caused by atmospheric neutrinos.

The atmospheric event sideband is shown in Figure 9.5.

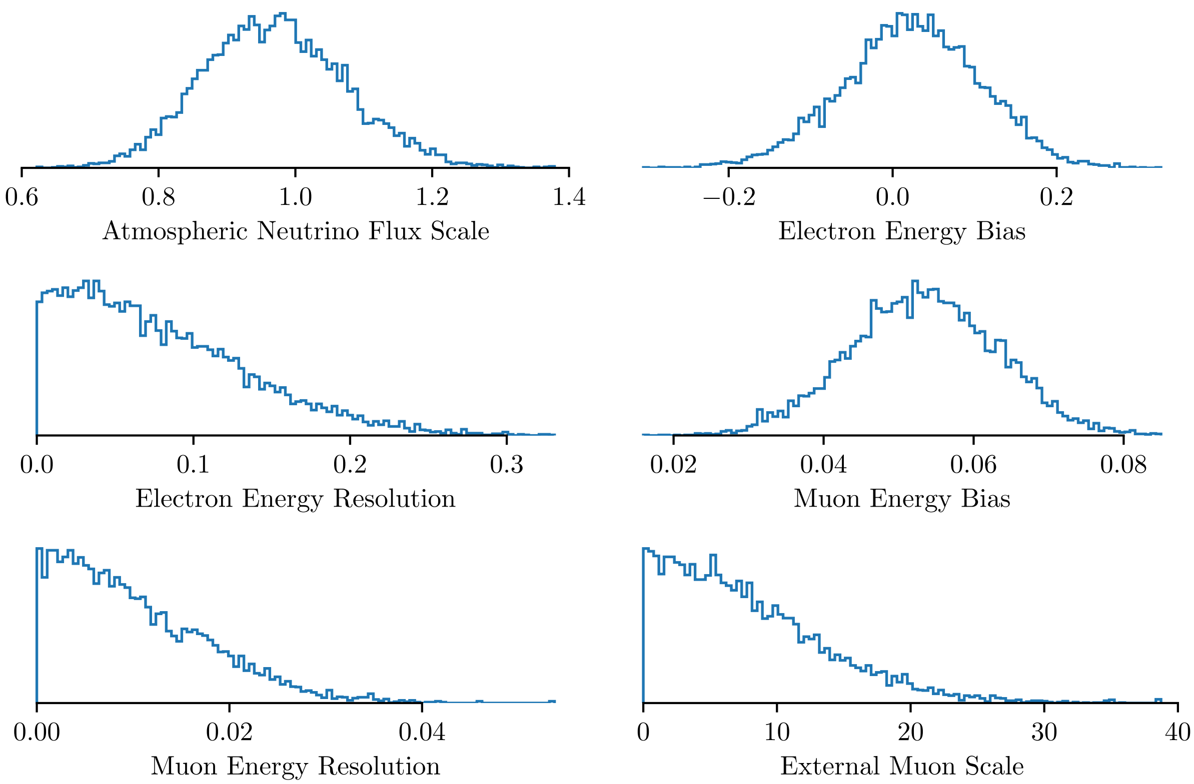

Figure 9.4: Posteriors for the fit parameters.

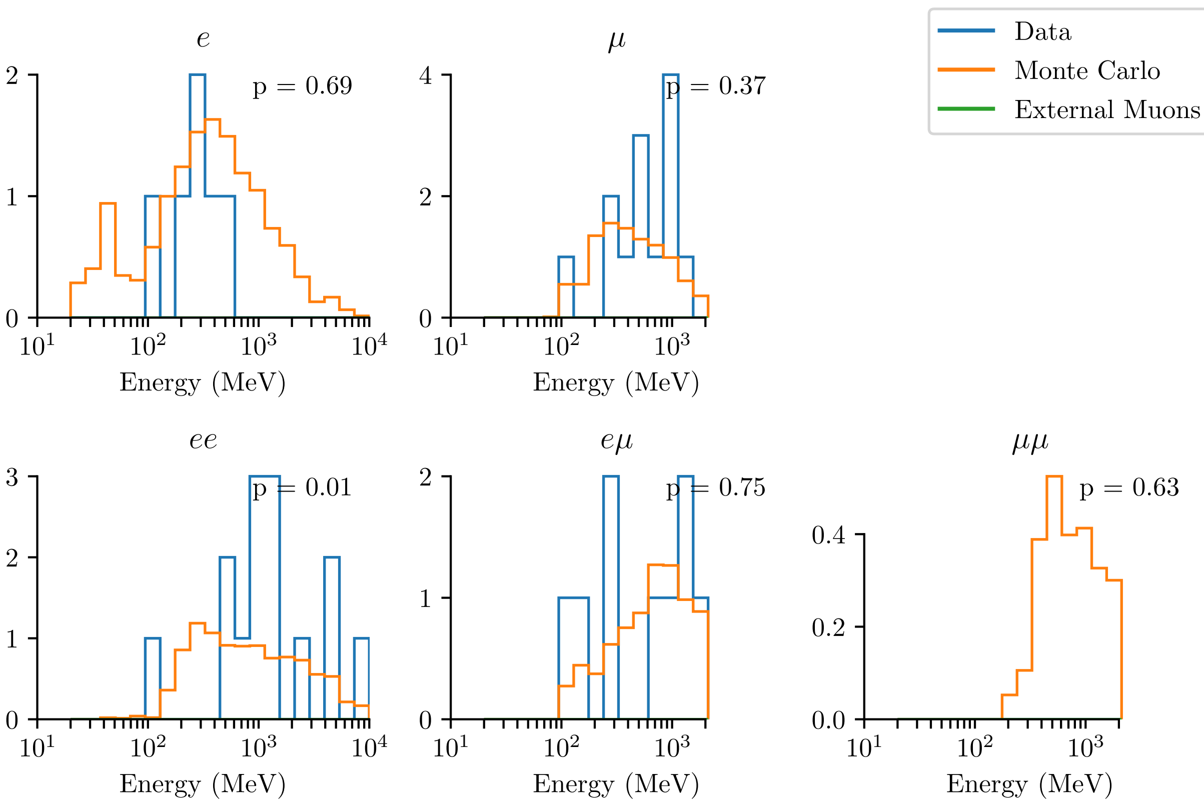

Figure 9.5: Energy distribution of events tagged with a neutron follower. The p-value shown represents the probability of obtaining a χ^2 value at least as extreme as the data.

9.2 Direct Dark Matter Search

To perform the direct search for self-destructing dark matter, we perform a fit almost identical to that described in the previous section, except we float an additional term for the dark matter. In this analysis, we assume the mediator is moving slowly enough that the daughter particles come out approximately back to back[2]. The 90% confidence limit is then obtained by taking the 90th percentile of the dark matter parameter samples from the MCMC.

To determine a "discovery threshold" we look at the best fit value of the dark matter term and compare it with a threshold designed to have a false positive rate of 5%. The calculation of the discovery threshold is discussed in Appendix I.

9.2.1 Results

Figure 9.8 shows the best fit for the number of dark matter events as a function of the dark matter mass along with an approximate 2σ discovery threshold. The results show no significant excess of events which can be attributed to a self-destructing dark matter signal. Figure 9.9 shows the 90% confidence limits for the event rate of self-destructing dark matter as a function of the dark matter mass.

Figures 9.6 and 9.7 show two signal events in the final sample which reconstructed as two muon-like and two-electron like rings respectively.

2. For a faster moving mediator the only difference is that a larger fraction of the events may be misreconstructed as a single particle and thus the limits would be slightly worse. In a future analysis of this data, it would be a good idea to perform this analysis by also looking at the opening angle of the particle pair to get a much better limit.

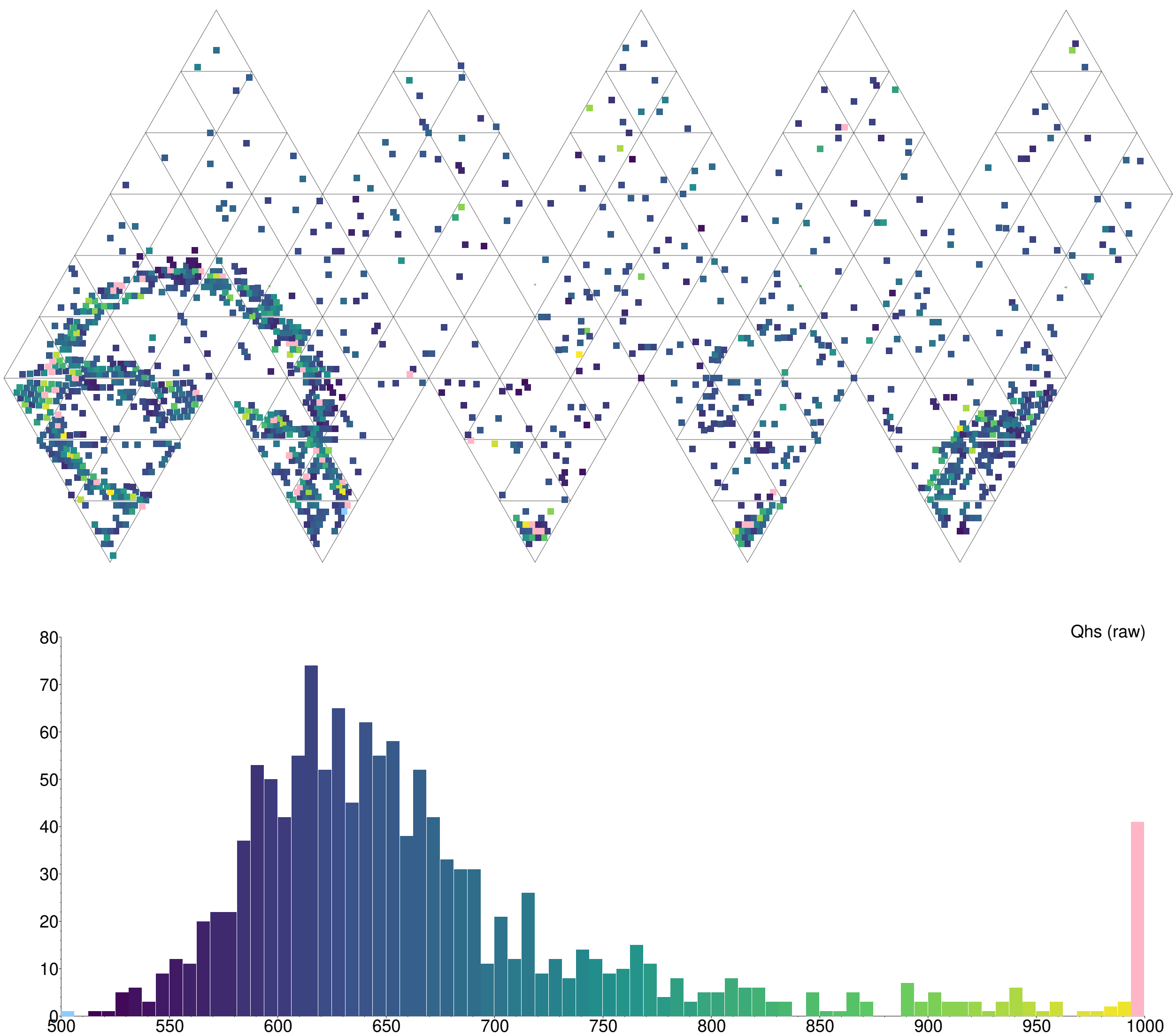

Figure 9.6: XSnoed event display showing a signal event that reconstructed with two muon-like rings. This event is from run 14190 and had the GTID number 4043274.

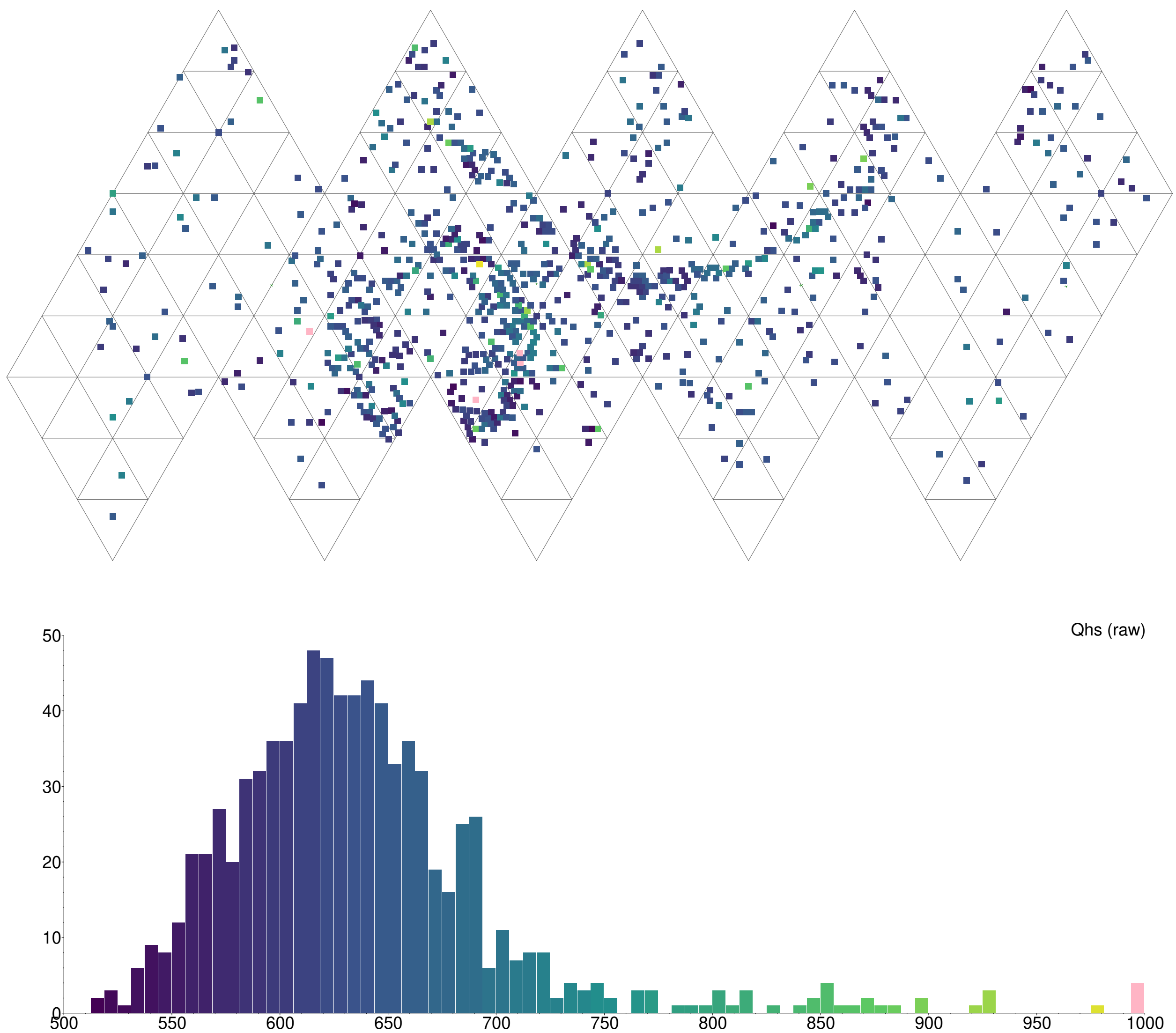

Figure 9.7: XSnoed event display showing a signal event that reconstructed with two electron-like rings. This event is from run 10536 and had the GTID number 734712.

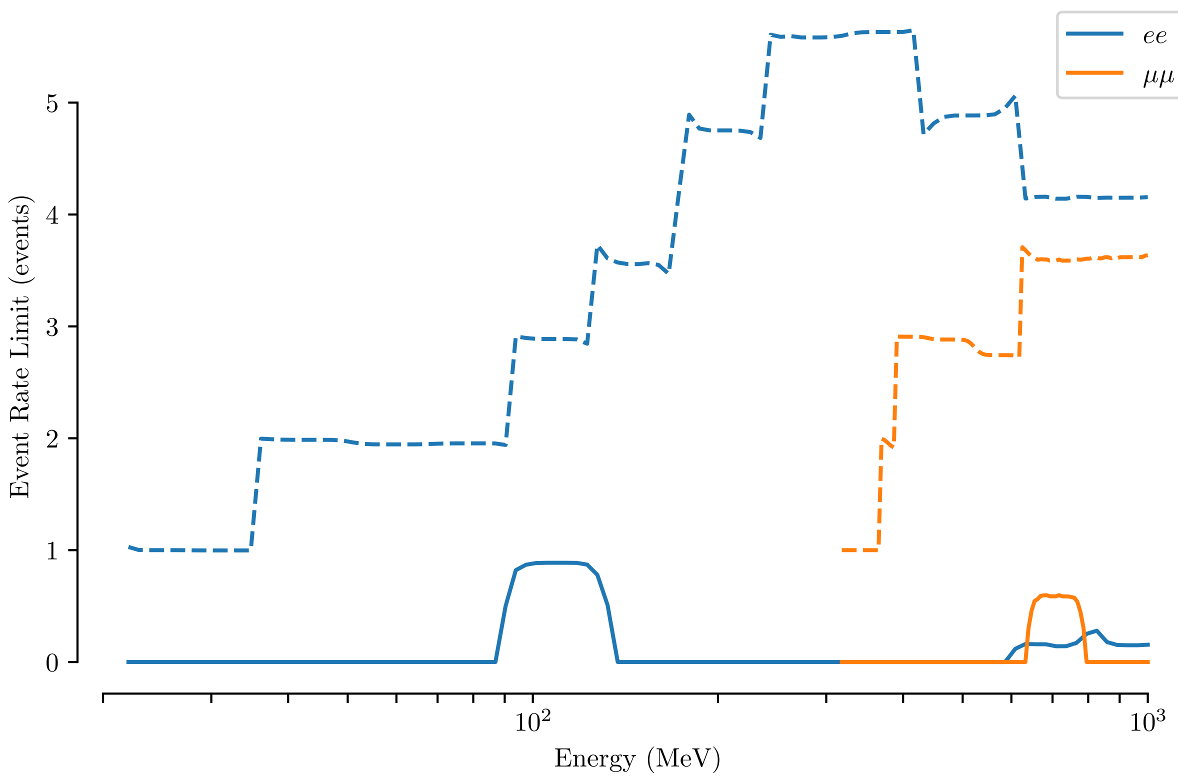

Figure 9.8: Best fit event rate for self-destructing dark matter as a function of the dark matter mass for a slow mediator. The two bumps come from the two bins which had a single event each.

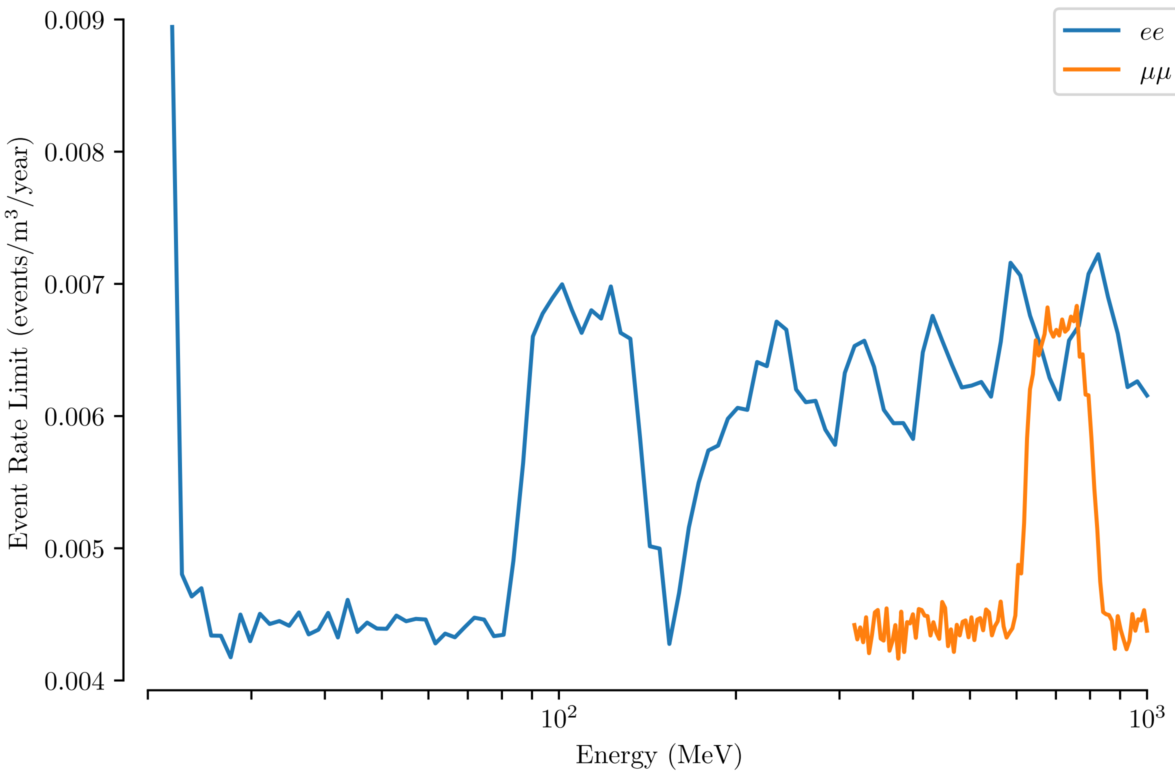

Figure 9.9: Event rate limit (90% CL) for self-destructing dark matter as a function of the dark matter mass for a slow mediator. The high frequency noise is due to sampling error from the MCMC. The five spikes in the electron positron limit on the right are an artifact of the binning that we chose, and will be corrected in a later publication.

CHAPTER 10

CONCLUSION

I performed the first search for a new class of dark matter models called self-destructing dark matter using data from the Sudbury Neutrino Observatory. To perform this search, I developed a new reconstruction algorithm which is able to fit for multi-track events from 20 MeV to 10 GeV and determine both the multiplicity and particle ID of events using Bayes factors.

The primary background for this search consisted of atmospheric neutrino events and detector instrumentals. To predict the events from atmospheric neutrinos, I simulated the oscillated atmospheric neutrino flux using GENIE and SNOMAN. The instrumentals were tagged by a new set of data cleaning cuts and the residual instrumental contamination was verified to be negligible.

To calibrate the reconstruction algorithm, I used two natural calibration sources: stopping muons and Michel electrons. No significant difference between data and Monte Carlo was seen for the particle ID probabilities and energy resolution. A small energy bias was detected in the stopping muons likely due to a mismodeling of the single photoelectron charge in the PMTs.

The final null hypothesis test was performed by doing a Bayesian analysis on the energy distribution of all the events between 20 MeV and 10 GeV (2 GeV for muons) for each reconstructed particle ID. No significant excess of multi-particle events consistent with the self-destructing dark matter model was found and I placed the first such limits on these events. In addition, I performed what I believe is the first broad band null hypothesis test in a large Water Čerenkov detector to search for any deviation from known physics. No new physics was found; all events were consistent with the atmospheric neutrino prediction. I hope that this analysis will motivate other experiments like Super Kamiokande to perform similar tests.

This analysis could be significantly extended by performing the same search as a function of both the dark mediator mass and energy instead of focusing on back to back events. To do this, more simulation would be required to parameterize the probability of correctly identifying a multi-particle pair as a function of how boosted the dark mediator is. It may also be possible to set better limits by performing a 2D analysis in both energy and opening angle for multi-particle events.

APPENDIX A

POISSON BINOMIAL

Suppose we have a Poisson process whose output is then subject to a binomial process. For example, we expect μ background events on average and we can detect them with probability p. What is the probability of detecting n background events?

$$
\begin{aligned}
p(n) &= \sum_{N=n}^{\infty} P(n|N)P(N) \\
&= \sum_{N=n}^{\infty} \frac{N!}{n!(N-n)!} p^n (1-p)^{N-n} e^{-\mu} \frac{\mu^N}{N!} \\
&= \sum_{N=n}^{\infty} \frac{1}{n!(N-n)!} p^n (1-p)^{N-n} e^{-\mu} \mu^N \\
&= e^{-\mu} \frac{p^n}{n!} \sum_{N=n}^{\infty} \frac{1}{(N-n)!} (1-p)^{N-n} \mu^N \\
&= e^{-\mu} \frac{(\mu p)^n}{n!} \sum_{N=n}^{\infty} \frac{(\mu(1-p))^{N-n}}{(N-n)!} \\
&= e^{-\mu p} \frac{(\mu p)^n}{n!}
\end{aligned}
$$

Therefore the end result is a Poisson distribution with mean μp.

APPENDIX B

CORNER PLOTS

Corner plots for events tagged as a muon, noise, neck, flasher, breakdown, or signal event are shown in Figures B.1-B.6.

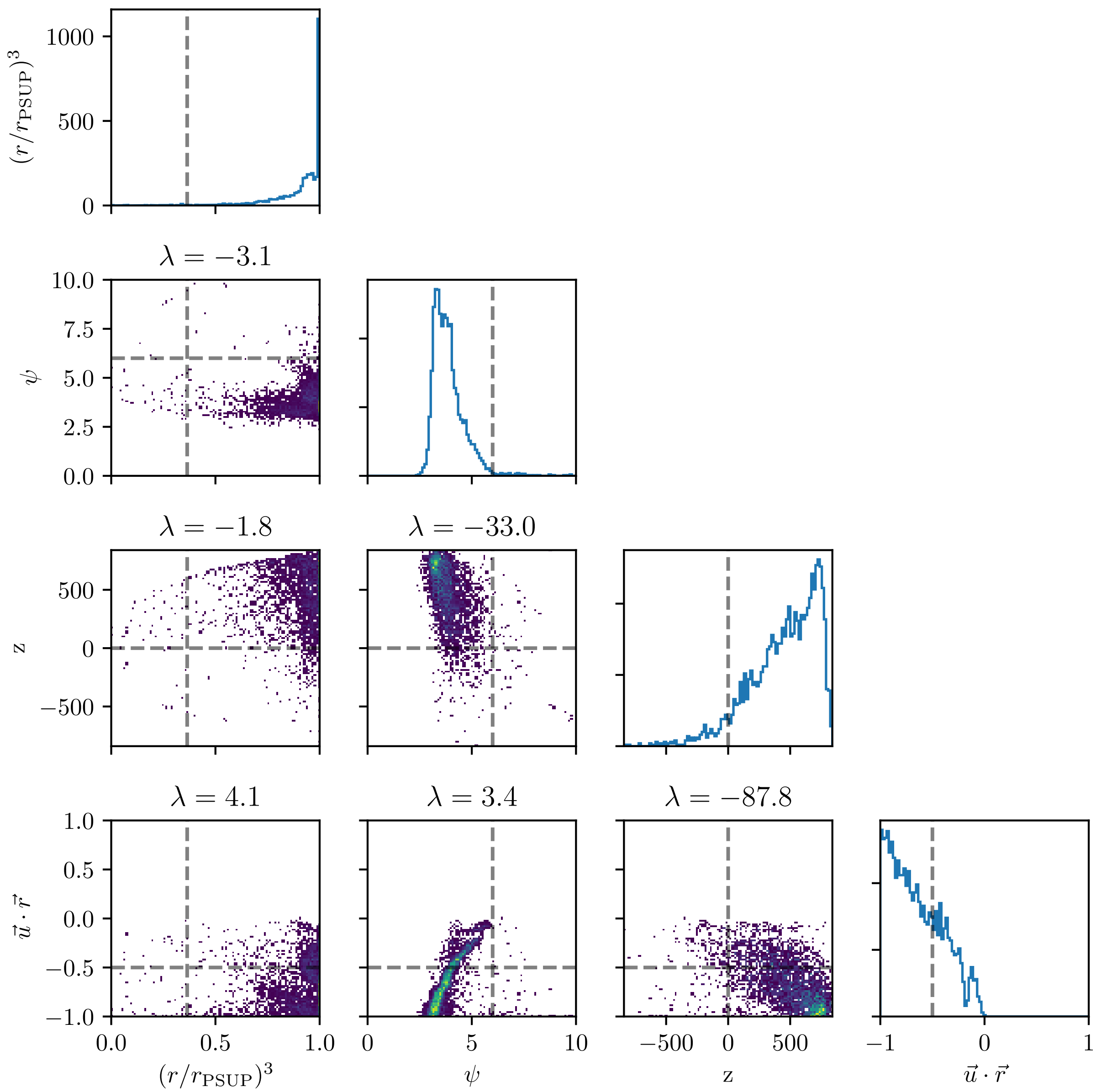

Figure B.1: High Level Variables for Events tagged with the Muon cut

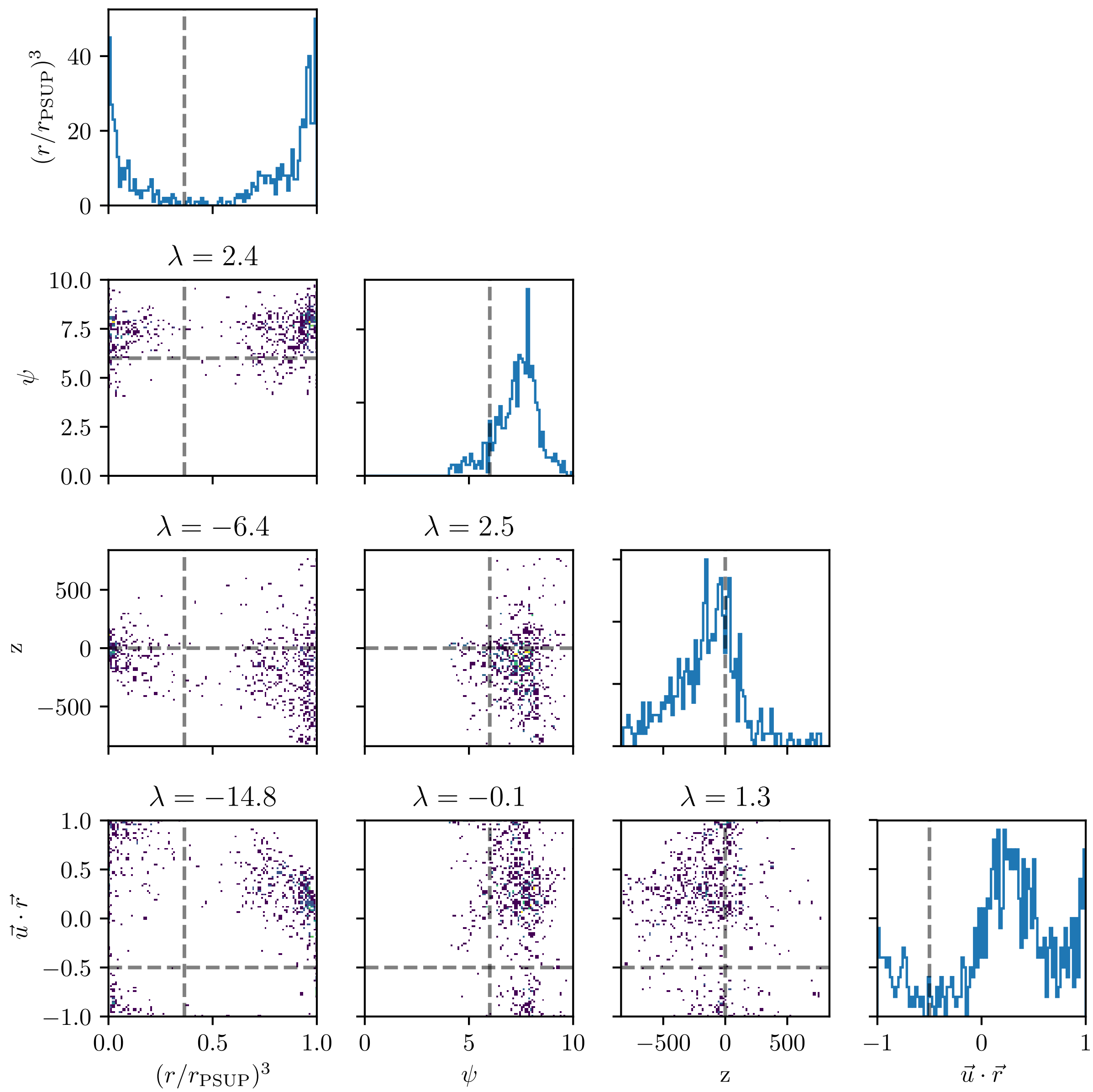

Figure B.2: High Level Variables for Events tagged with the Noise cut

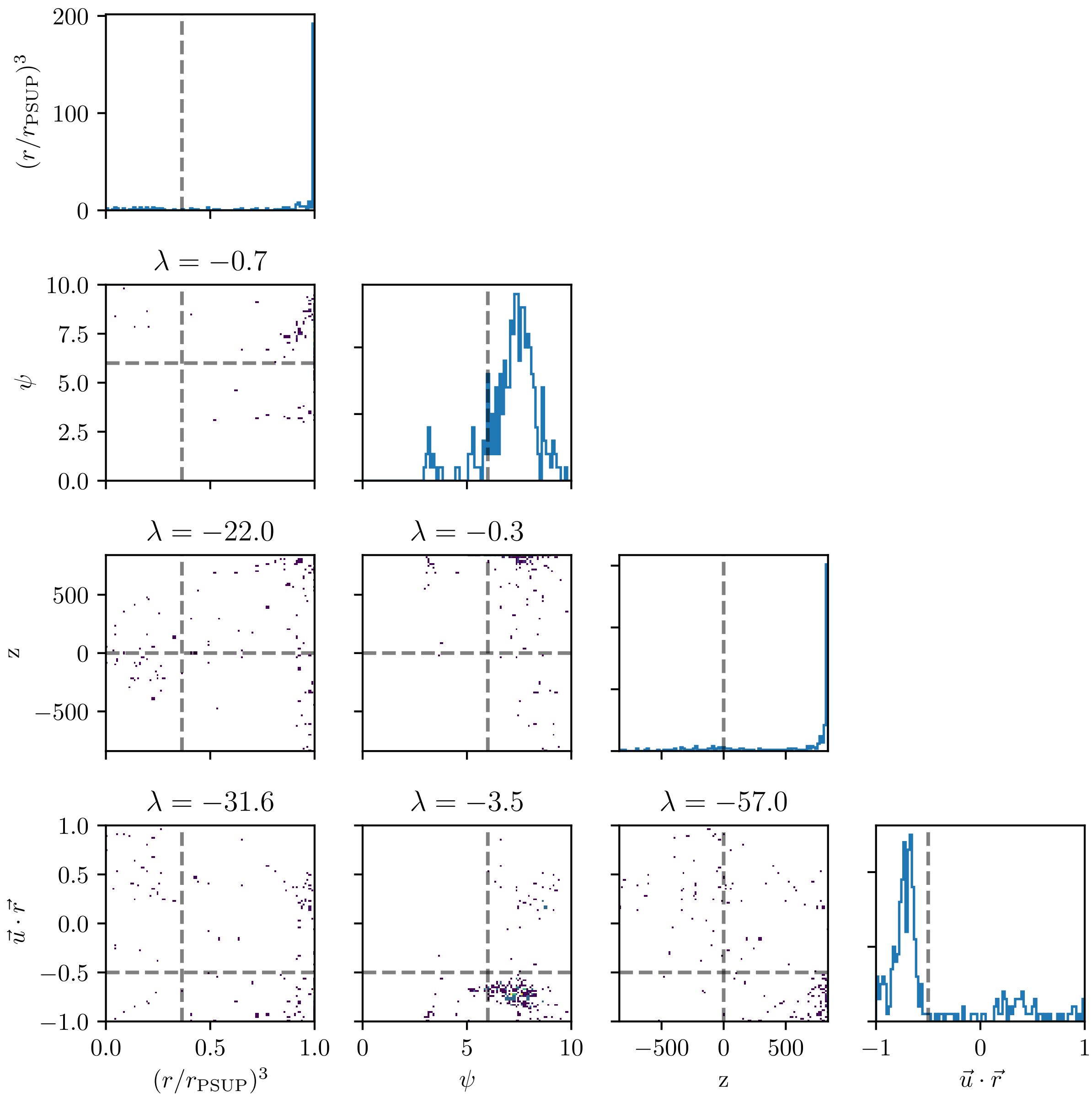

Figure B.3: High Level Variables for Events tagged with the Neck cut

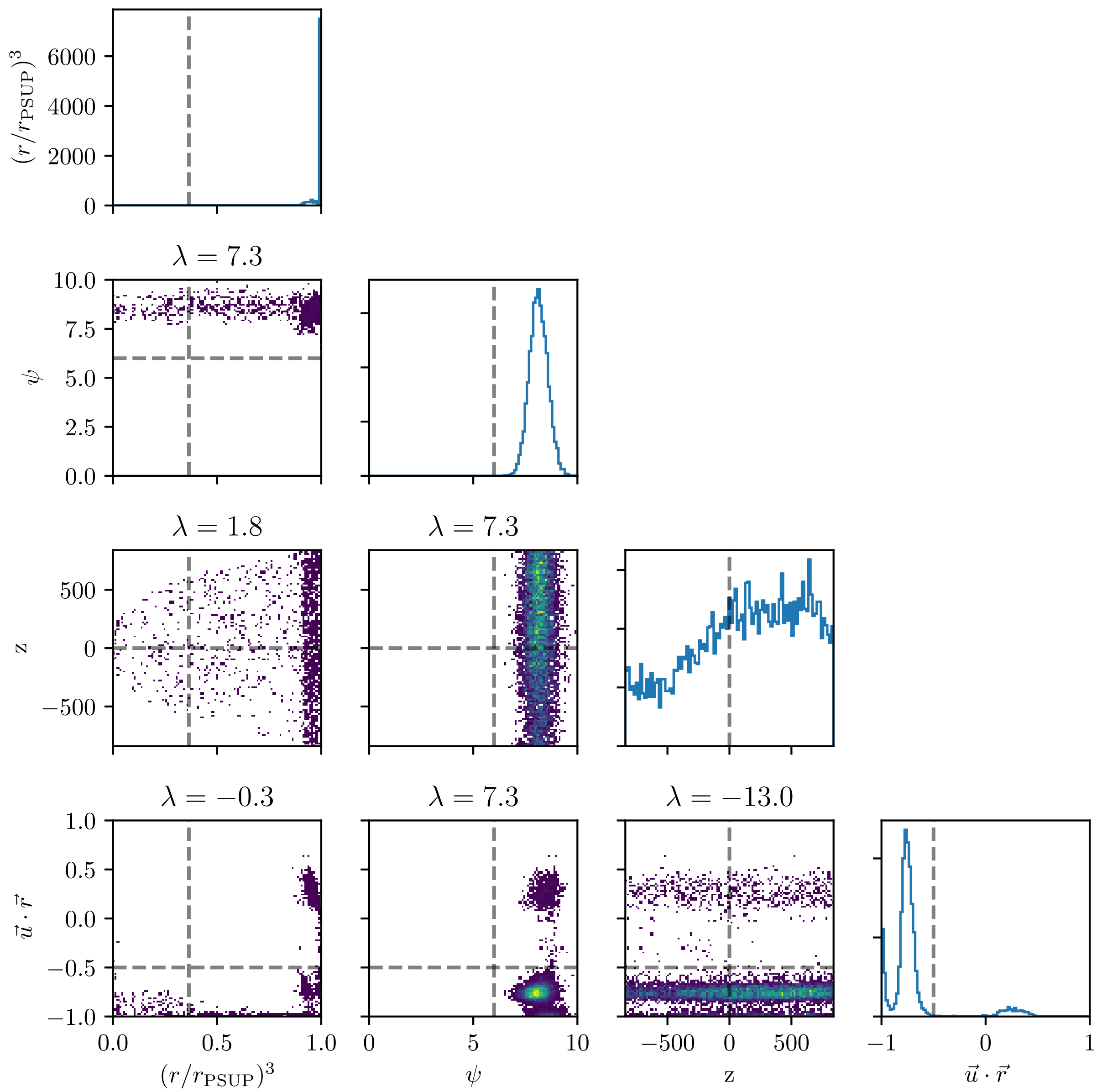

Figure B.4: High Level Variables for Events tagged with the Flasher cut

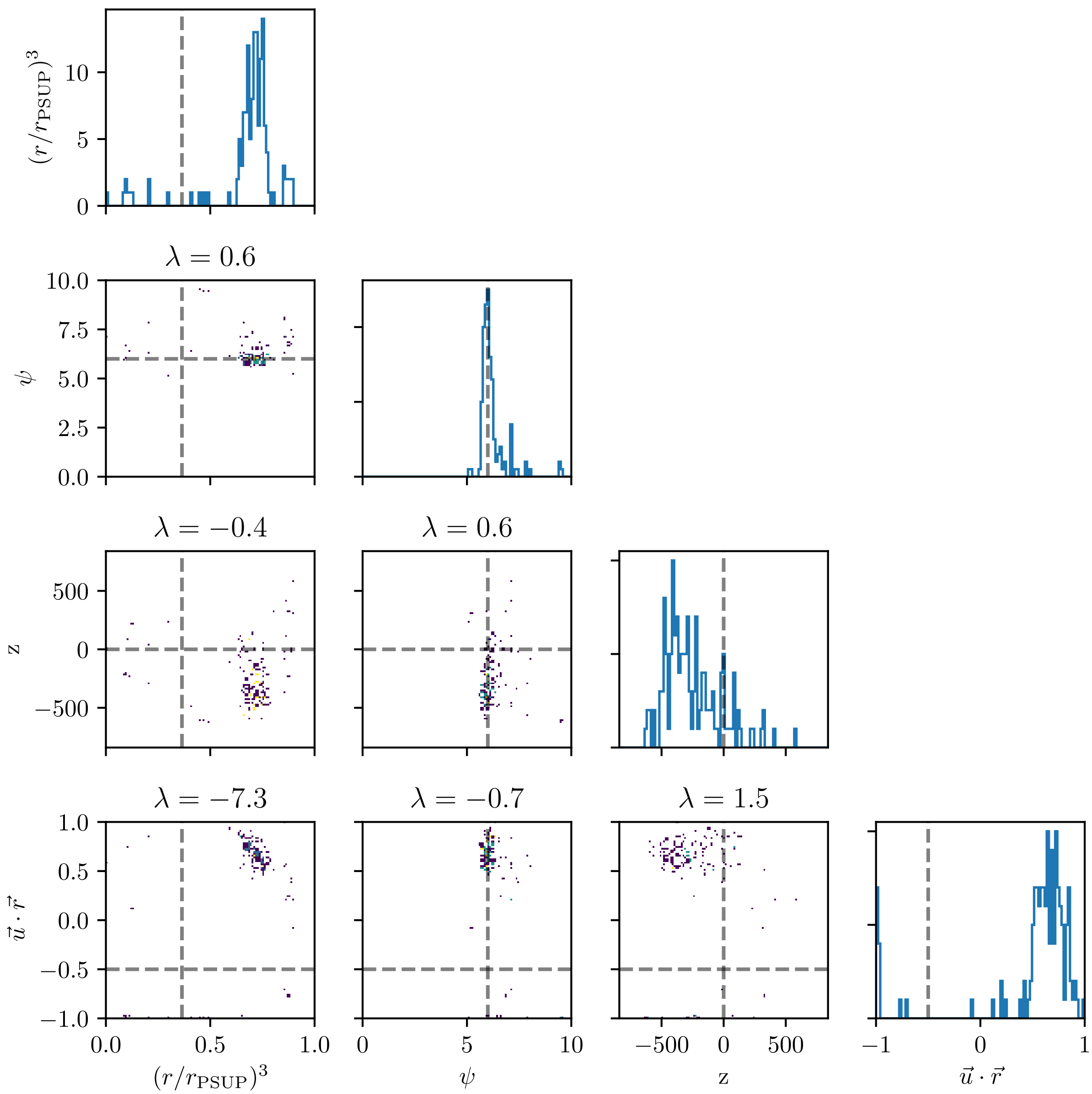

Figure B.5: High Level Variables for Events tagged with the Breakdown cut

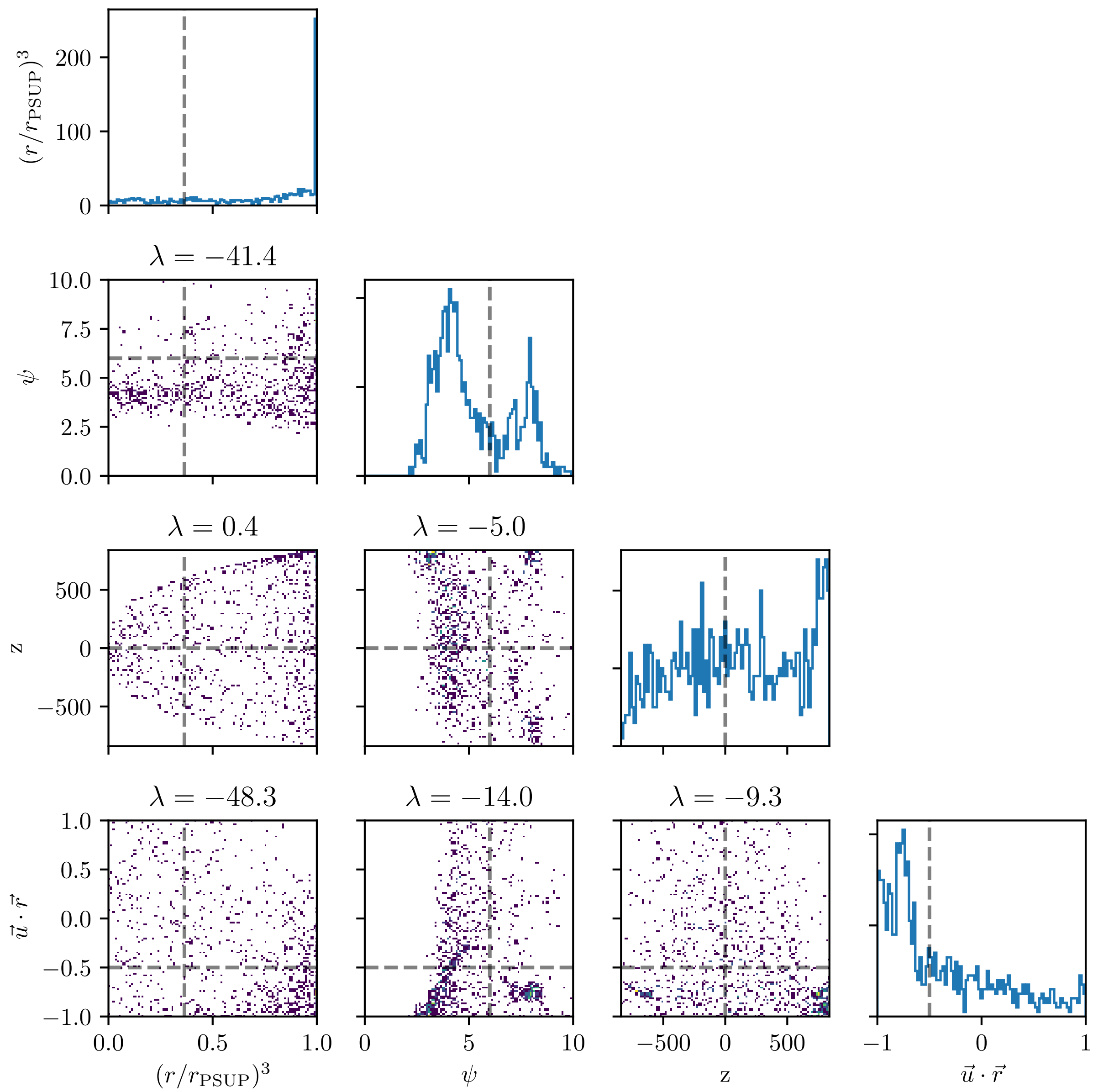

Figure B.6: High Level Variables for Events with No Tags

APPENDIX C

ORPHANS

After initially unblinding 33 runs and running the analysis, I found that there was a large group of instrumental events in run 10141 that were not getting cut by the data cleaning cuts (although they were cut by the ψ cut). I was able to track down the shift report which said:

> Mon Nov 15 15:56:12 1999 Walter went on deck to reflood the bubblers: we've got a lot of activity, and a large number of orphans afterward, things settled down after a couple of minutes. Walter gave us the levels of the bubblers estimated after having refloodeed them: so during a 5 day period

So, this run shouldn't have made it to the run list since the bubblers were active. I therefore decided to do a "poor man's" run selection by cutting runs with a high number of high nhit orphans. Figure C.1 shows the distribution of the number of orphans in a run with greater than 10 PMT hits. I decided to place the cut at 100 orphans since that appears to cover most of the main distribution (although not the tail). Run 10141 had 1199 orphans with at least 10 PMT hits.

Table C.1 shows the number of runs in the "muon neutrino" run lists before and after the orphan cut for the D2O and salt phases. For the D2O phase, this cuts almost 30% of the runs, while for the salt phase it only cuts 2%. In the future, it would be a good idea to study the instrumentals which cause these orphans so that we can expand the run list again.

Phase	Initial # of Runs	# of Runs After Orphan Cut
D2O	603	434
Salt	1628	1590

Table C.1: Number of runs in the run list before and after the orphan cut.

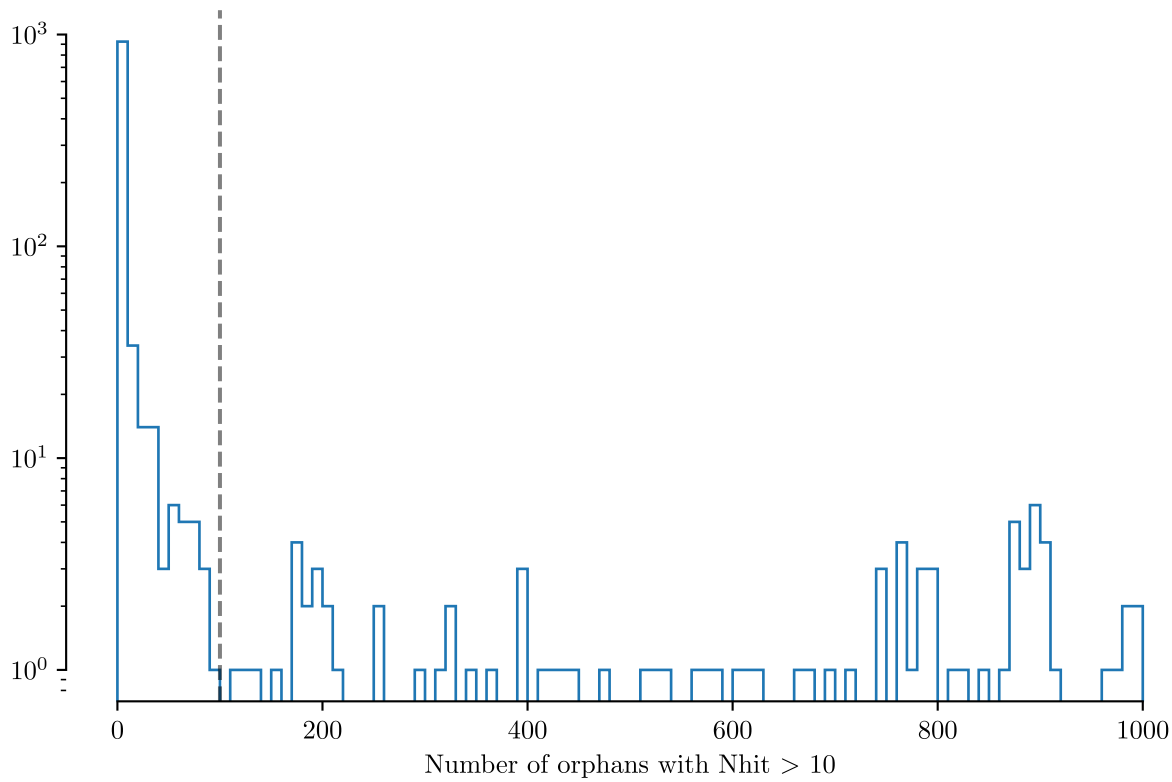

Figure C.1: Distribution of the Number of Orphans with Nhit > 10. The dashed line shows the cut value above which we discard runs from the run list.

APPENDIX D

D2O RUN LIST

10000	10002	10003	10005	10008	10015	10020	10023	10025	10030	10035	10036
10040	10124	10125	10129	10130	10133	10142	10149	10151	10161	10162	10163
10169	10170	10171	10172	10189	10197	10219	10221	10224	10236	10237	10534
10536	10549	10555	10650	10651	10675	10678	10680	10704	10708	10710	10736
10737	10740	10741	10742	10743	10744	10747	10748	10756	10762	10770	10773
10776	10782	10783	10784	10803	10804	10806	10811	10813	10815	10821	10841
10843	10869	10873	10878	10879	10881	10882	10883	10886	10891	10894	10923
10924	10925	10927	10935	10936	10942	10943	10949	10950	10951	10953	10954
10956	10962	10963	10975	11271	11272	11281	11289	11303	11310	11313	11347
11368	11377	11383	11390	11402	11406	11407	11415	11417	11443	11444	11446
11462	11481	11489	11502	11504	11506	11508	11512	11528	11532	11533	11537
11539	11541	11543	11544	11550	11553	11558	11561	11568	11570	11575	11591
11621	11650	11655	11657	11670	11676	11679	11681	11682	11703	11706	11783
11802	11804	11805	11816	11819	11820	11828	11829	11831	11859	11864	11867
11875	11890	11899	11901	11903	11924	11925	11928	11977	11981	11985	11988
11991	12038	12054	12059	12082	12125	12131	12150	12159	12165	12168	12178
12183	12190	12192	12197	12201	12224	12226	12227	12234	12237	12238	12289
12290	12329	12330	12506	12571	12575	12576	12577	12582	12588	12590	12598
12614	12615	12618	13121	13292	13294	13302	13331	13335	13341	13351	13389
13396	13401	13405	13408	13415	13418	13423	13426	13428	13431	13432	13434
13444	13446	13451	13729	13774	13886	13895	14006	14008	14031	14033	14077
14078	14080	14083	14177	14185	14186	14190	14196	14252	14255	14287	14289
14291	14293	14301	14304	14308	14386	14388	14394	14398	14402	14409	14410
14425	14429	14431	14451	14466	14493	14496	14652	14677	14680	14684	14685
14750	14757	14764	14768	14770	14775	14777	14781	14787	14814	14878	14883
14915	14958	14961	14969	14970	15005	15011	15012	15014	15018	15020	15021
15025	15028	15034	15058	15065	15067	15083	15111	15112	15117	15119	15120
15129	15132	15147	15153	15165	15180	15214	15228	15268	15269	15270	15271
15272	15276	15279	15309	15340	15352	15370	15538	15563	15567	15595	15598
15600	15601	15604	15610	15611	15612	15615	15617	15618	15620	15624	15625
15640	15641	15643	15647	15651	15652	15653	15654	15655	15656	15657	15662
15669	15670	15671	15672	15673	15679	15684	15696	15698	15724	15733	15745
15746	15748	15750	15752	15755	15762	15767	15768	15789	15791	15792	15794
15799	15802	15806	15808	15810	15819	15820	15821	15826	15828	15829	15830
15842	15843	15844	15862	15865	15869	15870	15871	15872	15874	15877	15884
15905	15907	15941	15943	15947	15948	15949	15958	15978	15997	15998	16002
16003	16013										

APPENDIX E

SALT RUN LIST

20684	20691	20694	20697	20699	20700	20704	20705	20706	20739	20741	20742
20744	20747	20750	20751	20752	20754	20757	20759	20773	20776	20777	20779
20781	20785	20786	20787	20789	20796	20800	20807	20808	20809	20814	20841
20846	20852	20854	20865	20866	20870	20873	20874	20877	20900	20906	20911
20930	20934	20936	20937	20947	20964	20965	20967	20968	20969	20978	20979
20980	20993	20994	20998	21002	21005	21263	21501	21541	21562	21566	21594
21598	21604	21609	21610	21611	21612	21615	21616	21617	21620	21621	21622
21624	21628	21629	21643	21647	21651	21652	21654	21659	21660	21663	21667
21668	21672	21674	21681	21682	21693	21697	21705	21706	21707	21708	21711
21713	21715	21717	21730	21736	21740	21758	21777	21781	21784	21785	21786
21793	21794	21795	21796	21797	21798	21807	21808	21809	21810	21828	21835
21838	21846	21854	21863	21864	21865	21870	21871	21873	21884	21898	21901
21903	21911	21912	21913	21922	22001	22006	22009	22011	22013	22023	22027
22029	22030	22031	22063	22065	22066	22078	22086	22088	22090	22092	22126
22329	22331	22361	22366	22369	22375	22378	22381	22382	22383	22397	22399
22400	22401	22402	22405	22406	22410	22414	22417	22418	22419	22420	22422
22423	22426	22430	22434	22435	22440	22441	22444	22452	22453	22469	22475
22482	22490	22491	22498	22500	22502	22509	22511	22515	22519	22520	22526
22527	22531	22532	22538	22555	22557	22558	22561	22563	22606	22607	22609
22622	22624	22626	22630	22631	22634	22645	22655	22658	22661	22676	22678
22706	22708	22710	22711	22712	22729	22731	22732	22733	22734	22735	22736
22737	22738	22742	22745	22751	22759	22761	22762	22769	22770	22771	22776
22777	22779	22780	22781	22782	22783	22801	22809	22817	22819	22820	22853
22858	22860	22863	22866	22878	22880	22881	22882	22886	22890	22893	22896
22900	22901	22903	22904	22907	22912	22933	22936	22949	22952	22981	22997
23007	23015	23017	23021	23031	23036	23080	23097	23098	23135	23137	23163
23164	23165	23169	23178	23179	23180	23181	23182	23192	23193	23194	23198
23200	23201	23202	23205	23208	23212	23213	23214	23215	23219	23221	23226
23230	23232	23237	23249	23263	23265	23293	23294	23316	23318	23324	23581
23582	23634	23645	23651	23653	23654	23655	23657	23663	23664	23689	23693
23695	23701	23710	23714	23715	23717	23718	23726	23727	23728	23730	23731
23734	23745	23748	23749	23750	23751	23780	23807	23826	23827	23828	23853
23870	23874	23877	23887	23893	23897	23899	23900	23901	23902	23903	23904
23917	23920	23925	23928	23930	23932	23933	23948	23949	23950	23961	23963
23965	23966	23970	23972	23974	23978	23988	23992	24005	24006	24007	24011
24016	24017	24018	24019	24049	24053	24054	24298	24299	24302	24305	24307
24311	24318	24319	24321	24322	24323	24324	24325	24326	24329	24333	24347

24349	24366	24367	24372	24374	24378	24379	24381	24384	24387	24388	24389
24396	24399	24407	24411	24414	24460	24465	24507	24516	24520	24525	24526
24527	24528	24530	24531	24535	24538	24539	24551	24552	24567	24572	24573
24574	24576	24580	24581	24583	24584	24590	24591	24593	24604	24768	24776
24777	24778	24781	24782	24783	24787	24788	24795	24799	24811	24812	24827
24836	24857	24861	24862	24877	24879	24885	24887	24888	24890	24896	24897
24898	25460	25463	25465	25468	25470	25488	25490	25492	25494	25497	25502
25506	25508	25509	25510	25511	25520	25527	25528	25551	25555	25557	25558
25559	25560	25561	25569	25583	25585	25609	25610	25611	25612	25617	25618
25622	25624	25630	25631	25638	25639	25644	25646	25650	25651	25680	25684
25685	25686	25687	25692	25698	25700	25701	25703	25873	25878	25897	25902
25905	25907	25914	25940	25950	25952	25953	25954	25955	25956	25958	25960
25965	25966	25974	25976	25979	25983	25987	25988	25990	25994	25996	25997
26012	26022	26023	26026	26032	26040	26041	26043	26057	26066	26067	26068
26069	26071	26077	26079	26080	26082	26083	26086	26098	26099	26101	26104
26108	26110	26122	26123	26124	26126	26128	26129	26130	26131	26135	26147
26149	26153	26155	26158	26159	26160	26163	26188	26192	26197	26200	26201
26207	26224	26225	26226	26227	26229	26230	26234	26236	26241	26242	26244
26245	26246	26248	26253	26254	26255	26259	26260	26267	26269	26270	26280
26281	26282	26286	26288	26289	26291	26296	26301	26302	26303	26304	26314
26321	26323	26324	26325	26326	26327	26328	26329	26330	26331	26332	26333
26334	26335	26336	26341	26343	26344	26346	26349	26351	26356	26374	26377
26379	26383	26384	26385	26387	26389	26390	26391	26392	26393	26394	26395
26401	26416	26510	26512	26514	26516	26517	26518	26519	26521	26522	26524
26530	26533	26534	26551	26553	26554	26558	26583	26586	26587	26591	26593
26597	26598	26608	26609	26610	26616	26618	26622	26623	26626	26628	26631
26632	26636	26640	26641	26647	26648	26649	26650	26654	26665	26669	26671
26674	26675	26696	26706	26722	26724	26726	26750	26753	26755	26756	26758
26762	26765	26772	26778	26779	26782	26784	26785	26786	26788	26789	26792
26797	26810	26813	26814	26816	26818	26819	26826	26831	26833	26836	26844
26857	26860	26861	26863	26864	26866	26881	26882	26883	26889	26892	26927
26929	26931	26933	26935	26937	26938	26946	26947	26948	26951	26957	26959
26962	26963	26977	26985	26987	26991	26992	26994	26997	27008	27013	27018
27022	27038	27040	27045	27050	27056	27057	27065	27073	27075	27076	27078
27081	27086	27089	27117	27119	27124	27126	27133	27134	27141	27142	27156
27158	27175	27180	27182	27187	27188	27189	27190	27194	27196	27201	27202
27227	27504	27505	27507	27508	27510	27516	27519	27522	27524	27526	27532
27533	27537	27539	27541	27544	27545	27549	27596	27606	27644	27656	27657
27659	27660	27661	27662	27663	27664	27677	27678	27683	27686	27695	27702
27705	27709	27711	27714	27715	27717	27718	27719	27720	27722	27723	27735

27736	27737	27738	27750	27752	27753	27755	27758	27759	27765	27771	27773
27774	27776	27782	27809	27814	27816	27818	27819	27822	27825	27829	27836
27837	27993	27995	28006	28008	28019	28025	28050	28059	28071	28073	28075
28078	28082	28084	28087	28088	28091	28092	28096	28098	28099	28100	28101
28104	28115	28116	28117	28123	28129	28131	28137	28145	28150	28152	28154
28155	28160	28163	28165	28171	28178	28184	28186	28188	28191	28193	28202
28204	28207	28209	28210	28212	28224	28225	28849	29789	29798	29800	29801
29836	29840	29841	29846	29847	29848	29849	29850	29871	29876	29877	29878
29879	29880	29881	29882	29883	29886	29887	29893	29894	29895	29896	29897
29898	29899	29900	29901	29906	29909	29910	29911	29913	29917	29918	29919
29933	29934	29938	29942	29948	29951	29952	29957	29987	29988	29989	29996
29998	30001	30002	30006	30013	30015	30021	30025	30029	30032	30091	30094
30095	30097	30098	30103	30111	30112	30117	30118	30120	30121	30122	30123
30124	30126	30127	30135	30138	30140	30148	30149	30150	30153	30172	30178
30179	30182	30185	30191	30235	30237	30404	30405	30408	30419	30424	30425
30426	30431	30432	30442	30444	30448	30477	30485	30486	30498	30499	30501
30502	30519	30535	30537	30546	30563	30565	30567	30600	30601	30615	30619
30620	30622	30629	30632	30633	30635	30636	30637	30639	30643	30644	30646
30648	30649	30650	30658	30699	30701	30702	30723	30727	30738	30748	30749
30751	30754	30755	30757	30758	30759	30762	30763	30764	30765	30768	30769
30770	30771	30772	30777	30779	30784	30789	30790	30791	30794	30797	30815
30816	30817	30818	31685	31686	31687	31688	31689	31691	31694	31695	31699
31700	31702	31705	31761	31762	31763	31772	31773	31776	31777	31778	31780
31781	31782	31783	31786	31787	31788	31789	31790	31791	31792	31793	31795
31799	31800	31801	31802	31803	31804	31805	31806	31807	31808	31809	31812
31813	31814	31816	31817	31819	31821	31822	31824	31825	31826	31827	31829
31830	31832	31833	31837	31838	31839	31840	31841	31844	31845	31851	31852
31853	31854	31859	31865	31866	31867	31868	31873	31874	31890	31896	31897
31898	31903	31904	31905	31906	31907	31932	31933	32155	32167	32168	32177
32178	32179	32180	32184	32185	32196	32202	32203	32204	32205	32209	32210
32211	32212	32213	32214	32215	32224	32227	32229	32230	32236	32237	32238
32239	32240	32241	32242	32248	32249	32258	32259	32260	32261	32262	32263
32264	32265	32266	32267	32268	32269	32270	32271	32272	32273	32274	32278
32279	32282	32283	32285	32286	32287	32288	32289	32292	32293	32294	32295
32299	32300	32301	32302	32390	32393	32394	32395	32396	32399	32400	32401
32402	32457	32459	32460	32461	32462	32463	32464	32465	32466	32467	32468
32477	32478	32479	32480	32481	32482	32483	32485	32486	32487	32488	32489
32490	32491	32492	32493	32494	32495	32496	32511	32512	32515	32528	32529
32530	32531	32532	32533	32539	32541	32542	32543	32548	32551	32553	32554
32557	32611	32612	32613	32614	32615	32616	32617	32618	32619	32623	32624

32631	32632	32633	32634	32635	32636	32638	32639	32640	32641	32643	32644
32645	32646	32647	32650	32656	32661	32662	32663	32664	32666	32667	32671
32672	32673	32674	32675	32676	32677	32682	32683	32684	32685	32689	32691
32692	32693	32694	32695	32696	32697	32698	32703	32706	32707	32710	32711
32712	32713	32714	32715	32716	32717	32718	32723	32727	32728	32729	32730
32741	32744	32745	32746	32747	32748	32749	32750	32751	32752	32763	32764
32767	32769	32773	32774	32775	32776	32777	32778	32781	32782	32783	32784
32785	32786	32788	32791	32792	32793	32794	32795	32797	32800	32801	32802
32803	32804	32809	32810	32811	32812	32813	32814	32815	32816	32817	32818
32819	32820	32821	32822	32823	32825	32826	32836	32837	32838	32849	32850
32855	32859	32873	32874	32875	32876	32877	32878	32879	32880	32881	32882
32883	32884	32885	32886	32888	32889	32894	32895	32896	32897	32898	32899
32900	32903	32904	32909	32910	32911	32912	32913	32914	32919	32920	32921
32922	32925	32926	32927	32930	32958	32961	32962	32963	32970	32971	32975
32977	33338	33340	33344	33347	33348	33349	33495	33501	33502	33504	33515
33516	33517	33520	33528	33529	33534						

APPENDIX F

MODEL COMPARISON

In this chapter, I will discuss how I perform a likelihood ratio test to determine whether two high-level variables are independent or not with respect to two cut values on either variable. If the two variables are independent, then we can model them with only two variables: $P(\text{pass 1})$ and $P(\text{pass 2})$. If, however, the two variables are not independent, then we need to use three variables: $P(\text{pass 1}, \text{pass 2})$, $P(\text{pass 1}, \text{fail 2})$, and $P(\text{fail 1}, \text{pass 2})$, where the final probability $P(\text{fail 1}, \text{fail2})$ is determined from the other three. For discussion I will refer to the independent hypothesis as M_I and the not independent hypothesis as M_C (the C is for correlated). To compare these two models I use the likelihood ratio test:

$$\lambda = \frac{P(M_I \mid DI)}{P(M_C \mid DI)} = \frac{P(D \mid M_I, I)P(M_I \mid I)}{P(D \mid M_C, I)P(M_C \mid I)} \tag{F.1}$$

where D stands for data and I for any prior information. This is the same as a traditional likelihood ratio test with the addition of an "Ockham Factor", i.e. we can write the same ratio as

$$\lambda = \frac{P(M_I \mid I)}{P(M_C \mid I)} \frac{(\mathscr{L}_I)_{\max}}{(\mathscr{L}_C)_{\max}} \frac{W_I}{W_C}, \tag{F.2}$$

where $\mathscr{L}$ is the likelihood function, and W for the Ockham factor. The Ockham factor in its most general form is

$$W = \int \frac{\mathscr{L}(\vec{\theta})}{\mathscr{L}_{\max}} P(\vec{\theta} \mid \mathrm{I}) \mathrm{d}\vec{\theta}, \tag{F.3}$$

where $\vec{\theta}$ represents all the variables being fitted for (position, energy, direction, etc.), $\mathscr{L}_{\max}$ represents the maximum value of the likelihood, and $P(\vec{\theta} \mid \mathrm{I})$ represent priors on the variables being fitted for[11].

F.1 Correlated Hypothesis

For the correlated hypothesis, our model has four parameters representing: $P(\text{pass } 1, \text{pass } 2)$, $P(\text{pass } 1, \text{fail } 2)$, $P(\text{fail } 1, \text{pass } 2)$, and $P(\text{fail } 1, \text{fail } 2)$. In addition, there is the constraint that the sum of these probabilities must be equal to one. If we assume no knowledge of these parameters beforehand, a flat prior is given by the Dirichlet distribution

$$P(\vec{p} \mid I) = \frac{1}{B(\vec{\alpha})} \prod_{i=1}^{4} p_i^{\alpha_i - 1} \tag{F.4}$$

with $\vec{\alpha} = (1, 1, 1, 1)$ and where $B(\vec{\alpha})$ is the multivariate beta function given by

$$B(\vec{\alpha}) = \frac{\prod_i \Gamma(\alpha_i)}{\Gamma(\sum_i \alpha_i)}. \tag{F.5}$$

The likelihood is given by the multinomial probability distribution function,

$$\mathscr{L}(\vec{p}) = \frac{N!}{n_1! \cdots n_k!} p_1^{n_1} \cdots p_k^{n_k} \tag{F.6}$$

Since the Dirichlet distribution is the conjugate prior of the multinomial distribution, the posterior will also be a Dirichlet distribution. The posterior in this case is given by the Dirichlet distribution with $\vec{\alpha}$ equal to the previous vector plus the observed counts in each box, i.e.

$$P(\vec{p} \mid D, I) = \frac{1}{B(\vec{\alpha})} \prod_{i=1}^{4} p_i^{\alpha_i - 1} \tag{F.7}$$

with $\vec{\alpha} = (n_1 + 1, n_2 + 1, n_3 + 1, n_4 + 1)$.

F.2 Independent Hypothesis

For the independent hypothesis we have only two parameters: $P(\text{pass } 1)$ and $P(\text{pass } 2)$ which we will denote by π_1 and π_2. The prior in this case is simply two flat distributions between 0 and 1. The likelihood is the same as in the previous section except we set $p_1 = \pi_1 \pi_2$,

$p_2 = \pi_1(1-\pi_2)$, $p_3 = (1-\pi_1)\pi_2$, and $p_4 = (1-\pi_1)(1-\pi_2)$. Plugging these in, we get

$$\mathscr{L}(\pi_1, \pi_2) = \frac{N!}{n_1!n_2!n_3!n_4!}(\pi_1\pi_2)^{n_1}(\pi_1(1-\pi_2))^{n_2}((1-\pi_1)\pi_2)^{n_3}((1-\pi_1)(1-\pi_2))^{n_4}. \quad \text{(F.8)}$$

Rearranging the terms we get

$$\mathscr{L}(\pi_1, \pi_2) = \frac{N!}{n_1!n_2!n_3!n_4!}\pi_1^{n_1+n_2}(1-\pi_1)^{n_3+n_4}\pi_2^{n_1+n_3}(1-\pi_2)^{n_2+n_4}. \quad \text{(F.9)}$$

Since the prior is flat, the likelihood is proportional to the posterior distribution. In this case, since the likelihood is proportional to the product of two beta distributions, we can immediately determine that

$$P(\pi_1, \pi_2 \mid M_I, D, I) = f(\pi_1; n_1+n_2+1, n_3+n_4+1)f(\pi_2; n_1+n_3+1, n_2+n_4+1), \quad \text{(F.10)}$$

where $f(x;\alpha,\beta)$ is the beta distribution,

$$f(x;\alpha,\beta) \propto x^{\alpha-1}(1-x)^{\beta-1}. \quad \text{(F.11)}$$

F.3 Likelihood Ratio

In order to compute the likelihood ratio λ between the two models we need to compute $P(M \mid D, I)$ for each of the models. One way to compute this quantity is to expand it over the model parameters $\vec{\theta}$

$$P(M \mid DI) = \int \mathscr{L}(\vec{\theta})P(\vec{\theta}|I)\mathrm{d}\vec{\theta}. \quad \text{(F.12)}$$

We can then approximate this integral by sampling values of $\vec{\theta}$ from the prior and computing

$$P(M \mid DI) \approx \frac{1}{n}\sum \mathscr{L}(\vec{\theta}) \quad \text{(F.13)}$$

where n is the total number of samples and the samples are drawn from the prior distribution,

$P(\vec{\theta}|I)$. However, this is an extremely inefficient means of computing the quantity and, for more than a few parameters, becomes computationally infeasible. If, however you can approximate the posterior with a function $f(\vec{\theta})$, i.e.

$$P(\vec{\theta} \mid D, I) \approx f(\vec{\theta}), \tag{F.14}$$

then we can avoid sampling very unlikely regions by computing

$$P(M \mid DI) \approx \frac{1}{n} \sum \frac{\mathscr{L}(\vec{\theta}) P(\vec{\theta} \mid I)}{f(\vec{\theta})} \tag{F.15}$$

where the sum is over samples drawn from $f(\vec{\theta})$[28].

In our case, however, since we already know the analytic form for the posterior, likelihood, and prior we can directly calculate

$$P(D \mid M, I) = \frac{P(D \mid \vec{\theta}, M, I) P(\vec{\theta} \mid M, I)}{P(\vec{\theta} \mid D, M, I)}. \tag{F.16}$$

For the correlated hypothesis:

$$\begin{aligned} P(D \mid M, I) &= \frac{P(D \mid \vec{\theta}, M, I) P(\vec{\theta} \mid M, I)}{P(\vec{\theta} \mid D, M, I)} \\ &= \frac{6}{(N+3)(N+2)(N+1)}. \end{aligned}$$

For the independent hypothesis:

$$\begin{aligned} P(D \mid M, I) &= \frac{P(D \mid \vec{\theta}, M, I) P(\vec{\theta} \mid M, I)}{P(\vec{\theta} \mid D, M, I)} \\ &= \frac{(n_1+n_2)!(n_3+n_4)!(n_1+n_3)!(n_2+n_4)!}{(N+1)(N+1)!n_1!n_2!n_3!n_4!}. \end{aligned}$$

The likelihood ratio is then given by

$$
\begin{aligned}
\Lambda &= \frac{P(M_I \mid D, I)}{P(M_C \mid D, I)} \\
&= \frac{P(D \mid M_I, I) P(M_I \mid I)}{P(D \mid M_C, I) P(M_C \mid I)} \\
&= \frac{(n_1+n_2)!(n_3+n_4)!(n_1+n_3)!(n_2+n_4)!}{(N+1)! n_1! n_2! n_3! n_4!} \frac{(N+3)(N+2)}{6} \frac{P(M_I \mid I)}{P(M_C \mid I)}.
\end{aligned}
$$

And, assuming both models are equally likely:

$$
\Lambda = \frac{(n_1+n_2)!(n_3+n_4)!(n_1+n_3)!(n_2+n_4)!}{(N+1)! n_1! n_2! n_3! n_4!} \frac{(N+3)(N+2)}{6}. \tag{F.17}
$$

Finally, we will use the log of the likelihood ratio to characterize the test:

$$
\lambda = \log \left(\frac{(n_1+n_2)!(n_3+n_4)!(n_1+n_3)!(n_2+n_4)!}{(N+1)! n_1! n_2! n_3! n_4!} \frac{(N+3)(N+2)}{6} \right). \tag{F.18}
$$

For values of the log likelihood ratio that are greater than zero, the ratio favors the hypothesis that the two models are independent and for values less than zero it favors the hypothesis that they are not independent.

APPENDIX G

TEST STATISTIC

The test statistic used when calculating p-values is based on calculating a likelihood ratio which is recommended in [29]. The likelihood of observing a histogram O_i from an expected histogram E_i is given by:

$$\mathcal{L}(O_i \mid E_i) = e^{-N}\frac{N^n}{n!}\frac{n!}{\prod_i O_i!}\prod_i \left(\frac{E_i}{N}\right)^{O_i} \tag{G.1}$$

where N is the total number of expected events, and n is the total number of observed events. The likelihood is just the product of the probabilities of observing n events given you expected N events and the probability of observing O_i events from a multinomial distribution with probabilities given by the expected distribution.

First, it is useful to calculate the log of the likelihood:

$$\log\left(\mathcal{L}(O_i \mid E_i)\right) = -N - \sum_i \log(O_i!) + \sum_i O_i \log(E_i) \tag{G.2}$$

The test statistic is created by taking the negative log of a likelihood ratio λ. The likelihood in the numerator is equal to the likelihood of observing the data given the expected distribution, while the likelihood in the denominator is equal to the likelihood of observing the data assuming the expected distribution is equal to the data.

$$\log\lambda = \log\left(\frac{\mathcal{L}(O_i \mid E_i)}{\mathcal{L}(O_i \mid O_i)}\right) = -N + \sum_i O_i \log\left(\frac{E_i}{O_i}\right) + n \tag{G.3}$$

Finally, to facilitate interpretation, it's convenient to return -2 times the log likelihood ratio since that asymptotically approaches the χ^2 of the two distributions:

$$-2\log\lambda = 2\left(n - N + \sum_i O_i \log\left(\frac{E_i}{O_i}\right)\right) \tag{G.4}$$

APPENDIX H

SIMULATION-BASED CALIBRATION

When doing a complex fit, it is useful to be able to verify that the likelihood function or posterior distribution is self-consistent and free of any logical or coding errors. When doing an analysis where all of the priors are flat and the errors on the parameters are assumed to be Gaussian, it is common in high energy physics to make a “pull plot”. This plot is produced by simulating the data with known true values, running the fit, and calculating the difference between the fit value and the true value divided by the error on the value returned by the fit. Assuming the likelihood has Gaussian errors, the resulting value should be distributed as a Gaussian with mean zero and variance one. By repeating this process over and over and histogramming the results you get what is called a “pull plot” which allows you to easily spot biases or problems with the error returned by the fit.

However, if the likelihood function or posterior does not have Gaussian errors or the posterior includes non-flat priors, the pull plots may appear biased or skewed simply because the errors are not Gaussian or because of the prior terms. In this case, we can still check the consistency of the likelihood by making an “SBC plot”. The procedure for making these plots is as follows[30]:

1. Draw truth values for all parameters in the fit from the priors
2. Simulate data based on these truth values
3. Run an MCMC to sample values from the posterior
4. Compute the percentile of the truth value in the MCMC samples for each of the parameters

Other considerations, like thinning the MCMC samples are discussed in Reference [30]. Then, you can plot a histogram of the percentiles for each parameter. Assuming the fit is

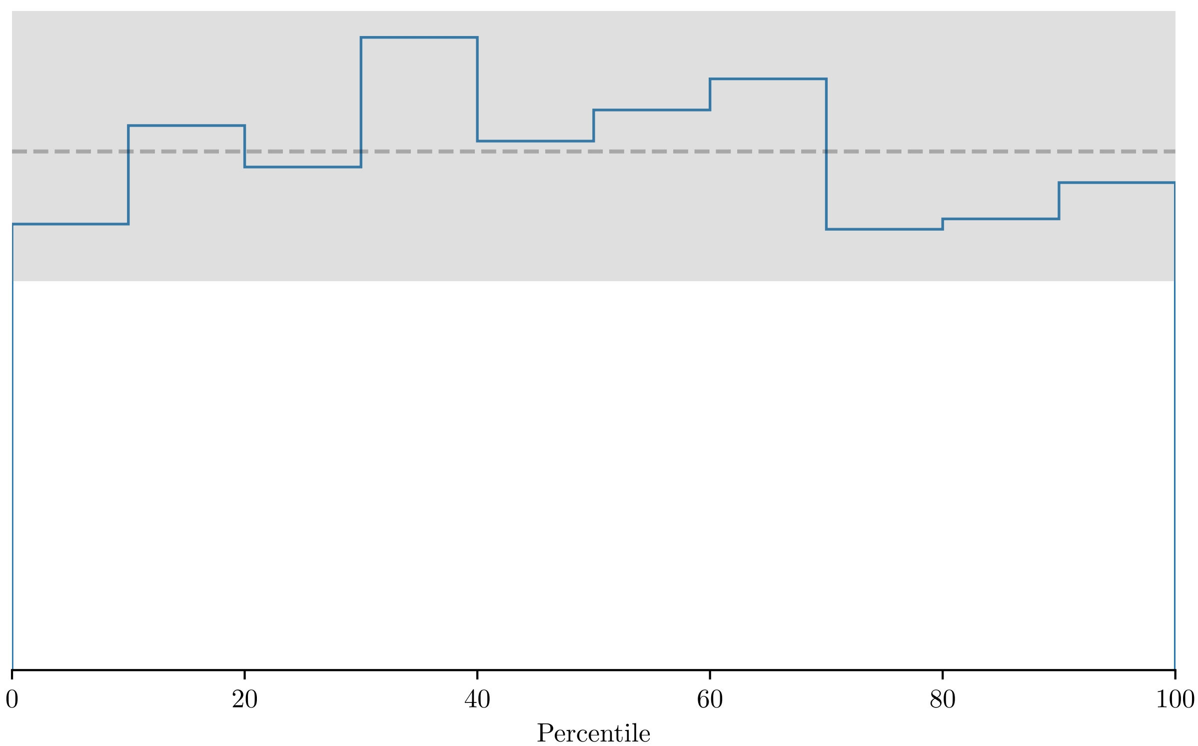

Figure H.1: Example of a properly calibrated SBC plot. The dashed line shows the expected flat distribution and the grey band shows where we expect 99% of the bin contents to fall within assuming a flat distribution.

unbiased, the resulting histograms should all be a flat distribution between 0% and 100%. This works because the data averaged posterior is equal to the prior distribution[30], i.e.

$$\pi(\theta) = \int \mathrm{d}\tilde{y}\ \mathrm{d}\tilde{\theta}\ \pi(\theta \mid \tilde{y})\pi(\tilde{y} \mid \theta)\pi(\tilde{\theta}). \tag{H.1}$$

Cook, Gelman, and Rubin then showed that the quantiles for each parameter will be uniformly distributed provided that the posteriors are absolutely continuous[31].

A properly calibrated SBC plot is shown in Figure H.1, while a poorly calibrated plot is shown in Figure H.2.

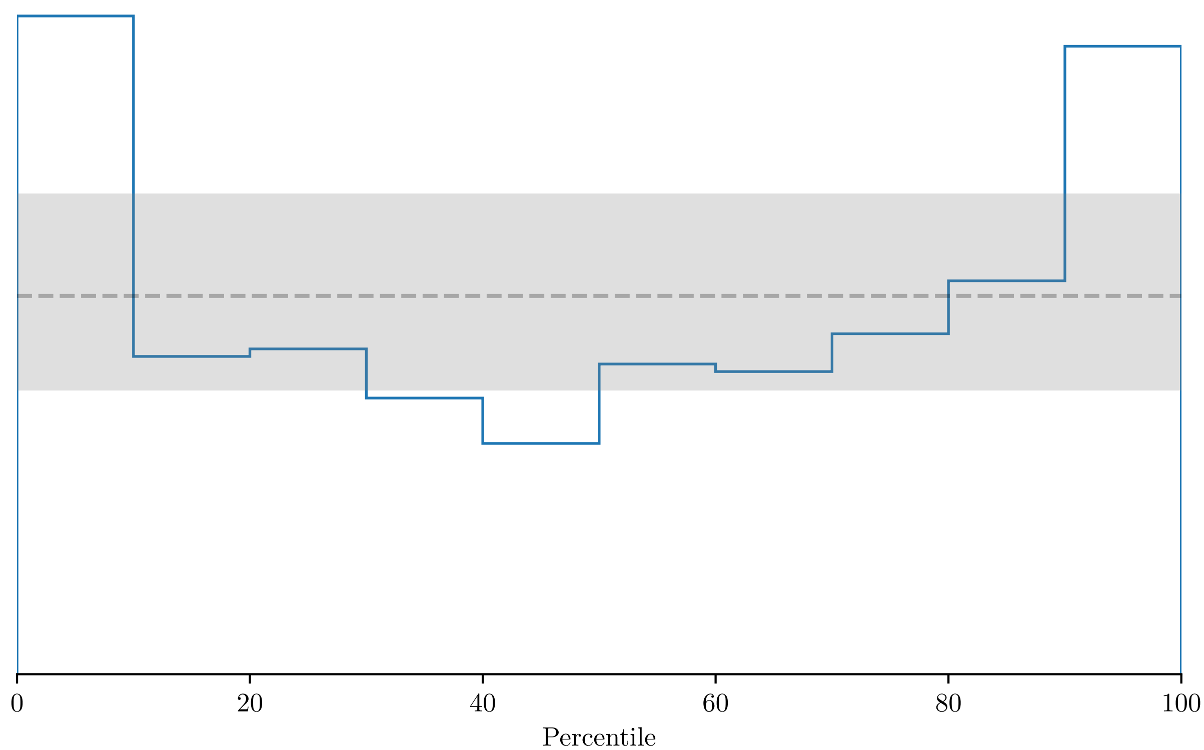

Figure H.2: Example of a poorly calibrated SBC plot. In this case, the posterior was modified to return the χ^2 instead of the negative log of the posterior. The dashed line shows the expected flat distribution and the grey band shows where we expect 99% of the bin contents to fall within assuming a flat distribution.

APPENDIX I

DISCOVERY THRESHOLD

When doing the direct dark matter search, I would like to come up with a "discovery threshold" prior to doing the fit. What we would like is to specify a probability of discovering dark matter under the null hypothesis, p, and then determine what that means in terms of our fit results. Ideally, for each dark matter mass, I would simulate the null hypothesis many times, look at the distribution of the best fit event rate for the dark matter, and then choose the value such that only p percent of the best fit values constitute a discovery. However, this would be very computationally expensive, and so I have come up with an alternative approach which is approximately correct in the limit that the energy resolution is smaller than the bin size.

The basic idea is that if the energy resolution is much smaller than the bin size, then we expect most of the dark matter events to fall within a single bin. In that case, under the null hypothesis, we expect the number of events in the bin to be distributed according to a Poisson distribution:

$$P(n) = e^{-\mu}\frac{\mu^n}{n!} \tag{I.1}$$

where μ is the expected number of MC events in the bin.

The likelihood for this bin including the dark matter term, λ will look like:

$$P(n) = e^{-(\mu+\lambda)}\frac{(\mu+\lambda)^n}{n!}. \tag{I.2}$$

Given n events, the best fit value for λ will be:

$$\lambda = \begin{cases} n-\mu & \text{if } n > \mu \\ 0 & \text{if } n \leq \mu \end{cases} \tag{I.3}$$

Since we expect n to be distributed according to a Poisson as shown in Equation (I.1), then we immediately know the distribution for λ since μ is a constant.

Therefore, the event rate limit for a false positive rate of p is

$$\lambda > Q(p, \mu) - \mu \tag{I.4}$$

where $Q(p, \mu)$ is the percent-point function of the Poisson distribution with mean μ.

There are two extra small caveats. First, to account for the look elsewhere effect, we divide p by the number of bins in the analysis, since each bin is approximately a different potential test. Second, since it is possible to have a dark matter mass at the boundary of two bins, we calculate μ as

$$\mu = \frac{\sum_i \vec{h}_{\mathrm{MC},i} \cdot \vec{h}_{\mathrm{DM},i}}{\sum_i \vec{h}_{\mathrm{DM},i}} \tag{I.5}$$

where $\vec{h}_{\mathrm{MC},i}$ is the expected number of atmospheric MC events in bin i, and $\vec{h}_{\mathrm{DM},i}$ is the expected distribution for a single dark matter event. This has the nice property of smoothly interpolating between two bins.

Lastly, when testing this out in a toy Monte Carlo program, I realized that I also have to divide the discovery threshold by the fraction of the normalized dark matter histogram which is in the range of the histogram. This corrects for the fact that near the edge of the histogram we expect to get a higher best fit event rate since only a fraction of the expected dark matter events are in the first or last bin.

So, in the end we calculate our discovery limit as

$$\lambda > \frac{Q(p/N, \mu) - \mu}{\sum_i \vec{h}_{\mathrm{DM},i}}. \tag{I.6}$$

where N is the number of bins.

APPENDIX J

BEST UNCALIBRATED CHARGE

When doing data cleaning cuts it is often necessary to use uncalibrated charges to make the cuts more robust. For example, in many instrumental events the TAC may be too low or high causing the PCA calibration to fail. It would be nice to have a single uncalibrated charge to use however so that it's not necessary to deal with QHS and QLX separately. Therefore, I defined a routine which returns the "best" uncalibrated charge which is used in my data cleaning cuts.

This routine does two main things. First, it checks if QHS is railed, in which case it uses QLX, and secondly it renormalizes the uncalibrated QLX into "QHS units". This renormalization is necessary since we are working in ECA calibrated counts above pedestal and the QHS and QLX paths have different gains.

The full algorithm is shown in Algorithm 2.

Algorithm 2 Best Uncalibrated Charge Algorithm

```
/* Average gain between QLX and QHS.
 *
 * Note: This is not an accurate number it's just something I got from the SNO
 * document "SNO Electronic Calibration Constants" and confirmed by looking at
 * a plot of QHS vs QLX. */
static double QLX_TO_QHS = 12.0;

/* Returns the "best" ECA calibrated charge (in units of EHS).
 *
 * This is sort of modelled after the SNOMAN routine cal_best_q() but here we
 * do it for uncalibrated charges (in counts above pedestal). One issue here is
 * that QHS and QLX have different gains, so to deal with that we just assume a
 * constant gain and multiply any QLX values by QLX_TO_QHS.
 *
 * This best uncalibrated charge is mainly useful for data cleaning cuts. */
double cal_best_q(float pihs, float pilx, float ehs, float elx)
{
    if (pilx < 300|| pilx > 4000|| pihs < 300|| pihs > 4000) {
        /* QHS or QLX is railed, so use QLX. */
        if (pilx < 300)
            return 4095.0*QLX_TO_QHS;
        return elx*QLX_TO_QHS;
    }

    return ehs;
}
```

APPENDIX K

BIFURCATED ANALYSIS

In SNO the instrumental contaminations were measured using a technique called the "bifurcated analysis"[32]. The LETA unidoc describes the general problem as:

> We have two independent cuts. Given that d events do not pass either cut, c events pass only the first cut and b events only pass the second cut, how can we determine the number of background events a that pass both cuts in a blind fashion?

A derivation of the results can be found in the LETA unidoc, but is also reproduced here. We can write the expected number of events as[1]:

$$\begin{pmatrix} N_s + a \\ b \\ c \\ d \end{pmatrix} = \begin{bmatrix} P(\text{pass 1 pass 2} \mid \text{signal}) & P(\text{pass 1 pass 2} \mid \text{background}) \\ P(\text{pass 1 fail 2} \mid \text{signal}) & P(\text{pass 1 fail 2} \mid \text{background}) \\ P(\text{fail 1 pass 2} \mid \text{signal}) & P(\text{fail 1 pass 2} \mid \text{background}) \\ P(\text{fail 1 fail 2} \mid \text{signal}) & P(\text{fail 1 fail 2} \mid \text{background}) \end{bmatrix} \begin{pmatrix} N_s \\ N_b \end{pmatrix} \tag{K.1}$$

where N_s and N_b are the true number of signal and background events respectively.

The first step is to assume that the two cuts are independent for both the signal and background events:

$$\begin{pmatrix} N_s + a \\ b \\ c \\ d \end{pmatrix} = \begin{bmatrix} P(\text{pass 1} \mid \text{signal})P(\text{pass 2} \mid \text{signal}) & \epsilon_1\epsilon_2 \\ P(\text{pass 1} \mid \text{signal})P(\text{fail 2} \mid \text{signal}) & \epsilon_1(1-\epsilon_2) \\ P(\text{fail 1} \mid \text{signal})P(\text{pass 2} \mid \text{signal}) & (1-\epsilon_1)\epsilon_2 \\ P(\text{fail 1} \mid \text{signal})P(\text{fail 2} \mid \text{signal}) & (1-\epsilon_1)(1-\epsilon_2) \end{bmatrix} \begin{pmatrix} N_s \\ N_b \end{pmatrix}.$$

where

1. Here we've already assumed that there is a negligible sacrifice since we set the number of events passing both to $N_s + a$. This assumption is made explicit later.

$$\epsilon_1 = P(\text{pass } 1 \mid \text{background})$$
$$\epsilon_2 = P(\text{pass } 2 \mid \text{background})$$

Now, we assume that the sacrifice for each cut is negligible, i.e.

$$\begin{aligned}
P(\text{pass } 1 \mid \text{signal}) &\simeq 1 \\
P(\text{pass } 2 \mid \text{signal}) &\simeq 1 \\
P(\text{ fail } 1 \mid \text{signal}) &\simeq 0 \\
P(\text{ fail } 2 \mid \text{signal}) &\simeq 0
\end{aligned}$$

and the equation therefore becomes:

$$\begin{pmatrix} N_s + a \\ b \\ c \\ d \end{pmatrix} = \begin{bmatrix} 1 & \epsilon_1\epsilon_2 \\ 0 & \epsilon_1(1-\epsilon_2) \\ 0 & (1-\epsilon_1)\epsilon_2 \\ 0 & (1-\epsilon_1)(1-\epsilon_2) \end{bmatrix} \begin{pmatrix} N_s \\ N_b \end{pmatrix}. \tag{K.2}$$

From this we can derive the main result[2] which is that the number of background events passing both cuts a, is

$$a = \frac{bc}{d}. \tag{K.3}$$

The most important assumption which enables this analysis to work is the assumption of independence between the low and high-level cuts. Although it's not specified directly, an important point to consider is whether the low and high-level cuts are assumed to be independent *for a given background source* or *globally*. The first case of having the low and high-level cuts be independent for a given background source is certainly not obvious[3], but

2. Technically one should treat the left hand side of Equation (K.2) as expected values for a Poisson distribution and do a likelihood fit in order to properly estimate the contamination and the associated uncertainty.

3. Flasher events are cut by looking for a cluster of PMT hits and looking for the majority of the other

can be argued for certain cuts and can also be explicitly demonstrated if you make some assumptions about how the cut might fail. For example, to check if the low-level cuts are independent of the high-level cuts for flasher events, one can take known flashers and remove the PMT hits in the flashing channel to simulate a blind flasher[4]. Then the independence between the high and low level cuts can be demonstrated by showing that the high-level variables do not change when removing the hits near the flashing channel (assuming that this is the most likely reason for a flasher to not get tagged). However even if we do assume the independence of the low and high-level cuts for a given class of backgrounds, I will show that the overall probability for the backgrounds combined is *not* independent.

For a simple example, consider flasher events and pickup events and assume the following:

$$\begin{aligned}
N_{\text{flasher}} &= 1000 \\
N_{\text{pickup}} &= 10 \\
P(\text{pass } 1 \mid \text{flasher}) &= 0.2 \\
P(\text{pass } 2 \mid \text{flasher}) &= 0.99 \\
P(\text{pass } 1 \mid \text{pickup}) &= 0.001 \\
P(\text{pass } 2 \mid \text{pickup}) &= 0.1
\end{aligned}$$

In this case, the expected number of background events that pass both cuts, a, is:

$$\begin{aligned}
a &= N_{\text{flasher}} P(\text{pass } 1 \mid \text{flasher}) P(\text{pass } 2 \mid \text{flasher}) + N_{\text{pickup}} P(\text{pass } 1 \mid \text{pickup}) P(\text{pass } 2 \mid \text{pickup}) \\
&= 1000 \times 0.2 \times 0.99 + 10 \times 0.001 \times 0.1 \\
&\simeq 198
\end{aligned}$$

PMT hits in the event to be farther away. The events which sneak past the cut occur because, for whatever reason, the flashing channel and the nearby channels don't show up in the event. This is called a "blind flasher". One reason why the low and high-level cuts might be correlated is that the high charge channel significantly affects the reconstruction.

4. A blind flasher is a flasher event where the channel producing the flasher is not read out for some reason and thus there is no high charge channel on the opposite side of the majority of the light to tag.

However, following the bifurcated analysis technique we would calculate:

$$\begin{aligned}
b &= N_{\mathrm{B}} P(\text{pass 1 } \overline{\text{ fail }} \text{ 2} \mid \text{background}) \\
&= N_{\text{flasher}} P(\text{pass 1} \mid \text{flasher}) P(\text{ fail } \text{ 2} \mid \text{flasher}) + N_{\text{pickup}} P(\text{pass 1} \mid \text{pickup}) P(\text{ fail } \text{ 2} \mid \text{pickup}) \\
&= 1000 \times 0.2 \times 0.01 + 10 \times 0.001 \times 0.9 \\
&\simeq 2
\end{aligned}$$

$$\begin{aligned}
c &= N_{\mathrm{B}} P(\text{ fail } \text{ 1 pass 2} \mid \text{background}) \\
&= N_{\text{flasher}} P(\text{ fail } \text{ 1} \mid \text{flasher}) P(\text{pass 2} \mid \text{flasher}) + N_{\text{pickup}} P(\text{ fail } \text{ 1} \mid \text{pickup}) P(\text{pass 2} \mid \text{pickup}) \\
&= 1000 \times 0.8 \times 0.99 + 10 \times 0.999 \times 0.1 \\
&\simeq 792
\end{aligned}$$

$$\begin{aligned}
d &= N_{\mathrm{B}} P(\text{ fail } \text{ 1 } \text{ fail } \text{ 2} \mid \text{background}) \\
&= N_{\text{flasher}} P(\text{ fail } \text{ 1} \mid \text{flasher}) P(\text{ fail } \text{ 2} \mid \text{flasher}) + N_{\text{pickup}} P(\text{ fail } \text{ 1} \mid \text{pickup}) P(\text{ fail } \text{ 2} \mid \text{pickup}) \\
&= 1000 \times 0.8 \times 0.01 + 10 \times 0.999 \times 0.9 \\
&\simeq 16
\end{aligned}$$

and therefore estimate a as

$$a = \frac{bc}{d} \simeq \frac{2 \times 792}{16} \simeq 93$$

Thus we would have underestimated the background by a factor of 2!

The reason the bifurcated analysis does not work is because **even if you assume independence between the low-level and high-level cuts for each background source independently, the low and high-level cuts will not in general be independent for all background sources together**. This is a very general result which can be seen directly if we consider the case of two example backgrounds: flasher events and pickup events. The probability to pass cuts A and B for both backgrounds is

$$P(AB \mid \text{background}) = P(\text{flasher})P(AB \mid \text{flasher}) + P(\text{pickup})P(AB \mid \text{pickup}). \tag{K.4}$$

Now, suppose that the low and high-level cuts are independent for each background, i.e.

$$\begin{aligned} P(AB \mid \text{flasher}) &= P(A \mid \text{flasher})P(B \mid \text{flasher}) \\ P(AB \mid \text{pickup}) &= P(A \mid \text{pickup})P(B \mid \text{pickup}). \end{aligned}$$

Then, Equation (K.4) becomes:

$$\begin{aligned} P(AB \mid \text{background}) = P(\text{flasher})&P(A \mid \text{flasher})P(B \mid \text{flasher}) \\ &+ P(\text{pickup})P(A \mid \text{pickup})P(B \mid \text{pickup}) \end{aligned} \tag{K.5}$$

which is **not** equal to the product of the probabilities

$$\begin{aligned} P(A \mid \text{background})&P(B \mid \text{background}) = \\ &(P(\text{flasher})P(A \mid \text{flasher}) + P(\text{pickup})P(A \mid \text{pickup})) \\ &\quad (P(\text{flasher})P(B \mid \text{flasher}) + P(\text{pickup})P(B \mid \text{pickup}))\,. \end{aligned} \tag{K.6}$$

Therefore,

$$P(AB \mid \text{background}) \neq P(A \mid \text{background})P(B \mid \text{background}).$$

References

[1] R. Pordes, D. Petravick, B. Kramer, D. Olson, M. Livny, A. Roy, P. Avery, K. Blackburn, T. Wenaus, F. Würthwein, I. Foster, R. Gardner, M. Wilde, A. Blatecky, J. McGee, and R. Quick, "The open science grid," in *J. Phys. Conf. Ser.*, 78, Vol. 78 (2007) p. 012057.

[2] I. Sfiligoi, D. C. Bradley, B. Holzman, P. Mhashilkar, S. Padhi, and F. Wurthwein, "The pilot way to grid resources using glideinwms," in *2009 WRI World Congress on Computer Science and Information Engineering*, 2, Vol. 2 (2009) pp. 428–432.

[3] M. Tanabashi *et al.* (Particle Data Group), "Review of particle physics," Phys. Rev. D **98**, 030001 (2018), https://link.aps.org/doi/10.1103/PhysRevD.98.030001.

[4] Y. Grossman, R. Harnik, O. Telem, and Y. Zhang, "Self-destructing dark matter," Journal of High Energy Physics **2019** (2019), 10.1007/jhep07(2019)017, https://arxiv.org/abs/1712.00455.

[5] B. Aharmim *et al.* (SNO), "Combined Analysis of all Three Phases of Solar Neutrino Data from the Sudbury Neutrino Observatory," Phys. Rev. C **88**, 025501 (2013), arXiv:1109.0763 [nucl-ex].

[6] "RAT-PAC," Git-sha1: 7ce0b72dbe2c500ea3ccdab52100b0d6c615239a, https://github.com/rat-pac/rat-pac.

[7] M. J. Berger, "ESTAR, PSTAR, and ASTAR: Computer programs for calculating stopping-power and range tables for electrons, protons, and helium ions," Unknown (1992).

[8] D. Groom and the PDG, "Atomic and nuclear properties of materials," http://pdg.lbl.gov/2018/AtomicNuclearProperties/.

[9] S. J. Brice, *Monte Carlo and Analysis Techniques for the Sudbury Neutrino Observatory*, Ph.D. thesis, Balliol College, Oxford University (1996), https://sno.phy.queensu.ca/sno/papers/brice_phd.pdf.

[10] R. Bonventre, *Neutron Multiplicity in Atmospheric Neutrino Events at the Sudbury Neutrino Observatory*, Ph.D. thesis, University of Pennsylvania (2014), https://repository.upenn.edu/edissertations/1213.

[11] E. T. Jaynes, *Probability Theory: The Logic of Science* (Cambridge University Press, 2003).

[12] J. A. Formaggio and G. P. Zeller, "From ev to eev: Neutrino cross sections across energy scales," Rev. Mod. Phys. **84**, 1307 (2012), https://arxiv.org/abs/1305.7513.

[13] G. Battistoni, A. Ferrari, T. Montaruli, and P. Sala, "The atmospheric neutrino flux below 100-MeV: The FLUKA results," Astropart. Phys. **23**, 526 (2005).

[14] G. D. Barr, T. K. Gaisser, P. Lipari, S. Robbins, and T. Stanev, "Three-dimensional calculation of atmospheric neutrinos," Phys. Rev. D **70**, 023006 (2004), https://link.aps.org/doi/10.1103/PhysRevD.70.023006.

[15] M. Wallraff and C. Wiebusch, "Calculation of oscillation probabilities of atmospheric neutrinos using nucraft," Computer Physics Communications **197**, 185189 (2015), https://arxiv.org/abs/1409.1387.

[16] D. H. Hathaway, "The solar cycle," Living Rev. Sol. Phys. (2015), https://arxiv.org/abs/1502.07020.

[17] C. Andreopoulos *et al.*, "The GENIE Neutrino Monte Carlo Generator," Nucl. Instrum. Meth. **A614**, 87 (2010), arXiv:0905.2517 [hep-ph].

[18] P. Harvey, "XSNOED," https://owl.phy.queensu.ca/%7Ephil/xsnoed/xsnoed.html.

[19] J. Park, *Neutrino-Electron Scattering in MINERvA for Constraining the NuMI Neutrino Flux*, Ph.D. thesis, U. Rochester (2013).

[20] C. Andreopoulos *et al.*, "The genie neutrino monte carlo generator physics & user manual," (2015).

[21] T. Sonley, *A Measurement of the Atmospheric Neutrino Flux and Oscillation Parameters at the Sudbury Neutrino Observatory*, Ph.D. thesis, Massachusetts Institute of Technology (2009), https://www.sno.phy.queensu.ca/sno/papers/Sonley_phd_physics_2009.pdf.

[22] S. G. Johnson, "The nlopt nonlinear-optimization package," http://github.com/stevengj/nlopt.

[23] T. Rowan, *Functional Stability Analysis of Numerical Algorithms*, Ph.D. thesis, University of Texas at Austin (1990).

[24] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, "emcee: The MCMC Hammer," The Astronomical Society of the Pacific **125**, 306 (2013), arXiv:1202.3665 [astro-ph.IM].

[25] B. Aharmim *et al.* (SNO Collaboration), "Measurement of neutron production in atmospheric neutrino interactions at the sudbury neutrino observatory," Phys. Rev. D **99**, 112007 (2019), https://link.aps.org/doi/10.1103/PhysRevD.99.112007.

[26] P. Antonioli, C. Ghetti, E. Korolkova, V. Kudryavtsev, and G. Sartorelli, "A Three-dimensional code for muon propagation through the rock: Music," Astropart. Phys. **7**, 357 (1997), arXiv:hep-ph/9705408.

[27] G. Barr, T. Gaisser, S. Robbins, and T. Stanev, "Uncertainties in Atmospheric Neutrino Fluxes," Phys. Rev. D **74**, 094009 (2006), arXiv:astro-ph/0611266.

[28] W. Jarosz, *Efficient Monte Carlo Methods for Light Transport in Scattering Media*, Ph.D. thesis, UC San Diego (2008).

[29] L. Demortier, "P values and nuisance parameters," PHYSTAT LHC Workshop on Statistical Issues for LHC Physics, PHYSTAT 2007 - Proceedings (2007).

[30] S. Talts, M. Betancourt, D. Simpson, A. Vehtari, and A. Gelman, "Validating Bayesian Inference Algorithms with Simulation-Based Calibration," arXiv e-prints , arXiv:1804.06788 (2018), arXiv:1804.06788 [stat.ME].

[31] S. R. Cook, A. Gelman, and D. B. Rubin, "Validation of software for bayesian models using posterior quantiles," Journal of Computational and Graphical Statistics **15**, 675 (2006).

[32] LETA WG, "Low energy threshold analysis unidoc," Access Date: Oct 22, 2019, https://www.snolab.ca/snoplus/private/DocDB/0052/005207/001/leta_unidoc.pdf.