The Search for Sterile Neutrinos with IceCube

By

Kevin Ghorbani

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

(Physics)

at the
UNIVERSITY OF WISCONSIN–MADISON
2019

Date of final oral examination: December 21st 2018

The dissertation was reviewed by the following members of the Final Oral Committee:

Francis Halzen, Gregory Breit Professor and Hilldale Professor, Physics
Albrecht Karle, Professor, Physics
Kael Hanson, Professor, Physics
Michael DuVernois, Senior Scientist, Wisconsin IceCube Particle Astrophysics Center

THE SEARCH FOR STERILE NEUTRINOS WITH ICECUBE

Kevin Ghorbani

Under the supervision of Professor Francis Halzen
At the University of Wisconsin-Madison

ABSTRACT

IceCube is a kilometer-cubed neutrino detector located at the South Pole. IceCube is sensitive to sterile neutrinos in the range of the masses and mixing angles of the LSND/MiniBooNE anomaly ($\Delta m^2 \sim 1\ eV$). In this thesis, I am looking at upgoing atmospheric neutrinos (that have traveled through the Earth from the Northern Hemisphere) with energies between 200 GeV to 20 TeV. I am trying to measure the changes in the distribution of neutrino energy versus the zenith angle, which reflects the distance that the neutrinos traveled through the Earth, due to the existence of sterile neutrinos. In the case of the $3 + 1$ sterile neutrino model, a strong disappearance of muon antineutrinos and a weak disappearance of muon neutrinos is anticipated, due to MSW-resonant oscillation (or matter effect). This thesis focuses on the neutrino events that interact with ice and produce secondary muons within the detector boundaries. By analyzing these types of events, we get a measurement of energy that is superior to previous analyses that looked at muons regardless of whether they started within the detector. This thesis describes the process of data selection and the methods used in the sterile neutrino search and shows that the sensitivity to sterile neutrinos with lower Δm^2 increases compared to previous analyses.

DISCARD THIS PAGE

TABLE OF CONTENTS

DISCARD THIS PAGE

LIST OF TABLES

DISCARD THIS PAGE

LIST OF FIGURES

Figure Page

Figure Page

Chapter 1

Introduction

In this thesis, I will be presenting a sterile neutrino search using data from the IceCube Neutrino Observatory. IceCube is a cubic kilometer neutrino detector at the South Pole that is sensitive to sterile neutrinos with masses and mixing angles at and around the range of the LSND/MiniBooNE anomaly. In this thesis, I measure the upgoing atmospheric neutrino flux for energies from approximately 200GeV to 20TeV as a function of zenith angle, which reflects the distance that the neutrinos traveled through the Earth. In the case of the 3+1 sterile neutrino model, I anticipate a strong matter resonance resulting in the disappearance of muon antineutrinos (and a weak disappearance of muon neutrinos) due to MSW-resonant oscillation.

In this analysis, I concentrate on contained neutrino events with secondary muons that start in the detector to obtain a superior measurement of energy compared to previous analyses [1]. In addition, this analysis has a completely new event selection using both statistical modeling and machine learning to classify the events, which gives us better statistics compared to the previous analyses that used traditional cuts on the data.

In Chapter 2 describes the IceCube detector and the methods that are used to observe and reconstruct the events in the detector. Chapter 3 describes atmospheric neutrinos and the origin of the events that will be used in our analysis. Chapter 4, I introduce the physics of neutrinos and neutrino oscillations. Chapter 5 thoroughly describes my event selection. Chapter 6 describes the systematics considered in this analysis, and finally, in Chapter 7 we present the result of this analysis and how it changes IceCube's sensitivity to sterile neutrinos.

Chapter 2

The IceCube Detector

IceCube is a cubic kilometer for neutrino detector deep in the ice of the South Pole [[2]]. IceCube was primarily designed but not limited to, the detection of astrophysical neutrinos.

IceCube detects photons emitted from high energy charged particles traveling through the ice. If a charged-particle travels with a speed higher than the speed of light in a medium, it emits radiation known as Cherenkov photons. These photons will travel at an angle of θ from the direction of particle:

$$\cos\theta = 1/n\beta \tag{2.1}$$

where 'n' is the medium's refractive index and 'β' is the ratio of the particle speed to the speed of light in the medium. The refractive index in the South Pole ice is about 1.32 and β is close to one, for these high energy particles ($\theta \approx 41\,\mathrm{deg}$).

IceCube consists of 78 strings (completed in 2011) with 60 Digital-Optical-Modules (DOMs) on each string with vertical separation of 17 m and horizontal separation of 125 m spreads over a volume of about one kilometer-cubed (Fig. 2.1). This makes IceCube sensitive to neutrinos in the energy range of $\mathcal{O}(100GeV)$ to $\mathcal{O}(PeV)$.

In addition, there are 8 more strings in IceCube with a higher density of DOMs, with an average of 72 m spacing horizontally and with 60 DOMs on each string with a spacing of 7 m apart, vertically. This part of IceCube is known as 'Deep Core' and is located at the center and the bottom half of IceCube where the ice is the clearest. Deep Core is sensitive to a lower neutrinos energy ($\mathcal{O}(10GeV)$ to $\mathcal{O}(100GeV)$). Deep Core is used in analyses such as

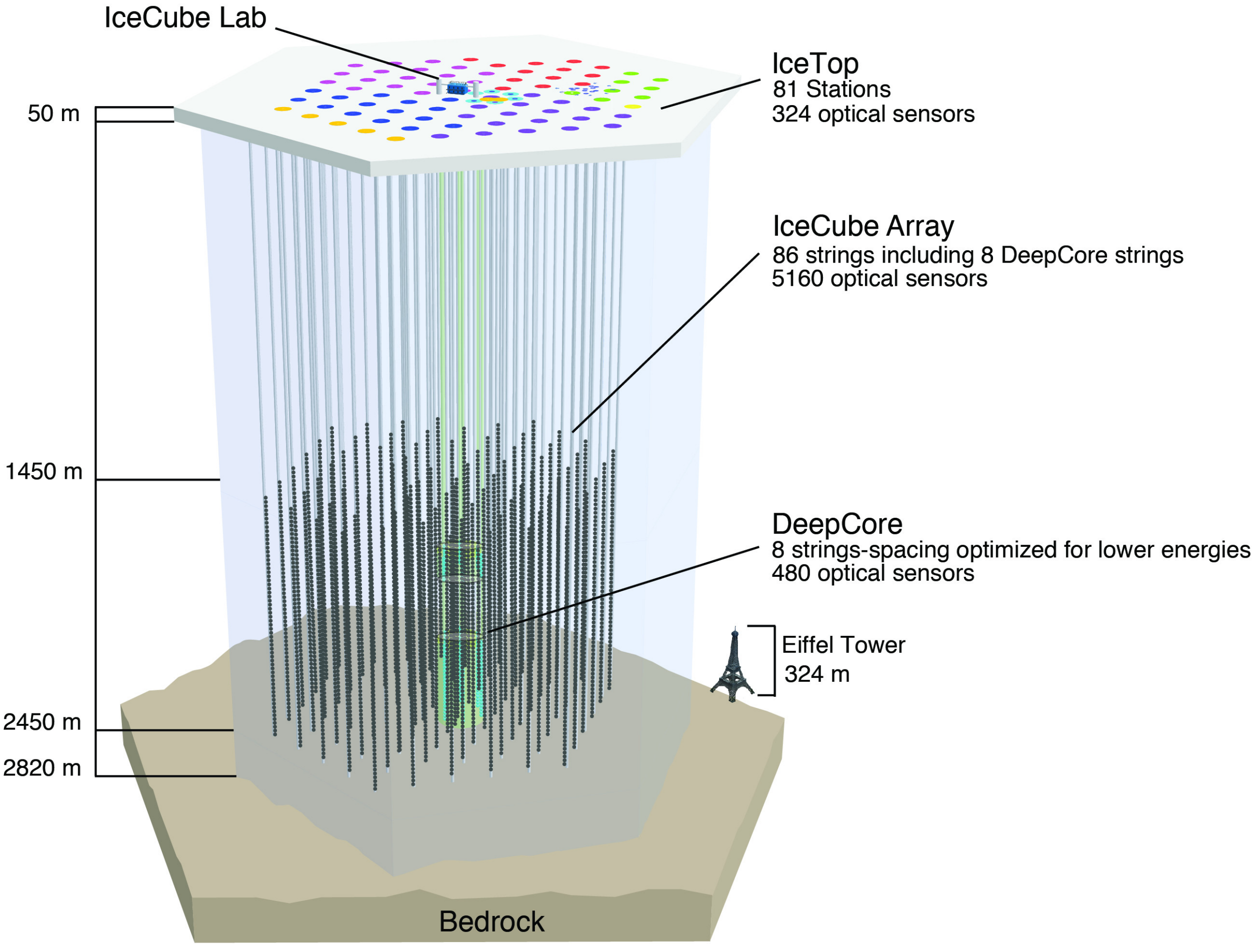


Figure 2.1 IceCube Neutrino Observatory is located from 1450 m to 2450 m deep in the ice of the south pole, less than 400 m above the Antarctic bedrock.

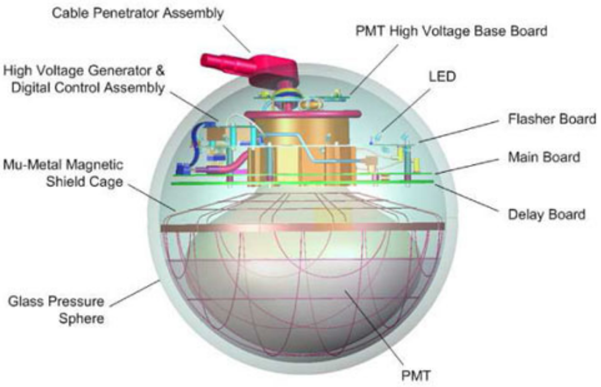


Figure 2.2 Mechanical layout of a Digital-Optical-Module (DOM).

standard neutrino oscillations, neutrinos from WIMP dark matter annihilation, and Galactic supernovae.[[3]]

On the surface of the ice, a cosmic ray air shower array known as IceTop [4] is located. The IceTop array is 162 ice tanks with DOMs to detect Cherenkov light from cosmic ray particles. These tanks are stationed at 81 stations above the IceCube strings. Currently, some of these tanks are buried under a few meters of snow, due to the snow drift. IceTop is sensitive to events with energies $\mathcal{O}(PeV)$ to $\mathcal{O}(EeV)$.

DOMs (Digital-Optical-Modules) are IceCube's light sensor and data acquisition unit (Fig. 2.2). Each DOM is a pressured spherical glass vessel that contains a 10-inch photomultiplier-tube (PMT) facing down and circuit boards. All the communications and power are transferred through a cable which travels to the surface of the ice.

2.1 Neutrino Detection

IceCube is a Cherenkov detector, which means that it can only detect charged particles. However, neutrinos can be detected indirectly. Neutrinos passing through ice may interact with ice and produce charged leptons that emit Cherenkov light in the ice that will be detected by DOMs.

There are two types of interactions (Fig. 2.3). *Neutral Current Interaction* (NC-interactions) proceed by exchanging a Z-boson:

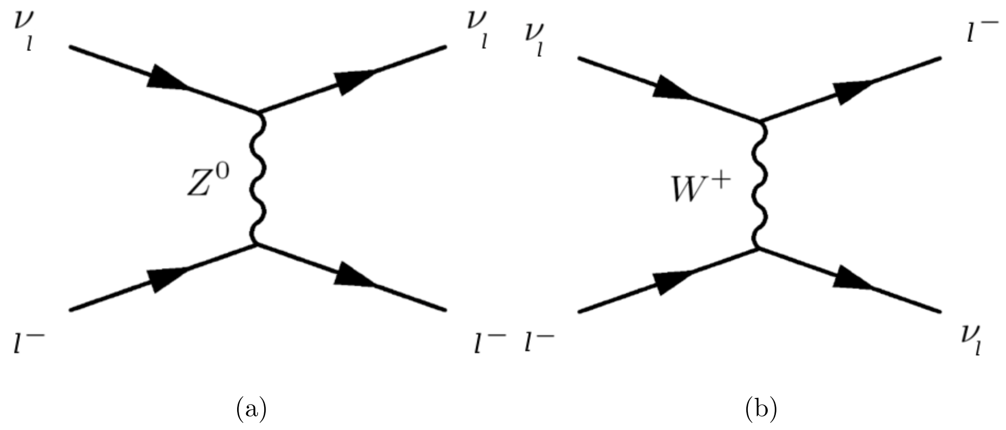


Figure 2.3 Neutral-current (a) and charged-current (b) interactions.

$$p/n + \nu_l \rightarrow p/n + \nu_l \tag{2.2}$$

$$p/n + \bar{\nu}_l \rightarrow p/n + \bar{\nu}_l \tag{2.3}$$

and *Charged Current Interactions* (CC-interactions) proceed by exchanging a W-boson:

$$n + \nu_l \rightarrow p + l^- \tag{2.4}$$

$$p + \bar{\nu}_l \rightarrow n + l+ \tag{2.5}$$

These two types of interactions are distinguishable in IceCube. In NC-interactions at high energies, the neutrino ejects a quark from the nuclei that initiates a shower. This can be seen in IceCube as a blob of light. However due to the momentum of the primary neutrino, and time measurement precision in DOMs (about 1 ns), direction of the primary neutrino can be calculated. The angular resolution for neutral currents (cascade-like events) is about 10 deg. CC-interactions, on the other hand, produce leptons and hadrons. Even though usually most of the neutrino energy and momentum transfers to the lepton, nuclei receive a large amount of energy which leave a starting cascade at the vertex of interaction (see Fig. 5.8). The out-going lepton travels at -almost- the speed of light at the same direction of the primary neutrino (at high energies) and produce Cherenkov light along its trajectory which looks like a track in the detector. The angular resolution of these track-like events are much better than cascade-like's, and is about 0.6 deg.

Length of tracks depend on the type of lepton and lepton energies. Electrons generally don't leave a track since their interaction rate is very high and can only travel a few meters in the ice. Taus particles are heavy so they decay and their track length can be about a few hundred meters and leaves a double bang signature.

2.1.1 Cherenkov Radiation

When charged particles move through a medium with a speed greater than the speed of light in that medium ($v_\gamma = c/n$ where n is the refractive index of the medium), an electromagnetic shock wave will be produced. Radiation travels with the angle of $\theta_C = \cos^{-1}(1/n\beta)$, where $\beta = v/c$ (see Fig. **??**). Ice's refractive index is about 1.3, which makes the angle of Cherenkov radiation about 41 degrees. In reality, the refractive index is a function of wavelength and varies between 1.31 and 1.33 for the photons to which IceCube PMTs are sensitive. The IceCube detector is sensitive to wavelengths between 300 and 600 nm. These Cherenkov photons are emitted with the spectrum of [5]:

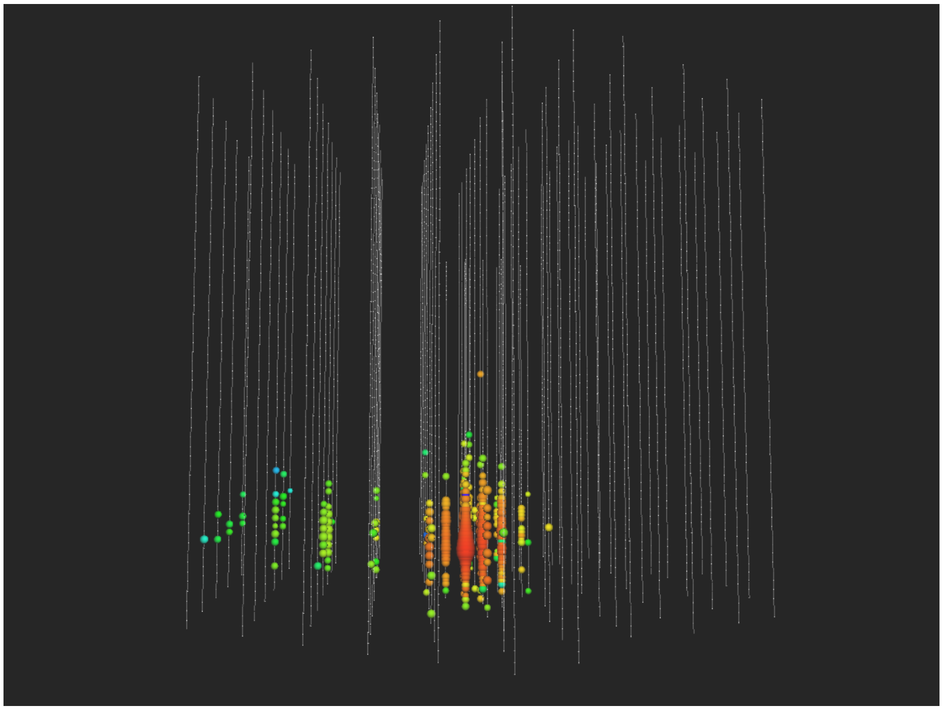
$$\frac{dN}{dxd\lambda} = \frac{2\pi\alpha}{\lambda^2}(1 - \frac{1}{\beta^2 n^2(\lambda)}) \tag{2.6}$$

where λ is the photon's wavelength, and α is the fine-structure constant.

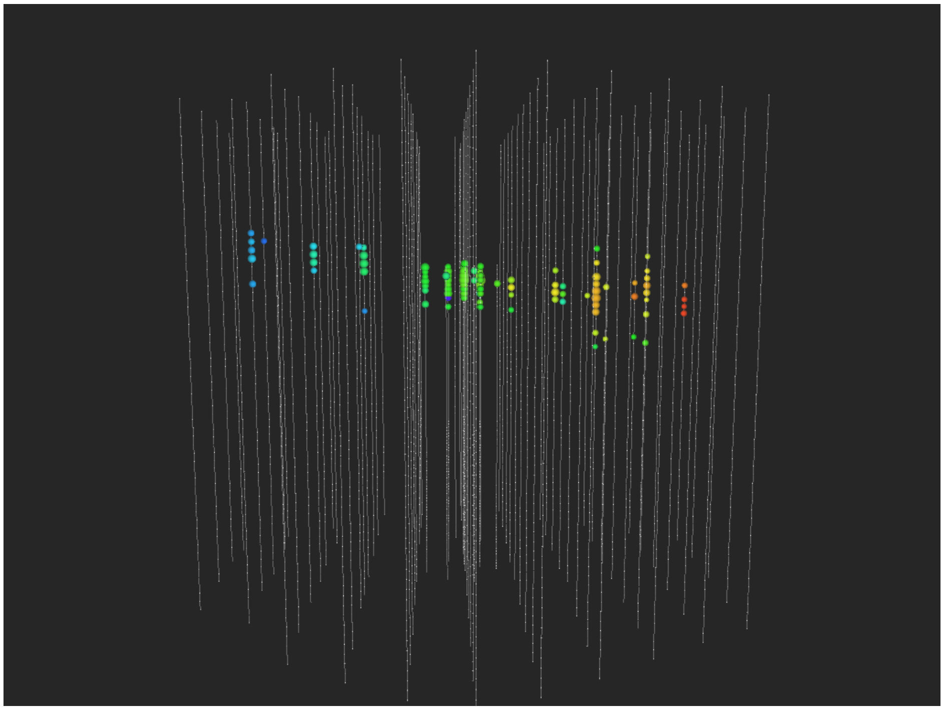
2.2 Reconstruction

Each event in IceCube consists of pulses recorded by DOMs. Fig. 2.4 shows what these events look like (noise is cleaned in those figures). From these pulses, we can reconstruct the direction and energy of each event. Even though, the events in Fig. 2.4 are very clear, there are many computational challenges for their reconstructions.

For directional reconstruction, we usually start with simple models called "first-guess algorithms". First-guess algorithms are very fast to calculate and often can be derived analytically. They use many assumptions which makes them less accurate than more complex



(a)



(b)

Figure 2.4 Two muon events in IceCube. Figure (a) shows a starting-track, where neutrino interacted inside the detector and creates a cascade and a muon-track which leave the detector from the left of the picture. Figure (b) shows a throughgoing muon track traveling leftward. Color show time of arrival, where red is early and blue is late. Size shows the amount of charge deposited in each DOM.

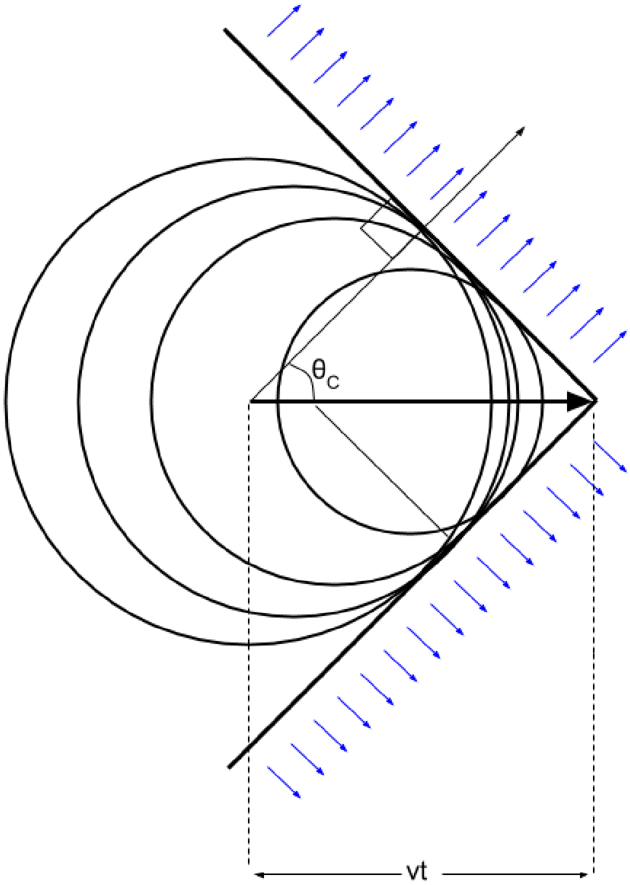


Figure 2.5 Cherenkov radiation illustrated. The particle travels toward the right. An electromagnetic shock wave is created and photons emitted at an angle of θ_C.

algorithms. The results from First-guess algorithms are often used as a starting point for more complex and accurate reconstruction algorithms. These algorithms usually try to maximize a likelihood function in order to find the best fit, which is more time consuming [6].

The simplest algorithm that reconstructs the tracks' directions is *line-fit* [7]. It is important to know that line-fit does not consider the geometry of the Cherenkov cone and the optical properties of the medium. In reality, photons can scatter in ice before reach any DOMs, therefore it will be more accurate if only the first hit in each DOMs is considered for this algorithm. *Line-fit* assumes that photons travel in a line with a constant speed of v. The location of each DOM, $\mathbf{r}_i$, can be calculated where λ is the photon's wavelength, and α is the fine-structure constant.

$$\mathbf{r}_i \approx \mathbf{r} + \mathbf{v}.t_i \tag{2.7}$$

To get the best result χ^2 is defined to be minimized.

$$\chi^2 = \sum_{i=1}^{N_{hit}} (\mathbf{r}_i - \mathbf{r} - \mathbf{v}.t_i)^2 \tag{2.8}$$

where N_{hit} is the number of hits. However, the result can be derived analytically.

$$\mathbf{r} = \langle \mathbf{r}_i \rangle - \mathbf{v}.\langle t_i \rangle \tag{2.9}$$

and

$$\mathbf{v} = \frac{\langle \mathbf{r}_i.t_i \rangle - \langle \mathbf{r}_i \rangle . \langle t_i \rangle}{\langle t_i^2 \rangle - \langle t_i \rangle^2} \tag{2.10}$$

From equation 2.9 and 2.10, direction and zenith angle of the track can be calculated.

$$\mathbf{e} = \frac{\mathbf{v}_{LF}}{|\mathbf{v}_{LF}|} \tag{2.11}$$

$$\theta_{LF} = -\arccos(v_z/|\mathbf{v}_{LF}|) \tag{2.12}$$

Due to the limitations of *line-fit*, it is usually used as a "first guess" algorithm for more detailed reconstructions.

Noise in the detector can cause mis-reconstruction in the direction of tracks because these pulses do not fit to the true track. Some of these errors can be avoided by cleaning the noise beforehand. "*Improved Line-fit*"[8] replaces the least-square model with a Huber fit [9], and it penalized farther hits (generally noise) so that they have a lesser effect on the result.

For a better reconstruction, the time distribution of each track's first hits can be considered and be fitted by a Pandel Function [6]. The Pandel function is an analytical model which is motivated by a laser light analysis in the BAIKAL experiment and predicts the distribution of photon arrival time affected by scattering over a given distance [10]. "*SPEFit*" is an IceCube's track reconstruction algorithms which takes *SPEFit* as a starting point and tries for fit the track using Pandel function. Similar to *line-fit*, *SPEFit* only uses the first photon of each DOM. *line-fit* is usually done in multiple iterations for better accuracy. The average directional resolution for IceCube is less than 1 degree for track-like events, using *SPEFit*. Zenith angle resolution for the final even selection (see Section 5.5) is shown in Figure 2.6.

Event energies are also calculated with DOM pulses. In section 5.7, I will explain more about muon and cascade energy reconstruction and how to combine them in order to get the muon's energy.

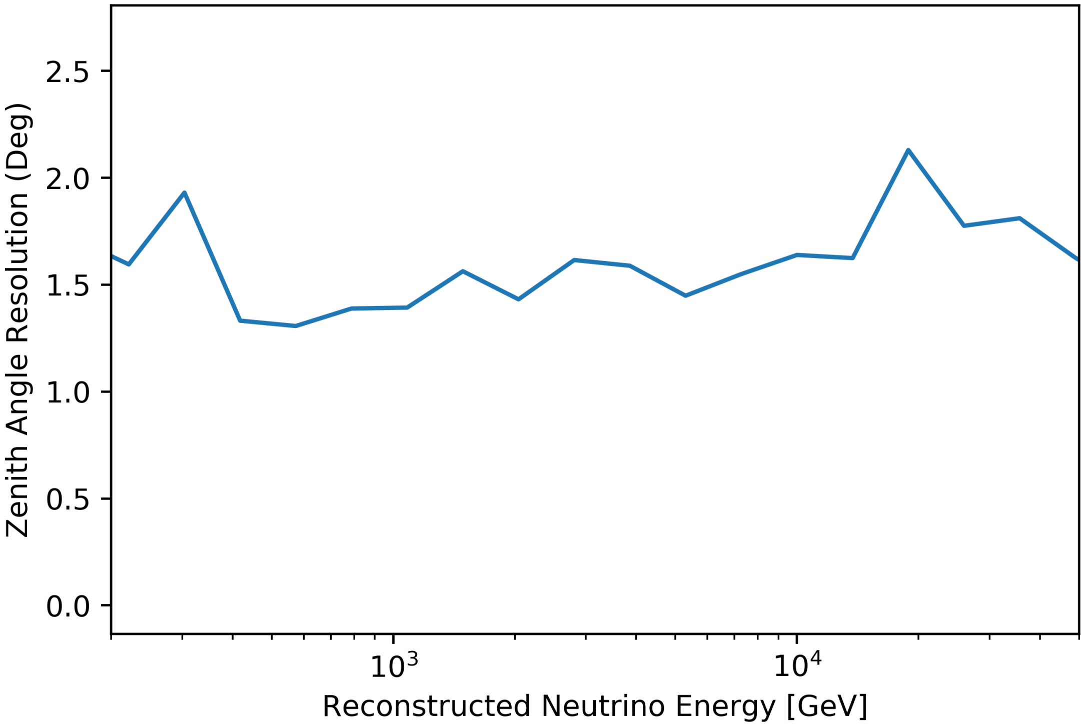


Figure 2.6 Zenith angle resolution in degrees as a function of reconstructed neutrino energy for the final event selection.

Chapter 3

Neutrino Production

Unlike other particles, neutrinos can travel across the universe without being absorbed by matter due to their small cross-sections, nor are they deflected by the magnetic field, since they have no charge. Because neutrinos have very small masses, they will travel essentially at the speed of light at the energies that are important to this thesis.

3.1 Cosmic Rays

Cosmic rays are high energy charged particles and nuclei arriving from space. Cosmic rays hit the Earth's atmosphere at a rate of 1000 per square meter per second. The vast majority of these particles are protons but the flux also contains electrons, helium nuclei (alpha particles) and heavier nuclei. Most of these particles originate from outside the solar system within the Galaxy, but low energy particles are associated with the solar flares. Low energy cosmic rays (<10 GeV) will be decelerated by expanding magnetized plasma from the sun [11]. Earth's magnetic field absorbs most cosmic rays at GeV energy.

The spectra of cosmic ray elements can be described as an inverse power law of energy

$$\frac{dN}{dE} = E^{-(\gamma+1)}. \tag{3.1}$$

However the spectral index (γ) slightly changes with the energy but remains between 2.5 and 3.5 [12, 13, 14]. Figure 3.1 shows that the cosmic ray spectrum becomes steeper above PeV energies. This is commonly known as the "knee" of the spectrum. A Change in the slope at EeV energies is known as the "ankle" of the spectrum. But all primary nuclei spectra (cosmic rays that have been accelerated at astrophysical sources) have less steep than secondary nuclei (cosmic rays that are produced via interaction of primary nuclei and the interstellar medium).

3.2 Conventional Atmospheric Neutrinos

Cosmic rays interact with N and O in the atmosphere and they produce pions and kaons that will decay into muons (and gamma which will become an electron/positron pair) and they eventually produce up to thousands of particles in a single cosmic ray shower. High energy neutrinos (a few GeV to about 100 TeV) can be produced primarily in the decay of charged kaons and pions. At lower energies (less than a few GeV), muon decay can also produce neutrinos and at energies higher than 100 TeV, prompt neutrinos (neutrinos produced in the semi-leptonic decays of charmed mesons and baryons) become important (see Fig. 3.2). However, IceCube is not sensitive to neutrinos with energies lower than tens of GeV and has yet to observe prompt neutrinos, because the astrophysical neutrinos rate (neutrinos produced from the astrophysical sources) can become as large as (or even larger than) prompt neutrino flux.

The Cosmic ray energy spectrum is about $dN/dE \propto E^{-2.7}$ and $dN/dE \propto E^{-3}$ above PeV (the knee) [12]. A second knee with $dN/dE \propto E^{-3.2}$ also appears at energies above 500 PeV [15] but the spectrum seems to be flatten with a kink at the energy about 3 EeV, to $dN/dE \propto E^{-2.7}$.

3.3 Astrophysical Neutrinos

The production of astrophysical neutrinos is no different from atmospheric neutrinos (see section 3.2). Interaction of cosmic rays from their sources with interstellar medium produces pions. Charged pions can decay into muons and muon neutrinos where those muons will produce electron, muon and electron neutrinos. These neutrinos will propagate outwards. Neutral pions that have been produced in the cosmic ray interaction will decay into gamma rays. If sources are not opaque to gamma rays, they will be a gamma-ray source as well as a

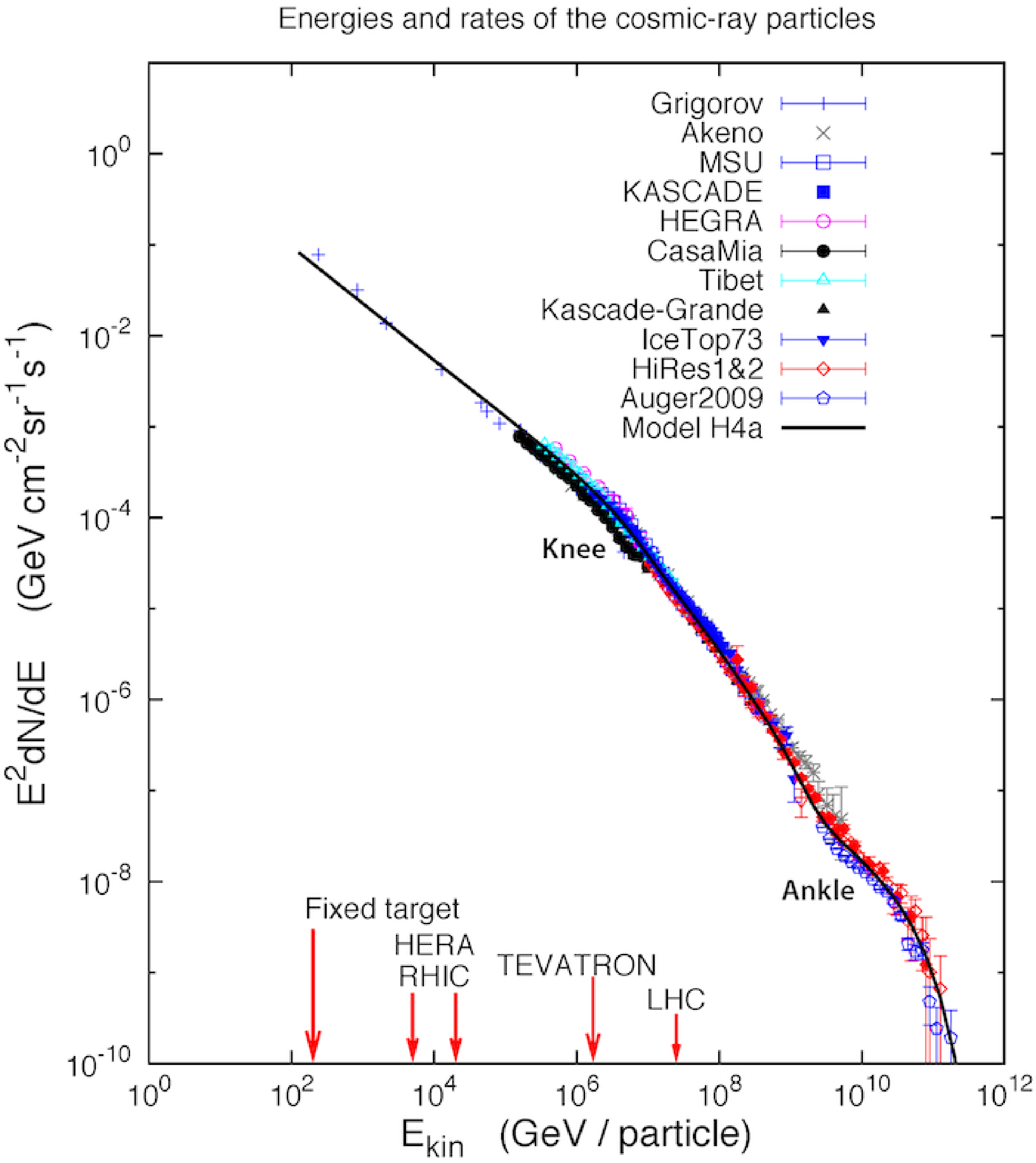


Figure 3.1 Flux of cosmic rays exponentially decrease with energy. Change in the slope of the flux at around PeV and EeV, is known as the "knee" and "ankle" of the spectrum.

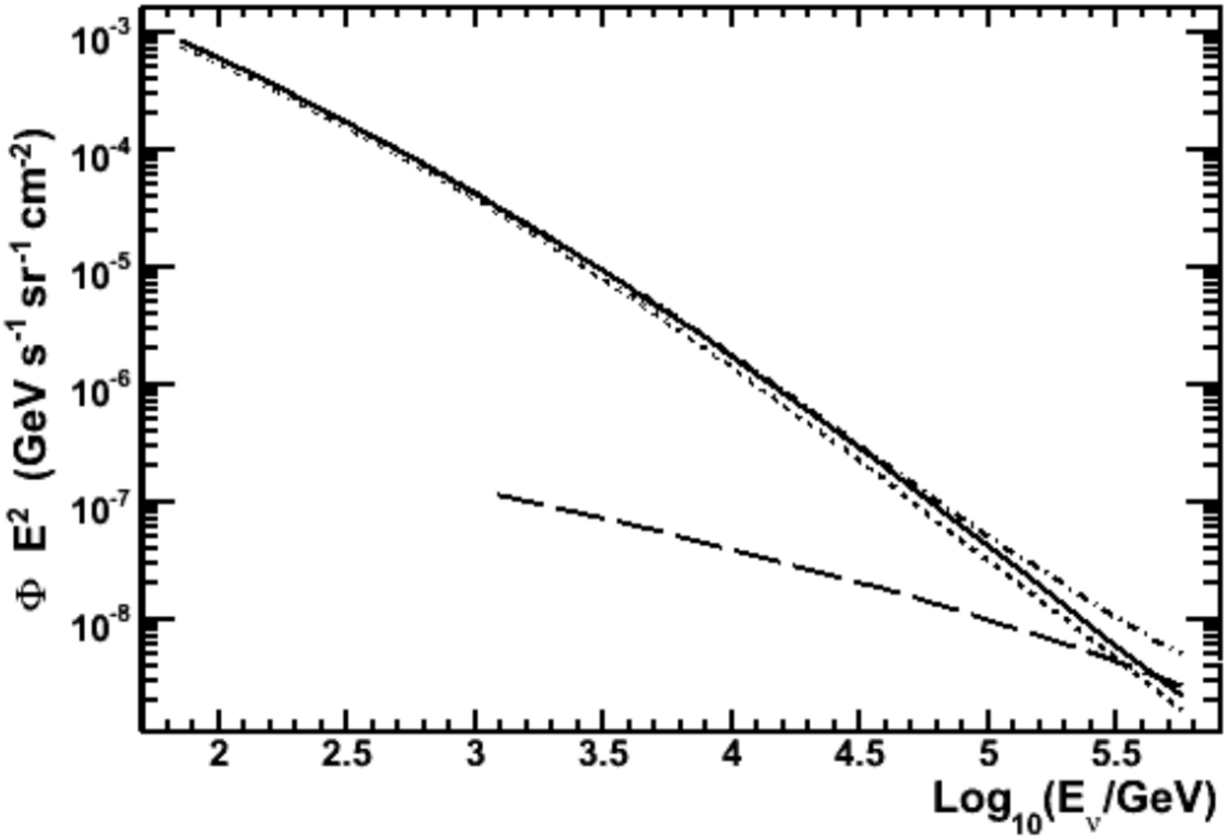


Figure 3.2 Atmospheric flux of upgoing muon neutrinos [14]. Solid-line is the conventional $\nu_\mu + \bar{\nu_\mu}$ flux [16], dashed-line is the prompt flux [17], and dotted-line is the sum of both.

neutrino source. The inverse is not necessarily true since gamma-ray sources can be a result of electromagnetic production.

Supernova remnants can be a source of neutrinos. Very massive stars at the end of their life collapse due to their own massive gravity. This results in an explosion which is called a supernova. The shock wave produced by the explosion which accelerates particle via Fermi acceleration. These particles can interact with the molecular cloud around the supernova and will produce neutrinos. This is what accelerator physicists call beam dump.

Pulsars, which are rapidly rotating neutrons stars or white dwarfs, can also be a candidate to neutrino productions. Supernovae are one of the ways that neutron stars are formed. Pulsars are usually very small with the radius of about 10s of kilometers and rotate with the frequency of hundreds of hertz.

Active Galactic Nuclei (AGNs) are galaxies with extremely large black holes at their center. These black holes have jets in along their rotational axes, which are outflows of material fallen onto the black hole. If an AGN's axis is pointed toward the Earth, it is called a Blazar. Accelerated particles around these jets are a source of neutrinos. Recently, IceCube has observed the first neutrino source that is believed to be a blazar (TXS 0506+056) [18].

Chapter 4

Neutrinos

4.1 Neutrino Oscillation

The standard model of particle physics accommodates three flavors of active neutrinos, ν_e, ν_μ and ν_τ. Neutrino oscillations have first been predicted by Bruno Pontecorvo in 1957 [[19, 20]]. Oscillation will result in appearance or disappearance of neutrino flavors which have been observed by experiments in solar neutrinos [[21, 22, 23, 24, 25, 26, 27, 28, 29]], KamLAND [[30, 31]] and Super-Kamokande [[32, 33]]. Neutrino oscillation is a consequence of neutrino mass. In this paradigm active neutrinos are superposition of three massive neutrinos ν_1, ν_2 and ν_3 with masses of m_1, m_2 and m_3 respectively:

$$\nu_{\alpha L} = \sum_{j=1}^{3} U_{\alpha j} \nu_{jL}(x) \tag{4.1}$$

where $\alpha = e, \mu, \tau$ and U is the 3×3 unitary neutrino mixing matrix [[19, 34]], which can be parametrized by:

$$U = \begin{bmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}s_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{i\frac{\alpha_{21}}{2}} & 0 \\ 0 & 0 & e^{i\frac{\alpha_{31}}{2}} \end{bmatrix} \tag{4.2}$$

where $c_{ij} = \cos\theta_{ij}$, $s_{ij} = \sin\theta_{ij}$, $\theta_{ij} = [0, \pi/2]$, $\delta = [0, 2\pi]$ in the Dirac CP violation phase and α_{21}, α_{31} are Majorana CP violation phases [[35, 36]].

State of neutrino at $t_0 = 0$ is:

$$|\psi(t_0)\rangle = \sum_i U_{\alpha i} |\nu_i\rangle e^{ipx} \tag{4.3}$$

where p is the neutrino momentum. According to the Schrödinger equation at $t = t_1$

$$|\psi(t_1)\rangle = \sum_i U_{\alpha i} |\nu_i\rangle e^{[ipx - iE_i(t_1 - t_0)]} \tag{4.4}$$

where:

$$E_i = \sqrt{p^2 + m_i^2} \tag{4.5}$$

and for $m_i^2 \ll p^2$:

$$E_i \simeq p_i + \frac{m_i^2}{2p} \tag{4.6}$$

In order to calculate the probability of a neutrino flavor at distance $x = L$, replace eqn. 4.6 into 4.4. We should also include $t = L/v$, but for relativistic particles $t = L$

$$|\psi(t_1)\rangle = \sum_i U_{\alpha i} |\nu_i\rangle e^{-i\frac{m_i^2}{2p}L}. \tag{4.7}$$

To have the equation in a flavor basis, eqn. 4.1 can be used; The wave function is

$$|\psi(t_1)\rangle = \sum_i \sum_\gamma U_{\alpha i} U^*_{\gamma i} |\nu_\gamma\rangle e^{-i\frac{m_i^2}{2p}L} \tag{4.8}$$

Now the probability of a neutrino with flavor α to oscillate into a neutrino with flavor β can be calculated as

$$P_{\alpha\to\beta} = |\langle\beta|\psi(t)\rangle|^2 = \sum_{i,j} U_{\alpha i} U^*_{\alpha j} U_{\beta j} U^*_{\beta i} e^{i\frac{\Delta m_{ij}^2 L}{2p}}. \tag{4.9}$$

The matrix below is the general mixing matrix for two neutrinos

$$U = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \tag{4.10}$$

by using this matrix into eqn. 4.9, the flavor change probability will be

$$P_{\alpha\to\beta} = |\delta_{\alpha\beta} - \sin^2(2\theta)\sin^2(\frac{\Delta m^2 L}{2E})| \tag{4.11}$$

Fig. 4.1 shows that for $L/E \ll 1$ almost no oscillation occurs but at large L/E very fast oscillation occurs.

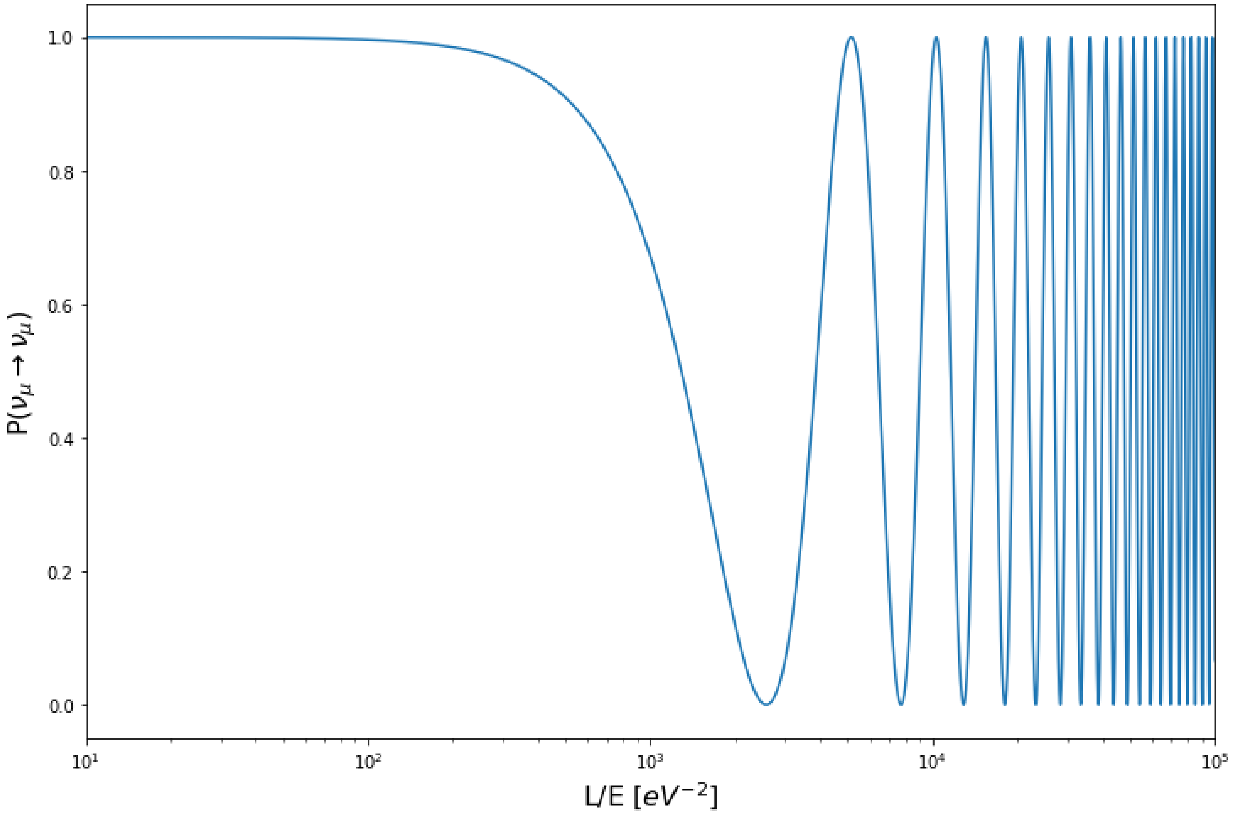


Figure 4.1 for $L/E \ll 1$ almost no oscillation occurs and for $L/E \gg 1$ very fast oscillation occurs. $\sin^2(2\theta) = 1$ for maximum oscillation and $\Delta m^2_{32} = 2.44^{-3} eV^2$.

This probability is for a neutrino traveling in vacuum. For neutrino disappearance probability can be calculated by the same equation if $\beta = \alpha$, i.e. $P_{\alpha\to\alpha}$. If the standard model's prediction is correct in predicting massless neutrino, from eqn. 4.11 it would be clear that there would not be any neutrino oscillation, or in another word $P_{\alpha\to\alpha}$ would be 1 if $\alpha = \beta$ and 0 if $\alpha \neq \beta$.

The values below obtained form 3-neutrino mixing scheme [11]
$\sin^2(\theta_{12}) = 0.304 \pm 0.014$
$\Delta m^2_{21} = (7.53 \pm 0.18) \times 10^{-5} eV^2$
$\sin^2(\theta_{23}) = 0.51 \pm 0.05$ (normal mass hierarchy)
$\sin^2(\theta_{23}) = 0.50 \pm 0.05$ (inverted mass hierarchy)
$\Delta m^2_{32} = (2.44 \pm 0.06) \times 10^{-3} eV^2$ (normal mass hierarchy)
$\Delta m^2_{32} = (2.51 \pm 0.0.6) \times 10^{-3} eV^2$ (inverted mass hierarchy)
$\sin^2(\theta_{13}) = (2.19 \pm 0.12) \times 10^{-2}$

4.2 The Sterile Neutrino

Anomalies in the result of some neutrinos experiments such as LSND[37, 38], Gallium neutrino experiments[39, 40, 41, 42, 43], GALLEX:[44], SAGE:[23] and Reactor antineutrino experiments [45, 46, 47] have raised questions for existence of a neutrino flavor other than the three standard model neutrinos. Since we know from LEP experiments [48] that there are only three active neutrinos (i.e. neutrinos that have a corresponding lepton), any other neutrino must a sterile neutrino (i.e. does not have a corresponding lepton and cannot undergo weak interaction but can interact with other matter through neutrino oscillation, see Section 4.1). Sterile neutrinos can potentially explain the anomalies mentioned above. Several null [49, 50, 51, 52, 53] and global fits [54, 55] constrain these light sterile neutrinos to mass of $O(1eV^2)$) and mixing angle of $\sin^2\theta_{24} = 0.1$.

In this analysis, we consider $3+1$ scenario. $3+1$ is referred to the scenario when the sterile neutrinos are more massive than the active neutrinos. Because cosmology has put a bound on the commutative mass of all active neutrinos ($\sum_i \nu_i < 0.23\ eV$), $1+3$ scenario (where active neutrinos are more massive than the sterile neutrinos) is not likely [56].

4.3 MSW effect

The MSW is a flavor-changing mechanism of neutrinos in matter, that was introduced by Mikheyev, Smirnov and Wolfenstein [[57, 58, 59]]. It comes from the fact that electron neutrinos and anti-neutrinos ($\nu_e/\bar{\nu_e}$) can undergo different interactions (neutral and charged current interaction) as opposed to other types of neutrinos (muon, tau or sterile) which can only have neutral current interactions with electrons in matter.

The total Hamiltonian in matter is

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_I \quad \text{with} \quad \mathcal{H}_I \left|\nu_\alpha(p)\right\rangle = V_\alpha \left|\nu_\alpha(p)\right\rangle \tag{4.12}$$

where V_α ($\alpha = e, \nu, \tau$) is the effective potential coming from active neutrino interaction with matter. There are two neutral-current and charged-current interactions (shown in figure 2.3) with potentials of

$$V_{CC} = \sqrt{2} G_F N_e, \quad V_{NC} = -\frac{\sqrt{2}}{2} G_F N_n \tag{4.13}$$

where G_F is the Fermi constant and N_e and N_n is number density of electron and neutron in matter. This Potential can be summarized at [60, 61]

$$V_\alpha = \begin{cases} V_{CC}\delta_{\alpha e} + V_{NC} & \text{for neutrinos} \\ \\ -(V_{CC}\delta_{\alpha e} + V_{NC}) & \text{for anti-neutrinos} \end{cases} \tag{4.14}$$

Since the most important flavor for IceCube to measure is muon neutrinos, we need to know the disappearance of those muon neutrinos, i.e. $P(\nu_\mu \to \nu_\mu)$ as a function of Energy and zenith angle, which is the length and density of matter that the neutrino has traveled through. This will come from solving the Schrödinger equation

$$P(\nu_\mu \to \nu_\mu) = |\psi_{\mu\mu}(x)|^2 = 1 - P_{\nu_\mu \to \nu_\mu}(x) = 1 - \sin^2 2\theta_M \sin^2(\frac{\Delta m_M^2 L}{4E_\nu}) \tag{4.15}$$

where:

$$\Delta m_M^2 = \sqrt{(\Delta m^2 \cos 2\theta - A)^2 + (\Delta m^2 \sin 2\theta)^2} \tag{4.16}$$

$$\theta_M = \frac{1}{2} \arctan(\frac{\tan 2\theta}{1 - \frac{A}{\theta m^2 \cos 2\theta}}) \tag{4.17}$$

$$A = \pm 2\sqrt{2} E G_F N \tag{4.18}$$

and N is the electron number density, and θ and Δm^2 represent mixing angle and mass of the neutrinos.

From the equations above, we can deduce that a resonance occurs at:

$$N = \mp \cos 2\theta \frac{\Delta m^2}{2E_\nu} \frac{1}{\sqrt{2} G_F} \tag{4.19}$$

This resonance is known as the MSW effect. For $1eV$ sterile neutrinos and density of the Earth, this resonance happens at energies of $O(TeV)$, which happens to be the energies that IceCube is sensitive to, the most.

Chapter 5

Event Selection

IceCube detects hundreds of billions of events each year. The majority of these events are downgoing atmospheric muons. In the sterile neutrino analysis, we are interested in upgoing atmospheric neutrinos to detect the resonant matter effect (see section 4.3). Our event selection is based on many cuts and reconstructions including machine learning.

5.1 Online Filters

After events triggered the detector, “Processing and Filtering” (PnF) will be performed almost simultaneously at the South Pole. Each event will go through sets of reconstructions. These reconstructions are usually simpler and faster at level 1 (which happens at the Pole) compared to the ones that are performed at higher levels that happen later either in Madison (level 2) or analysis level. Each event may be selected to pass sets of predefined filters such as “Muon Filter” which selects events that look like muon tracks. In addition, there are sets of unconventional filters such as “MinBias” which select random events regardless of any reconstructions mainly for calibration purposes. If an event has passed any of these filters, that event will be selected to be transferred to the North via satellite transmission.

In this analysis, all events that passed any of conventional filters at level 2 will be selected for further analysis.

5.2 Coincident Event Splitting

Given that the events readout time in the detector is 10 μs and the trigger rate is about 3 kHz, it is obvious that there will be many overlapping events. Also, other factors such as an excess of deposited charge in the detector due to the existence of multiple events raise the chance of coincident event pass the filters compared to single events, which raises this probability to about 10% of all events.

In order to have an accurate event reconstruction, we need to split these outputs into multiple events. "Topological Splitter" is one of the modules that will do this splitting [62]. The concept of topological splitter is based on the fact that photons detected by the photomultipliers are Cherenkov lights traveling with the speed of c/n (where 'c' is the speed of light and 'n' is the index of refraction in ice) emitted by the muon traveling at the speed of light (c). Ergo most of the photons coming from a single muon event have time-like spacetime separation from the photons coming from a different muon.

Of course, this method has its own limitations in reality. For example, noise in the detector can mislead the splitter, so splitting should be performed on clean pulses (from which most of the noise has already been removed). In this analysis, we will keep and add the noise to our pulses after splitting since, in one of our event reconstructions, noise will be needed.

5.3 Directional Selection

The median range of muon in ice is about 3 km for 1 TeV muons (and about 20 km for high energy $1PeV$ muon)[63]. Therefore, one of the best ways to eliminate atmospheric muon background is by selecting only upgoing muons in the detector. This will collect events where atmospheric neutrinos traveled through the earth and interacted near or at the detector.

Single muon tracks in IceCube have an average directional resolution of less than 1 degree. We are only interested in the event with zenith angles of more than $\pi/2$ (i.e. upgoing events). But this is not enough to make sure there are no downgoing atmospheric muons in the sample left that are mis-reconstructed as an upgoing event. Mis-reconstructions happen due to many factors including not having enough hits, timely noise or coincident and timely events that failed to be split in the previous process. In fact, by only cutting on zenith angle, there are more than two orders of magnitude more atmospheric –mis-reconstructed– muons in the sample than neutrinos. It is obvious that events that have less deposit energy in the detector are more likely to be mis-reconstructed. Therefore, in this selection, we cut out all the event that hit 15 or less DOMs.

In addition, a variable called 'paraboloid sigma' describes the shape of the likelihood space around the maximum point in order to estimate the statistical uncertainty on the location of that point [64], and 'reduced log likelihood' tries to describe how good the directional fit is. These variables will be used later in machine learning, but at this point, very poorly reconstructed events with a large 'reduced log likelihood' will be cut out of the sample.

5.4 Neutral Current Events

There are mainly two types of neutral-current event observable by IceCube, $\nu_\mu - NC$ and ν_e. From the detector's point of view, these events are indistinguishable and will give a very bad directional reconstruction due to their spherical shape (they have a resolution of about 10 degrees). It will be to our advantage in the analysis to cut out all the neutral-current events.

Although, selecting pure neutral-current sample may not be trivial in practice (some charged-current starting event cascades can resemble neutral current) cutting them out is fairly easy. We can define a track length for each event by looking at its direct hits based on its directional reconstruction, known as 'LDir'. Neutral-currents have a very short track length (tens of meters in reality) as shown in Fig. 5.1. By cutting, out short events ($LDir < 200m$), we can guarantee that the rate of ν_e will be about 0.2% of all event in the final sample. $\nu_\mu - NC$ is expected to have a similar rate at the final level. In addition, as is shown in Fig. 5.1, a large number of atmospheric neutrinos have a direct length of less than 200 m and will be cut out at this level. Cutting out short $\nu_\mu - CC$ events are not necessarily a bad thing for the analysis since shorter events have worse energy reconstructions and will reduce the sample's energy-resolution.

5.5 Neutrino Event Selection

The main cut for selecting background muons from neutrino events was machine learning algorithms. After testing many different algorithms, supervised AdaBoost [A.4] showed the best result for classifying these events, as well as being one of the fastest ones. For training the machine, about 40 days of atmospheric muon simulations were used as background and 2 years of atmospheric muon simulations were used as foreground. 90% of these simulation events were randomly selected to be used for training the machine, and the rest (10%) were used for testing the result for over-fitting. Due to the lack of cosmic ray simulations, we need to keep as much of the simulation events as we can in training sample.

For each event, 10 variables are fed to the machine. Some of these variables can determine the quality of reconstructions, such as "log likelihood", "reduced log likelihood" and "corrected paraboloid sigma" (see section 5.3). Some are associated with the amount of charge in the detector, the length of the tracks, as well as skewness of the direct hits, are also used as input variables of the machine. Finally, the zenith angle at which the particle entered the detector does have a high correlation with the type of events.

Correlation matrix (Fig. 5.2) shows correlations between each two variables that was used for training. Although having a very strong correlation between two variables may not negatively affect the result of the classifier, by removing one of them the outcome will not change. Strong correlation among input variables (usually over 0.9) may increase the chance of overfitting. The strongest correlation between these ten variables are 0.77.

In this two-class AdaBoost, a decision function between 1 (certainly atmospheric neutrino) and -1 (certainly atmospheric muon/background) is given to each event. Cuts based on the decision function determine the purity of the final selection. Figure 5.3 shows the distribution of atmospheric neutrinos and muons. Fluctuations in the decision function histogram are features of AdaBoost. Since AdaBoost is a form of a decision tree, different sizes

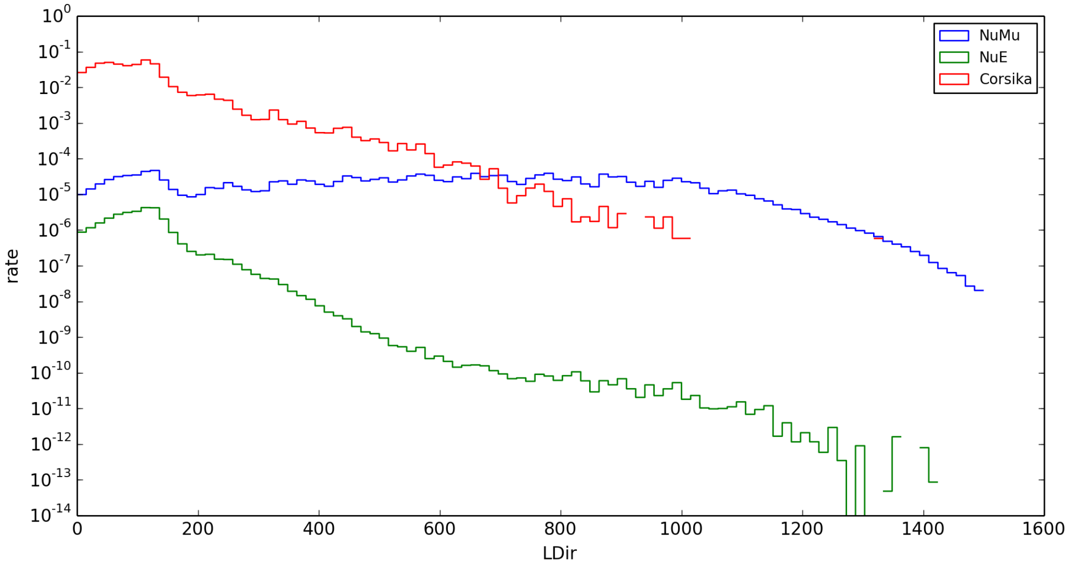
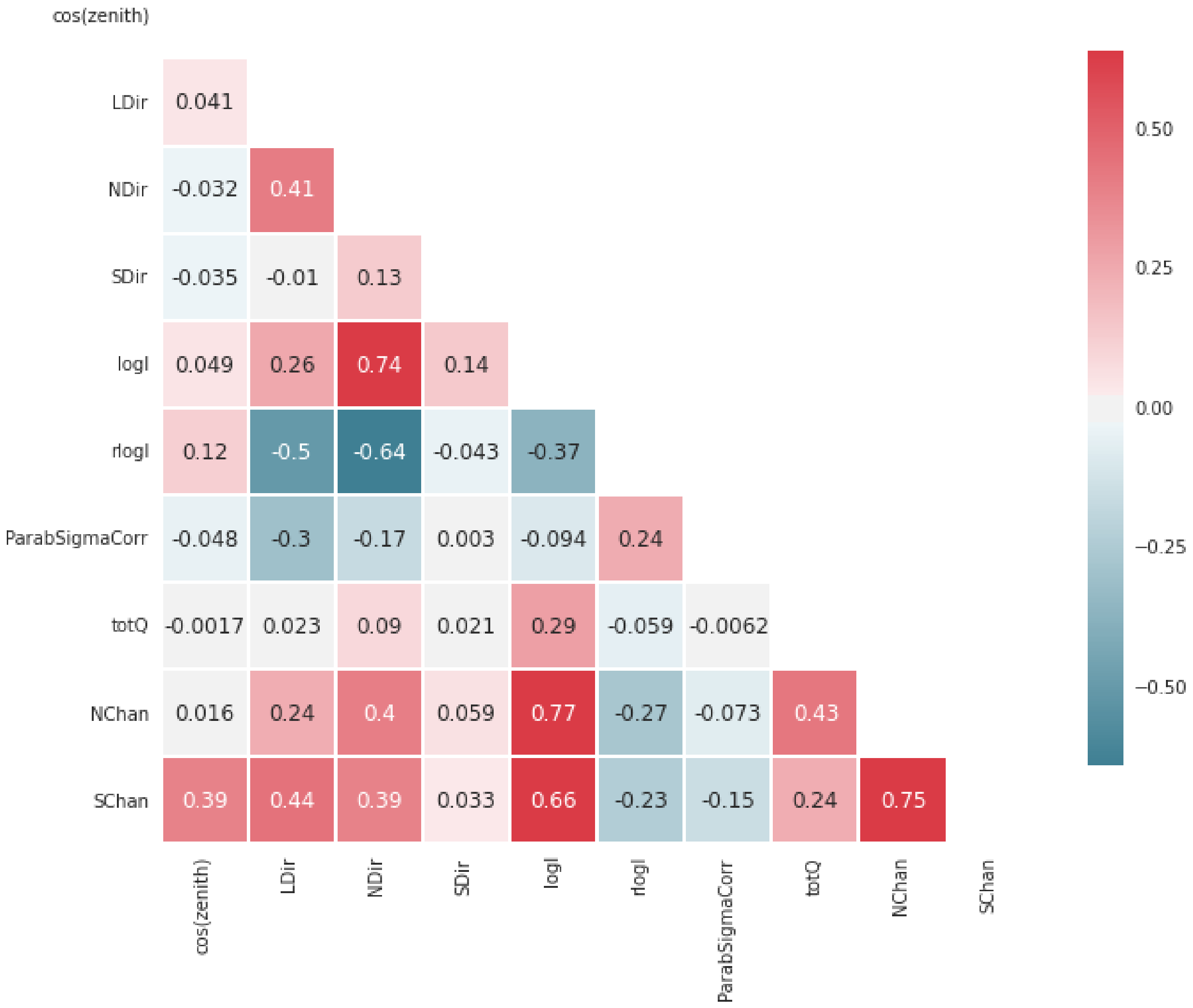


Figure 5.1 Histogram of 'LDir' in meters is shown for three different events. Blue line shows ν_μ events, green shows ν_e events, and red shows atmospheric muon events.



	cos(zenith)	LDir	NDir	SDir	logl	rlogl	ParabSigmaCorr	totQ	NChan	SChan
cos(zenith)										
LDir	0.041									
NDir	-0.032	0.41								
SDir	-0.035	-0.01	0.13							
logl	0.049	0.26	0.74	0.14						
rlogl	0.12	-0.5	-0.64	-0.043	-0.37					
ParabSigmaCorr	-0.048	-0.3	-0.17	0.003	-0.094	0.24				
totQ	-0.0017	0.023	0.09	0.021	0.29	-0.059	-0.0062			
NChan	0.016	0.24	0.4	0.059	0.77	-0.27	-0.073	0.43		
SChan	0.39	0.44	0.39	0.033	0.66	-0.23	-0.15	0.24	0.75	

Figure 5.2 Correlation between every two variables that are used for machine learning are shown in this plot. These variables have been chosen because they are important to determine if the event is reconstructed with high accuracy. Some test and trial-and-error have determined which of the variable give the best result and should be considered as an input in the analysis.

of branches of the tree (which each assigned a particular decision function) may contain different amount of events which will cause the fluctuations in the histogram. Histograms of variables in the adjacent event (on peaks and valleys) show no major physical differences between those events.

By choosing the events with a decision function of more than 0.343, we can guarantee that there is less than 1 percent background in the final sample. 1 percent muon in the sample will not change the sensitivity of the analysis and is consistent with the previous analyses in IceCube ([1, 62]).

Overtraining (or overfitting) is fitting the trained data to noise in the sample. If overtraining occurs, trained events will not agree with the untrained events. One of the easiest ways to look for any overtraining is to keep a portion of our sample as a testing sample (which the machine has not been trained on) and compare them to the trained sample. Since these two samples are randomly selected from the same events, they must agree with each other. As mentioned above, we have kept 10% of the event for testing. Zenith angle and neutrino energy distributions (which are the variables important for the sterile analysis) for both trained and test sample are shown in Fig. 5.6. Agreement within their statistical limits indicated lack of overtraining. Distribution for 9 other variables which were used in AdaBoost shows no overtraining either (Figure. 5.7). It is worth mentioning because we are using different sets of simulations for event selection and the final analysis, overfitting would not be an issue. However, no overfitting has been identified.

5.6 Starting Events

At this point, events in the sample are upgoing atmospheric neutrinos that pass through the detector, which include neutrinos that interact inside the detector volume ("starting events") with an outgoing muon track, and neutrinos that interacted outside of the detector volume and only the outgoing muon track will reach the detector ("throughgoing events"). In the previous sterile analyses [[1]] a mixture of both types of events was used. In this analysis, we are going to attempt to separate these two types of events.

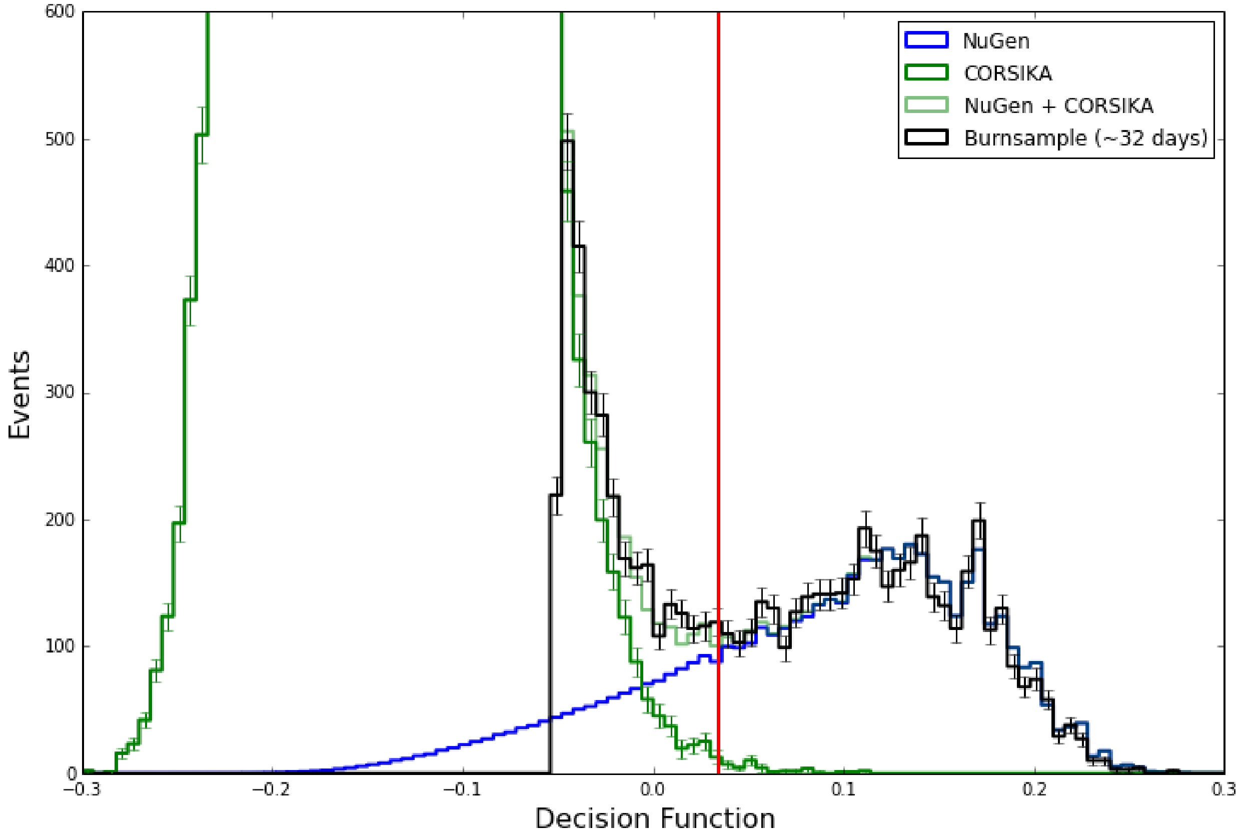


Figure 5.3 "Decision Function" histogram shows the distribution of atmospheric neutrinos (NuGen simulation) and atmospheric muons (CORSIKA simulation) for 1 year of data. The red line shows the cut which the events are classified in this analysis. Events on the right side of the red line, are kept. Black is 32 days of data as the burnsample.

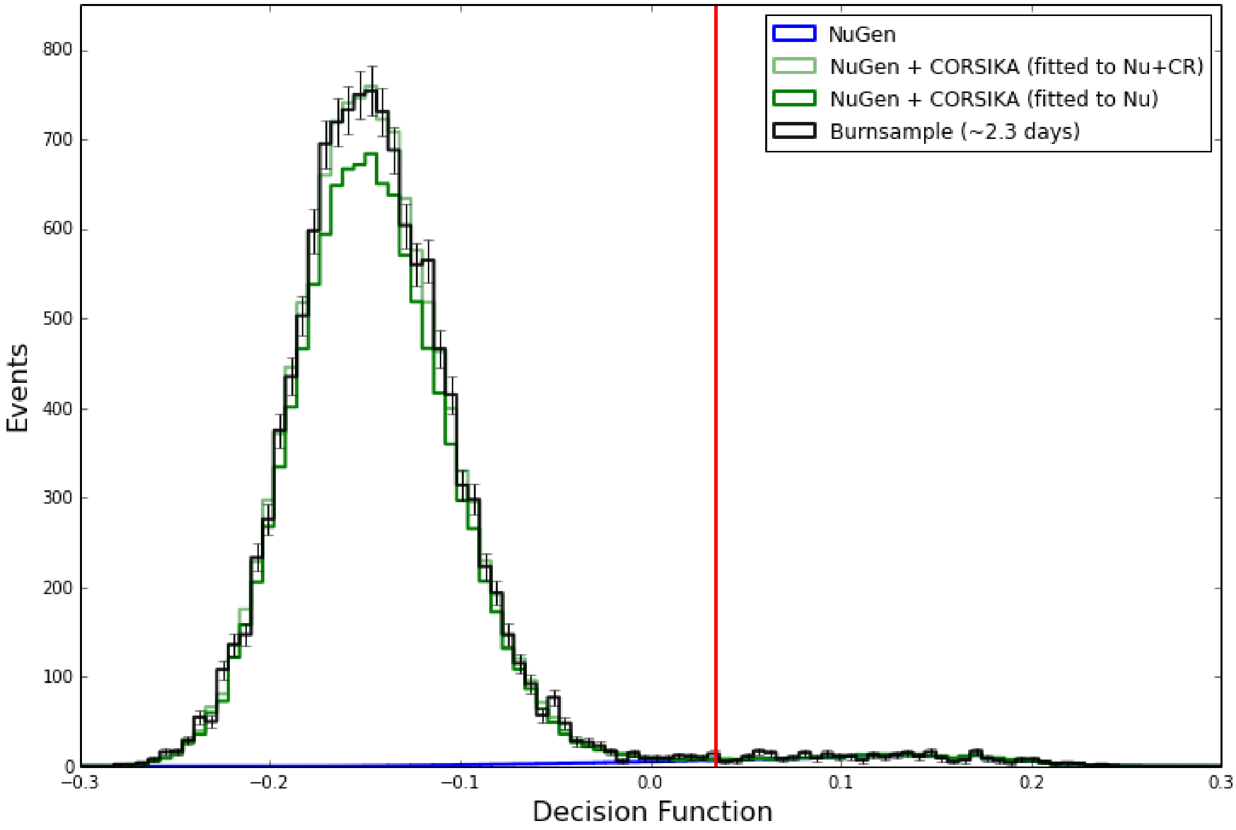


Figure 5.4 "Decision Function" histogram shows the distribution of atmospheric neutrinos (NuGen simulation) and atmospheric muons (CORSIKA simulation) for 1 year of data. The red line shows the cut which the events are classified in this analysis. Events on the right side of the red line, are kept. Black is 2.3 days of data from the burnsample from 2012 data. Simulation (neutrino + cosmic ray) is normalized to the data. There is a known disagreement between cosmic ray simulations (CORSIKA) and data.

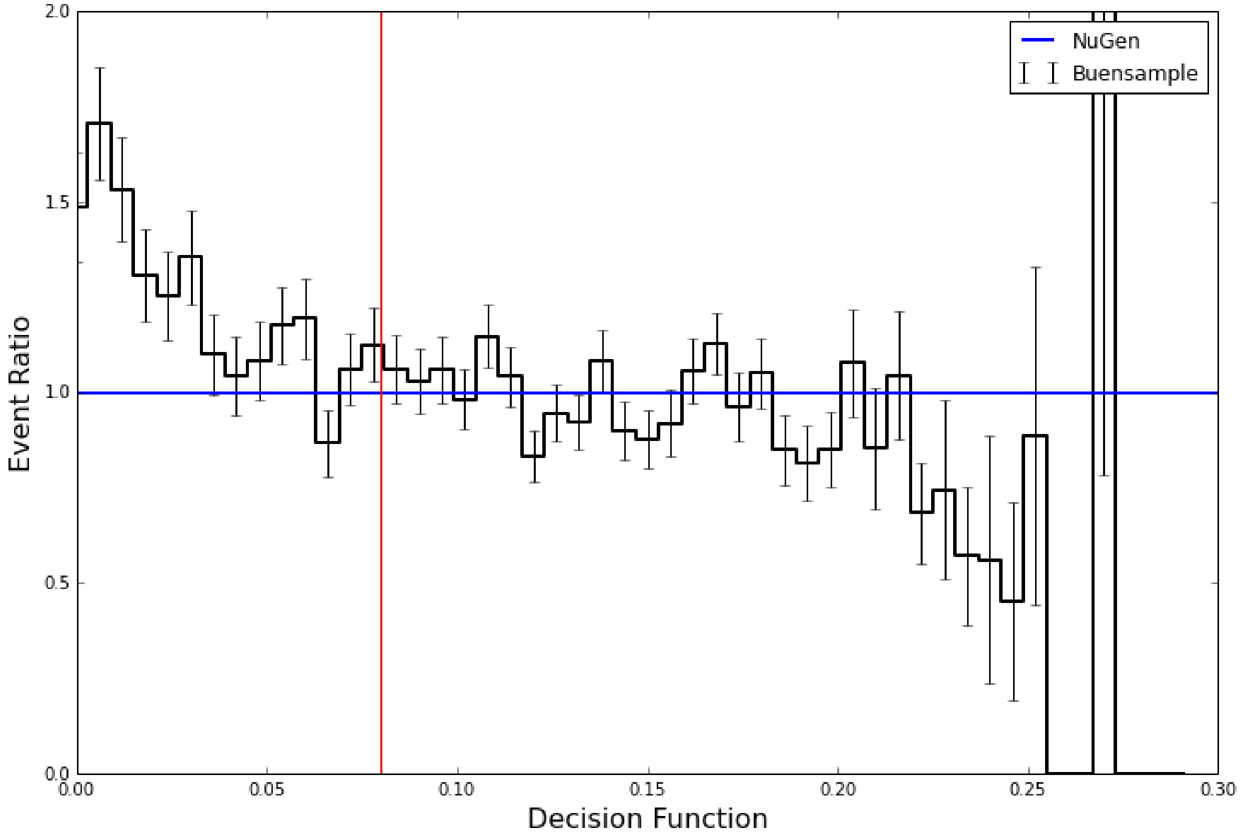
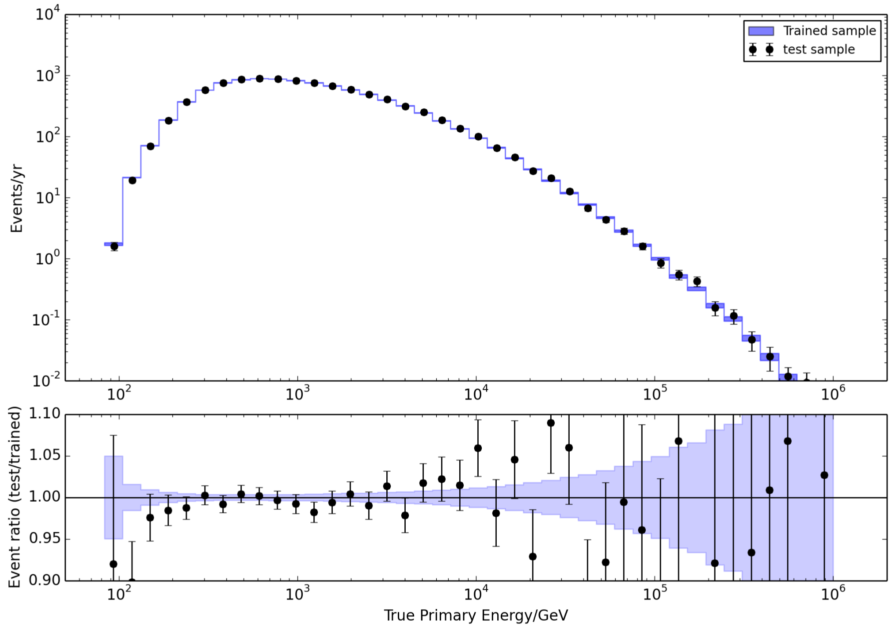
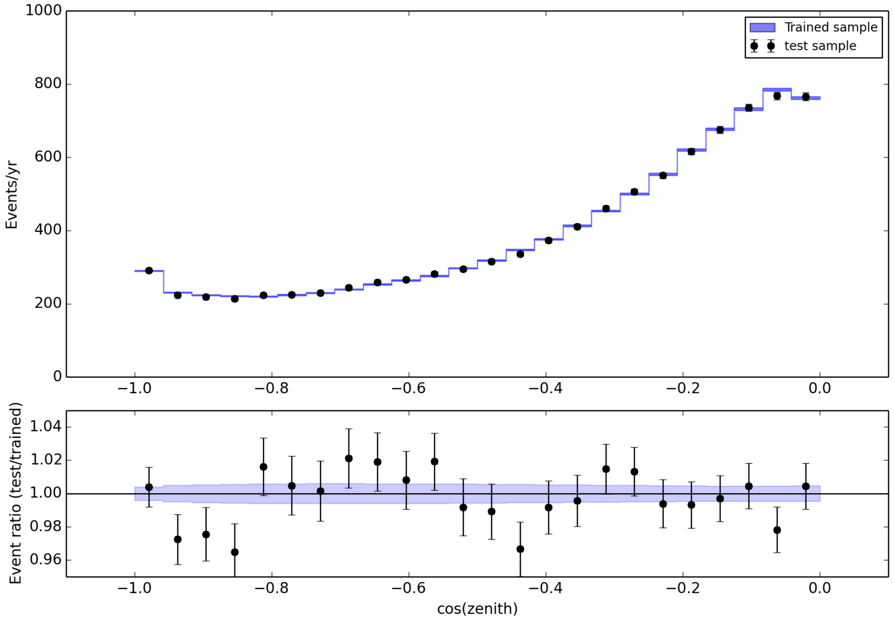


Figure 5.5 Ratio of data to simulated neutrino events is shown. Data agree with neutrino simulation on the right-hand side of the cut (red line). Increase in the data events at lower decision functions come from muon that is not included in this particular simulations. Those events are cut out and will not be used in the analysis.

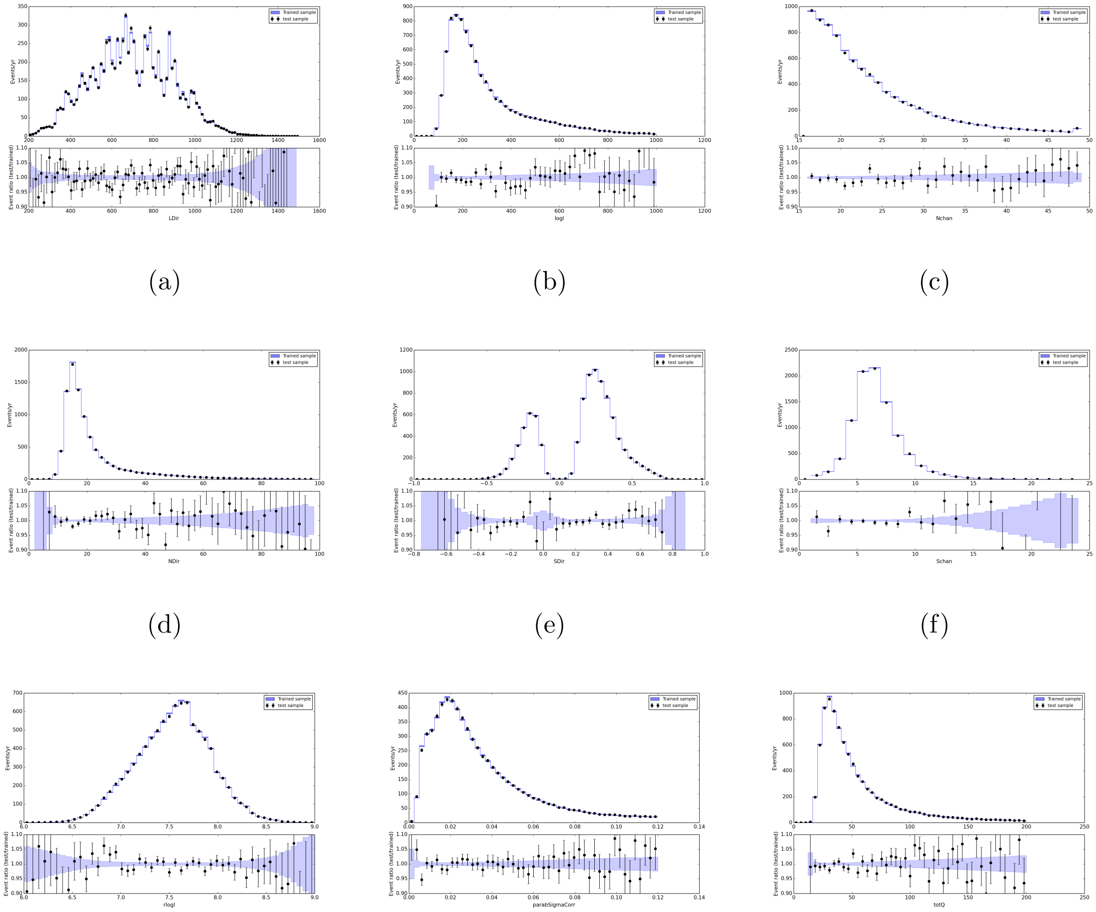


(a)



(b)

Figure 5.6 Comparison between distribution of trained and test samples are shown here. Figure (a) shows the true neutrino energy distribution for each sample and figure (b) show the distribution of cos(zenith angle). Agreement of both sample indicates a lack of overtraining.



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Figure 5.7 Comparison of distribution of trained and test samples for 9 variables used in AdaBoost are shown here. The tenth variable, cos(zenith), is shown in Fig.5.6(b). The plots show events' histogram and ratio (test to trained distribution). The sequence of figures show the number of events on the vertical axis and the observables on the horizontal axis. These are sequentially, (a) LDir, (b) LogL, (c) NChannel, (d) NDir, (e) SDir, (f) SChannel, (g) RLogL, (h) Corrected Paraboloid Sigma, and (i) total charge.

Since neutrino disappearance depends on its zenith angle (baseline) and energy (see chapter 1), the precisions with which we reconstruct these two properties have a direct effect in the sensitivity of the analysis to muon neutrino disappearance, and consequently to sterile neutrinos. After neutrino interaction a portion of its energy will be released at the in interaction vertex ("cascade energy"); The rest of the energy will be carried by the outgoing muon track. In general starting events have a much better energy resolution than throughgoing ones. There are three reasons for this. First, the ratio of the cascade to muon energy is not constant (see Fig. 5.8); Therefore for throughgoing events, which the cascade is outside of the detector, there will be an uncertainty on the primary neutrino energy since only the muon can be seen by the detector. Second, IceCube cannot determine the interaction vertex locations, if it is outside the detector, and since muons lose energy as they travel through ice, throughgoing muons have less energy at the detector than at the vertex. This difference is unknown to us since it depends on the location of the vertex. Finally, at these energies ($\sim TeV$) the length of muons id a few kilometers (from creation to decay) [[63]], therefore part of the muons are not in the kilometer-cubed detector, and given the randomness of stochastical losses of energy (which will be detected), there will be large uncertainties on energy reconstruction results. The last point has a negative effect on muon energy resolution, and cannot be removed in the starting tracks but can be suppressed by introducing other better-known properties of the particle.

Having a good energy resolution is a key advantage in this analysis, so it makes sense that we work with starting events for the analysis. However, we are not going to throw away throughgoing muons since they still contain some information about the sterile neutrinos, but we are going to separate them from the starting events, and we can subsequently fit our hypothesis on both sets of events simultaneously. This combination has not been done in the final analysis, but they can be added to improve the sensitivity for larger Δm^2.

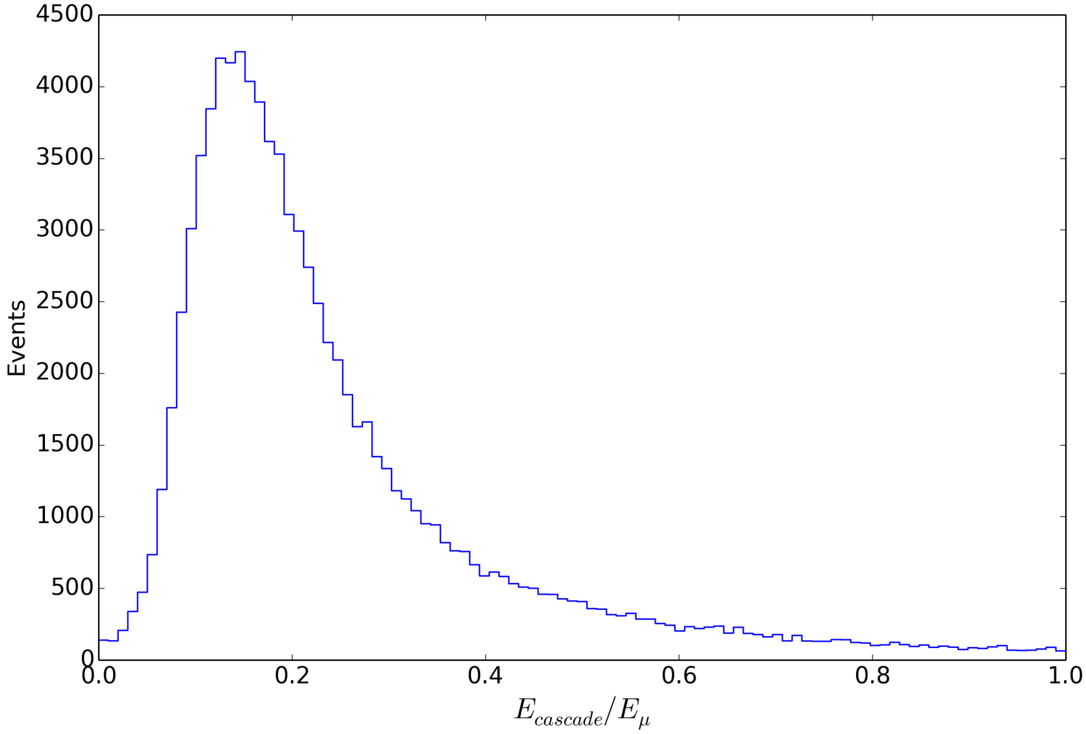


Figure 5.8 Histogram show the ratio of cascade to muon energy for high energy neutrinos in ice. Vertical axis has an arbitrary unit.

5.7 Energy Reconstruction

Energy reconstruction is based on energy losses along the muon track in the detector. These reconstructions can generally be divided into two segments, energy coming from the hadronic showers (which is reconstructible for starting tracks only, since the throughgoing track's interaction occurs outside of the detector volume) and electromagnetic losses of the muon tracks.

The average energy loss of the track can be simplified as [[65, 11]]:

$$-dE/dx = a(E) + b(E)x \tag{5.1}$$

where $a(E)$ is the ionization energy loss and $b(E)$ is the sum of e^+e^- pair production, bremsstrahlung and photo-nuclear contributions. The rates are shown in Fig. 5.9.

Each event will be divided into 10 m segments and for each segment energy loss will be estimated using "millipede" software [63]. Since Cherenkov light travels at an angle from the track before illuminating the DOMs, estimated energy losses along the track for throughgoing tracks, when falls outside of the detector volume. Therefore, if the first energy loss of the event is outside or very close to borders of the detector, it is probably a throughgoing event and the first energy loss along the track falls within the detector volume, the event is probably a starting event (see Section 5.8).

For throughgoing tracks, the analytical interpretation of Eq. 5.1 will be used for energy reconstruction [66, 67]. This method cannot be used for starting tracks since it uses the assumption of having an infinite track. It also does not consider any hadronic loss.

For starting track reconstruction, the energy losses obtained from "millipede" will be used. The first 3 losses (30 m) on the track, will be associated with hadronic loss. The rest of energy losses, from 30 m of the interaction vertex to the edge of the detector where the muon exits, are associated with the muon energy loss and are used to calculate the out-going muon energy. Primary neutrino energy is equal to the cascade plus muon energies, but in practice adding these two is not trivial.

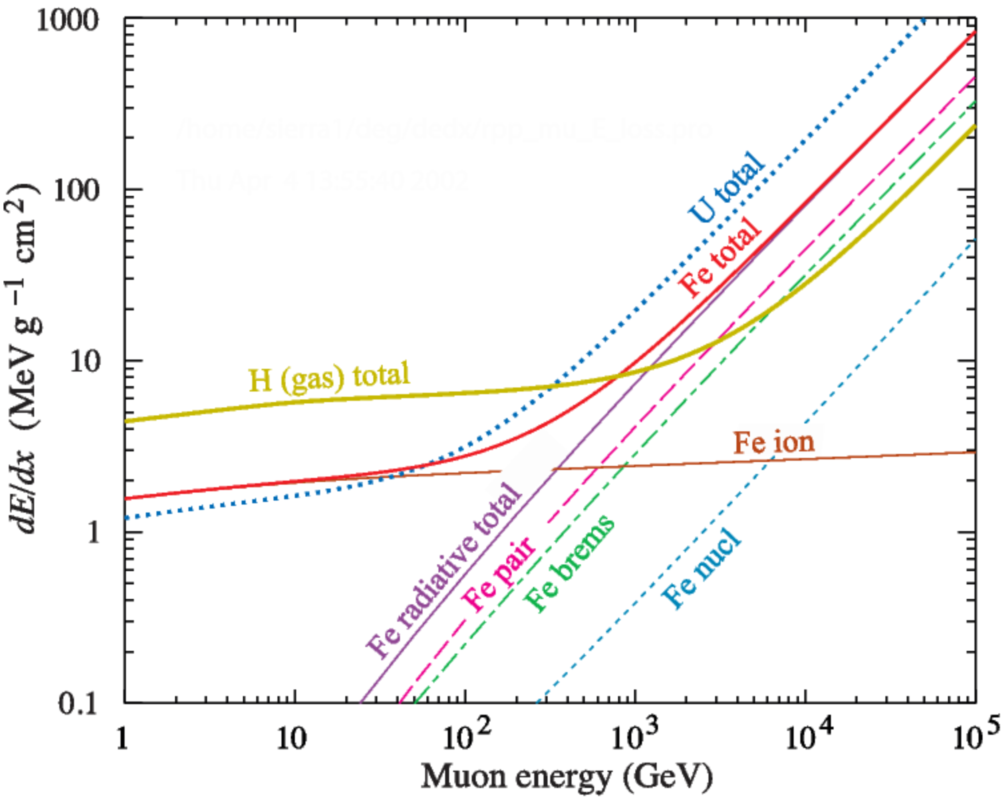


Figure 5.9 The average energy loss of muon in different elements as a function of muon energy [[11]]

I have tried to reconstruct the energy of neutrinos from starting tracks in two different ways. First by using an AdaBoost regressor (machine learning). Inputs for this regressor are cascade energy from millipede (explained above), dE/dx of the track using millipede losses, as well as the track length (from the cascade vertex to the border of the detector where the muon exits), and accumulated muon energy along the track as a function of percentile of track length saved in 100 variables (each for one percentile from 1 to 100) as shown in Fig. 5.10.

The second method is by unfolding muon energy from dE/dx calculated from millipede losses into the true muon energy, and add them to the cascade energy. Muon energies in IceCube are usually calculated analytically from Eqn. 5.1 and represent an energy proxy rather than muons' actual energy (unlike cascade energies); However, by unfolding dE/dx

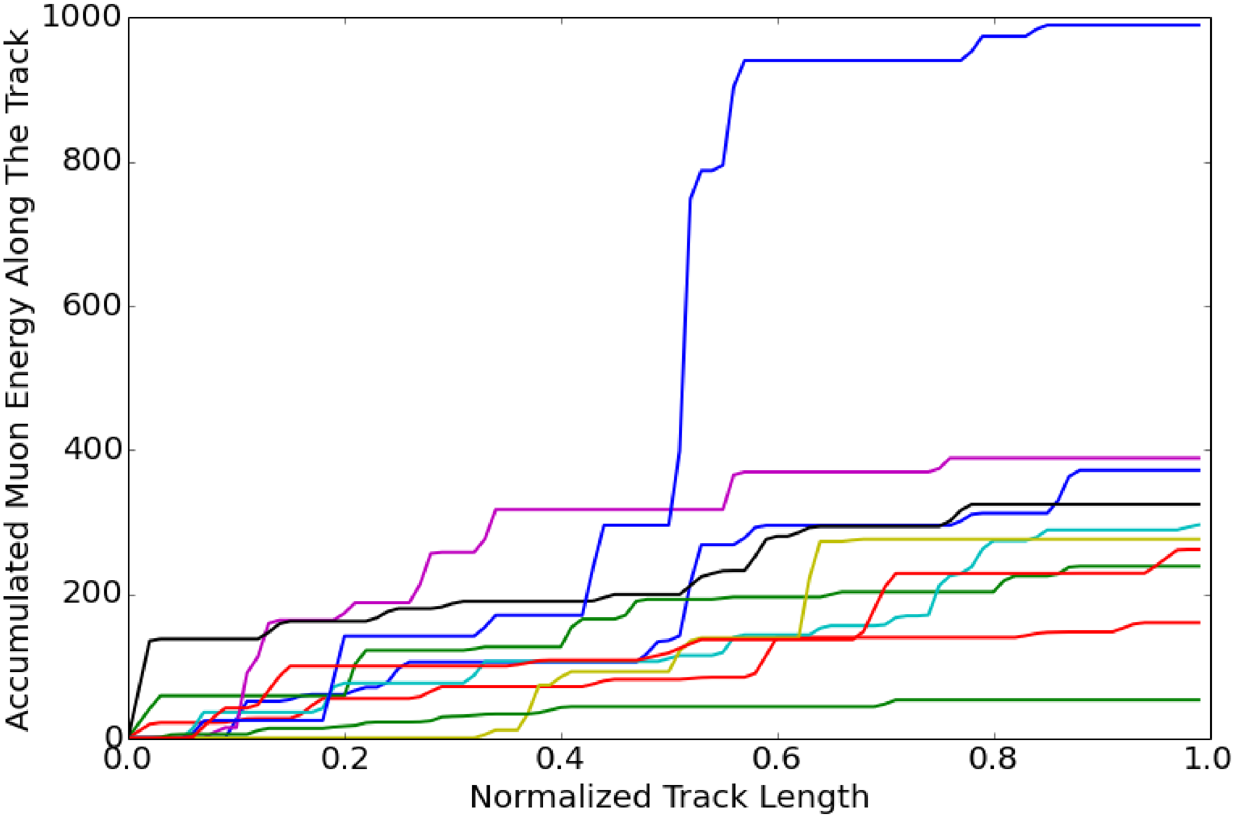


Figure 5.10 Accumulated energy of muons along their track as a function of normalized track length, for several simulated tracks. This information is used for neutrino energy reconstruction using AdaBoost regressor. Each line represents a sampled event. Y-axis has a unit of GeV.

into true muon energy, the output of muon energy reconstruction has the same unit as the cascade energy and can be added to each other. Our reconstructed cascade energy is a good representation of actual energy, so there is no need to unfold the cascade energy. Fig. 5.11 show a 2-dimensional histogram of reconstructed neutrino energy obtained by the unfolding method as a function of true neutrino energy.

Fig. 5.12 and 5.13 compares energy resolution of these two neutrino reconstruction methods. Both methods show a large improvement in the energy-resolution of neutrinos compared to muon energy-resolution. This is about 0.4 for starting muons (this will be much larger for throughgoing muons since those muons lose an unknown amount of energy before they reach the detector). As is shown in the figures, the unfolding method has a better resolution at almost all neutrino energies compared to the AdaBoost method. Even though AdaBoost has a better resolution for higher reconstructed energies, it is important to keep in mind that most of the events in this analysis are around 1 TeV. We have also seen that the unfolding method gives a better sensitivity for the final analysis. Ergo, the unfolding method is used from here on in the analysis as the primary method for neutrino energy reconstruction.

5.8 Starting Events Veto

In this section, I will discuss how starting tacks will be separated from throughgoing tracks. As was discussed earlier in the chapter, starting track have a better energy resolution and will give us an edge on the sensitivity to sterile neutrinos.

To perform a starting track veto, new variables need to be introduced. “MinDist” is defined as a point’s minimum distance to the detector borders in three dimensions. A positive value means the point is outside of the detector volume and negative means it is inside. Detector side borders are defined as the broken hexagon i.e. a concave octagon, which is the shape of the detector strings from above, and its bottom and top are defined as the height of upper and lower DOMs in the middle string (number 36), which is very close

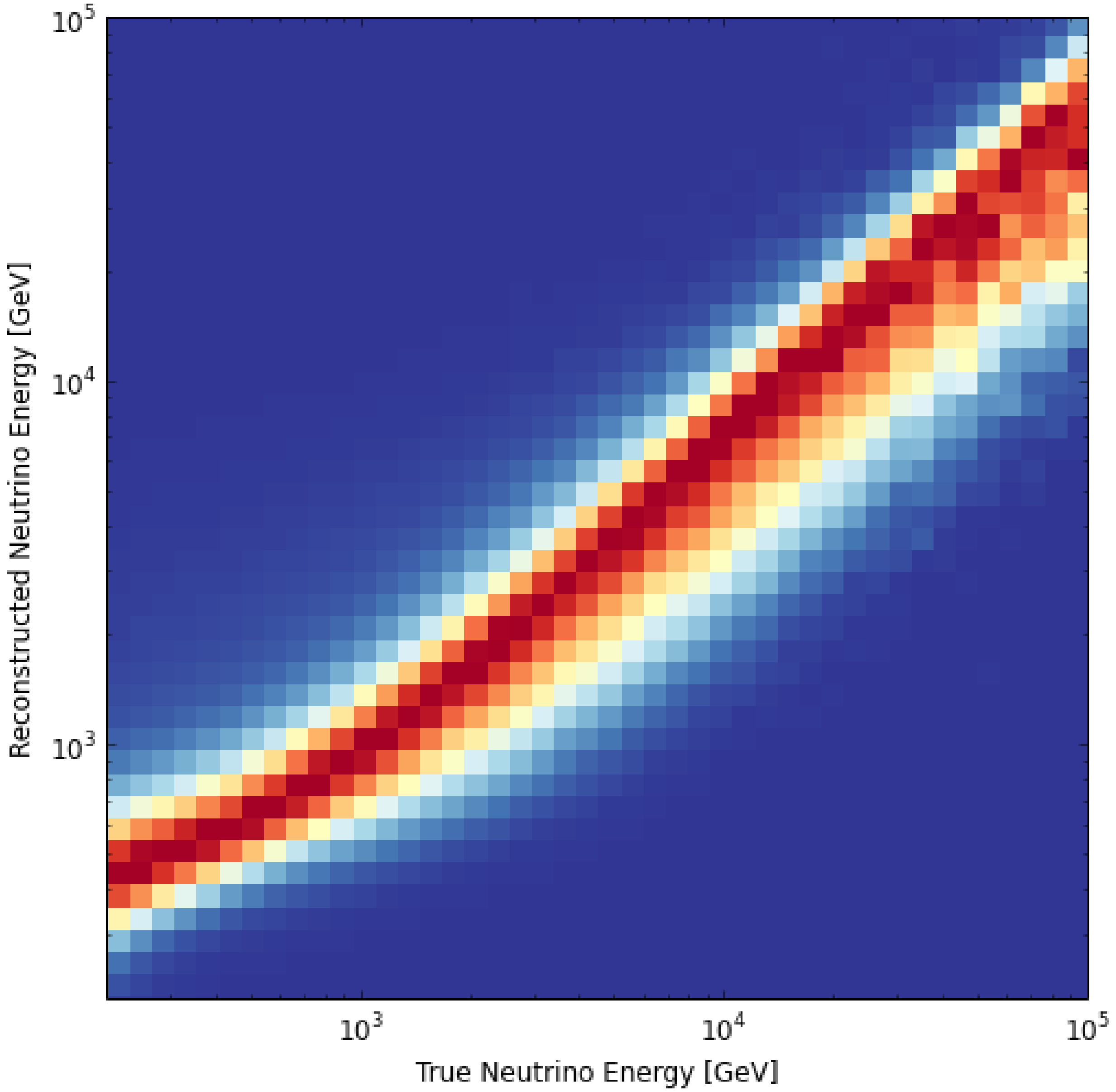


Figure 5.11 This plot shows the 2-dimensional histogram of reconstructed neutrino energy obtained by the unfolding method as a function of true neutrino energy from the simulation. Bins in this plot are vertically normalized and the color scale is linear.

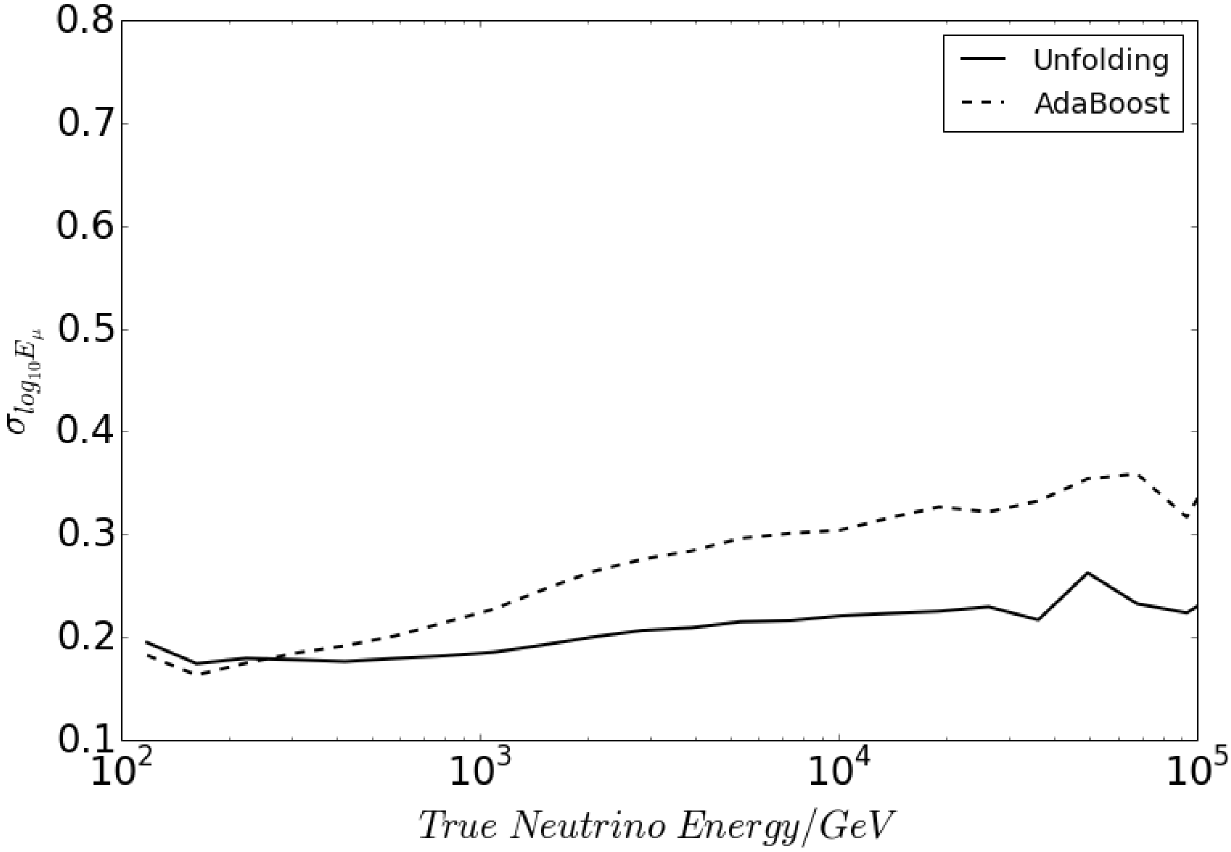
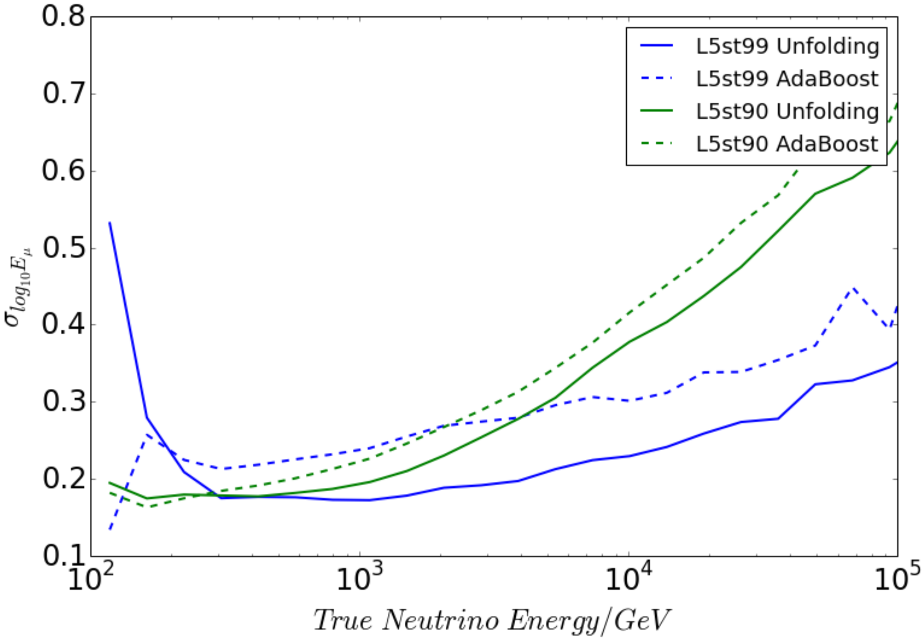
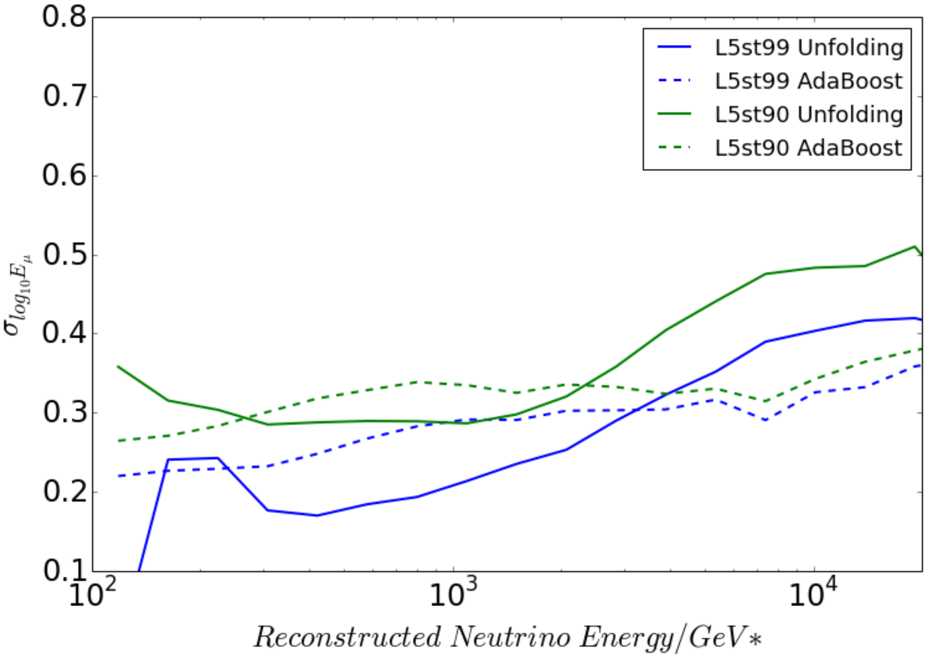


Figure 5.12 This plot compares the energy resolution of starting events for different neutrino reconstruction methods as a function of true neutrino energy.



(a)



(b)

Figure 5.13 These plots show the energy resolution for two different selections starting sample with 90 and 99 percent starting-track purity, L5st90 and L5st99 (see Sec. 5.8) for two different methods of neutrino energy reconstructions. The peak on the left hand side of 'L5st99 Unfolding' is due to the fact that there are very few events at those energies in the sample. The difference in these plots with Fig. 5.12 is that all events in the sample are included i.e. impurity of the starting sample (throughgoing events in the sample). Figure (a) shows the resolution as the function of true neutrino energy and figure (b) show the resolution as a function of reconstructed energy (for the specific method used).

to DOMs heights on other strings. Now "Cascade_MinDist" can be defined as the *MinDist* of the event vertex which is calculated in millipede from the energy loss reconstruction.

It is obvious that starting tracks should have a non-positive *Cascade_MinDist*. However in reality about 25 percent of events with negative *Cascade_MinDist* are throughgoing events. This is due to the fact that muon energy loss is stochastic and some events may produce no light over a portion of their length, which by accident may be at the detector borders. Ergo they may appear as starting events. The easiest way to reduce the factor error is to change the starting event cut from *Cascade_MinDist* of less than zero to a lower limit, since the probability of a muon event being dark in larger lengths is smaller. However, this probability depends on the energy of the events, so it is best to make *Cascade_MinDist* cut based on the reconstructed energy of the event (see Section 5.7).

After cutting out any events with a positive *Cascade_MinDist* (cascades outside of the detector), AdaBoost [[68]] is used with three input variables viz. *Cascade_MinDist*, cascade energy measured by millipede, and dE/dx of muon track coming from millipede's energy losses. Fig. 5.14 shows how well AdaBoost can separate starting from throughgoing events. In this plot the comparison of simulation with data (32.6 days of burnsample) and their agreement shows that the features of the histogram that are not coming from simulation, are AdaBoost features (the burnsample is a random test sample from the data which will only be used for tests and will not be used for the final results). Fig. 5.17 shows agreement between MC and data in all three input variables of AdaBoost at the final level which indicates that there is no overfitting. Table 5.8 shows the p-values for each input variable of starting-track events, and table 5.8 shows the p-values for the throughgoing events. Throughgoing events have not been used in this analysis, but they can be added in the future. Table 5.8 shows the p-values for the variables used in the machine to distinguish starting-track from throughgoing events.

To get a sample of starting tracks, we need to cut out events that have less than a certain decision function value. The larger the limit, the better the purity of our sample, but the smaller the statistics (see Fig. 5.16). Purity is defined as the ratio of the number of truly

Variables	**Chi2_stat**	**Critical Value 0.95**	**p-value**
Rec Zenith	9.6	28.9	0.94
Rec Nu Energy	10.0	28.9	0.93
NDir	63.5	64.0	0.05
SDir	137.7	55.8	0.85
N Channel	13.4	26.3	0.64
N String	36.4	62.8	0.84
logL	53.9	65.2	0.26
RlogL	29.4	65.2	0.98
total charge	45.8	62.8	0.48
corrected Paraboloid Sigma	18.8	41.3	0.90

Table 5.1 P-value of test/train comparison for each input used in the machine. These are the starting-track events only.

”Variables”	**Chi2_stat**	**Critical Value 0.95**	**p-value**
Rec Zenith	19.3	28.9	0.38
Rec Nu Energy	24.7	25.0	0.05
NDir	40.8	64.0	0.73
SDir	28.1	56.9	0.94
N Channel	15.7	26.3	0.47
N Stirng	35.5	64.0	0.89
logL	24.9	40.1	0.58
RlogL	78.4	65.2	0.004
total charge	37.5	62.8	0.81
corrected Paraboloid Sigma	21.8	41.3	0.79

Table 5.2 P-value of test/train comparison for each input used in the machine. These are the throughgoing track events only.

”Variables”	**Chi2_stat**	**Critical Value 0.95**	**p-value**
minDistDetCas	44.3	60.5	0.46
Rec dEdx	15.1	28.9	0.66
Rec Cas Energy	8.0	29.9	0.98

Table 5.3 P-value of test/train comparison for each input used in the machine, in order to predict starting-track from throughgoing track events.

starting events to all events in the sample. It is obvious that purity can only be measured in MC simulations.

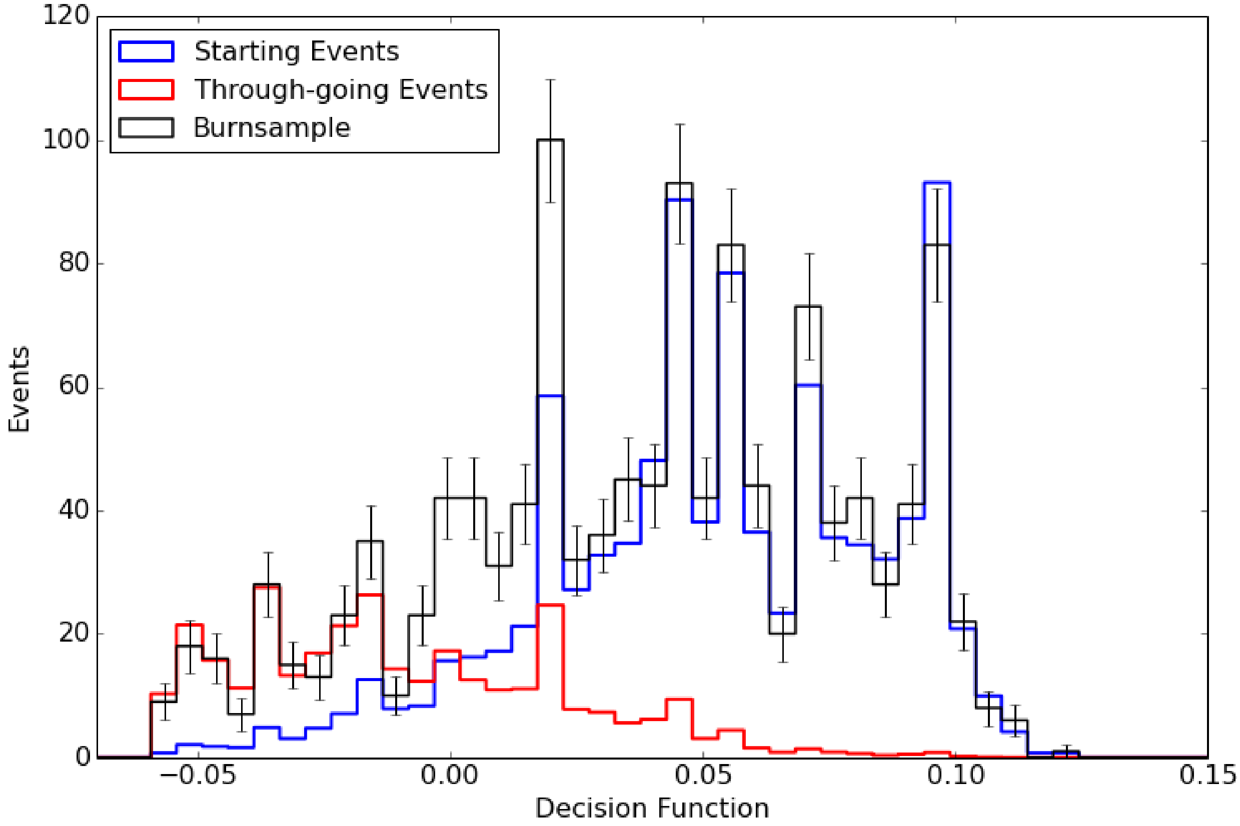
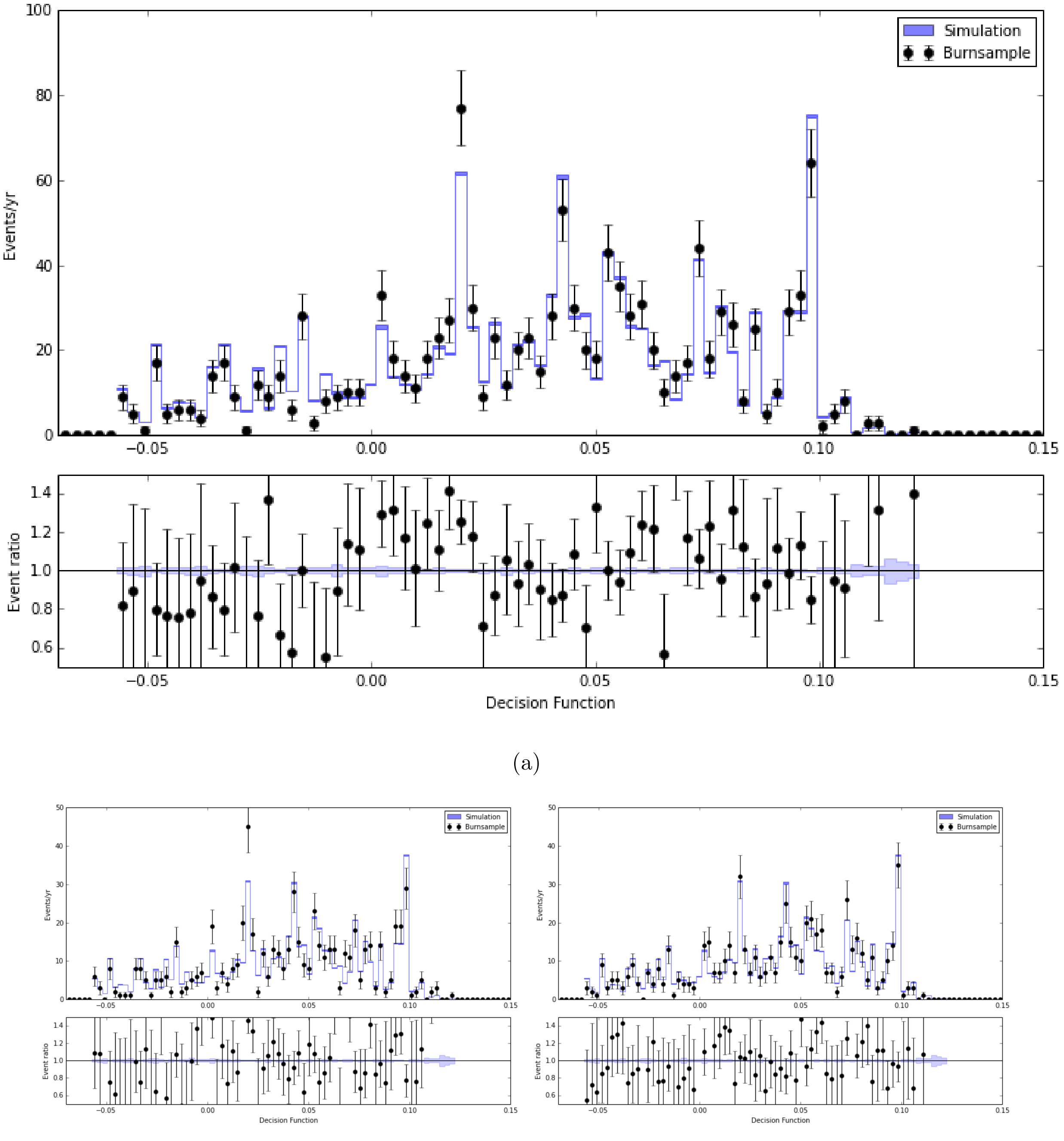


Figure 5.14 This plot shows histogram of decision functions from AdaBoost to classify starting tracks. Blue is for starting-tracks, red is for throughgoing tracks, and black is 32.6 days of burnsample which agrees with the sum of the other two histograms.



(a)

(b)

(c)

Figure 5.15 (a) Histogram of decision functions from AdaBoost used for classifying starting tracks. Blue is the sum of starting tracks and throughgoing tracks, and black is 32.6 days of burnsample which agrees with the sum of the other two histograms (similar to the plot above). The rise in around -0.01 to 0.02 in the plot above seems to be statistical only. It is more obvious if we split the burnsample into two samples (b) and (c).

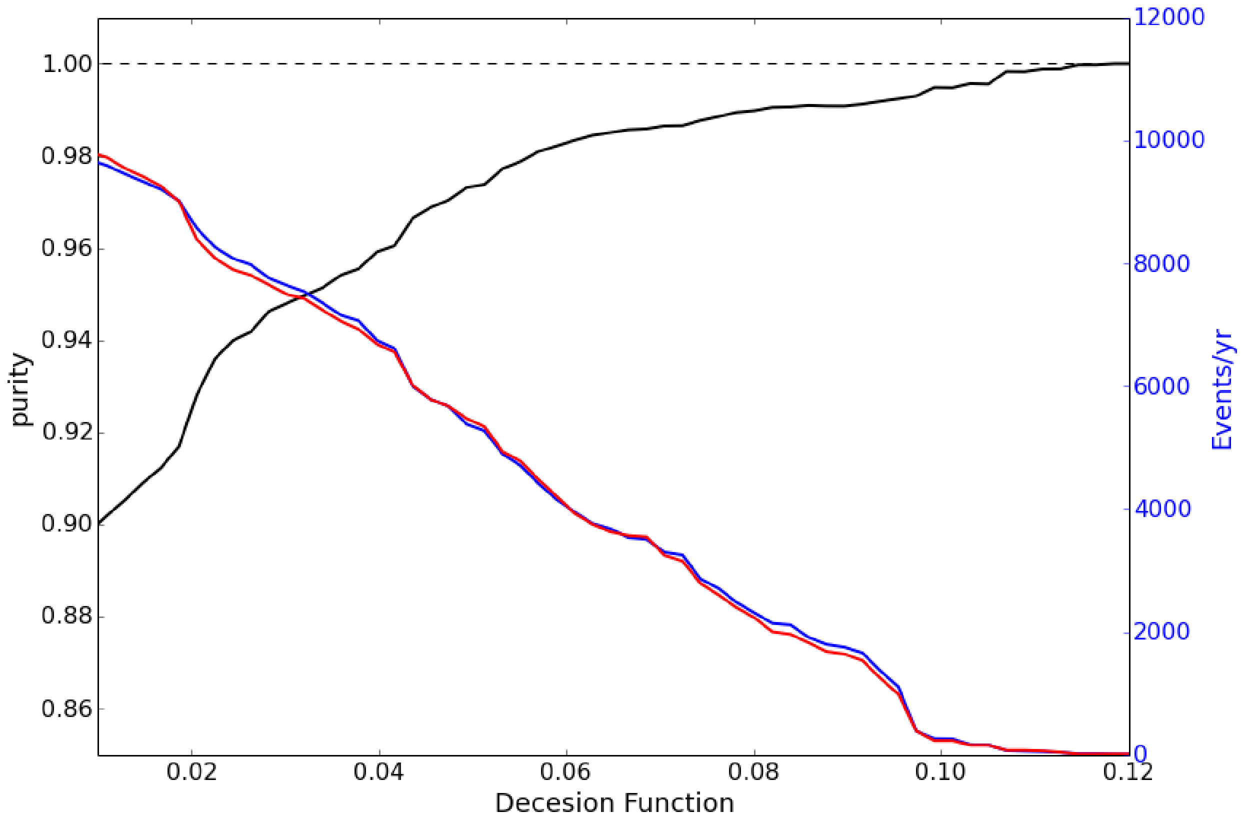
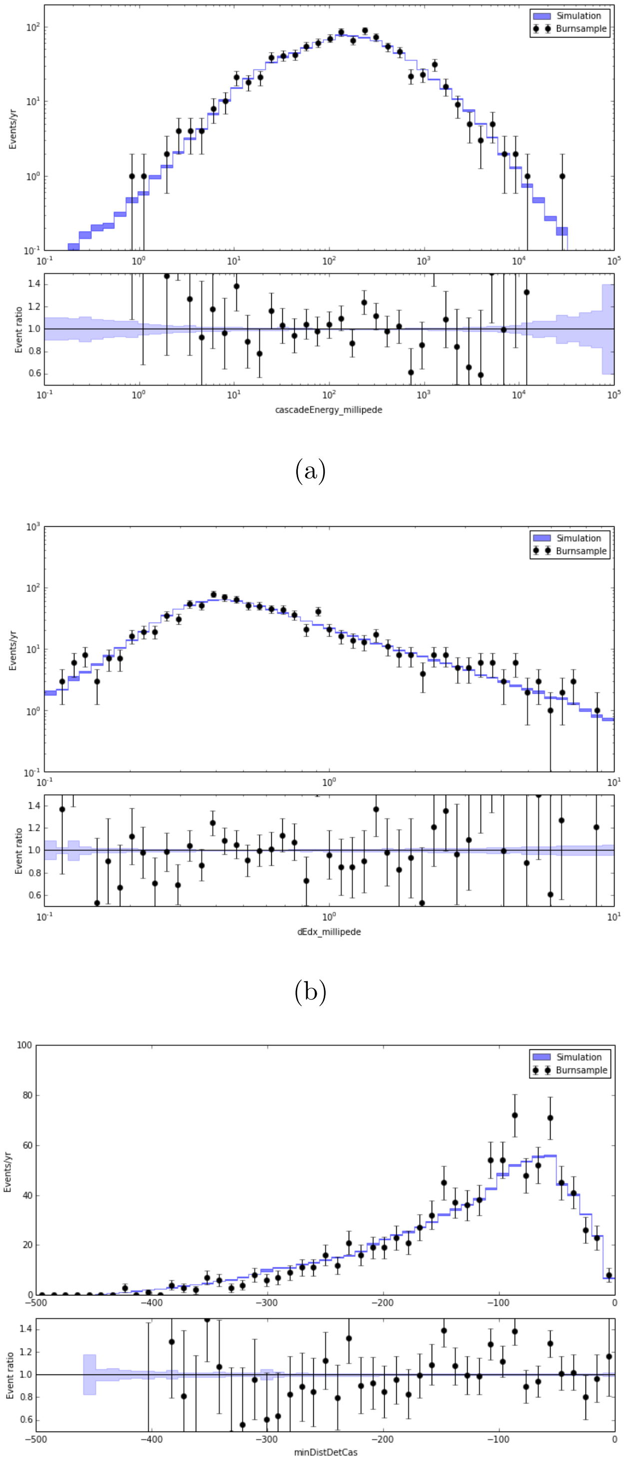


Figure 5.16 This plot shows purity versus the number of events per year in the sample when cutting at different decision functions e.g. for 99% purity, we need to keep events with only decision function of more than 0.8, which means having about 2500 starting event per year, in which there is about 25 event, i.e. 1%, throughgoing events. Blue represents MC event rate, red is the 32-day burnsample, and the black line shows the purity of MC.

In this analysis, samples with 90 percent starting-track purity are being used (decision function cut is at 0.01). This sample contains about $12,000$ events per year of data. This level of purity will give us the optimal event selection for the sterile neutrino analysis. Should we choose to tighten the cut to achieve a 99 percent starting-track purity, the sample will have about $2,500$ events.

Figures 5.18 and 5.19 show histogram of final event selection. Plot 5.20 show the ratio of simulated to experimental data. The high ratio guarantees us that simulation statistical uncertainty is significantly lower than the experimental data.

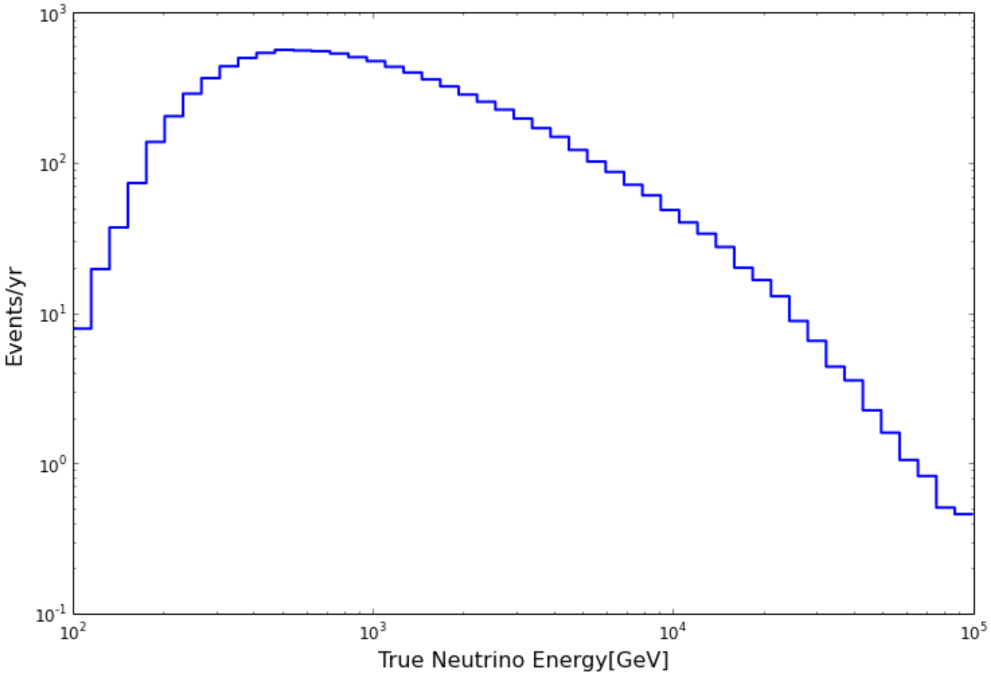


(a)

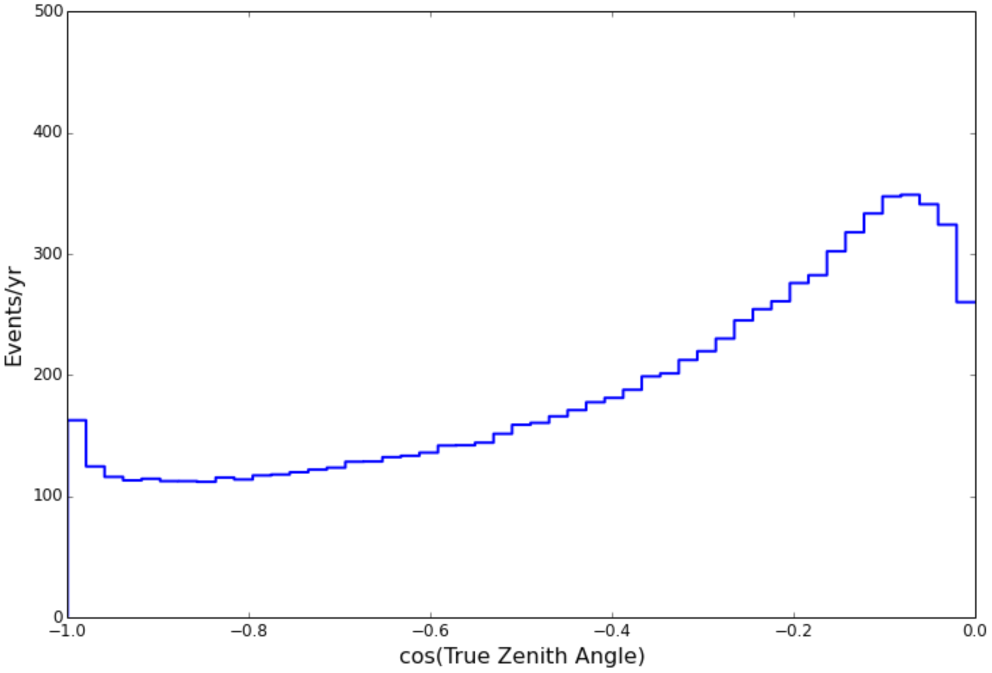
(b)

(c)

Figure 5.17 Agreement between MC and data in all three input variables of AdaBoost at final level indicates that there is no overfitting. (a) Cascade Energy, (b) dE/dx, and (c) MinDistDetCas are plotted.

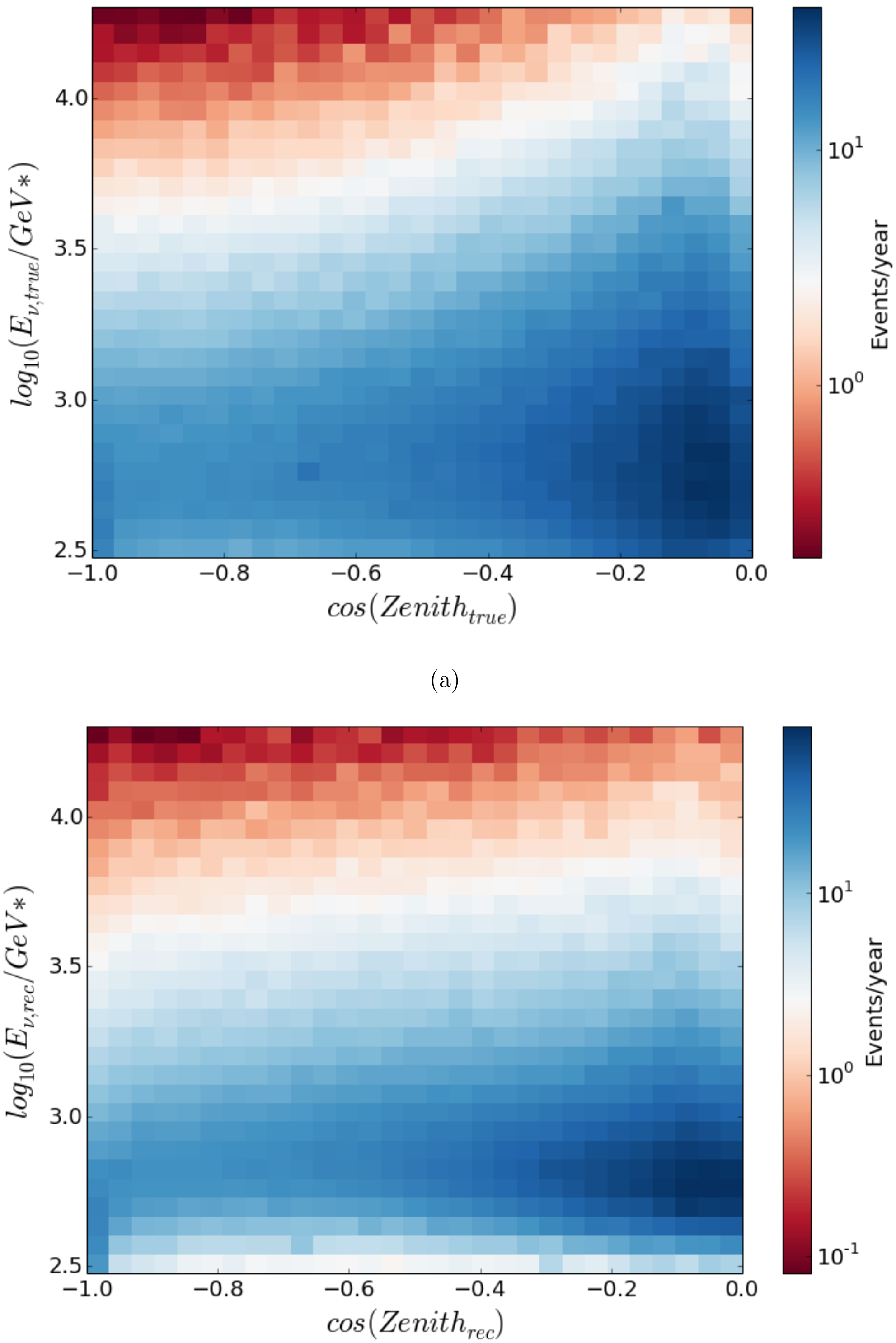


(a)



(b)

Figure 5.18 The plots show the 1-dimensional histogram of events at L5st90 (final) level. Histogram (a) plots the true energies of the simulated events. Events with energies between 300 GeV and 20 TeV have been used for the analysis. Histogram (b) plots the cosine of the true zenith angles of simulation events.



(a)

(b)

Figure 5.19 The plots show the 2-dimensional histogram of events at L5st90 (final) level, as a function of log(energy) and cos(zenith). Histogram (a) plots the true variable of the simulated events and histogram (b) plots the reconstructed variables.

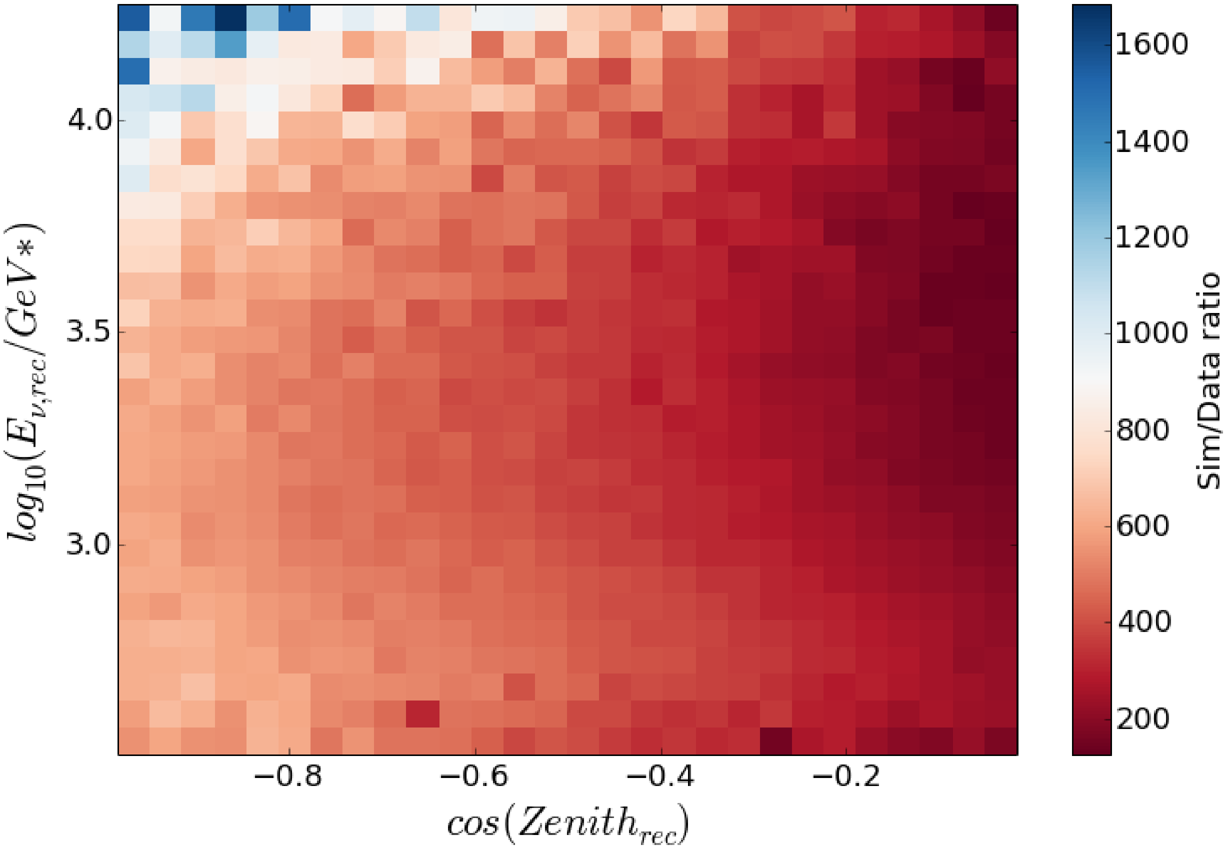


Figure 5.20 Simulated to experimental data ratio of one year of data is shown in this plot.

Chapter 6

Systematics

Systematic errors are the errors associated with uncertainties in measurements caused by equipment or uncertainty of some physical values. In this chapter, I will discuss the systematics in this sterile neutrino search analysis.

6.1 Nuisance Parameters

Five nuisance parameters have been used to fit the data to simulated events in this analysis. We can divide them into two categories of "atmospheric flux systematics" and "detector systematics." Table 6.1 shows the input values for the parameters.

- Atmospheric Flux Systematics
  - Normalization (no prior)
  - Atmospheric flux slope (Gaussian prior)
  - Pion/Kaon ratio (Gaussian prior)
  - Neutrino/Antineutrino ratio (Gaussian prior)
- Detector Systematics
  - DOM efficiency (Gaussian prior)

Atmospheric flux systematics come from our uncertainty of these physical values. In order to fit the data to the analysis, we float these continuous values with or without priors, to get the maximum likelihood. See Chapter 7 for details on the likelihood analysis.

Parameters	**Central Value**	**Gaussian Prior Width**
Normalization	1	no prior
Cosmic Ray Spectral Shift	0	0.05
Pion/Kaon ratio	1	0.1
Neutrino/Antineutrino ratio	1	0.025
DOM efficiency	0.99	no prior

Table 6.1 Nuisance parameters used to fit the data.

DOM efficiency refers to the corrected efficiency of the DOMs (see Chapter 2). Systematics rising from the uncertainty of DOM efficiency is one of the largest systematics for this analysis. Although, it may be small for other analyses, have its effect could be similar to a sterile neutrino signal. We have simulated three different simulation sets with DOM efficiencies of 0.9, 0.99, and 1.089. And we created splines on final histograms (logarithm of energy vs. cosine of zenith angle) to interpolate the values in between. This way, we can consider this as a continuous nuisance parameter in order to fit the data using a likelihood analysis. Figure 6.1 shows the distribution of nuisance parameters using sampled simulated data.

Some of IceCube simulations use oversized DOMs in order to save computer time and then scale it back at the end of the simulation. Our simulations have not used oversized DOMs therefore, this is not one of our systematics.

Variations in Earth models can also be a factor since different densities in the Earth's interior can cause different oscillation pattern of neutrinos as they travel through Earth. But previous analyses have shown that this effect is virtually nonexistent [1].

Systematics can also arise from discrete parameters as opposed to continuous parameters. And discrete parameters can be considered in a similar manner in the likelihood analysis. Even though our software programs have already had the capability of taking discrete parameters into account. We have not used any discrete nuisance parameters due to their marginal effect. Ice model and cross-section model uncertainties are two examples of discrete systematics.

6.2 Ice Uncertainty

Ice properties are important inputs that are needed in order to have a reliable simulation and reconstruction of events. These properties are calculated by using flasher data information. Flashers refer to LEDs installed on each DOM for calibration purposes.

SPICE Mie has considered the default ice model used when the simulations were developed because SPICE Lea (which includes anisotropy) requires full propagation of all photons and is computationally more expensive and needs a lot of GPU time. We know that SPICE Lea has a slightly lower normalization and a little shape discrepancy. These discrepancies are much smaller than one-year statistical errors, shown in the previous chapter. Models with different absorption and scattering length of the ice have the largest ice-related discrepancies but as you can see in Figures 6.2 and 6.3, they are still smaller than one-year statistical errors. Therefore, due to their marginal effect on this dataset, they are not included in the systematics.

Another potential ice systematic errors could come from hole-ice, which is part of the ice that has been melted and refrozen during construction of the detector in order to lower the strings. However, different hole-ice models have very small discrepancies (even compared to no-hole-ice). And I did not see any need to consider it as a systematic error. It is important to note that creating simulations with different hole-ice models, this would have been a very expensive computational process.

6.3 Cross Section

Scattering between a particle and the quarks inside the nucleus is called deep inelastic scattering, which happens at very high energies such for as the muons in this analysis. Cross sections can be obtained from structure functions, in which their distributions are calculated by theoretical models.

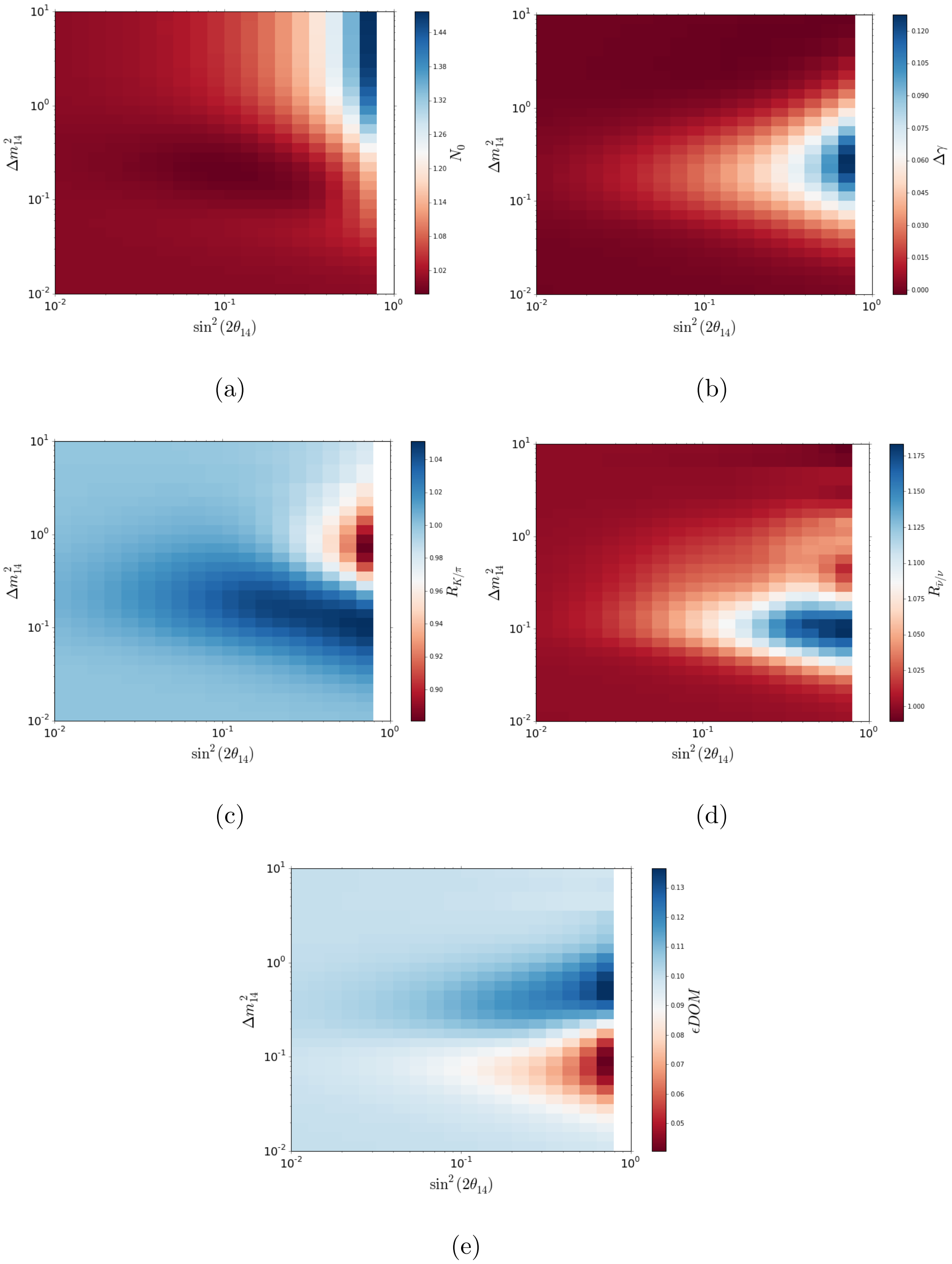
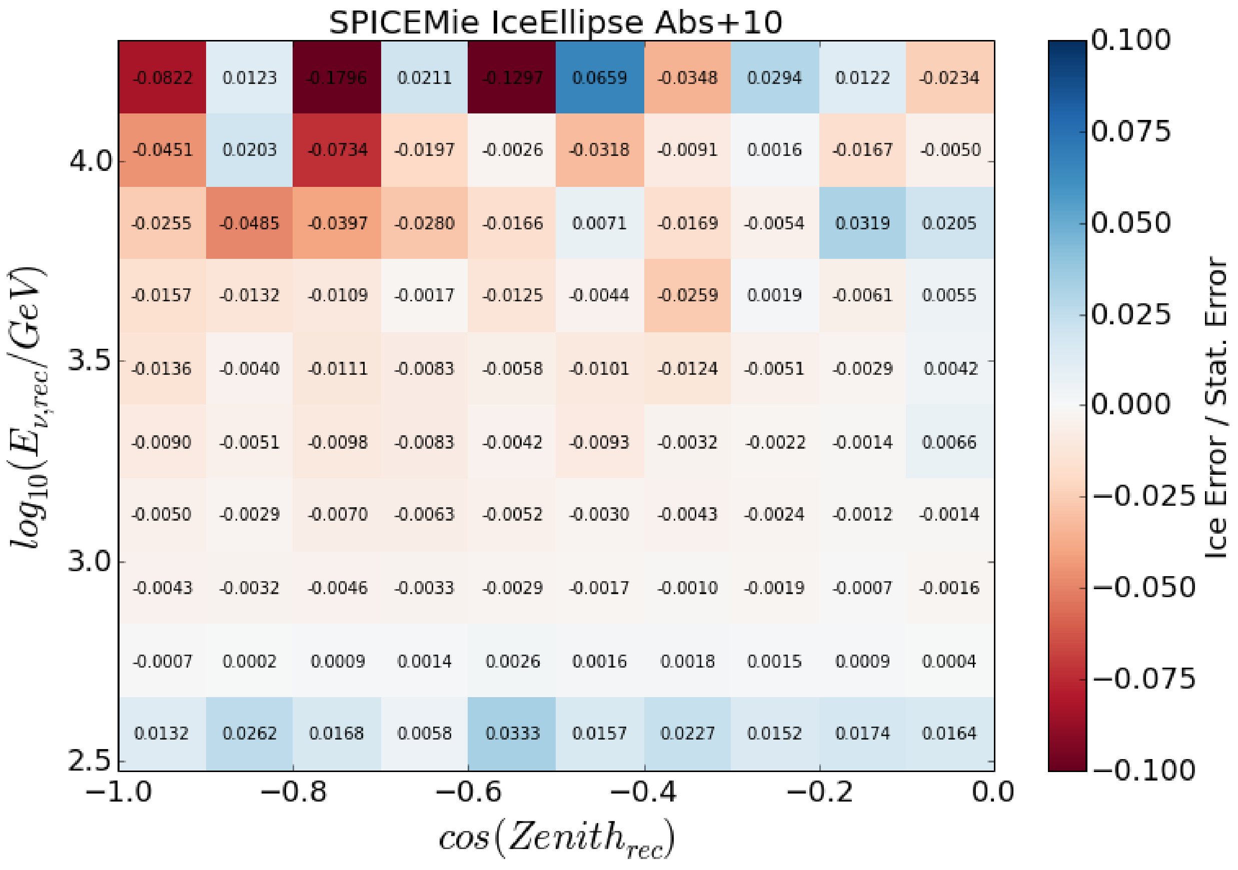


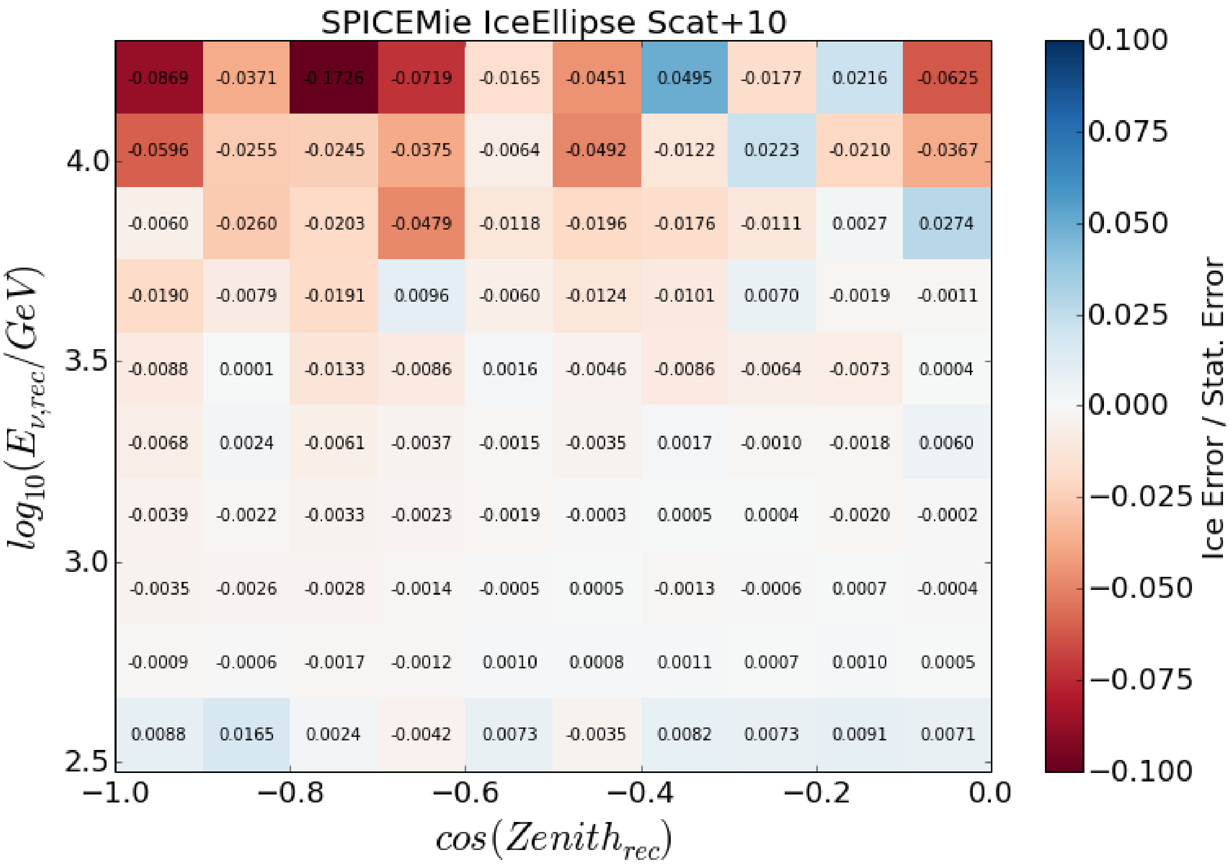
Figure 6.1 Distribution of nuisance parameters for a null hypothesis is shown, using sampled simulated data. (a) Normalization, (b) $\delta\gamma$, (c) kaon to pion ratio, (d) antineutrino to neutrino ration, and (e) ϵDOM are plotted.



-0.0822	0.0123	-0.1796	0.0211	-0.1297	0.0659	-0.0348	0.0294	0.0122	-0.0234
-0.0451	0.0203	-0.0734	-0.0197	-0.0026	-0.0318	-0.0091	0.0016	-0.0167	-0.0050
-0.0255	-0.0485	-0.0397	-0.0280	-0.0166	0.0071	-0.0169	-0.0054	0.0319	0.0205
-0.0157	-0.0132	-0.0109	-0.0017	-0.0125	-0.0044	-0.0259	0.0019	-0.0061	0.0055
-0.0136	-0.0040	-0.0111	-0.0083	-0.0058	-0.0101	-0.0124	-0.0051	-0.0029	0.0042
-0.0090	-0.0051	-0.0098	-0.0083	-0.0042	-0.0093	-0.0032	-0.0022	-0.0014	0.0066
-0.0050	-0.0029	-0.0070	-0.0063	-0.0052	-0.0030	-0.0043	-0.0024	-0.0012	-0.0014
-0.0043	-0.0032	-0.0046	-0.0033	-0.0029	-0.0017	-0.0010	-0.0019	-0.0007	-0.0016
-0.0007	0.0002	0.0009	0.0014	0.0026	0.0016	0.0018	0.0015	0.0009	0.0004
0.0132	0.0262	0.0168	0.0058	0.0333	0.0157	0.0227	0.0152	0.0174	0.0164

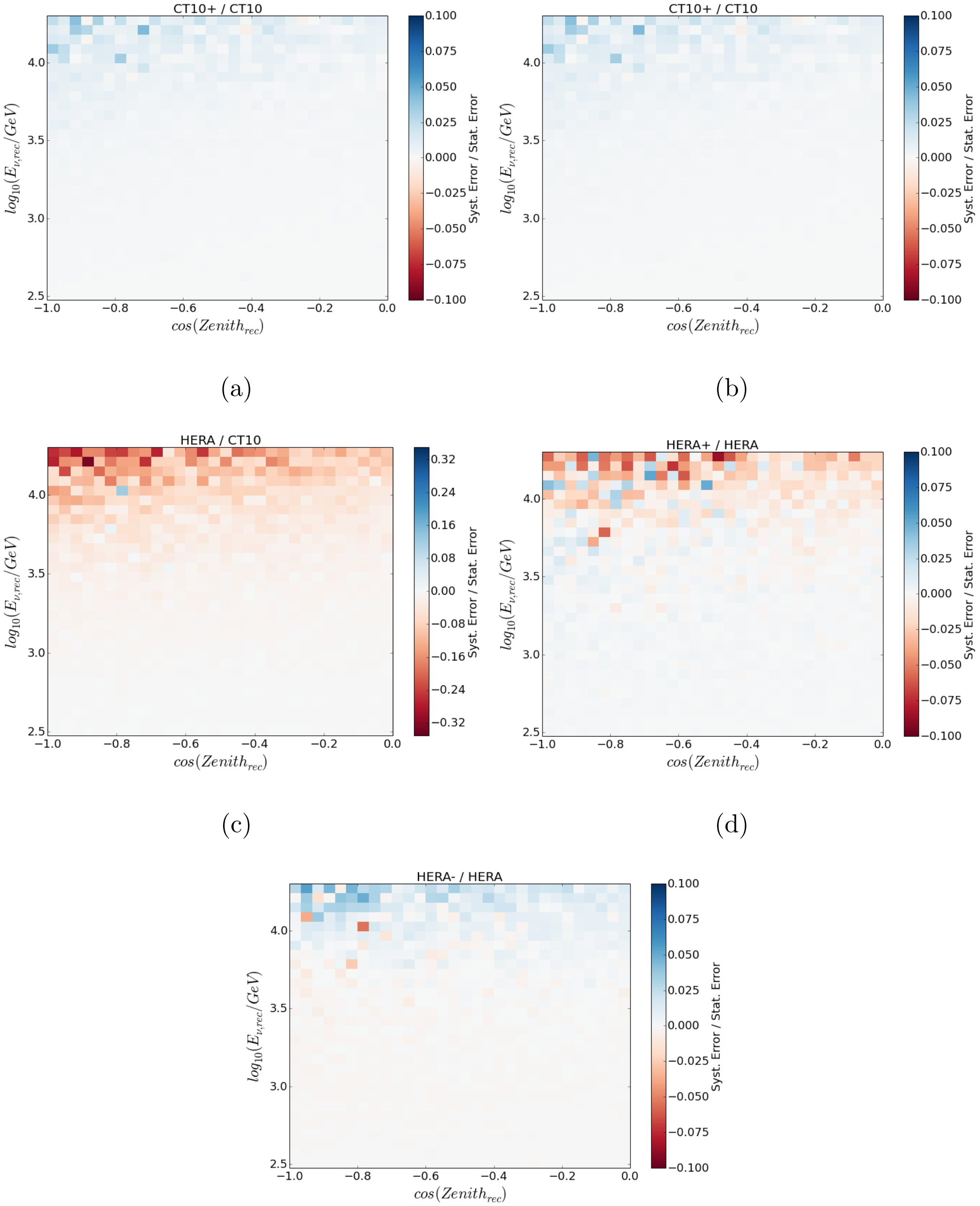
Figure 6.2 Ice model uncertainty compared with statistical uncertainty. This plot compares an ice model with the absorption coefficient increased by %10 compared to SPICEMie base model.

We have looked at two different models, HERAPDF [69] and CT10 [70], at their central values and at ±1σ of their uncertainty. Fig. 6.4 shows that the effect of cross-section uncertainty on the distribution of events is relatively small compared to the statistical uncertainty from one year of data; Hence, we have not included the cross section as a systematic error in this analysis. HERAPDF is used in this analysis for having a smaller uncertainty (see [71] for more information).



-0.0869	-0.0371	-0.1726	-0.0719	-0.0165	-0.0451	0.0495	-0.0177	0.0216	-0.0625
-0.0596	-0.0255	-0.0245	-0.0375	-0.0064	-0.0492	-0.0122	0.0223	-0.0210	-0.0367
-0.0060	-0.0260	-0.0203	-0.0479	-0.0118	-0.0196	-0.0176	-0.0111	0.0027	0.0274
-0.0190	-0.0079	-0.0191	0.0096	-0.0060	-0.0124	-0.0101	0.0070	-0.0019	-0.0011
-0.0088	0.0001	-0.0133	-0.0086	0.0016	-0.0046	-0.0086	-0.0064	-0.0073	0.0004
-0.0068	0.0024	-0.0061	-0.0037	-0.0015	-0.0035	0.0017	-0.0010	-0.0018	0.0060
-0.0039	-0.0022	-0.0033	-0.0023	-0.0019	-0.0003	0.0005	0.0004	-0.0020	-0.0002
-0.0035	-0.0026	-0.0028	-0.0014	-0.0005	0.0005	-0.0013	-0.0006	0.0007	-0.0004
-0.0009	-0.0006	-0.0017	-0.0012	0.0010	0.0008	0.0011	0.0007	0.0010	0.0005
0.0088	0.0165	0.0024	-0.0042	0.0073	-0.0035	0.0082	0.0073	0.0091	0.0071

Figure 6.3 Ice model uncertainty compared with statistical uncertainty. This plot compares an ice model with the scattering coefficient increased by %10 compare to SPICEMie base model.



(a)

(b)

(c)

(d)

(e)

Figure 6.4 Error caused by the cross section compared with the statistical error from one year of data.

Chapter 7

Sensitivity Results

7.1 Likelihood Analysis

The analysis is being done on reconstructed neutrino energies between 3×10^2 and 2×10^4 GeV* for starting tracks and zenith angles of $\pi/2$ and π, and for 30×30 bins.

Data will be fit to the MC by maximizing the likelihood over the nuisance parameters.

$$\max(L) = \max(L(\theta; d)) \tag{7.1}$$

where θ corresponds to continuous nuisance parameters and d is the set of discrete systematic parameters. The likelihood L can be written as

$$L(\theta, d) = \left[\prod_{t}\prod_{i}^{n_t}(\mu_i^t(\hat{\theta}, d); \theta_i^t)\right]\prod_{\eta} d_\eta, \tag{7.2}$$

where the index t runs over the topologies, i iterates the bins in each topology, s over the nuisance parameters, μ is the expected number of events, $\hat{\theta}$ is the number of expected events from Monte Carlo, and θ is the number of observed events.

We use a minimizer to minimize the negative of log likelihood.

$$\min(-LLH) = \min(\sum_{i=0}^{N_{bins}}[x_i \log\lambda_i(\theta, d) - \lambda_i(\theta, d)] + \sum_{\eta}\frac{(\theta_\eta - \hat{\theta}_\eta)^2}{2\sigma_\eta^2} + const.; d) \tag{7.3}$$

and the test statistics is defined as:

$$TS = -2\log\frac{L_{N_{null}}}{L_{N_{alt}}} = -2\Delta LLH \tag{7.4}$$

Subsequently, we calculate a 2D histogram of cos(zenith angle) and the log of neutrino energy. For the plots that include both starting tracks (which I'm not using in this analysis but I have calculated and can be added late on). I fit two 2D histograms (one for starting and one of throughgoing) at the same time, using the minimizer in Eqn. 7.3. No burnsample data is used to calculate the sensitivity plots but instead, we are using data challenge which is bootstrapping from Monte Carlo with an expected weight of null hypothesis (to test sensitivity to the null hypothesis). Here, I use 30×30 bins, and as long as I have enough Monte Carlo events in each bin (which is the case here), the size of the bins should not matter, but smaller bins are the better in general. Although smaller bin size (than 30×30) did not give me a better sensitivity but made the code slower. We are also assuming a $3+1$ model of sterile neutrinos (i.e. normal hierarchy).

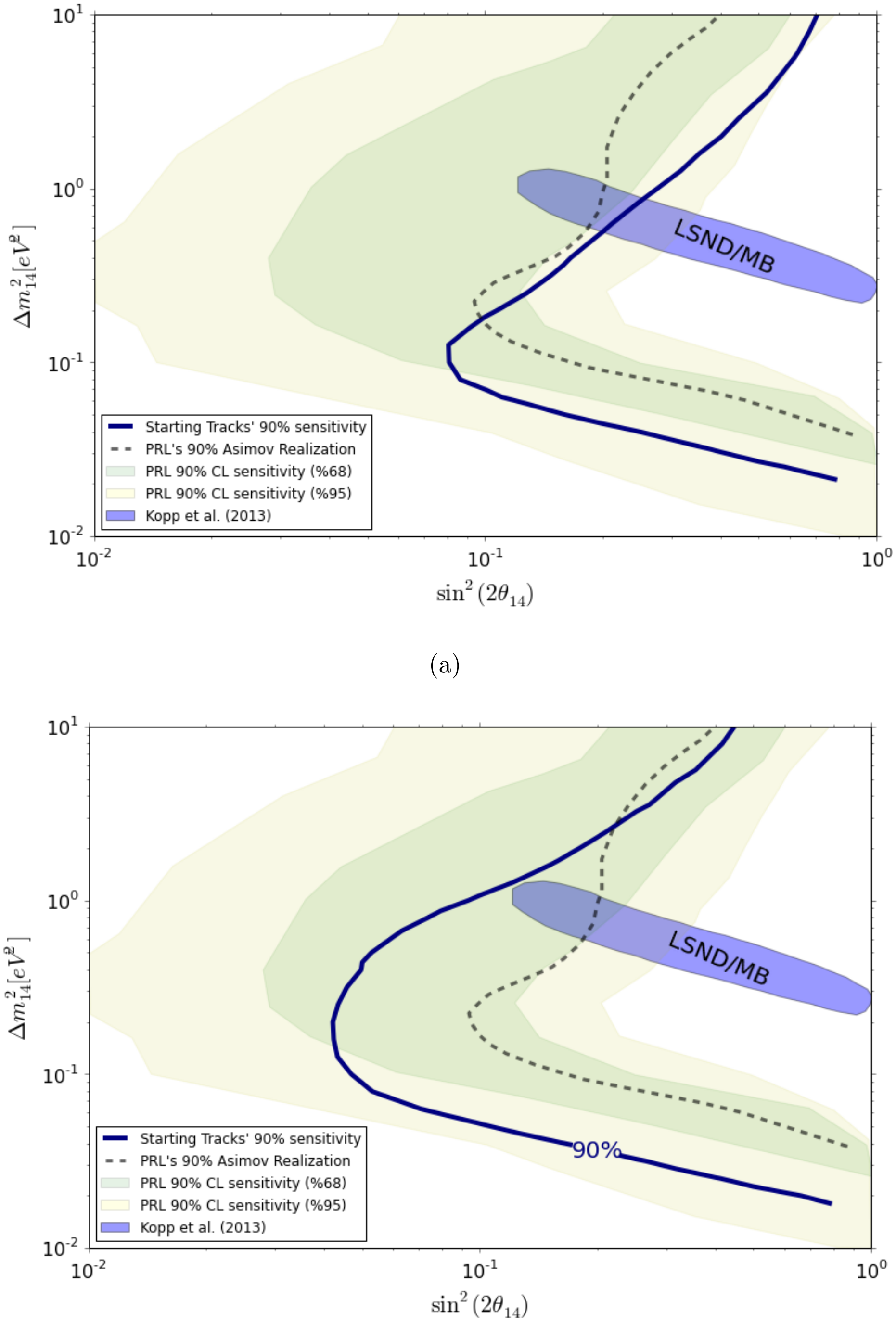
7.2 Chi-Squared

Although it may not be reasonable to just assume that the test statistics behave like a χ^2 distribution, it is a good place to start (see Section 7.3). In this case, we have two degrees of freedom which are the mass difference and mixing angle of the sterile neutrino.

As it is shown in Figures 7.1 and 7.2, starting-tracks gain sensitivity at small Δm^2. But we gain sensitivity at a larger Δm^2 due to the increase in statistics compare to the previous analysis. However, in this thesis, we are going to focus on starting-tracks only and throughgoing tracks are not used in this final sterile analysis.

7.3 Bootstrapping

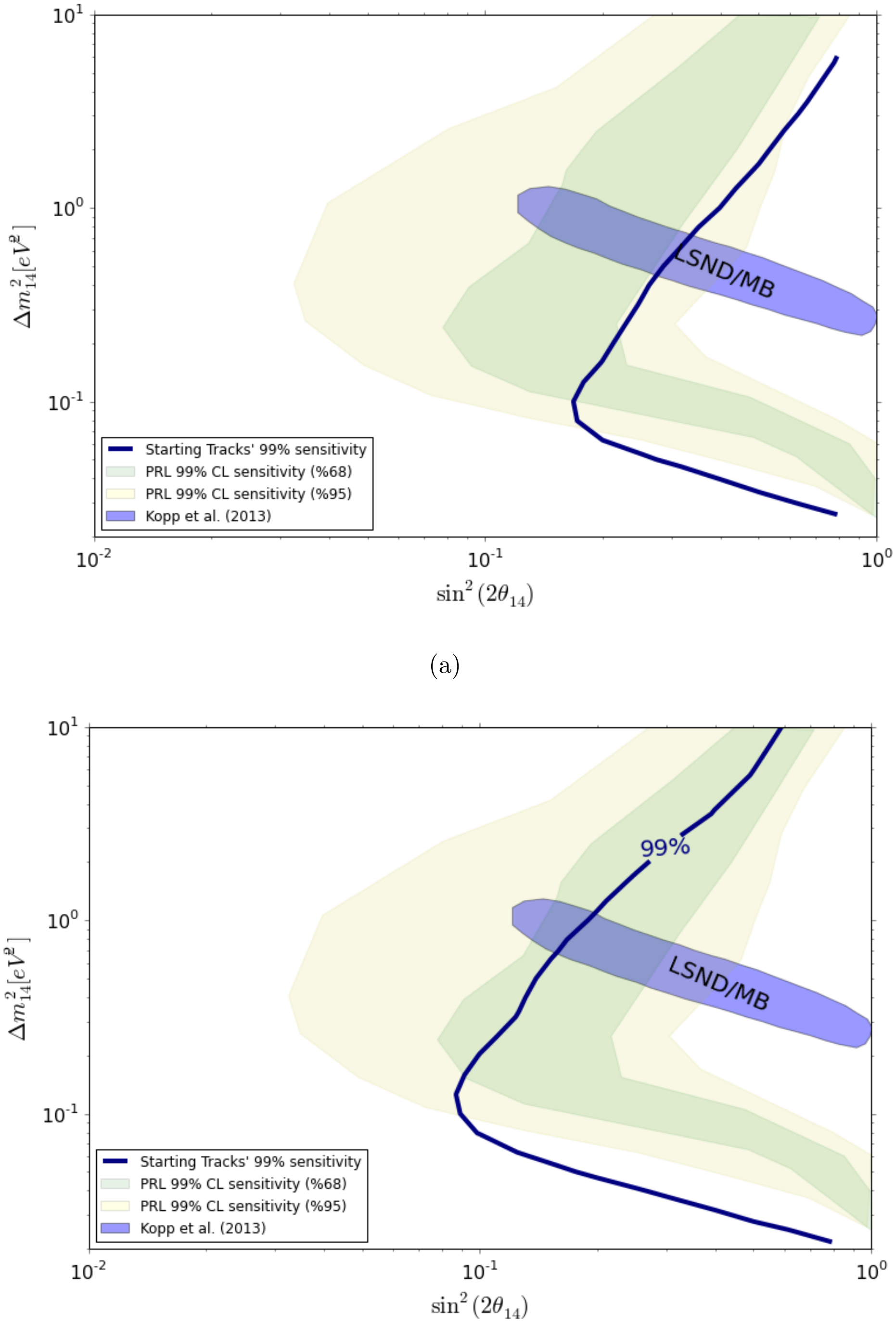
As mentioned above, our model does not have to behave like χ^2 distribution. The most accurate way to determine the confidence level from test statistics is to run trials for each point in the space. This method is computationally expensive, that is why we did most of



(a)

(b)

Figure 7.1 The plots show the difference between using (b) only starting-tracks and using (b) both starting and throughgoing tracks at 90% confidence level. Solid blue line is the sensitivity of this analysis, and dotted blue line represents the sensitivity of the previous analysis. Both calculated using Asimov realization. The green-yellow band is the Brazilian sensitivity of the previous analysis. Brazilian plots are explained in Section 7.4.



(a)

(b)

Figure 7.2 The plots show the difference between using (a) only starting-tracks and (b) using both starting and throughgoing tracks at 99% confidence level. Solid blue line is the sensitivity of this analysis. Both calculated using Asimov realization. The green-yellow band is the Brazilian sensitivity of the previous analysis.Brazilian plots are explained in Section 7.4.

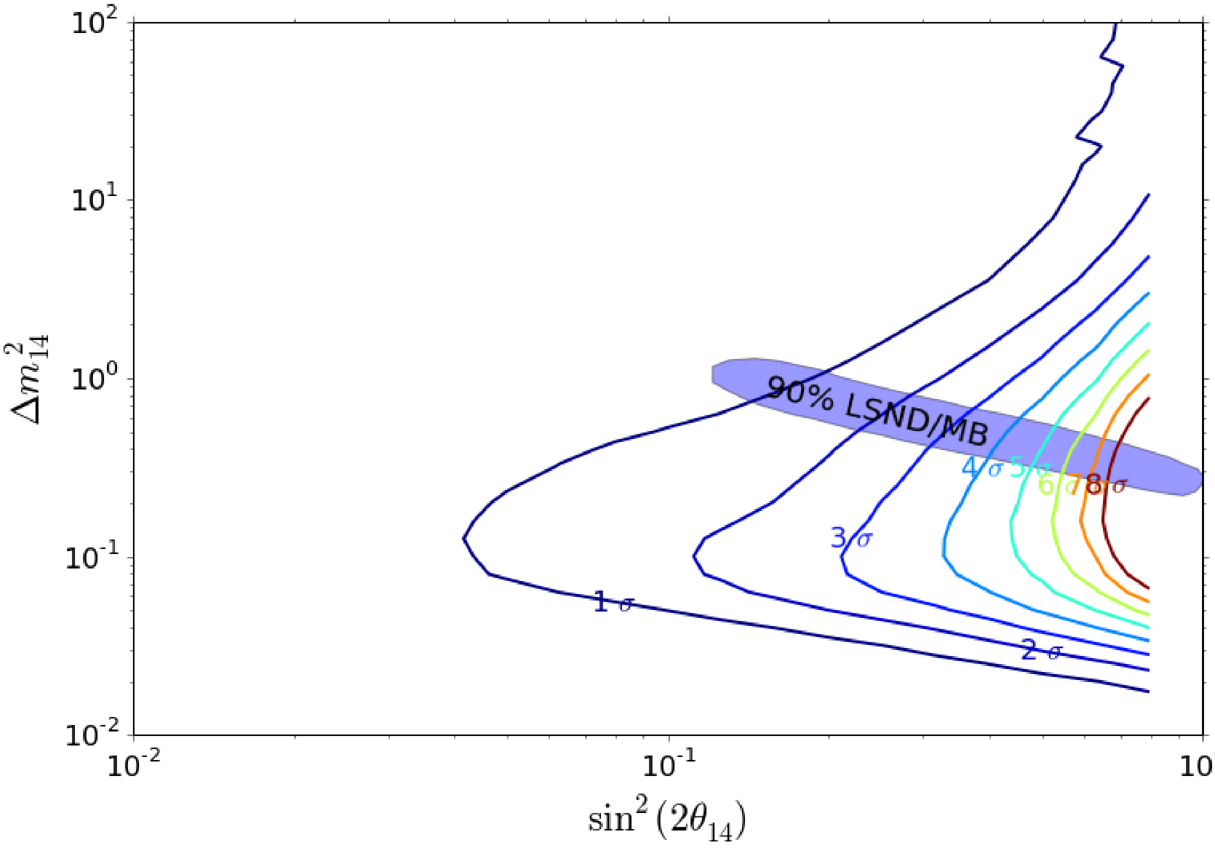
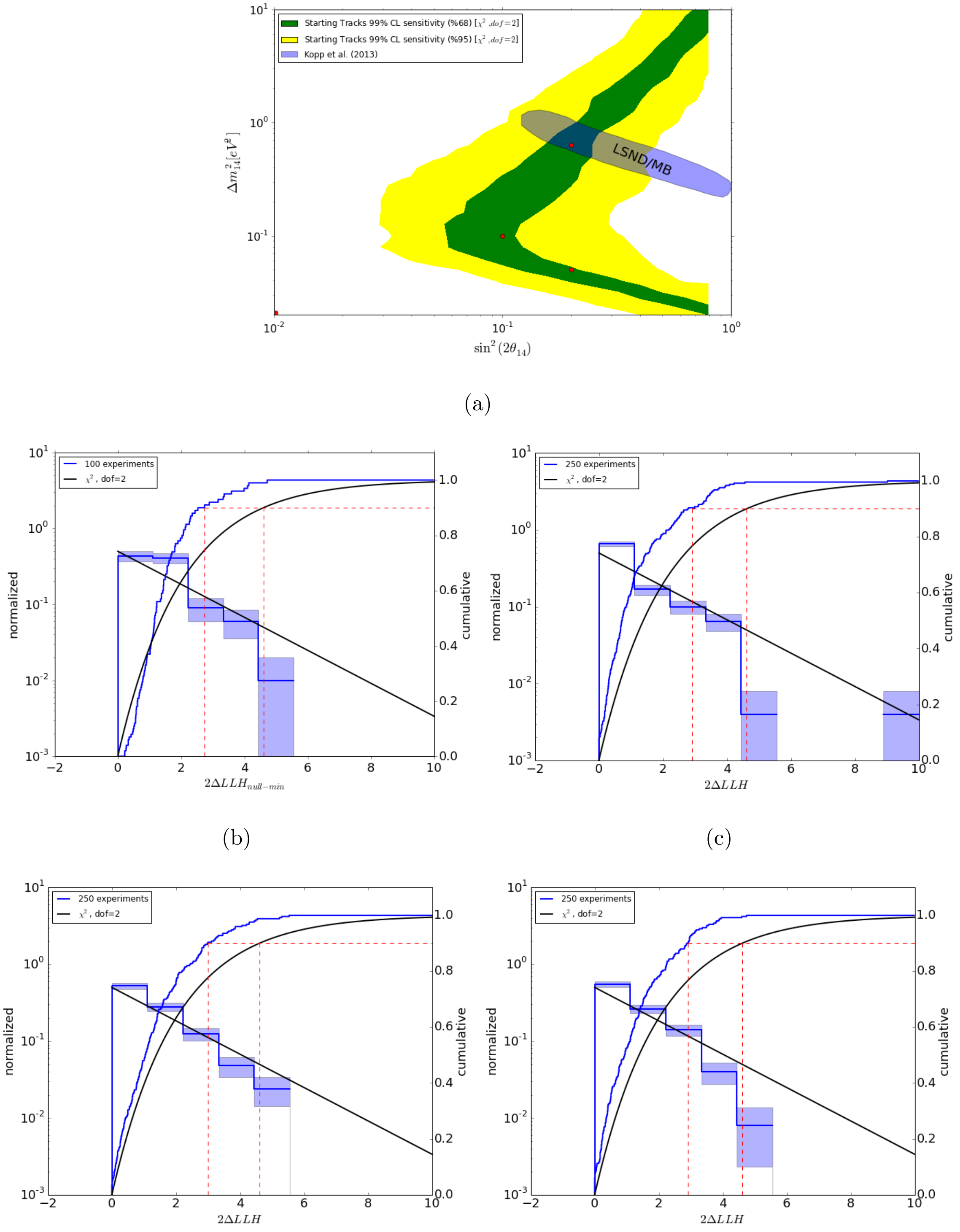


Figure 7.3 In the plot, one can see that, as we approach higher mixing angles, we are more sensitive in ruling out the null hypothesis.

our tests using χ^2 distribution which can give us a good idea of the sensitivity. Figure 7.4 shows some of those distributions, and as it is shown in those distributions this analysis is more sensitive to sterile neutrinos than previously thought under χ^2 distribution.

7.4 Starting-Track Sensitivity

Sensitivity plots are often shown in Brazilian plots (referred to the color of the Brazilian flag). The green and yellow bands in the plots are the uncertainty of the sensitivity, calculated from trials. Usually, green band represents the 1σ (68%) confidence level and yellow band, the 2σ (95%).



(a)

(b)

(c)

(d)

(e)

Figure 7.4 The plots show the distribution of test statistics in multiple experiments for four points in the sterile neutrino space shown as red dots in plot (a).

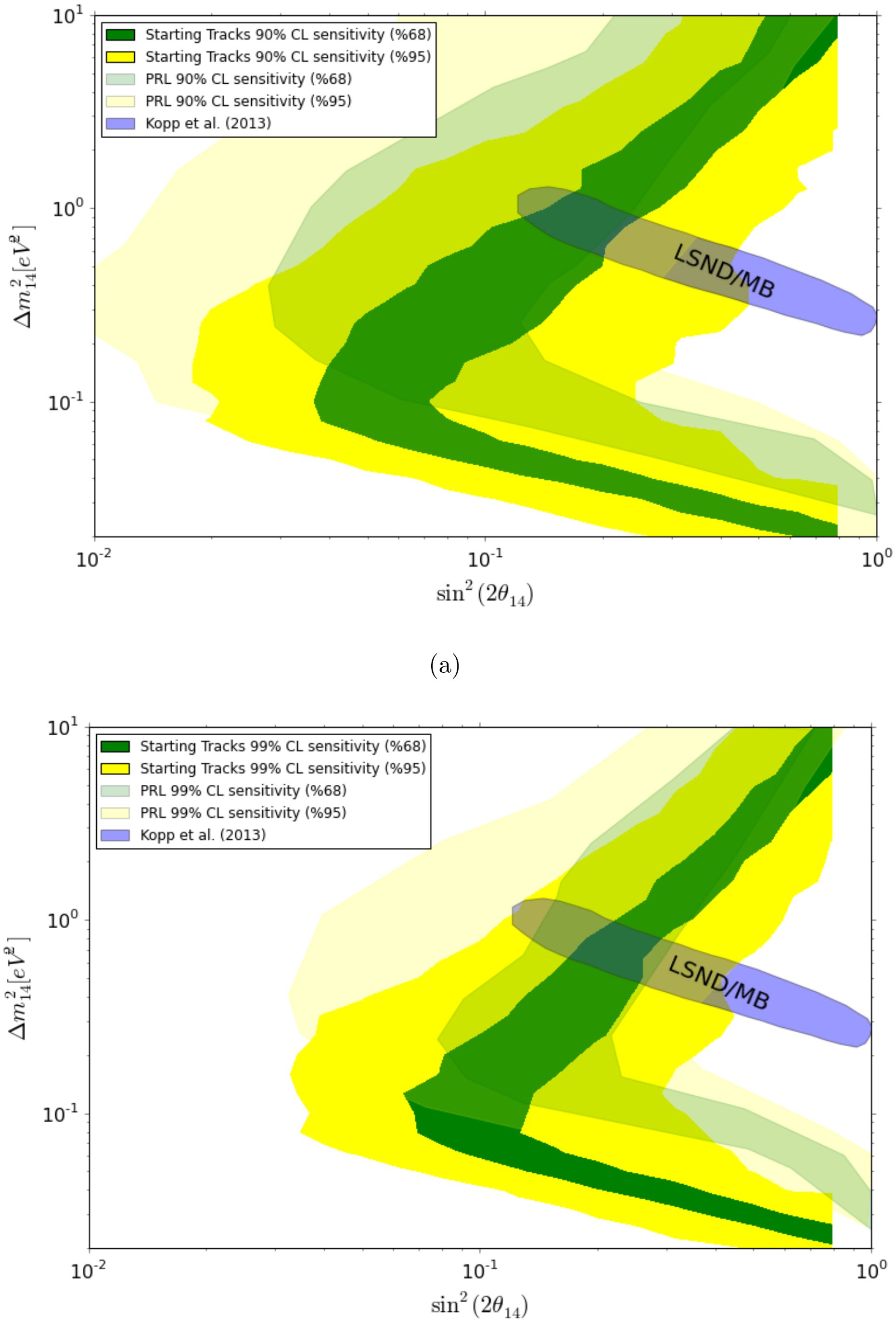
Considering the distributions form the trials in Section 7.3, the new Brazilian plots for 90% and 99% sensitivity will now can be calculated more accurately. Figure 7.5 shows the sensitivity of this analysis in a Brazilian plot and compare it to the previous analysis. It is important to know that this data selection has about twice as many throughgoing events as the previous analysis, therefore, combining those data can give us even a better sensitivity to eV sterile neutrinos.

However, it is clear that we are more sensitive to smaller Δm^2 by using only starting-tracks.

As an example, for sterile neutrinos with $\Delta_{14}^2 = 0.630957$ and $\theta_{14} = 0.393293$ (near the LSND/MiniBooNE anomaly), sterile neutrinos can be detected with a good significance as shown in Figure 7.6.

The main reason for this improvement is the increased energy resolution of the neutrinos. Although it is not entirely correct to directly compare resolution in the two analyses, because we are looking at different energies (neutrinos in this analysis and muons in the previous one), the resolution in this thesis is about 20% for neutrinos and the best muon energy resolution one can get for IceCube in these energies is more than 40%. This number gets worse (especially at higher energies) since this is the resolution of true muon energies not observed energy.

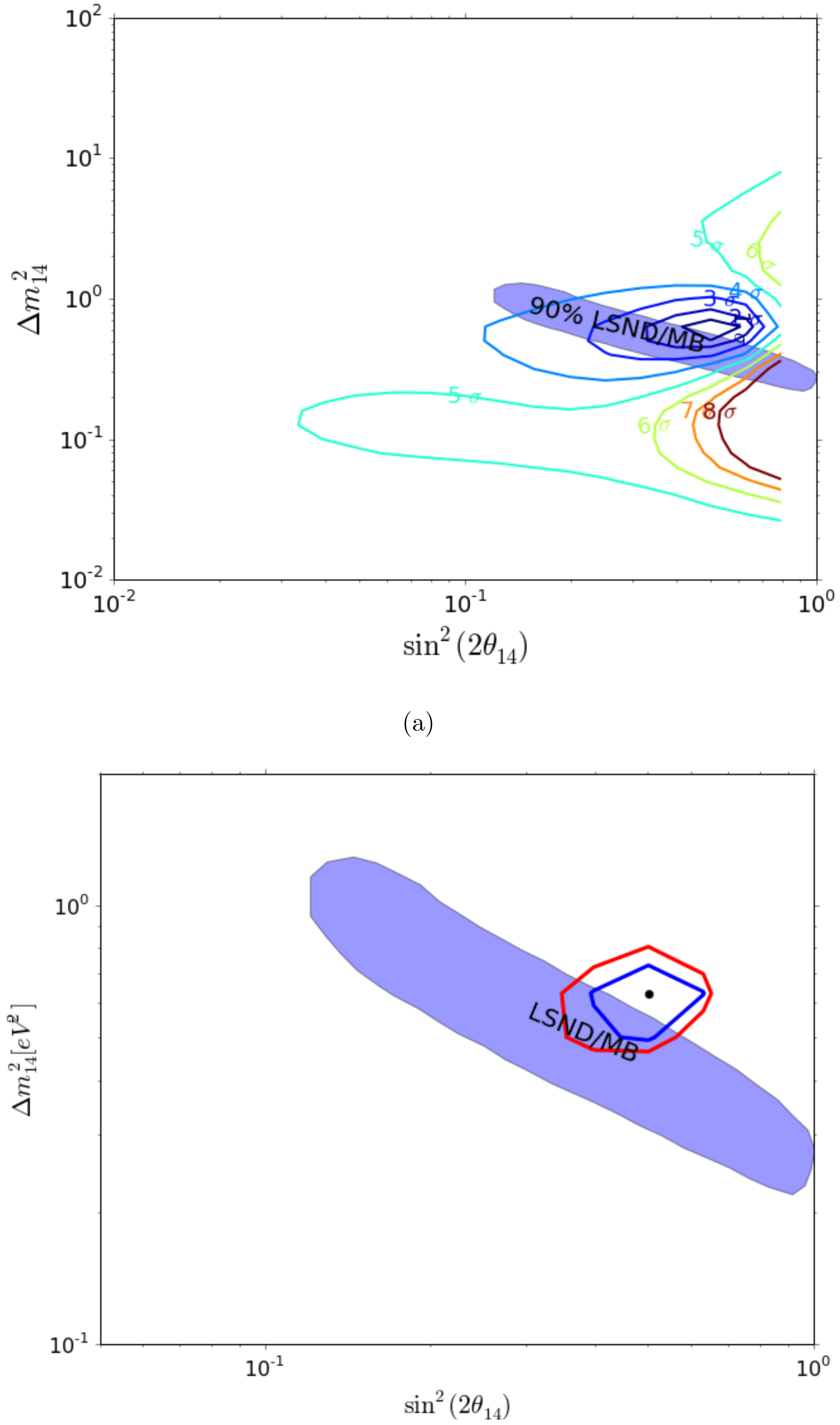
We have also a very good angular resolution for tracks in IceCube (less than 0.5 degree) which is much larger than our bin size and any improvement should not play any major role in the sensitivity. I am using SPEFit8 (SPEFit with 8 iteration) for this analysis and SPEFit2 (SPEFit with 2 iteration) was used previously, but the improvement is marginal at the final level. However, the resolution may play a larger role in the event selection process because we should have less mis-reconstructed events.



(a)

(b)

Figure 7.5 The plots show the final sensitivity of this analysis to the sterile neutrinos using starting-tracks only.



(a)

(b)

Figure 7.6 Sensitivity to a sterile neutrino with $\Delta_{14}^2 = 0.630957$ and $\theta_{14} = 0.393293$. Plots (a) show sensitivity with sigma significance and plot (b) shows 90% (blue) and 99% (red) confidence level. The black dot as the hypothetical sterile neutrino and the best fit (they are on top of each other).

Chapter 8

Conclusion

The sensitivity of the analysis in this thesis for one year of IceCube data is currently better than the one-year result that was published by IceCube for sterile neutrinos with smaller Δm^2 (see Figure 7.5) [1]. Previous analyses have been limited by background contamination of cosmic rays in their data selection at lower energies, which limits their sensitivity for smaller Δm^2 sterile neutrinos. Consequently, their sensitivity to smaller Δm^2 with seven years of data at those regions is predicted to be comparable to this analysis's sensitivity with one year of data.

In addition, as discussed in section 5.8, we have more throughgoing track events compared to the previous analyses. These events can be added to our starting-track event dataset to gain better sensitivity. We have not used those events for this thesis. Unlike starting-track events, we cannot reconstruct the neutrino energies for the throughgoing track events, and we should use reconstructed muon energy. Therefore, we do not have the advantage of a better energy resolution that starting-track events do. However, those extra events can give us more sensitivity at higher Δm^2. It is important to note that these events (starting and throughgoing track events) are fitted simultaneously, but in two different histograms, and measuring a global fit result will not produce the best combined sensitivity. Figures 7.1 and 7.2 show that, compared to the previous analysis, this can give us a better sensitivity at high Δm^2 as well.

In the end, this analysis can and should be followed up by measuring the result for the additional years of data and by fitting those simulated data to the experimental data from the detector, so we can get a superior measurement for sterile neutrinos.

LIST OF REFERENCES

[1] M. G. Aartsen et al. Searches for Sterile Neutrinos with the IceCube Detector. *Phys. Rev. Lett.*, 117(7):071801, 2016. `arXiv:1605.01990`, `doi:10.1103/PhysRevLett.117.071801`.

[2] MG Aartsen, M Ackermann, J Adams, JA Aguilar, M Ahlers, M Ahrens, D Altmann, K Andeen, T Anderson, I Ansseau, et al. The icecube neutrino observatory: Instrumentation and online systems. *arXiv preprint arXiv:1612.05093*, 2016.

[3] et. al R. Abbasi. The design and performance of icecube deepcore. *Astroparticle Physics*, 35(10):615 – 624, 2012. URL: `http://www.sciencedirect.com/science/article/pii/S0927650512000254`, `doi:https://doi.org/10.1016/j.astropartphys.2012.01.004`.

[4] R. Abbasi et. al. Icetop: The surface component of icecube. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 700:188 – 220, 2013. URL: `http://www.sciencedirect.com/science/article/pii/S016890021201217X`, `doi:https://doi.org/10.1016/j.nima.2012.10.067`.

[5] John David Jackson. *Classical electrodynamics*. Wiley, New York, NY, 3rd ed. edition, 1999. URL: `http://cdsweb.cern.ch/record/490457`.

[6] et. al J. Ahrens. Muon track reconstruction and data selection techniques in amanda. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 524(1):169 – 194, 2004. URL: `http://www.sciencedirect.com/science/article/pii/S0168900204001871`, `doi:https://doi.org/10.1016/j.nima.2004.01.065`.

[7] V. J. Stenger. Track fitting for dumand-ii octagon array. *External Report HDC-1-90,. University of Hawai'i at Manoa, Manoa, Hawaii, USA*, 1990.

[8] et. al M.G. Aartsen. Improvement in fast particle track reconstruction with robust statistics. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 736:143 – 149, 2014. URL: `http://www.sciencedirect.com/science/article/pii/S0168900213014654`, `doi:https://doi.org/10.1016/j.nima.2013.10.074`.

[9] L. Vandenberghe S. Boyd. Convex optimization. *Cambridge University Press*, page 298300, 2009.

[10] D. Pandel. *Bestimmung von Wasser- und Detektorparametern und Rekonstruktion von Myonen bis 100 TeV mit dem Baikal-Neutrinoteleskop NT-72.* PhD thesis, 1996.

[11] C. Patrignani et al. Review of Particle Physics. *Chin. Phys.*, C40(10):100001, 2016. `doi:10.1088/1674-1137/40/10/100001`.

[12] T. K. Gaisser. *Cosmic rays and particle physics.* 1990. URL: `http://www.cambridge.org/uk/catalogue/catalogue.asp?isbn=0521326672`.

[13] M. G. et. al. Aartsen. Measurement of the cosmic ray energy spectrum with icetop-73. *Phys. Rev. D*, 88:042004, Aug 2013. URL: `https://link.aps.org/doi/10.1103/PhysRevD.88.042004`, `doi:10.1103/PhysRevD.88.042004`.

[14] Rasha Abbasi, Yasser Abdou, T Abu-Zayyad, J Adams, JA Aguilar, M Ahlers, K Andeen, J Auffenberg, X Bai, M Baker, et al. Measurement of the atmospheric neutrino energy spectrum from 100 gev to 400 tev with icecube. *Physical Review D*, 83(1):012001, 2011.

[15] Jörg R Hörandel and Jonathan F Ormes. The composition of cosmic rays at the knee. In *AIP Conference Proceedings*, volume 1516, pages 185–194. AIP, 2013.

[16] Morihiro Honda, T Kajita, K Kasahara, S Midorikawa, and T Sanuki. Calculation of atmospheric neutrino flux using the interaction model calibrated with atmospheric muon data. *Physical Review D*, 75(4):043006, 2007.

[17] Rikard Enberg, Mary Hall Reno, and Ina Sarcevic. Prompt neutrino fluxes from atmospheric charm. *Physical Review D*, 78(4):043005, 2008.

[18] Aartsen et. al. Multimessenger observations of a flaring blazar coincident with high-energy neutrino icecube-170922a. 361(6398), 2018. `doi:10.1126/science.aat1378`.

[19] B. Pontecorvo. Mesonium and anti-mesonium. *Sov. Phys. JETP*, 6:429, 1957. [Zh. Eksp. Teor. Fiz.33,549(1957)].

[20] B. Pontecorvo. Neutrino Experiments and the Problem of Conservation of Leptonic Charge. *Sov. Phys. JETP*, 26:984–988, 1968. [Zh. Eksp. Teor. Fiz.53,1717(1967)].

[21] B. T. Cleveland, Timothy Daily, Raymond Davis, Jr., James R. Distel, Kenneth Lande, C. K. Lee, Paul S. Wildenhain, and Jack Ullman. Measurement of the solar electron neutrino flux with the Homestake chlorine detector. *Astrophys. J.*, 496:505–526, 1998. `doi:10.1086/305343`.

[22] Y. Fukuda et al. Solar neutrino data covering solar cycle 22. *Phys. Rev. Lett.*, 77:1683–1686, 1996. `doi:10.1103/PhysRevLett.77.1683`.

[23] J. N. Abdurashitov et al. Measurement of the solar neutrino capture rate with gallium metal. III: Results for the 2002–2007 data-taking period. *Phys. Rev.*, C80:015807, 2009. `arXiv:0901.2200`, `doi:10.1103/PhysRevC.80.015807`.

[24] P. Anselmann et al. Solar neutrinos observed by GALLEX at Gran Sasso. *Phys. Lett.*, B285:376–389, 1992. `doi:10.1016/0370-2693(92)91521-A`.

[25] W. Hampel et al. GALLEX solar neutrino observations: Results for GALLEX IV. *Phys. Lett.*, B447:127–133, 1999. `doi:10.1016/S0370-2693(98)01579-2`.

[26] M. Altmann et al. Complete results for five years of GNO solar neutrino observations. *Phys. Lett.*, B616:174–190, 2005. `arXiv:hep-ex/0504037`, `doi:10.1016/j.physletb.2005.04.068`.

[27] S. Fukuda et al. Determination of solar neutrino oscillation parameters using 1496 days of Super-Kamiokande I data. *Phys. Lett.*, B539:179–187, 2002. `arXiv:hep-ex/0205075`, `doi:10.1016/S0370-2693(02)02090-7`.

[28] Q. R. Ahmad et al. Measurement of the rate of $\nu_e + d \to p + p + e^-$ interactions produced by ^8B solar neutrinos at the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 87:071301, 2001. `arXiv:nucl-ex/0106015`, `doi:10.1103/PhysRevLett.87.071301`.

[29] Q. R. Ahmad et al. Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 89:011301, 2002. `arXiv:nucl-ex/0204008`, `doi:10.1103/PhysRevLett.89.011301`.

[30] K. Eguchi et al. First results from KamLAND: Evidence for reactor anti-neutrino disappearance. *Phys. Rev. Lett.*, 90:021802, 2003. `arXiv:hep-ex/0212021`, `doi:10.1103/PhysRevLett.90.021802`.

[31] T. Araki et al. Measurement of neutrino oscillation with KamLAND: Evidence of spectral distortion. *Phys. Rev. Lett.*, 94:081801, 2005. `arXiv:hep-ex/0406035`, `doi:10.1103/PhysRevLett.94.081801`.

[32] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81:1562–1567, 1998. `arXiv:hep-ex/9807003`, `doi:10.1103/PhysRevLett.81.1562`.

[33] Y. Ashie et al. Evidence for an oscillatory signature in atmospheric neutrino oscillation. *Phys. Rev. Lett.*, 93:101801, 2004. `arXiv:hep-ex/0404034`, `doi:10.1103/PhysRevLett.93.101801`.

[34] Ziro Maki, Masami Nakagawa, and Shoichi Sakata. Remarks on the unified model of elementary particles. *Prog. Theor. Phys.*, 28:870–880, 1962. `doi:10.1143/PTP.28.870`.

[35] Samoil M. Bilenky, J. Hosek, and S. T. Petcov. On Oscillations of Neutrinos with Dirac and Majorana Masses. *Phys. Lett.*, B94:495–498, 1980. `doi:10.1016/0370-2693(80)90927-2`.

[36] M. Doi, T. Kotani, H. Nishiura, K. Okuda, and E. Takasugi. CP Violation in Majorana Neutrinos. *Phys. Lett.*, B102:323–326, 1981. `doi:10.1016/0370-2693(81)90627-4`.

[37] C. Athanassopoulos et al. Candidate events in a search for anti-muon-neutrino — ¿ anti-electron-neutrino oscillations. *Phys. Rev. Lett.*, 75:2650–2653, 1995. `arXiv:nucl-ex/9504002`, `doi:10.1103/PhysRevLett.75.2650`.

[38] A. Aguilar-Arevalo et al. Evidence for neutrino oscillations from the observation of anti-neutrino(electron) appearance in a anti-neutrino(muon) beam. *Phys. Rev.*, D64:112007, 2001. `arXiv:hep-ex/0104049`, `doi:10.1103/PhysRevD.64.112007`.

[39] J. N. Abdurashitov et al. Measurement of the response of a Ga solar neutrino experiment to neutrinos from an Ar-37 source. *Phys. Rev.*, C73:045805, 2006. `arXiv:nucl-ex/0512041`, `doi:10.1103/PhysRevC.73.045805`.

[40] Marco Laveder. Unbound neutrino roadmaps. *Nucl. Phys. Proc. Suppl.*, 168:344–346, 2007. `doi:10.1016/j.nuclphysbps.2007.02.037`.

[41] Carlo Giunti and Marco Laveder. Short-Baseline Active-Sterile Neutrino Oscillations? *Mod. Phys. Lett.*, A22:2499–2509, 2007. `arXiv:hep-ph/0610352`, `doi:10.1142/S0217732307025455`.

[42] Carlo Giunti and Marco Laveder. Statistical Significance of the Gallium Anomaly. *Phys. Rev.*, C83:065504, 2011. `arXiv:1006.3244`, `doi:10.1103/PhysRevC.83.065504`.

[43] C. Giunti, M. Laveder, Y. F. Li, Q. Y. Liu, and H. W. Long. Update of Short-Baseline Electron Neutrino and Antineutrino Disappearance. *Phys. Rev.*, D86:113014, 2012. `arXiv:1210.5715`, `doi:10.1103/PhysRevD.86.113014`.

[44] F. Kaether, W. Hampel, G. Heusser, J. Kiko, and T. Kirsten. Reanalysis of the GALLEX solar neutrino flux and source experiments. *Phys. Lett.*, B685:47–54, 2010. `arXiv:1001.2731`, `doi:10.1016/j.physletb.2010.01.030`.

[45] G. Mention, M. Fechner, Th. Lasserre, Th. A. Mueller, D. Lhuillier, M. Cribier, and A. Letourneau. The Reactor Antineutrino Anomaly. *Phys. Rev.*, D83:073006, 2011. `arXiv:1101.2755`, `doi:10.1103/PhysRevD.83.073006`.

[46] Th. A. Mueller et al. Improved Predictions of Reactor Antineutrino Spectra. *Phys. Rev.*, C83:054615, 2011. arXiv:1101.2663, doi:10.1103/PhysRevC.83.054615.

[47] Patrick Huber. On the determination of anti-neutrino spectra from nuclear reactors. *Phys. Rev.*, C84:024617, 2011. [Erratum: Phys. Rev.C85,029901(2012)]. arXiv:1106.0687, doi:10.1103/PhysRevC.85.029901,10.1103/PhysRevC.84.024617.

[48] Precision electroweak measurements on the z resonance. *Physics Reports*, 427(5):257 – 454, 2006. URL: http://www.sciencedirect.com/science/article/pii/S0370157305005119, doi:https://doi.org/10.1016/j.physrep.2005.12.006.

[49] et. al. Armbruster. *Phys. Rev. D*, 65:112001, Jun 2002. URL: https://link.aps.org/doi/10.1103/PhysRevD.65.112001, doi:10.1103/PhysRevD.65.112001.

[50] K. et. al. Abe. Limits on sterile neutrino mixing using atmospheric neutrinos in super-kamiokande. *Phys. Rev. D*, 91:052019, Mar 2015. URL: https://link.aps.org/doi/10.1103/PhysRevD.91.052019, doi:10.1103/PhysRevD.91.052019.

[51] et. al. Adamson. Active to sterile neutrino mixing limits from neutral-current interactions in minos. *Phys. Rev. Lett.*, 107:011802, Jun 2011. URL: https://link.aps.org/doi/10.1103/PhysRevLett.107.011802, doi:10.1103/PhysRevLett.107.011802.

[52] G. et. al. Cheng. *Phys. Rev. D*, 86:052009, Sep 2012. URL: https://link.aps.org/doi/10.1103/PhysRevD.86.052009, doi:10.1103/PhysRevD.86.052009.

[53] F. Dydak, G.J. Feldman, C. Guyot, J.P. Merlo, H.-J. Meyer, J. Rothberg, J. Steinberger, H. Taureg, W. von Rden, H. Wachsmuth, H. Wahl, J. Wotschack, H. Blmer, P. Buchholz, J. Duda, F. Eisele, K. Kleinknecht, J. Knobloch, B. Pszola, B. Renk, R. Belusevic, B. Falkenburg, T. Flottmann, J.G.H. de Groot, C. Geweniger, H. Keilwerth, K. Tittel, P. Debu, A. Para, P. Perez, B. Peyaud, J. Rander, J.P. Schuller, R. Turlay, H. Abramowicz, and J. Krlikowski. A search for oscillations in the m2 range 0.390 ev2. *Physics Letters B*, 134(3):281 – 286, 1984. URL: http://www.sciencedirect.com/science/article/pii/0370269384906889, doi:https://doi.org/10.1016/0370-2693(84)90688-9.

[54] Carlo Giunti and Marco Laveder. 3 + 1 and 3 + 2 sterile neutrino fits. *Phys. Rev. D*, 84:073008, Oct 2011. URL: https://link.aps.org/doi/10.1103/PhysRevD.84.073008, doi:10.1103/PhysRevD.84.073008.

[55] Joachim Kopp, Pedro A. N. Machado, Michele Maltoni, and Thomas Schwetz. Sterile neutrino oscillations: the global picture. *Journal of High Energy Physics*, 2013(5):50, May 2013. URL: https://doi.org/10.1007/JHEP05(2013)050, doi:10.1007/JHEP05(2013)050.

[56] P. A. R. Ade et al. Planck 2015 results. XIII. Cosmological parameters. *Astron. Astrophys.*, 594:A13, 2016. arXiv:1502.01589, doi:10.1051/0004-6361/201525830.

[57] L. Wolfenstein. Neutrino Oscillations in Matter. *Phys. Rev.*, D17:2369–2374, 1978. `doi:10.1103/PhysRevD.17.2369`.

[58] S. P. Mikheev and A. Yu. Smirnov. Resonance Amplification of Oscillations in Matter and Spectroscopy of Solar Neutrinos. *Sov. J. Nucl. Phys.*, 42:913–917, 1985. [Yad. Fiz.42,1441(1985)].

[59] S. P. Mikheev and A. Yu. Smirnov. Resonant amplification of neutrino oscillations in matter and solar neutrino spectroscopy. *Nuovo Cim.*, C9:17–26, 1986. `doi:10.1007/BF02508049`.

[60] Carlo Giunti and Marco Laveder. Neutrino mixing. 2003. `arXiv:hep-ph/0310238`.

[61] Paul Langacker, Jacques P. Leveille, and Jon Sheiman. On the detection of cosmological neutrinos by coherent scattering. *Phys. Rev. D*, 27:1228–1242, Mar 1983. URL: `https://link.aps.org/doi/10.1103/PhysRevD.27.1228`, `doi:10.1103/PhysRevD.27.1228`.

[62] M. G. Aartsen et al. Evidence for Astrophysical Muon Neutrinos from the Northern Sky with IceCube. *Phys. Rev. Lett.*, 115(8):081102, 2015. `arXiv:1507.04005`, `doi:10.1103/PhysRevLett.115.081102`.

[63] Jakob van Santen. *Neutrino Interactions in IceCube above 1 TeV: Constraints on Atmospheric Charmed-Meson Production and Investigation of the Astrophysical Neutrino Flux with 2 Years of IceCube Data taken 2010–2012.* PhD thesis, U. Wisconsin, Madison (main), 2014-11-18. URL: `http://search.proquest.com/docview/1637727437/abstract/CE0FBEF172B24F9APQ/1`.

[64] Christopher Weaver. *Evidence for Astrophysical Muon Neutrinos from the Northern Sky.* PhD thesis.

[65] Paul H Barrett, Lowell M Bollinger, Giuseppe Cocconi, Yehuda Eisenberg, and Kenneth Greisen. Interpretation of cosmic-ray measurements far underground. *Reviews of Modern Physics*, 24(3):133, 1952.

[66] D. Chirkin. MuEx. URL: `http://code.icecube.wisc.edu/svn/projects/mue/trunk`.

[67] MG Aartsen, R Abbasi, M Ackermann, J Adams, JA Aguilar, M Ahlers, D Altmann, C Arguelles, J Auffenberg, X Bai, et al. Energy reconstruction methods in the icecube neutrino telescope. *Journal of Instrumentation*, 9(03):P03009, 2014.

[68] Yoav Freund and Robert E Schapire. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *Journal of Computer and System Sciences*, 55(1):119 – 139, 1997. URL: `http://www.sciencedirect.com/science/article/pii/S002200009791504X`, `doi:http://dx.doi.org/10.1006/jcss.1997.1504`.

[69] F. D. Aaron et al. Combined Measurement and QCD Analysis of the Inclusive e+-p Scattering Cross Sections at HERA. *JHEP*, 01:109, 2010. `arXiv:0911.0884`, `doi:10.1007/JHEP01(2010)109`.

[70] Jun Gao, Marco Guzzi, Joey Huston, Hung-Liang Lai, Zhao Li, Pavel Nadolsky, Jon Pumplin, Daniel Stump, and C. P. Yuan. CT10 next-to-next-to-leading order global analysis of QCD. *Phys. Rev.*, D89(3):033009, 2014. `arXiv:1302.6246`, `doi:10.1103/PhysRevD.89.033009`.

[71] Carlos Alberto Arguelles Delgado. *New Physics with Atmospheric Neutrinos*. PhD thesis.

[72] Sebastian Raschka. *Python Machine Learning*. Packt Publishing, Birmingham, UK, 2015.

APPENDIX
Machine Learning

Machine Learning techniques allow computers to learn from data and to develop predictive algorithms. These methods are especially useful when data is complex and large. Statistical modeling is very challenging when we have very high dimensional data (i.e. with many input variables) or when the size of the data sample is very large. Traditional statistical models are much simpler but do not include all relations between features and target output of the data, compared to machine learning models.

A.1 Supervised and Unsupervised Learning

Machine learning tasks can be put into the two different categories of *supervised* and *unsupervised* learning. In supervised learning, the target output is known from training data sets. In IceCube data, we often use simulation data as a training dataset. Unlike experimental data, the properties of simulation data are known and can be used to train the machine to learn the relation from input features and output targets, which we can later use on experimental data to predict outputs that we do not know. For example, taus are very difficult to distinguish from other particles (muons and electrons) in experimental data, but there are separate simulation datasets for these particles. We can give these data to the machine to learn the difference between their properties over billions of events, so it can predict their type in the experimental data.

In contrast, in *unsupervised* learning the output is not known. In this case, the machine will learn the properties of data in order to classify or categorize them. This method is often used to learn about data that may not be well understood, or when supervised learning is not possible. For example, if there were no simulation in IceCube, we could use unsupervised learning to categorize the data into three different types. We should keep in mind that there's no guarantee that the three types will be electrons, muons, and taus. In fact, the machine tries to find the most obvious differences (maybe something like tracks and cascades).

A.2 Classification and Regression Learning

Two of the most important applications of machine learning which I have used in this analysis are *classification* and *regression.* In classification, data is divided into multiple classes. In this analysis, I have used statistical classification to classify my data into atmospheric muons and upgoing atmospheric neutrinos (see Section 5.5).

In regression analysis, the output is continuous. Regressors try to find a relation between input features (predictors) and an output. In section 5.7, I explain how I have trained my model to be able to predict the energy of primary neutrinos.

A.3 Overfitting and Underfitting

An underfitted model is referred to a model that does not predict trained or untrained sample very well. Underfitting occurs if the model is too simple compared to what it wants to predict. For example, if a power-law data is predicting with a linear model, this model would be an underfitted model. This is also referred to as a bias in the model.

Overfitting or model variance is basically the opposite of undefitting. In this case, the model is too complex for the data it is trying to predict. The main way to detect overfitting is when the model predicts the trained data very well but it cannot predict the additional test data as accurate. Overfitting happens when the machine tries to fit the model to the noise in data. There are different methods for different models that can be used in order to avoid overfitting. Making a less complex model or removing some of the complexity of the model randomly can make overfitting to disappear. Regularization is another way to avoid overfitting in neural networks, by penalizing nodes with a high weight that can dominate the model.

A.4 AdaBoost

AdaBoost is short for “Adaptive Boosting” which was first introduced by Robert Schapire and Yoav Freund in 1996 [68]. In order to have a strong classifier, an ensemble of weak classifiers (or weak learners) are used. These are typically very simple base classifiers such as a decision tree stump and are slightly better than a random guess. A classifier learns from its mistakes as it iterates through weak learners [72].

In the original boosting method classifiers train on samples from the training sample (without replacement) to train a classifier. These samples (with the exception of the first sample) also include half of the misclassified events from the previous classifier. There also needs to be a classifier from a sample which previous classifiers disagree. In the end, a majority vote of these weak classifiers create a strong classifier [72].

In contrast with the original boosting, in AdaBoost, each sample will be trained with replacement. In each round, the classifier learns from its previous mistakes by re-weighting the data. Misclassified events will be given a high weight for the next round. The code below is an example of how AdaBoost’s taring and weighting works.

```
def AdaBoost(data, y, weights=[], n_estimator,
training_method=DecisionTree(depth=1)):
# Define equal weights if not pre-defined
if weights is not []:
weights = numpy.ones(len(data))

# Normalize weights
weights /= sum(weights)

for i in range(n_estimator):
# Train classifier
clf[i] = train(training_method, data, y, weights)
```

```
# Predict y_pred
y_pred = predict(clf, data)

# Calculate error rate
error = numpy.dot(weights, (y == y_pred))

# Calculate coefficient
alpha[i] = numpy.log10(1. - error / error) / 2.

# Update and normalize weights
weights = numpy.cross(weights, numpy.exp(-1 *
numpy.numpy(numpy.cross(alpha[i], y_pred), y)))
weights /= sum(weights)
```

After training, a "decision function" can be calculated like this:

$$decision\_function = \sum_{i=1}^{n_{estimator}} \alpha_i \times predict(clf_i, data) \tag{A.1}$$

By default, the predicted value is 1 when the decision function is positive and 0 when negative i.e. splitting decision function on zero. But, in order to get a better purity on one side, we may change the threshold to a non-zero value. This method is useful in our selection because we only need to select neutrinos from our sample and not muons (see section 5.5).

Fig. A.1 shows a step-by-step process for a simple AdaBoost training with only 2 variables. At each step, misclassified events with that classifier will get a bigger weight (larger circles). In the end, a combination of all the weak classifiers will make the final classifier which is a tree. Each branch of the tree (areas in the final plot) will be assigned a specific decision function, and multiple events will be assigned the same decision function. It is important to note that branches do not have the same size, this explains why the histogram of

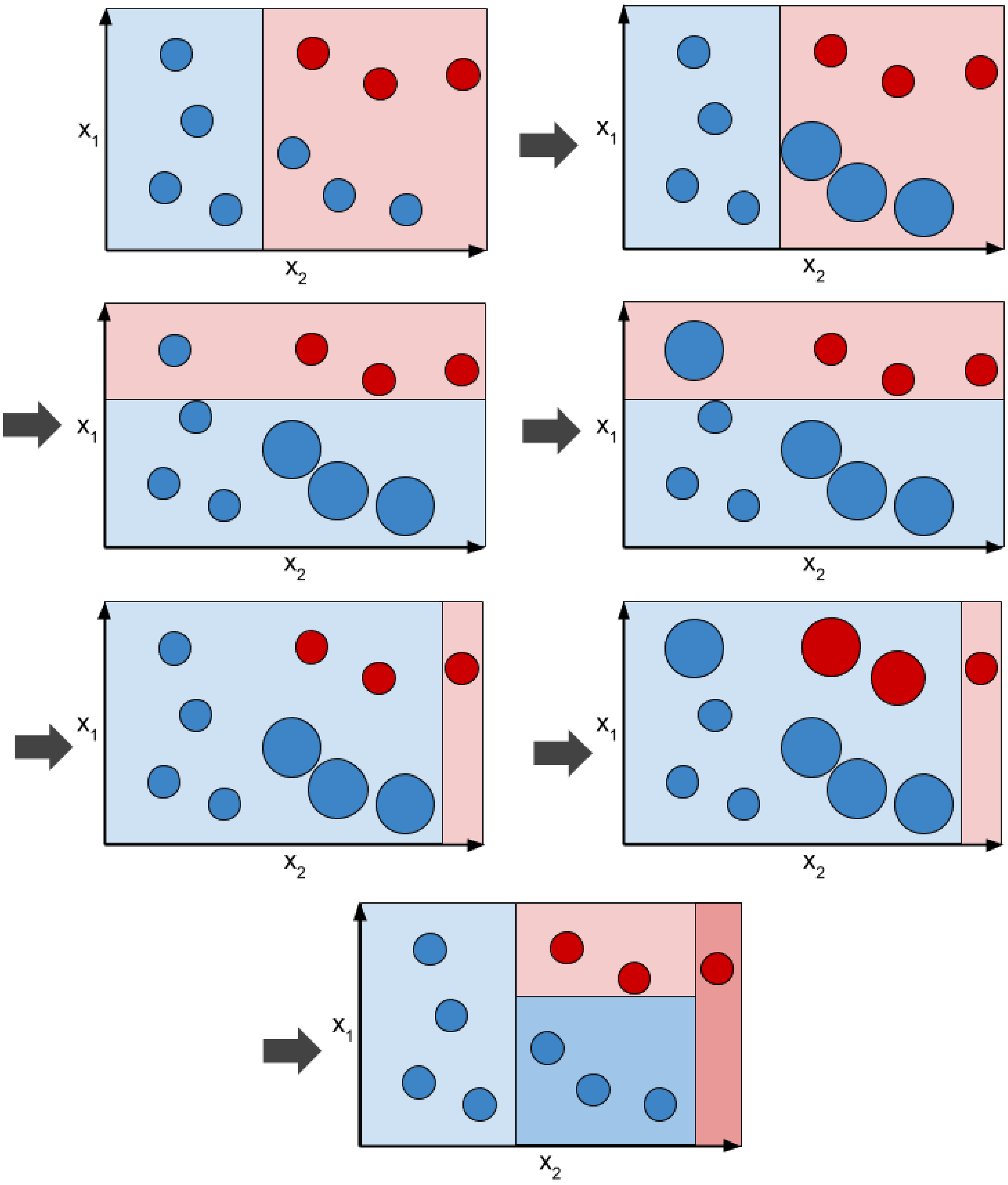


Figure A.1 This plot shows a step-by-step AdaBoost training.

decision functions can have peaks and fluctuations. These fluctuations go away if there are a large number of branches, or in other words, there are no thick branches in the classifier. However, this may result in overfitting of the sample.

APPENDIX
Summary of Cuts

Summary of all cuts are shown here. I start from the IceCube's offial Level 2.

Level 3:

- Using all conventional filters from L2
- Splitting coincident events using Topological Splitter
- NChannel >15
- Reconstructing SPEFit8
- Zenithrec $>\pi/2$
- LDir /textgreater 200 m (to clean ν_e)
- Cleaning and calculating the needed variables

Level 4:

- precuts
- AdaBoosts decision function (for ν selection) calculated
- Running millipede

Level 5 (neutrino level):

- Final neutrino selection
- Muon and cascade energies are calculated from millipede

Level 5 is the neutrino energy which contain both starting and throughgoing upgoing atmospheric neutrinos (with <1% background)

L5st90 (final level):

- electing starting-tracks with 90% purity using AdaBoost classifier
- Neutrino energy is calculated

Level L5St90 is the final level which contain starting tracks (90% truely starting with 10% background). L5st90 is the level which is determined to have the best sensitivity for the sterile analysis.