Muon Neutrino Contained Disappearance in NOvA

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Susan M. Lein

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

PROFESSOR KENNETH HELLER

September, 2015

Acknowledgements

I would like to thank the NOνA collaboration for its contribution to my graduation. Prof. Kenneth Heller helped immensely on my journey to a degree. Jody Kaplan helped to facilitate my path through the twisted networks of bureaucracy. Dr. Matt Strait spent many hours of his life on mentoring me and teaching me best practices of coding and science. Dr. Zeb Krahn protected me from many potential pitfalls and time sinks. Dr. Chris Backhouse has been ever patient with my numerous emails sporting the subject line "Help!" The graduate students of PAN 332 were my rock; I would not have made it to the end without their help. This help included distracting me from pain by frequent walks to Espresso Royale. And thus, I end with Rex, the world's best barista. Thank you to all!

Dedication

I would like to dedicate this thesis to my husband, Zach. I do not function well without Zach; he brings out the best version of Susan. He is ever patient, ever kind, ever encouraging. He is my constant North Star, which is about as much as one will ever hear me say about astronomy.

> Next door, there's an old man who lived to his 90's
> And one day, passed away in his sleep
> And his wife, she stayed for a couple of days
> And passed away
>
> I'm sorry, I know that's a strange way
> To tell you that I know we belong
> That I know that I am the luckiest
>
> -Ben Folds, ''The Luckiest''

Abstract

The NOνA experiment studies neutrino oscillations in the NuMI neutrino beam from Fermilab. NOνA consists of two liquid scintillator tracking calorimeters placed 14 milliradians off-axis from the beam and 810 km apart. The NOνA experiment started taking data in 2014. This thesis establishes the neutrino energy estimation procedures used to determine the oscillation parameters $\sin^2\theta_{23}$ and Δm^2_{32}.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Neutrinos are fundamental particles of the Standard Model. They have no electric charge; they do not interact via the strong force. They have almost no mass, so usually travel at almost the speed of light. Neutrinos can easily pass through the entire earth and never interact. These ghost-like particles are incredibly interesting in their own right; they can also shed light on the wider field of particle physics.

The NOνA (NuMI Off-axis ν_e Appearance) experiment is designed to measure some of the most important properties of neutrinos, such as the differences between mass states and amount of oscillation between flavor states. It consists of two detectors placed 810 km apart: the near detector, located at Fermilab in Batavia, IL, and the far detector, located at Ash River, MN. They detect the neutrinos from the Neutrinos at the Main Injector (NuMI) beam created at Fermilab.

Neutrinos are created at Fermilab in a state of definite flavor but a mixture of mass states. As they travel through time and space, quantum mechanics allows the mixture of mass states to change, giving a probability of observing a different flavor state. This behavior is called *neutrino oscillation*.

The NOνA near detector measures the initial state of the neutrinos in the NuMI beam. The beam travels 810 km to the NOνA far detector. Over this distance, the neutrinos oscillate. By comparing the difference between what is seen in the near detector and the far detector, one can measure the parameters which govern this oscillation. This thesis presents the first results from NOνA which measure two of these parameters, θ_{23} and Δm^2_{32}, from studying muon flavor neutrinos changing into tau flavor neutrinos.

Chapter 2 introduces the Standard Model, the framework particle physicists use to describe the fundamental particles and forces that govern the universe. Neutrinos are one of these fundamental particles.

Chapter 3 describes the historical developments, theoretical and experimental, that have lead to the current understanding of neutrinos and provides a framework for this experiment.

The physics that governs neutrino oscillations is the subject of Chapter 4. Section 4.1.2 reports the current measurements of the parameters of neutrino oscillation, as well as what NOνA hopes to contribute to this understanding.

Chapter 5 details the NuMI beam and the NOνA detectors. It also describes the topologies of interactions seen in the NOνA detectors.

Chapter 6 describes the simulation used to model physics interactions in the NOνA detectors.

Chapter 7 details the process used to reconstruct data from the detector signals or simulation into higher-level information, such as tracks and neutrino energies.

Chapter 8 describes the calibration process used to translate signals seen in the NOνA detectors into energy units of GeV. This process also removes channel-by-channel variations and dependence of the signal on distance to readout.

Chapter 9 details how this analysis reconstructs neutrino energies from the visible charged particle energy in the detector, as well as the resolutions achieved for these energies.

Chapter 10 lists the selection criteria used to define the sampled used for analysis, such as those required for ensuring data quality, beam quality, containment and cosmic ray rejection. The selection criteria used to identify ν_μ charged current events and to classify them as quasielastic and non-quasielastic interactions are also described.

Chapter 12 presents comparisons of data and simulation distributions in the near detector for reconstructed quantities relevant to the analysis.

Chapter 13 describes the extrapolation procedure used. Extrapolation uses the data seen in the near detector to alter the predicted spectra for the far detector. This procedure allows many systematic errors to cancel.

Chapter 14 details the systematic errors considered for this analysis.

Chapter 15 presents the final results of the analysis.

Finally, Chapter 16 discusses possible future analysis improvements.

Chapter 2

The Standard Model

The Standard Model is a physics theory that describes the fundamental building blocks of the universe and how they interact. It is the primary framework used by high-energy physicists to understand the world. See Figure 2.1 for a diagram of the Standard Model components.

Elementary particles are those believed to be the building blocks of all matter. Within the Standard Model, these elementary particles are not composed of smaller things. They are the smallest divisions of matter currently known to physics. These elementary particles are grouped into three types: the quarks, the leptons and the bosons. Quarks are the building blocks of nuclear particles, such as protons and neutrons. The most well-known lepton, the electron, orbits nuclear particles to make an atom. Bosons are the force-carrying particles, like the photon.

Each particle also has an antiparticle. Antiparticles have the same mass as the particle but opposite electrical charge and other properties, such as spin or color. For instance, the antiparticle of an electron would be a position, which has the same mass as the electron, but a positive electric charge instead of a negative electric charge. As far as we know, all ordinary matter in the Universe is composed of particles, not antiparticles. Antiparticles can be produced in radioactive decays or at particle accelerators, but are not the building blocks of a human body. Physicists do not know why there are so many more particles vs. antiparticles.

There are six types of quarks and six types of leptons. Physicists call these different types flavor - that is, that there are six flavors of quark and six flavors of lepton.

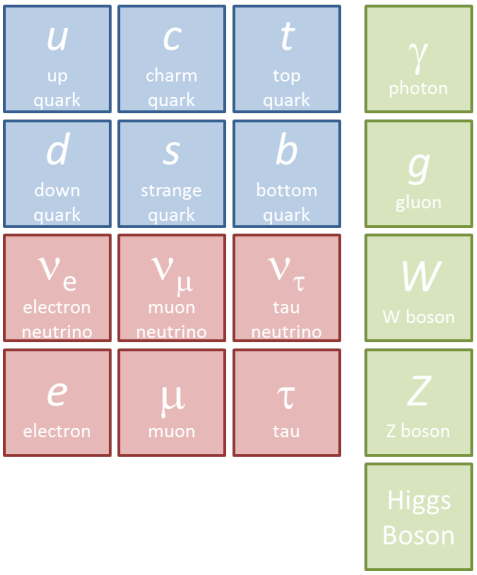


Figure 2.1: The Standard Model describes the fundamental particles of the universe.

Each group of six is actually organized into three pairs, called generations. The first generation of particles is the lightest and most stable. The second generation has higher mass and is less stable (shorter lived) than the first. The final generation, the third, is the heaviest and least stable of all. Because the higher generations decay quickly into lighter particles, stable matter is composed of the first generation.

Physicists do not know why there are three generations of matter. Since ordinary matter is composed of only the first generation of matter, why do any other generations exist? And if other generations exist, why are there three, instead of some other number? Physicists hope to one day answer these questions.

For quarks, the first generation consists of the up quark and down quark. The second generation is composed of the charm quark and the strange quark. The third generation consists of the top quark and the bottom quark.

The mass of the first generation of quarks is on the order of 1 MeV/c^2. The mass of the second generation of quarks is on the order of 100 to 1,000 MeV/c^2. The mass of the bottom quark is about 4,000 MeV/c^2 and the top quark has a mass of almost 200,000 MeV/c^2.

For each pair of quarks, the first (up, charm, and top) has an electric charge equal

to +2/3 the electric charge of an electron, written as 2/3 e. The second of the pair (down, strange, and bottom) has an electric charge of -1/3 e.

Quarks have another property called color charge. This color is not the visual color experienced in everyday life; instead, it is simply a property physicists find convenient to name color. This color charge has three possible values: red, green, and blue. Antiquarks carry anticolor, like antired. A composite particle made up of a red quark and an antired antiquark would be colorless. A composite particle made up of a red, a green, and a blue quark would also be considered colorless (as would one made up of an antired, an antigreen, and an antiblue quark).

For leptons, each generation pair consists of one electrically charged particle and one electrically neutral particle. The electrically neutral particle is called a neutrino (ν). Neutrinos have flavors that correspond to the paired lepton. For the first generation, the charged lepton is the electron (e) and it is paired to the electron neutrino (ν_e). The second generation consists of a muon (μ) and a muon neutrino (ν_μ). The third generation is composed of a tau particle (τ) and a tau neutrino (ν_τ).

The mass of the electron is 0.5 MeV/c^2, lighter than the lightest quark. The muon has a mass of 100 MeV/c^2 and the tau particle has a mass of almost 2,000 MeV/c^2. The mass of the neutrinos is much lighter than even the electron, but although limits on the values exist, the precise values are unknown. The mass of the neutrino is probably on the order of 1×10^{-7} MeV/c^2. The mass of the neutrino will be discussed further in Chapter 4.

Physicists do not understand why the fundamental particles of the universe have mass values over such a wide range. Why should they not be more similar? Why is the neutrino so much lighter than the rest of the particles? Physicists hope to one day have insight into these mysteries.

The charged leptons of each generation (e, μ, and τ) all have an electrical charge of -1 e. All neutrinos are electrically neutral.

Besides fundamental particles, there are also fundamental forces in the universe. In the Standard Model, these are the electromagnetic force, strong force, and the weak force. Although gravity is also a fundamental force, it does not currently fit into the framework of the Standard Model. Physicists certainly hope to one day have a theory that successfully combines the concepts of the Standard Model and gravity.

The electromagnetic force is the same one encountered in every day life; it is responsible for electricity which powers our homes and magnets sticking to the fridge. The strong force is responsible for binding quarks together into structures, like a proton. The weak force is responsible for things like radioactive decay and the fusion reaction in the sun.

Each force represented in the Standard Model has corresponding force carrier particles. Matter, made up of constituent particles like quarks and leptons, experiences these forces by exchanging force carrier particles. Imagine two people sitting on rolling chairs. If one threw a basketball at the other, the chairs would move. Force carrier particles are similar to the basketball; by transferring them, constituent matter particles change energy and momentum.

The force carrier particle for the electromagnetic force is the photon (γ). It is electrically neutral and has no mass. For particles to experience the electromagnetic force, they must have an electric charge. Electrons, with an electrical charge of -1 e, do experience the electromagnetic force, while the electrically neutral neutrinos do not.

The strong force has the gluon (g) as its force carrier particle. Gluons have no electric charge and are massless. They do have color charge; in fact, they carry both a color and an anticolor, like green and antiblue. For a particle to experience the strong force, it must have color. Therefore, quarks experience the strong force and leptons do not.

The weak force has two force carrier particles, the W and the Z. The W and Z have masses of almost 100,000 MeV/c^2. The W particle has an electric charge of either +1 or -1 e, written as either the W^+ or W^- particle. The Z particle has no electric charge and is sometimes written as Z^0. The weak force acts on particles that have flavor, quarks and leptons.

The four fundamental forces have different strengths and different ranges. Both gravity and electromagnetism can act across infinite distance. The strong force and the weak force, however, only act over short distances. 10^{-15} meters is called a Fermi and written as fm. The radius of an atomic nucleus is roughly 1 fm. The strong force has a typical range of a few fm, while the weak force has a range of 10^{-3} fm.

The strong force only acts over nuclear distances, but it is the strongest of the forces. The electromagnetic force is the next strongest, with a strength 1000 times less than

the strong force. The weak force, as the next strongest force, is 10^{16} times less than the strong force. Finally, gravity has a strength 10^{41} times less than the strong force.

These radical differences in strength and range of forces often allows one to simplify problems involving elementary particles. For instance, gravity is so much weaker than the other forces that it can often be ignored. This allows physicists to solve many problems using the Standard Model framework, despite the fact that it doesn't include gravity.

Chapter 3

History of Neutrino Research

Neutrinos were first proposed by Wolfgang Pauli in 1930 as a solution to the mystery of beta decay[12]. In beta decay, a neutron decays into a proton and, to conserve electric charge, also emits an electron. If this is all that were emitted, then experimental evidence showed both conservation of energy and angular momentum were violated by this process. Instead, Pauli suggested that another particle was also emitted, what we now call a neutrino[1], that happened to be very hard to detect experimentally. This would constrain a neutrino to being a lightweight, electrically neutral particle. Since its interactions would be so weak, Pauli speculated that it would never be detected.

Nevertheless, in 1956, the Cowan-Reines experiment found the first direct evidence for the existence of neutrinos[13]. In their experiment, the source of neutrinos was beta decays taking place in a nuclear reactor. They placed nuclei in the path of these neutrinos and were able to detect neutrons and positrons from the resulting interaction[2].

It was discovered that three flavors of neutrinos existed. The previously observed neutrinos turned out to be what we now call electron neutrinos. In 1962 Leon Lederman, Melvin Schwartz, and Jack Steinberger worked on an experiment at Brookhaven which discovered the muon neutrino[14]. They were awarded a Nobel Prize in 1988 for this work. In 1989, LEP (Large Electron-Positron Collider) at CERN[3] began making precise

[1] Pauli initially called this particle a neutron.

[2] Neutron decay generated the neutrinos through $n \rightarrow p^+ e^- \bar{\nu}_e$. The proton interaction with the escaping neutrino took the form $p^+ \bar{\nu}_e \rightarrow e^+ n$. The positron annihilated with an electron in the material, creating two photons. Meanwhile, a nucleus captured the neutron, also emitting a photon. It was this double photon followed by a single photon signature that was detected.

[3] CERN stands for the European Organization for Nuclear Research

measurements of the mass of the Z boson, which led to the conclusion that only three light neutrinos could exist[15]. Finally, in 2000, the DONUT collaboration at Fermilab made the first detection of tau neutrinos[16]. Parallel to the direct detection of three neutrino flavors, measurements of the width of the Z boson have led to limit on the number of light neutrinos[4]. Over time, the measurements of the width of the Z boson and its partial decays have become incredibly precise. The current limit on the number of light neutrinos is 2.984 ± 0.008[1][17].

In parallel to the discovery of three flavors of neutrinos, efforts were being made to determine the neutrino mass. The Standard Model, in its strictest form, assumes massless neutrinos. If this were the case, a neutrino of one flavor would always be a neutrino of that particular flavor. If neutrinos do have mass, then a neutrino of given flavor could become a neutrino of a different flavor (called *neutrino oscillations*).

In 1957, Bruno Pontecorvo first hypothesized the possibility of neutrino mixing and proposed that it could occur between electron neutrinos and electron anti-neutrinos[18]. Of course, at this time, the other flavors of neutrinos had not been discovered. Since these were the only two types of neutrinos known to exist, they were also the only two types that one could imagine oscillating into each other. As knowledge of multiple flavors of neutrinos developed, Ziro Maki, Masami Nakagawa, and Shoichi Sakata, in 1962, proposed instead that oscillations occur between different flavors of neutrinos (as opposed to neutrino and antineutrino)[19]. This work eventually developed into the theory of three neutrino flavor oscillations.

Experimentally, hints pointing to neutrino oscillation first arose in the 1960's when fewer than predicted electron neutrinos were found to be radiating from the sun by Ray Davis and John Bahcall at the Homestake Experiment[20]. The question of neutrino oscillations began to be addressed more directly in the latter part of 1990's with experiments such as Super-Kamiokande (Super-K) and Sudbury Neutrino Observatory (SNO), that found direct evidence that solar neutrinos do oscillate and solved the problem of

[4]By measuring the total decay width of the Z boson and comparing it with the measured visible partial decay widths of the Z boson to charged leptons and quarks, one can determine the invisible particle decay width. The invisible particle decay width is attributable to decays of the Z boson to neutrinos; the width of the decay determines the number of neutrinos. Any neutrinos with mass greater than that of the Z boson wouldn't be counted; however, the neutrinos known to exist all have masses much less than the Z boson.

missing neutrinos from the sun[21, 22]. Kamioka Liquid scintillator Anti-Neutrino Detector (KamLAND) investigated these oscillations terrestrially, using nuclear reactors as the source of the neutrinos[23].

Chapter 4

Neutrino Physics

4.1 Neutrino Oscillation

The weak force has three neutrino flavor states (electron (e), muon (μ), or tau (τ) flavor). The neutrino also has three mass states (m_1, m_2, and m_3). These states, flavor or mass, are *eigenstates*[1]. Eigenstates are states with a quantifiable characteristic. For example, a flavor eigenstate has a definite, determined flavor.

The flavor states and mass states do not correspond one-to-one; instead, they are a superposition of each other. This means that whenever a neutrino interacts weakly, it must be in one of the three flavor states. But whenever a neutrino is traveling through time and space, it evolves as energy eigenstates, which are the mass states. Thus, a neutrino is created from a weak force interaction in a definite state of flavor, then travels a long distance, during which the superposition of mass states are evolving separately (changing the relative probabilities of being found in a certain flavor state), and then interact again through the weak force, now possibly finding a different state of flavor.

What does this mean concretely? We can write each flavor eigenstate as a superposition of mass eigenstates, where ν_α is a flavor eigenstate, c_i is a numerical coefficient, and ν_i is a mass eigenstate.

$$|\nu_\alpha\rangle = \sum_{i=1}^{3} c_i |\nu_i\rangle$$

[1]The word eigenstate is derived from the German word "eigen," which means "characteristic."

We require for normalization that $\sum_{i=1}^{3} c_i^2 = 1$. Let us rewrite this equation by replacing the coefficients c_i with the matrix element $U_{\alpha i}^*$, where the star indicates complex conjugation.

$$|\nu_\alpha\rangle = \sum_{i=1}^{3} U_{\alpha i}^* |\nu_i\rangle \tag{4.1}$$

This matrix, $U_{\alpha i}$, is referred to as the Pontecorvo–Maki–Nakagawa–Sakata (PMNS) matrix in deference to those who first formulated it. This matrix is 3x3 if there are only three types of neutrinos in any basis. There is no prediction of what values should fill in this matrix. Instead, experimental results must determine them for now. If experiments found that neutrinos did not oscillate, then the PMNS matrix would be the identity matrix. A particular mass state would always correspond to a particular flavor state. However, this is not the case.

How many free parameters does the PMNS matrix have? A complex 3x3 matrix would in general have 18 parameters, but luckily, many of these are not free in our case. Between constraints such as normalization and unitarity, and factors that will be ignored because they do not affect oscillations, we can reduce the PMNS matrix to three independent angles, θ_{12}, θ_{13}, θ_{23}, and one phase, δ, that, if non-zero, would indicate charge-parity (CP) violation. It is conventionally written as:

$$U_{\alpha i} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \tag{4.2}$$

where $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$.

Leaving for now the question of what these four values are, let us first ask how to use this matrix. From quantum mechanics, it is easy to write down the plane-wave solution for a free particle of given mass propagating. It is:

$$|\nu_i(t)\rangle = e^{-i(E_i t - \vec{p}_i.\vec{x})} |\nu_i(0)\rangle, \tag{4.3}$$

where t is the time, E_i is the energy of the neutrino, $\vec{p}_i$ is the momentum of the i^{th} mass state, and $\vec{x}$ is the position of the neutrino. The neutrino is created in a state of definite energy, so we can write $E_i = E$. Next, we will take the ultra-relativistic limit. This

limit is appropriate because the current limit on neutrino mass is below 2 eV[24]. Since experiments like NOνA can only detect neutrinos of energies at least above 0.1 GeV, neutrino velocities essentially equal c. We can then equate $t \simeq L$ using natural units, where L is the distance the neutrino travels before being detected. This allows us to write:

$$|\nu_i(L)\rangle = e^{-iL(E-p_i)}|\nu_i(0)\rangle. \tag{4.4}$$

Since the neutrinos are traveling at almost the speed of light, we can also make the approximations $p_i >> m_i$ and $p_i \simeq E$. When we take a Taylor's expansion of the momentum, we obtain:

$$E - p_i = E - \sqrt{E^2 - m_i^2} \simeq E - E\left(1 - \frac{m_i^2}{2E^2}\right) = \frac{m_i^2}{2E}.$$

Plugging these approximations into Equation 4.4 we find:

$$|\nu_i(L)\rangle = e^{\frac{-im_i^2 L}{2E}}|\nu_i(0)\rangle. \tag{4.5}$$

A neutrino is created from a weak force interaction in a state of definite flavor, α. This flavor state is a linear superposition of mass states, give by Equation 4.1. Equation 4.5 tells us how each of these mass states propagates as the neutrino moves through space. When we detect a neutrino, it is projected into a state of definite flavor, β. We now have the pieces required the construct $P_{\alpha\to\beta}$, the probability of observing flavor state β given initial flavor state α after traveling a distance L:

$$P_{\alpha\to\beta} = \left|\langle\nu_\beta\,|\mathcal{O}(L)|\nu_\alpha\rangle\right|^2,$$

where $\mathcal{O}(L)$ describes the evolution of the state. We can now insert identity, to write:

$$P_{\alpha\to\beta} = \left|\sum_{i=1}^{3}\langle\nu_\beta\,|\mathcal{O}(L)|\nu_i\rangle\langle\nu_i|\nu_\alpha\,\rangle\right|^2.$$

Using Equation 4.5, we can rewrite the evolution of the mass state as:

$$P_{\alpha\to\beta} = \left| \sum_{i=1}^{3} \langle \nu_\beta \, | \nu_i \rangle \langle \nu_i | \nu_\alpha \, \rangle \, e^{\frac{-im_i^2 L}{2E}} \right|^2 .$$

Using Equation 4.1, we now can write:

$$P_{\alpha\to\beta} = \left| \sum_{i=1}^{3} U_{\alpha i}^* U_{\beta i} e^{\frac{-im_i^2 L}{2E}} \right|^2 .$$

Squaring this result, we get:

$$P_{\alpha\to\beta} = \sum_{i=1}^{3} \sum_{j=1}^{3} U_{\alpha j} U_{\beta j}^* U_{\alpha i}^* U_{\beta j} e^{\frac{im_j^2 L}{2E}} e^{\frac{-im_i^2 L}{2E}}.$$

Finally, we will use the notation $\Delta m_{ij}^2 \equiv m_i^2 - m_j^2$, so we can write the more familiar form:

$$P_{\alpha\to\beta} = \sum_{i=1}^{3} \sum_{j=1}^{3} U_{\alpha j} U_{\beta j}^* U_{\alpha i}^* U_{\beta j} e^{\frac{-i\Delta m_{ij}^2 L}{2E}}. \tag{4.6}$$

From here, we can go on to write this expression in terms of the real and imaginary parts, expanding the exponential into sine and cosine components as well as using trigonometry identities to obtain[2] [25]:

$$\begin{aligned} P_{\alpha\to\beta} = \delta_{\alpha\beta} &- 4 \sum_{i>j} \mathfrak{Re} \left(U_{\alpha j} U_{\beta j}^* U_{\alpha i}^* U_{\beta j} \right) \sin^2 \left(\frac{\Delta m_{ij}^2 L}{4E} \right) \\ &+ 2 \sum_{i>j} \mathfrak{Im} \left(U_{\alpha j} U_{\beta j}^* U_{\alpha i}^* U_{\beta j} \right) \sin \left(\frac{\Delta m_{ij}^2 L}{2E} \right). \end{aligned} \tag{4.7}$$

The first term is simply the Kronecker delta. Regarding the third term, the only complex phase in the PMNS matrix is $e^{i\delta}$. If $\delta = 0$, there is no imaginary part, the third term is zero, and it corresponds to no CP violation. The second term is of primary importance to neutrino oscillation. For each term in the sum, $\mathfrak{Re}\left(U_{\alpha j} U_{\beta j}^* U_{\alpha i}^* U_{\beta j}\right)$is just a constant

[2]The difference between denominators in the two sine terms results from taking the first term, originally a cosine, and using the "half angle" identity $2\sin^2 \frac{u}{2} = 1 - \cos u$.

that depends on the properties of neutrinos. However, the value of $\sin^2\left(\frac{\Delta m_{ij}^2 L}{4E}\right)$ alters based on the neutrino energy as well as the distance of the measurement point from the creation location. We can now clearly see that as one's detector moves farther away, the probability of measuring a certain flavor change.

Let us write the oscillation phase in the form usually used by long-baseline neutrino experiments, like NOνA . It is currently written in natural units, but let us restore the c and $\hbar$ factors. Then

$$\frac{\Delta m_{ij}^2 L}{4E} \rightarrow \frac{\Delta m_{ij}^2 L c^3}{4E\hbar}.$$

We want to input our variables in convenient units. This means that we want to measure Δm_{ij}^2 in eV^2, E in GeV, and L in km. Using the appropriate unit conversions and numerically evaluating $\hbar c$, we find that

$$\frac{\Delta m_{ij}^2 L c^3}{4E\hbar} \sim 1.27 \frac{\Delta m_{ij}^2 L}{E} \equiv \Delta_{ij} \tag{4.8}$$

for variables of the listed units.

4.1.1 Two Neutrino Approximation

In many experimental cases, one need not consider the full effect of three neutrino flavor mixing, but instead can consider the approximation of two neutrino mixing. If neutrinos had only two flavor states and two mass states, the mixing matrix could be written as:

$$U_{\alpha i} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}. \tag{4.9}$$

The probability of observing flavor state β given initial flavor state α (where $\beta \neq \alpha$) that has evolved distance L is:

$$P_{\alpha\rightarrow\beta,\alpha\neq\beta} = \sin^2 2\theta \sin^2 \frac{\Delta m^2 L}{4E}. \tag{4.10}$$

Similarly, the probability of observing flavor state α given initial flavor state α that has evolved distance L is:

$$P_{\alpha\rightarrow\alpha} = 1 - \sin^2 2\theta \sin^2 \frac{\Delta m^2 L}{4E}. \tag{4.11}$$

For a experiment like NOνA , the probability of observing a muon neutrino given a beam composed of initial state muon neutrinos can be described, to first order, by Equation 4.11. In this approximation, θ and Δm^2 can be parametrized as [26][27]:

$$\sin^2 2\theta = 4\sin^2\theta_{23}\cos^2\theta_{13}(1 - sin^2\theta_{23}\cos^2\theta_{13}) \tag{4.12}$$

and

$$\Delta m^2 = \Delta m^2_{32} + \Delta m^2_{21}\sin^2\theta_{12} + \Delta m^2_{21}\cos\delta_{CP}\sin\theta_{13}\tan\theta_{23}\sin 2\theta_{12}. \tag{4.13}$$

This approximation allows one to see, to first order, how the probability depends on the θ_{ij}, Δm^2_{ij}, and δ_{CP} values. For the analysis presented in this thesis, the full probability from three neutrino oscillation is used. θ_{23} and Δm^2_{32} are the measured quantities; all other parameters are fixed. Table 4.2 displays the fixed values used.

4.1.2 Parameter Values

Now that we understand the physics framework of neutrino oscillations, we can finally return to the question of what has already been experimentally determined. Table 4.1 lists the best-fit values for the parameters which govern neutrino oscillation. The two mass differences are measured. The three mixing angles are measured, although more precise values could help to illuminate underlying symmetries to theorists. If $\sin^2(2\theta_{23})$ doesn't equal exactly 1, physicists would want to measure if θ_{23} is greater than or less than 45°. This is called *determining the octant*. The CP violating phase factor, δ_{CP}, is still relatively unknown, although recent results have lead to a weak constraint on its value.

The actual masses of neutrinos, besides just the mass differences, need to be determined experimentally. Upper limits have been placed on neutrino masses[24]. Even without an absolute mass value, one would like to know the relative order in magnitude of the mass states. Experiment has determined that mass state ν_1 is lighter than mass state ν_2 and that the difference between these states is small relative to the third state. However, it is still undetermined if ν_3 is much heavier or much lighter than the other states. The case where ν_3 is much heavier is called *normal hierarchy*, while the other

Parameter	Best-Fit Value
Δm^2_{21} [10^{-5} eV^2]	7.53 ± 0.18
$\lvert\Delta m^2_{32}\rvert$ [10^{-3} eV^2]	2.44 ± 0.06 (2.52 ± 0.07)
$\sin^2(2\theta_{12})$	0.846 ± 0.021
$\sin^2(2\theta_{23})$	$0.999^{+0.001}_{-0.018}$ ($1.000^{+0.000}_{-0.017}$)
$\sin^2(2\theta_{13})$	0.093 ± 0.008
δ_{CP}/π	$1.39^{+0.38}_{-0.27}$ ($1.31^{+0.29}_{-0.33}$)

Table 4.1: Best-fit values for the parameters of neutrino oscillation from the Particle Data Group[1]. When the value is different under the assumptions of normal hierarchy ($m_1 < m_2 < m_3$) and inverted hierarchy ($m_3 < m_1 < m_2$), the normal assumption is given first and the inverted is listed next in parenthesis. The best-fit value is listed with $\pm 1\sigma$ errors.

case is *inverted hierarchy*. The two hierarchies are diagrammed in Figure 4.1.

NOνA was primarily designed to measure θ_{13} and determine the mass hierarchy. NOνA will also eventually make precise measurements of θ_{23} and Δm^2_{32}; this thesis presents the first measurements of these two parameters. If $\sin^2(2\theta_{23})$ doesn't equal 1, NOνA could determine the octant. NOνA will also limit the allowed CP phase values.

4.2 Weak Interactions

Neutrinos are created and destroyed through weak force interactions. An understanding of these physics processes is crucial for performing a neutrino oscillation experiment.

A weak force interaction consists of an exchange of either a W or Z boson. A W boson has an electrical charge of ± 1, while a Z boson is electrically neutral. Interactions that involve a W boson are therefore called a *charged-current interaction* (CC interaction), while those that are mediated by a Z boson are called a *neutral-current interaction* (NC interaction).

In CC interactions, a charged lepton (e, μ, τ) can interact with a W and produce a neutrino, which is electrically neutral. The flavor of the neutrino created will correspond to the type of charged lepton destroyed. This process can also happen in reverse, where

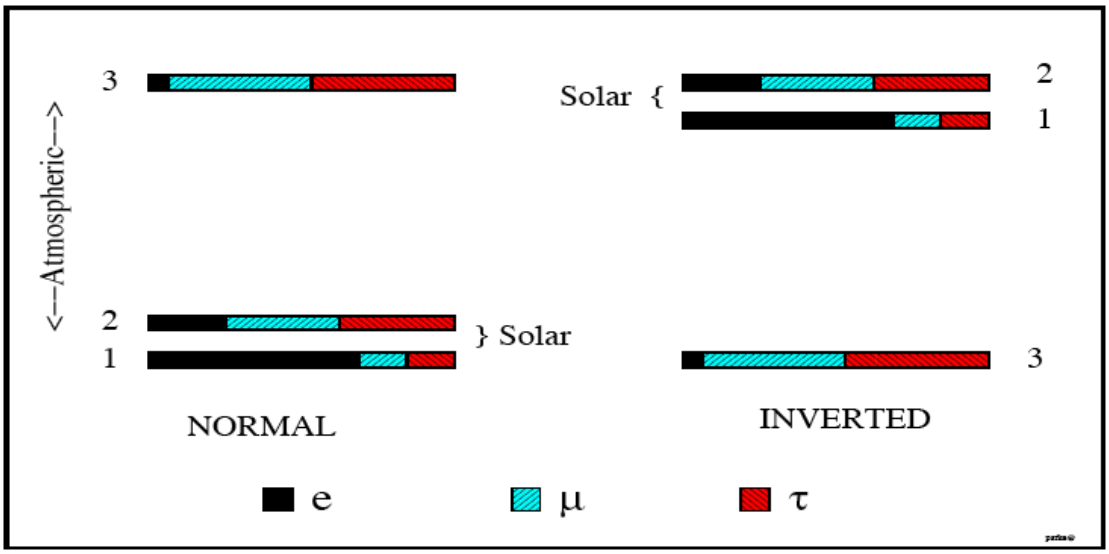
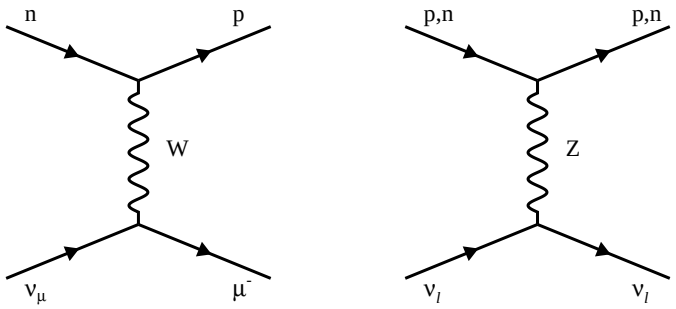


Figure 4.1: The two possible mass orderings for neutrinos. The normal hierarchy ($m_1 < m_2 < m_3$) is on the left; the inverted hierarchy ($m_3 < m_1 < m_2$) is on the right. The vertical axis is increasing mass. The relative proportion of flavor states in each mass state is indicated by v_e corresponding to the solid black, ν_μ is teal with right-leaning hashes, and ν_τ is red with left-leaning hashes. The v_e proportion in the third mass state is exaggerated[5].

Parameter	**Value**
Δm^2_{21} [10^{-5} eV^2]	7.53
$\sin^2(2\theta_{12})$	0.846
$\sin^2(2\theta_{13})$	0.084
δ_{CP}/π	0
L [km]	810
ρ [g/cm^3]	2.84
Hierarchy Assumption	Normal

Table 4.2: Parameter values used in NOνA ν_μ disappearance analysis. The values of Δm^2_{21} and $\sin^2(2\theta_{12})$ match the values from the Particle Data Group. The value of $\sin^2(2\theta_{13})$ was chosen to match the Neutrino 2014 Daya Bay results[2]. δ_{CP} has little impact on this analysis; zero was picked for simplicity. L is the distance the neutrinos travel between the near detector and far detector. ρ is the density of the earth; this parameter is necessary for the matter effects discussed in Section 4.3. Appendix B discusses the value of ρ. The hierarchy assumption used makes no difference to this analysis; for simplicity, the normal assumption was used.



(a) ν_μ Charged Current (b) Neutral Current

Figure 4.2: The left figure is a diagram of one type of CC interaction. In this example, a ν_μ interacts with a neutron. They exchange a W, creating a proton and a muon. The right figure is a diagram of a NC interaction. In it, a neutrino of any flavor exchanges a Z with either a proton or a neutron. The neutrino leaves with the same flavor and the particle type in the nucleus remains unchanged.

a neutrino absorbs a W to produce a corresponding charged lepton. See Figure 4.2a for a diagram.

In NC interactions, only energy and momentum is exchanged and charged leptons (e, μ, τ) can not create neutrinos or vice versa. See Figure 4.2b for a diagram.

In CC interactions, the NOνA detector would typically observe the charged lepton created by the incoming neutrino, as well as hadronic energy from the interacting nucleus. By determining the type of charged lepton, one can measure the flavor composition of the neutrino beam at that location. On the other hand, in a NC interaction, one would only expect to see the hadronic energy resulting from the nuclear interaction with the Z, while the outgoing neutrino would invisibly carry away much of the energy of the initial neutrino. NC interactions do not depend on the flavor composition of a neutrino beam; they depend only on the number of neutrinos and their momentum. NC interactions can be useful in understanding the total flux of neutrinos.

For CC interactions between neutrinos and nuclear matter, one can further subdivide the interactions into roughly elastic and inelastic types. The roughly elastic case is called a *quasielastic* (QE) interaction, while the other type of interaction will be referred to as a *non-quasielastic* (non-QE) interaction.

For a QE interaction, the nucleus stays relatively intact from the interaction. QE

interactions are the primary type of CC interaction at lower energies, up to roughly 1 GeV. This is because there isn't enough energy available to create more complicated final states. Once an energy threshold is reached, resonance production becomes available and contributes significantly to the total cross section. For resonance production, enough energy is transferred to create particles like a Δ. At slightly higher energies, deep inelastic scattering from individual quarks takes over as the predominant type of CC interaction. This is likely to cause nuclear breakup and create many particles, often referred to as a hadronic shower. The NOνA experiment measures neutrinos at the energy where all three types of CC interaction contribute. It is a particularly challenging, but exciting, region to understand, simulate, and interpret.

QE interactions are prized because they are easier to model. Also, because of the relatively simple topology of usually one or two final state particles, one can reconstruct the initial neutrino energy with higher precision. This helps to make a precise measurement of the oscillation parameters.

Since QE interactions are the predominant CC interaction only in NOνA's lowest energy region, non-QE interactions are also important to the oscillation analysis. Although harder to model and with a worse neutrino energy resolution, these interactions represent a significant amount of statistics. Especially in early analyses where the number of events is small, it would be foolish to ignore this category of events.

4.2.1 Final State Interactions

There is an important complication to the situation described. The interactions described take place within a nucleus. However, experiments are sensitive only to particles that leave a nucleus. The interactions that can happen between the initial, weak interaction and particles leaving the nucleus are called *final state interactions*. They occur through the strong force and can significantly alter the interaction.

In a general QE interaction, one expects to see a proton in addition to the outgoing charged lepton. However, the proton often does not have enough energy to emerge from the nucleus and isn't visible to the experimentalist. On the other hand, a non-QE interaction could create a pion, but this particle might be absorbed before it exists the nucleus. If one is not able to detect elevated hadronic activity, one might mistakenly classify it as "QE." Because different detector technologies have differing abilities to

detect and distinguish low energy final state hadrons, the results from different detectors are sensitive to the simulations of low energy hadron production and the detector's response.

Final state interactions are difficult to simulate. They involve strong force interactions that are not well-understood. This is a challenge for the experimentalist; in general, the simulation should not be overly trusted in the precise details of the hadronic system for non-QE interactions. Data-driven methods to understand the detector response are crucial.

For a more complete picture of neutrino interactions, see the excellent thesis by Scully[28].

4.3 Matter Effects

When neutrinos travel through a vacuum, they simply oscillate in the manner described in Section 4.1. However, when neutrinos travel through a large amount of matter, like the earth or the sun, their oscillation is not so straightforward. These complications are called *matter effects*. Matter is composed almost entirely of electrons, protons, and neutrons. When traveling though matter, ν_e can have charged current interactions with the electrons. But ν_μ and ν_τ cannot participate in similar interactions since muons and tau particles are not components of the earth. Instead, ν_μ and ν_τ can only interact through the neutral current interactions that ν_e also participate in. This asymmetry leads to an alteration of the oscillation probabilities. From the vacuum probability given in equation 4.7 and remembering the notation of equation 4.8, the matter effect gives [5]:

$$\sin \Delta_{ij} \xrightarrow[matter]{} \frac{\Delta_{ij}}{\Delta_{ij} \mp aL} \sin\left(\Delta_{ij} \mp aL\right) \tag{4.14}$$

where the top sign refers to neutrinos and the bottom to antineutrinos. Also,

$$a = \frac{G_F \rho_e}{\sqrt{2}}$$

where G_F is the Fermi constant and ρ_e is the electron density in the matter traversed. For the normal mass hierarchy, this would increase the oscillation probability of $\nu_\mu \to \nu_e$ while simultaneously decreasing the oscillation probability of $\bar{\nu}_\mu \to \bar{\nu}_e$. However, for the

inverted mass hierarchy, the opposite is true and the neutrino probability would be decreased while the antineutrino probability would be increased.

This effect alters the vacuum expectations by about 30% in the NOνA experiment[5], which amplifies the mass hierarchy signal. However, the relatively-unknown CP phase complicates the picture. Depending on what values are measured for neutrino and antineutrino oscillation probabilities, an ambiguity might exist where each mass hierarchy could be responsible for the observed values if it corresponding to a different CP phase. If this is the case, another independent experiment will be needed to resolve the issue. Figure 4.3 demonstrates this ambiguity.

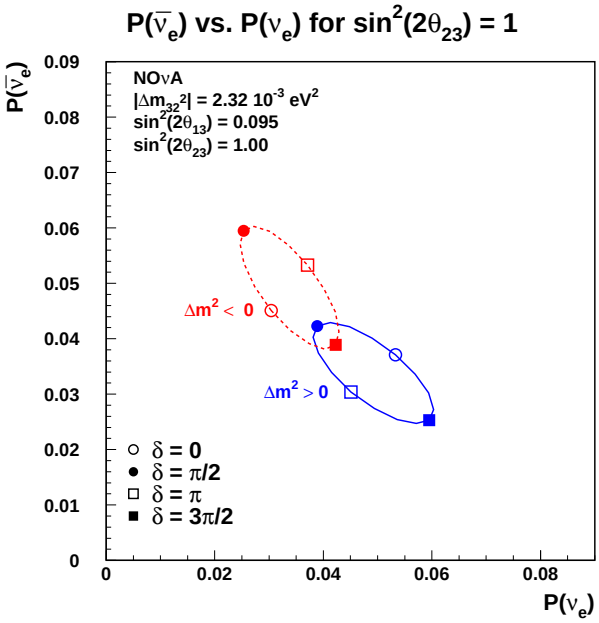


Figure 4.3: Plot comparing the appearance probability of ν_e, shown on the horizontal axis, with the appearance probability of $\bar{\nu}_e$, shown on the vertical axis. The blue ellipse traces out the probabilities for the normal hierarchy; the red ellipse is for the inverted hierarchy. Locations on the ellipse correspond to different values of δ_{CP}. The open circle corresponds to $\delta_{CP} = 0$, the solid circle corresponds to $\delta_{CP} = \pi/2$, the open square corresponds to $\delta_{CP} = \pi$, and the solid square corresponds to $\delta_{CP} = 3\pi/2$. Since the ellipses cross, for some combinations of hierarchy and δ_{CP}, NOνA will be unable to resolve the parameters. However, if nature happens to correspond to the normal hierarchy and $\delta_{CP} = 3\pi/2$, for instance, then NOνA would be able to determine both the hierarchy and the value of δ_{CP}. The probabilities are evaluated for 2 GeV neutrinos and antineutrinos which have traveled 810 km. The oscillation parameters used are listed on the plot.

Chapter 5

The NOνA Experiment

The NOνA (NuMI Off-axis ν_e Appearance) experiment consists of two detectors. The near detector is located at Fermilab in Batavia, IL, and the far detector is located at Ash River, MN near the Canadian border and Voyageurs National Park. The detectors are composed of a repeated cellular structure and are almost identical, except in overall size (the near detector is much smaller than the far detector) and a few other details, discussed further in Section 5.1. They are placed to intercept the neutrinos from the Neutrinos at the Main Injector (NuMI) beam created at Fermilab. See Figure 5.1 for a map of the detector locations and beam center.

The NuMI neutrinos are created by smashing protons from Fermilab's accelerator ring onto a fixed graphite target, shown in Figure 5.3. This collision creates a myriad of particles, many of which are pions and kaons and often carry an electric charge. Magnetic horns are placed to focus one sign of these charged particles into a beam. The particles then decay in flight, creating neutrinos that travel in essentially the same direction as the parent particle. Of these neutrinos, the majority are muon neutrinos from pion decay. These pass through the near detector, which can then determine the initial composition of the neutrino beam. The neutrino beam continues to the NOνA far detector in northern Minnesota. NOνA's far detector and near detector both sit 14 milliradians off-axis from the beam line, as this provides an approximately monoenergetic beam more amenable to a θ_{13} measurement (described in Section 5.2).

The near detector sits 105 m underground in a cavern. The far detector site also had to be excavated, but instead of being underground, the top of the far detector is

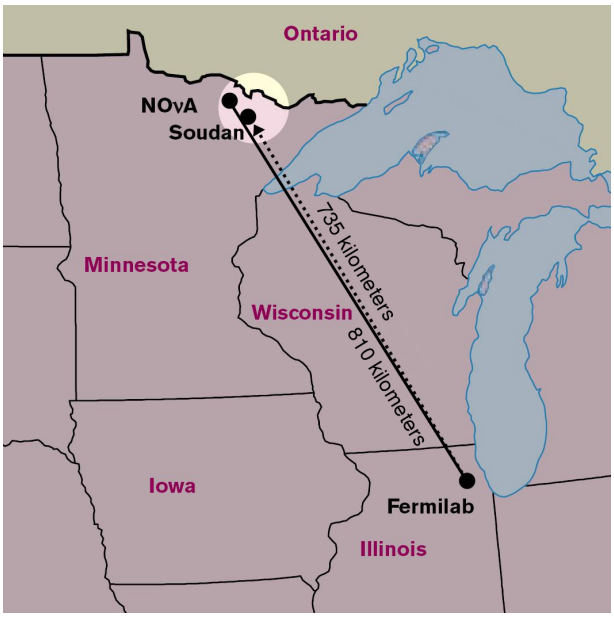


Figure 5.1: Map that shows the location of the NOνA detectors. The Near Detector is located at Fermilab, near Batavia, Illinois. The Far Detector is located at Ash River, Minnesota, near the Canadian border. It is 810 km from Fermilab. The NuMI beamline points at Soudan, Minnesota, where the MINOS Far Detector is located. The NOνA detectors are placed 14 milliradians off-axis from the beam center.

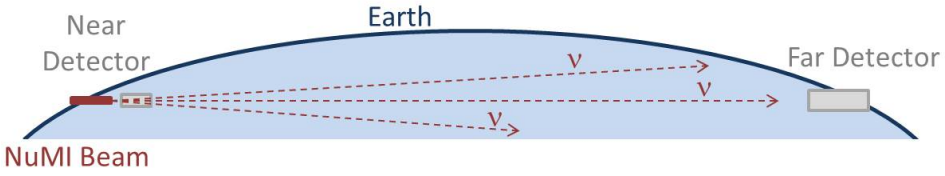


Figure 5.2: Diagram that shows the location of the NOνA detectors. The NuMI beamline is created at Fermilab, near Batavia, Illinois. The neutrinos travel through the Earth and emerge at the NOνA far detector, in northern Minnesota. The NOνA near detector is located near the source of the NuMI beamline to measure the initial properties of the beam.

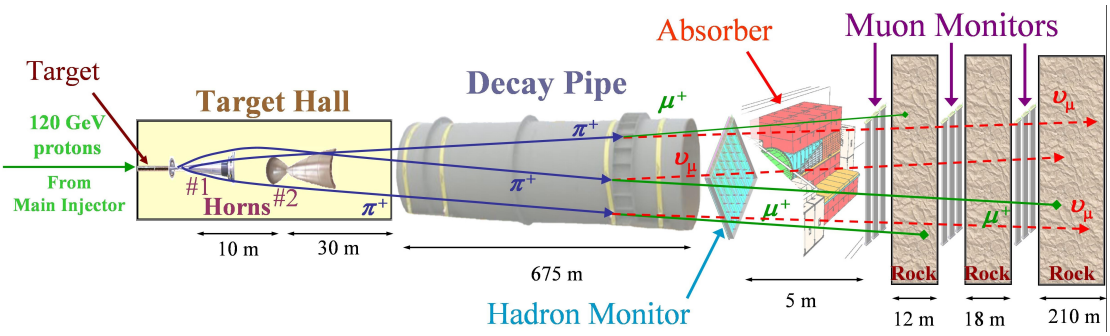


Figure 5.3: The NuMI beam decay pipe. Protons from the Main Injector are directed at graphite target. Charged pions and kaons are created and one sign is focused, using the magnetic focusing horns. These charged particles then decay in the decay pipe, producing primarily muons and muon neutrinos. The muons are absorbed in the rock, while the neutrinos travel onwards, creating the NuMI beam.[9].

	Far Detector	**Near Detector**
Angle Off-Axis [mrad]	14.6	14.6
Distance from Beam [km]	810	1
Depth Underground [m]	0	105
Effective Depth Underground [m of rock]	3	105

Table 5.1: Locations of the NOνA detectors relative to the NuMI beam and the surface of the earth. Values obtained from [3], [4], [5].

approximately level with the ground. The roof of the far detector hall is covered with concrete and barite, which is effectively equivalent to having 3 m of rock above it. Table 5.1 lists the relative positions of the NOνA detectors. A consequence of the differing depths of the two detectors is that the near detector sees significantly fewer cosmic rays than the far detector. This is discussed further in Sections 5.3 and 10.3.

5.1 The NOνA Detectors

The NOνA detectors are composed of many cells. Each cell is a reflective PVC tube that is filled with liquid scintillator (mineral oil doped with 5.23% pseudocumene[1] [29]).

[1] 1,2,3-Trimethylbenzene

Within the cell is a loop of 0.7 mm wave-length shifting fiber whose ends are connected to an Avalanche Photodiode (APD). This is illustrated in Figure 5.4. These cells are contained in PVC extrusions making up flat modules of 32 cells. As shown in figure 5.5, these modules are then glued together into blocks.

The far detector contains 12 modules per plane and 32 planes create a block. The far detector has 28 blocks. The near detector has a slightly more complicated structure. The fully-active region of the near detector consists of 3 modules per plane. There are 6 blocks of 32 planes in the fully-active region. The near detector also has a muon catcher. Because of its relatively short length, it can be hard for muons of interesting energies to be contained within the fully-active region of the near detector . The muon catcher allows for muons of a higher energy to be contained. It consists of 4 inch steel planes interleaved with two active planes[2]. For more information about the steel plates, see this document[30]. The muon catcher is also shorter than the fully-active region of the near detector; it is 3 modules wide but only 2 tall. The muon catcher has 10 steel planes intermixed with 22 active planes. For more details about the muon catcher dimensions and spacing, see the technical drawings[31]. Table 5.2 lists parameter values for the near and far detectors.

The detector planes are always orientated approximately perpendicular to the direction of the neutrino beam, but each plane alternates between being horizontally or vertically aligned. This allows for three-dimensional track reconstruction.

When a neutrino interacts within the detector volume, visible charged particles are usually created. These then travel through the detector and have further interactions. These interactions, although varied, often result in the creation of photons. These photons are frequently created within the cell (since 63% of the detector mass is liquid scintillator [32][33]) and bounce around, reflecting off the PVC walls. Eventually some photons hit the wave-length shifting fiber running the length of the cell. The fiber absorbs light primarily in the blue and UV range and re-emits photons in the green range isotropically. Some of these green photons are then trapped inside the fiber by internal reflection and travel to the fiber end where they are measured by the APD.

[2]This helps to contain muons because they lose proportionally much more energy in steel than in the active detector. Interleaving with active planes allows one to still track the muon somewhat, although at coarser resolution than in the fully-active region.

	Far Detector	**Near Detector**
Cells per Module	32	32
Modules per Plane	12	3 (3, 2)
Cells per Plane	384	96 (96, 64)
Planes per Block	32	32 (n/a)
Number of Blocks	28	6 (2)
Number of Cells	344,064	18,432 (1,760)
Number of Planes	896	192 (22 + 10 steel)
Cell Depth [cm]	5.64	5.64
Cell Width [cm]	3.6	3.6
Cell Length [cm]	1550	399 (399, 274)
Detector X Dimension Extents [cm]	-780 to 780	-200 to 200
Detector Y Dimension Extents [cm]	-780 to 780	-200 to 200 (-200 to 70)
Detector Z Dimension Extents [cm]	0 to 5,962	0 to 1,280 (1,280 to 1,560)
Detector Mass [ton]	14,363	293
Liquid Scintillator [gal]	2,674,000	41,140
Wavelength-shifting Fiber [km]	11,116	188

Table 5.2: Detector parameter values obtained from [3], [6]. In cases where the muon catcher has different values from the active region of the near detector, values in parenthesis correspond to muon catcher values. In cases where the vertically-orientated planes in the muon catcher have a different value from the horizontally-orientated planes in the muon catcher, the value for the vertically-orientated planes is listed first, followed by a comma and the value for the horizontally-orientated planes.

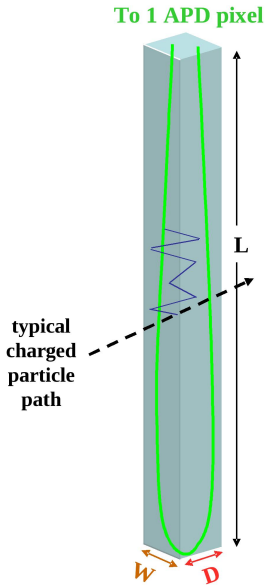


Figure 5.4: Diagram of a NOνA cell. The cell walls are composed of PVC. A loop of wave-length shifting fiber is read out by one APD pixel. The fiber is contained in a bath of liquid scintillator. When a charged particle travels through the cell, the liquid scintillator is excited and emits photons, primarily in the blue and UV range. Some of these photons are absorbed by the fiber, which re-emits photons isotropically in the green range. Some are captured within the fiber and travel its length, eventually being read out by the APD. The dimensions of a cell are 3.9 cm by 6.6 cm by a length of 15.6 m.

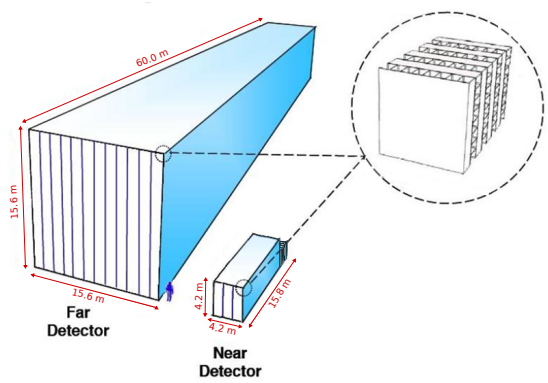


Figure 5.5: Diagram of the Far and Near NOνA Detectors. The far detector is 15.6 m by 15.6 m on the front face, with a depth of 60.0 m. The near detector has a front face of 4.2 m by 4.2 m. The fully active region is 12.7 m long. This is followed by a muon catcher, which has a front face with a width of 4.2 m and a height of 2.6 m. It has a depth of 3.1 m. The near detector has a total depth of 15.8 m. The inset figure illustrates that each plane has a rotated orientation relative to the one next to it. This allows for 3D reconstruction.

The APD reads these relatively small levels of photons[3] and amplifies the signal by a factor of one hundred into a level capable of being read by sensitive, low-noise electronics. The amplified signal is read by a Front End Board (FEB). There is one FEB for each APD. The FEB reads the analog signal from the APD and converts it into digital hits above threshold. The hit information from the FEB is collected by a Data Concentrator Module (DCM). One DCM reads 64 FEBs. The system encompassing the FEBs, DCMs, and buffer nodes (described next) is called a data acquisition (DAQ) system. For a diagram of this system, see Figure 5.6.

There are many DCMs on each detector; they send all the collected hit information for a 5 millisecond time window to a single buffer node. The buffer node then executes triggering decision code that decides if any portion of the hits should be written out to a file. NOνA employs a *round-robin* system with its buffer nodes. This means that there are many buffer nodes, arranged in a ring. The first set of 5 millisecond data is sent to the first buffer node in the ring. The next set of 5 millisecond data is sent to the second buffer node and so forth, until all the buffer nodes have received a set of 5 millisecond data. Then the first buffer node gets its second set of 5 millisecond

[3] An average value in the far detector is about 30 photoelectrons[34]

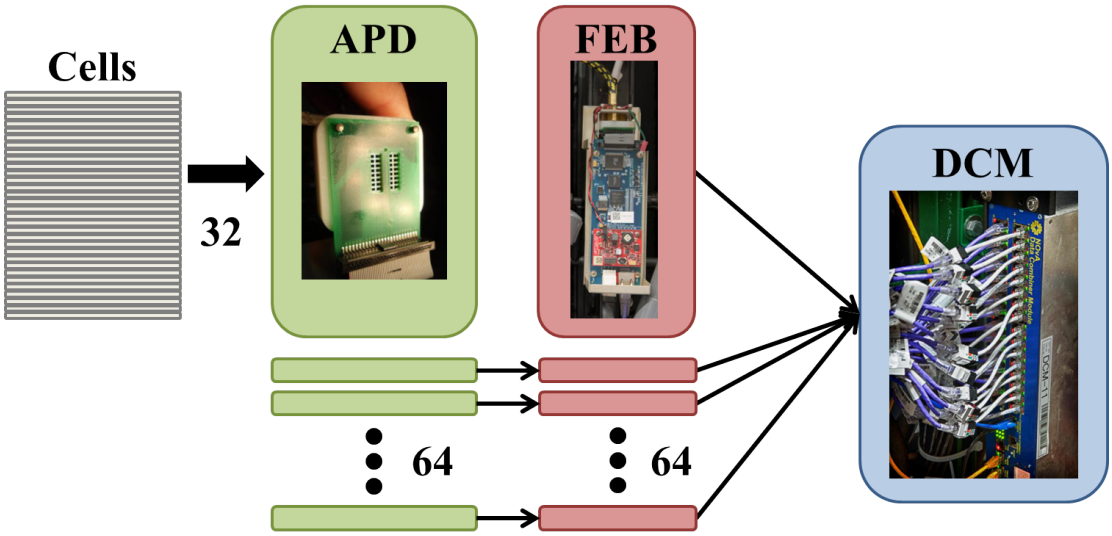


Figure 5.6: A PVC module consists of 32 cells. Each cell is read out by a pixel of an APD. Each APD has 32 pixels. Each APD has a corresponding FEB, which digitizes the signal from the APD. A DCM reads 64 FEBs and sends the information to the buffer node farm.

data. This second set overwrites the first; therefore, the triggering decision code has a limited amount of time to decide if hits should be save to file before the hits are lost. Currently, the software has about 3 seconds to decide. This will be longer in the future. Without the round-robin system, this time would be much shorter. 3 seconds is enough processing time for fairly complicated data-driven trigger algorithms to complete. For the oscillation experiment, this time isn't necessary, since trigger decisions are driven by a beam signal. However, it is crucial for detector abilities such as self-triggering on supernova.

Triggered readout windows are organized by runs and subruns, which are subdivisions of a run. During a run, manually-set detector configurations do not change. A run ends when it has 64 subruns, when the total run duration is 24 hours, or when the detector stops taking data. A subrun ends when it has a duration of 1 hour, the file size is 1 GB, or the detector stops taking data. For the near detector , subruns are often 1 hour; for the far detector, the actual duration is about 3 minutes.

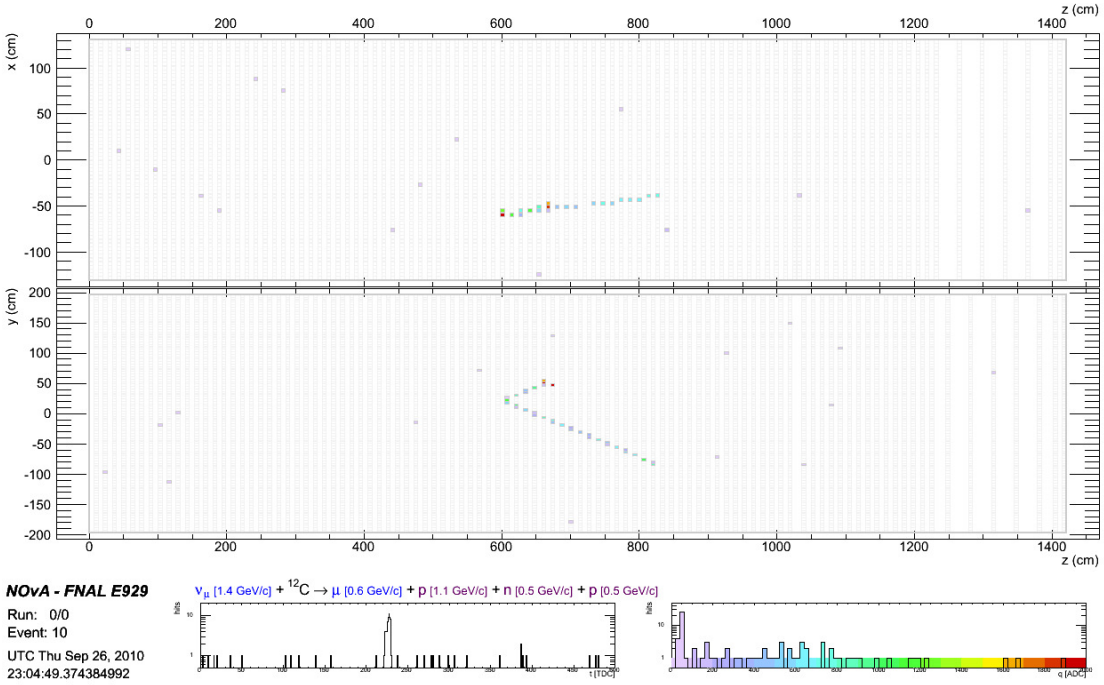


Figure 5.7: NOνA event display of a charged current muon neutrino event from Monte Carlo simulation. The longer track is a muon; the shorter track is a proton. Muons tend to make long, straight tracks in a NOνA detector.

5.1.1 The Event Display

An event display visualizes collected hit information by overlaying hits on a schematic of the detector. A typical event display shows two views. The top view shows a view of the detector as seen from above; the bottom view shows the detector from the side. Each hit is represented in only one view and can be colored by signal intensity (charge) or by time. The example event displays included are colored by charge. Event displays also include two small plots on the bottom of the display. The left plot shows the hit distribution in time across the span of the event. The right plot shows the charge distribution of the hits. See Figures 5.7, 5.8, and 5.9 for three representative event displays. Since the cellular orientation alternates between planes, if a charged particle (like a muon) travels through the detector and excites photons in multiple cells, a three dimensional track can be recreated for the particle. Based on track length, energy depositions and other characteristics, particle identifications and energies can be assigned.

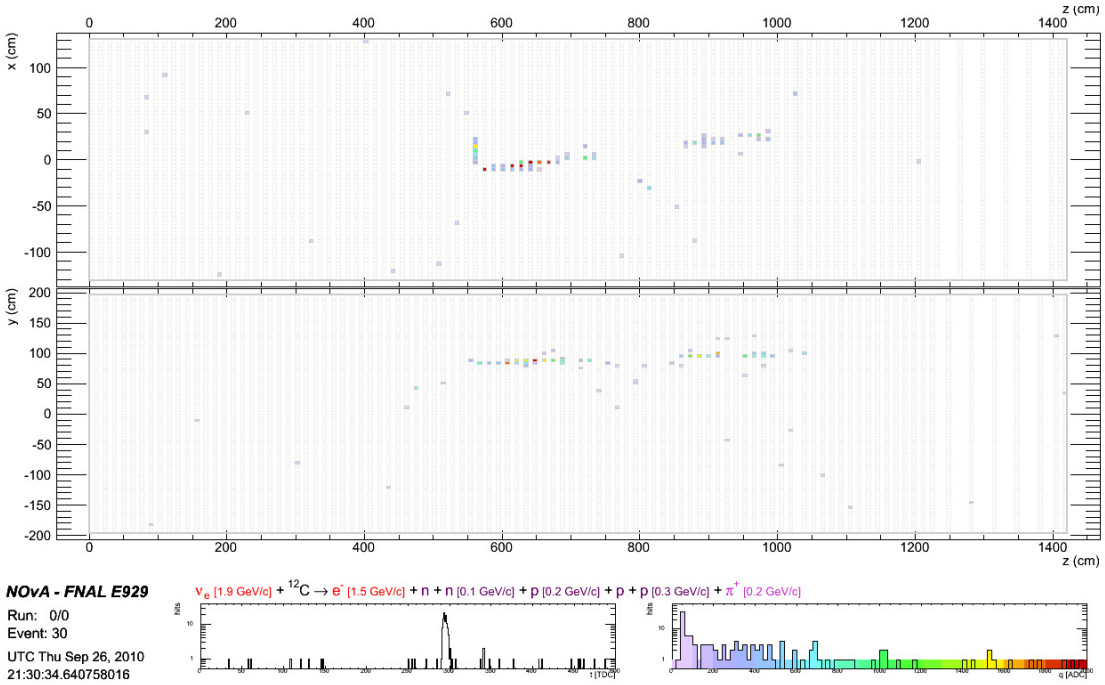


Figure 5.8: NOνA event display of a charged current electron neutrino event from Monte Carlo simulation. The primary activity is from an electron shower. In general, electron showers in a NOνA detector are much more diffuse than a muon track.

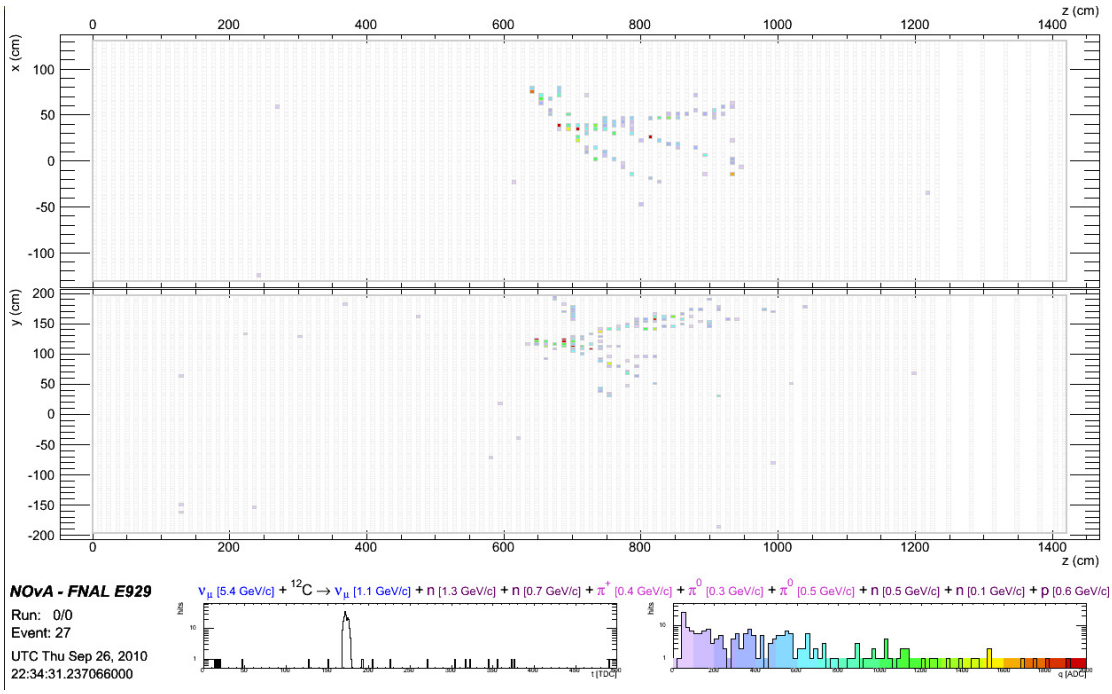


Figure 5.9: NOνA event display of a neutral current neutrino event from Monte Carlo simulation. Most of the activity is from π^0 showers. π^0 showers can look similar to electron showers and are a primary background for the electron neutrino appearance analysis.

Finally, for each potential neutrino interaction, the particle information can indicate the flavor and energy of the parent neutrino. By comparing near detector and far detector data, NOνA will be able to determine not only how many muon neutrinos oscillated to another flavor (seen as a deficit in the number of muon neutrinos) but also how many electron neutrinos appeared from oscillations.

5.2 An Off-Axis Detector

The far detector , as well as the near detector, is placed 14.6 milliradians off-axis from the NuMI beam line, which determines the energy spectrum of neutrinos it detects[5]. NOνA's neutrinos are created primarily from pion and kaon decays. Within the rest frame of the parent particle, the two-body decay[4] happens isotropically. However, in the lab view, the parent particle is moving and this boosts the neutrinos into a cone-shaped distribution. The neutrino energy, E_ν, from the decay of a pion with energy, E_π, with angle θ between the neutrino and pion is given by:

$$E_\nu = \frac{0.43 E_\pi}{1 + \frac{1}{m_\pi^2} E_\pi^2 \theta^2}. \tag{5.1}$$

We can see that if $\theta = 0$, E_ν is linearly proportional to E_π. Also, the highest possible values of E_ν will be found when $\theta = 0$. If $\theta \neq 0$, then the factor of E_π^2 in the denominator competes with the E_π of the numerator. This leads to a relatively flat relationship between E_π and E_ν. Figure 5.10 plots neutrino energy versus pion energies for a few values of θ.

From Figure 5.11, one can see that by being off-axis 14 milliradians, one gains about five times as many events near the 2 GeV energy range compared to on-axis, while significantly limiting the number of events with energy higher than 3 GeV. This is crucial - the oscillation maximum (seen in Figure 5.11 as a dip in the ν_μ energy distribution) is expected to be located at about 1.6 GeV[5]. To increase our ability to measure the oscillation parameters, we want to maximize the number of events we will see near this energy.

[4] π or $K \rightarrow \mu\, \nu_\mu$

[5] For $\Delta m^2_{23} = 2.4 \times 10^{-3}\text{eV}^2$

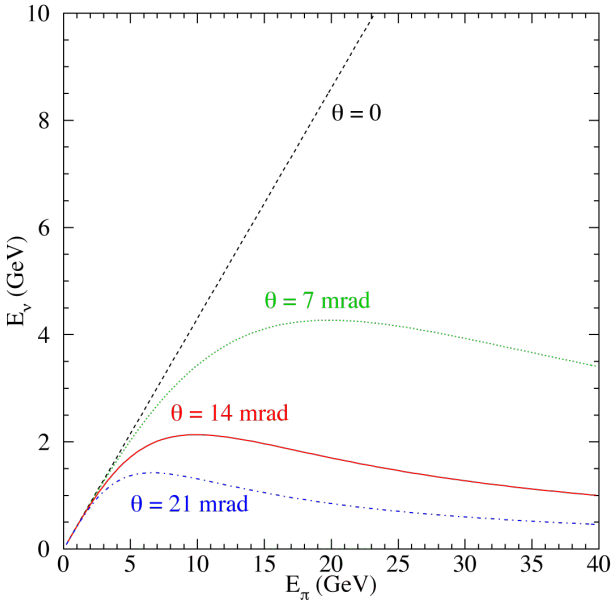


Figure 5.10: A plot of neutrino energy as a function of pion energy for different values of θ, the angle between the neutrino direction and the pion direction. For $\theta = 0$ (on-axis), there is a direct correlation between pion energy and neutrino energy. When $\theta = 14$ mrad (the off-axis location of NOνA), neutrino energy is relatively independent of pion energy. Plot from [5].

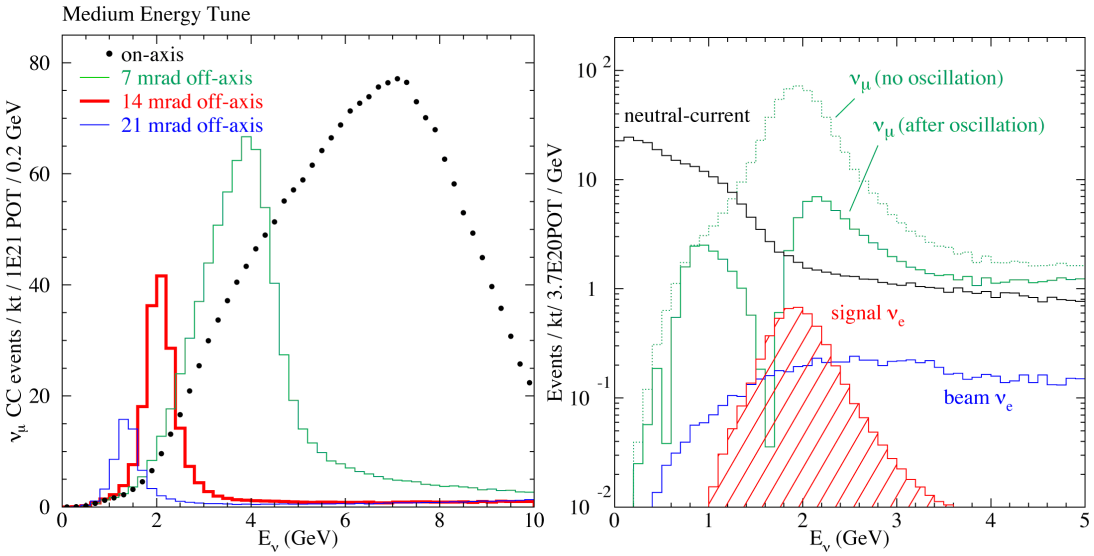


Figure 5.11: Left: The event rate of charged current ν_u interactions without oscillations for Far Detectors at 810 km and varying radial positions. This plot was created using the medium energy NuMI beam configuration[5]. Right: Energy distributions for the ν_μ CC events both with and without oscillation assumed, for ν_e events from the initial beam makeup and from oscillations, and for NC events. These distributions are from simulations and use $\Delta m^2_{32} = 2.5 \times 10^{-3}\text{eV}^2$, $\sin^2(2\theta_{23}) = 1$, and $\sin^2(2\theta_{13}) = 0.10$[5].

Being off-axis also helps with NC background suppression. Since the outgoing neutrino carries away energy that will not be detected, NC events look like lower-energy events than they really are, called the *feed-down effect*. High-energy NC events are the main background to identifying lower-energy electron neutrino CC events. By severely reducing the number of high-energy neutrinos that our detector will see, we will cut down on the number of NC events that can appear as electron neutrino CC events in our signal region. Therefore, by placing the NOνA detector off-axis, we increase our signal by a factor of five, while reducing our background and overall improving our signal/background ratio. Limiting high-energy NC events also helps reduce background for the muon neutrino analysis, by reducing the charged pion background. For more information on NC backgrounds, see the next section.

5.3 Interactions in NOνA

There are two types of charged current (CC) events that we need to distinguish - those of the muon neutrino and those of the electron neutrino. Also, we need to identify neutral current (NC) events because they are a primary background for the oscillation analyses.

Muon neutrino CC events are relatively easy to spot in the NOνA detector; they consist of a long, straight track from the muon, often with some shower activity at the initial vertex. Electron neutrino CC events do not have a long muon track and look like an electromagnetic shower, often accompanied by other shower activity. Differentiation between long tracks and electromagnetic showers is fairly easy. However, a NC event can sometimes look like an electron neutrino CC event. For a NC, no corresponding electron or muon is created. Instead, it simply leaves dispersed shower activity in the detector. Occasionally, a NC event contains a π^0 particle that immediately decays into two photons. These photons then each create electromagnetic showers, much like an electron. Careful analysis is able to differentiate between the shower activity of NC events and CC ν_e events in many cases, notably by looking for two showers displaced from the initial vertex (for NC) instead of one shower beginning at the vertex (ν_e CC). Sometimes a charged pion is created in a NC event. Charged pions can create long tracks in the detector that mimic a muon; this is a background for the ν_μ analysis.

Charged pions can experience hard scatters, which look like sharp kinks in the track. This can help to distinguish them from muon tracks.

The ν_μ disappearance analysis divides its signal more finely than simply a ν_μ CC event. Instead, we attempt to divide the quasielastic (QE) and non-quasielastic (nonQE) populations. The physics underlying each population is different, as is the model used to simulate these events (see Chapter 6 for more information about NOνA simulation). This can lead to different systematic errors and uncertainties in each population; differentiating the populations helps to understand these errors. Perhaps more importantly, though, the neutrino energy resolution of the two populations is different. By dividing populations, one is able to take advantage of the better energy resolutions of the QE events while not losing valuable statistics.

QE interactions typically consist of a muon track and a small to zero amount of other activity at the interaction vertex. This other activity can be a visible proton track, but often, the proton does not have enough energy to produce a reconstructed track. NonQE interactions, on the other hand, have more activity at the interaction vertex that is not part of a muon track. Characterizing the vertex activity allows the analysis to define a primarily QE population.

Because the far detector is near the surface of the earth, it sees a large number of cosmic rays. These cosmic rays are primarily muon tracks, which could be a problem for the ν_μ analysis. However, we know the time of the beam spills, so a timing cut requiring the activity to be within the beam spill reduces the cosmic ray flux. The muon track direction gives another important way to eliminate these cosmic rays. Cosmic rays are the result of high energy interactions in the atmosphere so they primarily travel from the top to the bottom of the detector, while muons created from beam neutrino interactions are relatively aligned with the beam direction, from the front of the detector toward the back. For more details on cosmic ray rejection in this analysis, see Section 10.3.1.

Bad detector conditions can cause cosmic ray background to be more of a problem. If a part of the detector doesn't read out properly, it can make cosmic rays traveling through the entire detector appear to be well-contained. The detector can fail to read out hits normally if certain DCMs have a wrong time or if the DCMs became overwhelmed and failed to report all of the hits received to the buffer node. It is important to have data quality cuts that remove problematic detector conditions.

Chapter 6

Simulation

To understand what physics interactions look like in a detector, physicists create simulation files. These are often called *Monte Carlo* files, in reference to the famous Monte Carlo Casino in Monaco. Monte Carlo files use random sampling to mimic the complexity and randomness that occurs when processes happen in a detector. In NOνA, we simulate neutrino interactions and what happens to the daughter particles as they travel through the detector. These signals are turned into simulated detector signals that imitate what we expect to see in real data-taking. The benefit of simulation files, as opposed to data from the detector, is that we know the "truth" information from our simulated files. Truth information includes things like the flavor of the neutrino which interacted, what energy each daughter particle started with, and how much energy each particle lost in each section of the detector. This truth information can be used to understand phenomenon such as how the visible energy left by a hadronic shower corresponds to the true energy of the hadronic shower.

Creating Monte Carlo files for the NOνA experiment is a multi-step process. A simulation of the neutrino beam is created. Then neutrinos from the beam are simulated interacting initially with the NOνA detector. Next, the daughter particles of the neutrino interaction must be propagated through the detector. The computed energy depositions are turned into simulated photon signals. These signals are then interpreted as they would have been if read by the DAQ system. Finally, one has a collection of detector hits[1] (just as our detector data in its initial stages is just a collection of detector

[1]These are also known as *raw digits* and are discussed further in Section 7.1.

hits), along with truth information about the underlying, simulated physics interactions.

To simulate the neutrino beam, NOνA uses FLUKA and FLUGG[35][36]. The NuMI beamline is simulated, including the target and horns. The output of the simulation is a flux of simulated neutrinos, with parent information preserved. These flux files are handed to the next step of the neutrino simulation.

To generate the initial neutrino interactions, NOνA has multiple simulators that it uses, depending on the type of physics to study. The most commonly used is GENIE (currently using v2.8.0i)[37], which is used to generate all of the neutrino physics interactions that NOνA studies. To study cosmic ray interactions, NOνA uses CRY[38]. Finally, for specialized studies, one can generate single particles at user-defined energies, directions, and locations in the detector.

Next, whatever the initial source of particles from a generator, the particles propagate through the detector using Geant4 (currently using v4.9.6.p03(c), with physics list QGSP_BERT_HP)[39][40]. This step includes the NOνA detector geometry and material composition. Geant4 steps particles through the detector environment, while simulating the energy loss and particle interactions. The primary outputs from this step are called FLShits, which are the computed energy depositions, and the hierarchy of particle daughters, mothers, and what physics process created each daughter. For more information on the Geant4 simulation step, see Section 6.1.

Next, a NOνA-specific module, called PhotonTransport, takes FLShits and determines how many photons would be seen at the APD readout. It takes into account the fiber capture probability, signal attenuation in the fiber, edge effects at the ends of the cells, and Poisson statistics.

Lastly, another NOνA-specific module, ReadoutSim, takes the photon signals as input. In channels with physics-signal photons, it overlays the photon signal with background noise fluctuations. In channels without physics-signals, it can create noise hits by sampling a histogram of noise hits from the physical detector. This total electronic signal is then parsed through the algorithm our DAQ system uses to decide if a channel had a hit or not. These hits are the starting point that NOνA uses to begin its reconstruction, discussed in Chapter 7.

6.1 Geant4 Simulation Details

To simulate the propagation of particles through the NOνA detector, we use Geant4. This analysis uses version 4.9.6.p03(c) with physics list QGSP_BERT_HP. This configuration uses the quark gluon string (QGS) model[41] for high energy interactions of protons, neutrons, pions, kaons and nuclei. Here, high energy means greater than 20 GeV. The Bertini cascade (BERT) model[42][43] is used to simulate primary protons, neutrons, pions and kaons with energies below 10 GeV. The low energy parametrized (LEP) model is used for intermediate energies and for low energy particles not modeled by BERT. Using BERT instead of LEP for all lower energy hadronic modeling results in better agreement to data. In general, it creates more secondary protons and neutrons than LEP. To simulate the de-excitation of nuclei, the G4Precompound (P) model is used. To simulate proton and neutron inelastic interactions, the Axen-Wellisch parametrization of cross sections is used. To simulate pion inelastic interactions, a table of Barashenkov cross sections is used. A data-driven high precision neutron (NeutronHP) simulation is used to model the interactions of neutrons with energies less than 20 MeV down to thermal energies. For more information about the Geant4 simulation, see the Geant website[44].

6.2 Accessing Truth Information

Having Monte Carlo files that simulate data is nice; being able to access the truth information contained within the files is nicer. To this end, a service called BackTracker was developed. BackTracker connects the various types of truth information properly and efficiently, allowing the user to get answers to common physics questions. BackTracker functions are documented in the code viewer[45].

There are few subtleties that the user should be aware of when asking questions of the BackTracker. Sometimes the NOνA implementation of Geant4 decides processes are too complicated. Instead of giving individual particle ids to each daughter particle in, say, an electron shower, it gives the whole chain the mother particle id. The PDG value, however, matches that of the depositing particle. That means that one can look at many FLShits from the same particle id and they might have different PDG values. This isn't a bug; rather, it is considered a feature.

Naively, one would think that each detector hit could be tied in a perfect way to the FLShits which created it. However, because NOνA simulates the noise on the channel and turns the complicated, simulated electronic trace into detector hits, it isn't quite that simple. Instead, we try to relate the time of the photon signals to the time of the hit. This is almost always correct, but not always. Technically, every photon signal that came before the detector hit could in some small way contribute to its creation. The answer depends on the interplay between the random noise, the previous history of the channel, and the cuts we use to determine that a detector hit has occurred. For most users, the simple cut on time should be sufficient. For more information, see [46].

Chapter 7

Event Reconstruction

Reconstructing an interaction event in data or simulation is a multi-step process. Each step has an associated *module*, with each module run in sequence. A module is a C++ class which performs a specific task on the input file. In order, the required modules to perform the ν_μ CC contained analysis presented in this thesis are:

1. DAQ2RawDigit (data only) - converts raw data to a convenient format for further processing
2. CalHit - writes hit information in a form more convenient for analysis
3. Slicer - groups hits in time and space
4. KalmanTrack and KalmanTrackMerge - takes slice hits and makes tracks
5. CosmicTrack - an alternative method that takes slice hits and makes tracks
6. ReMId - produces a metric for how muon-like a track is
7. NumuEnergy - determines the neutrino energy of a slice under a ν_μ CC assumption
8. QePId - produces a metric for how quasi-elastic the interaction was
9. CAFMaker - writes a summary file of physics variables

Each of these processes will be explained in a following section.

7.1 DAQ2RawDigit

This module is only run over data files from the NOνA detectors and is unnecessary for simulation files. For data from a detector, the information recorded is initially saved in a format that is a convenient and compact format for the data acquisition system. To make the data accessible to the data analysis software, one must run the module DAQ2RawDigit over an input .raw file to get out a file with a .root extension. The primary information contained in the output file is a collection of raw digits for each event. Raw digits are hits which know only their location in electronic coordinates, a pulse-height determined by each cell's analog-to-digital converter (ADC), and a time determined by each cell's time-to-digital converter (TDC). Simulation files are created in the form of raw digits and this extra step of translation is not required.

7.2 CalHit

This is a module that takes in raw digits and creates cell hits. It removes hits that come from channels deemed bad by the BadChannel service, software that monitors detector performance. A channel is considered bad for a period of detector running if it has either too many or too few hits. Too many hits can indicate that the channel has excessive noise; too few hits can indicate that the electronics system is not properly reading out. For each channel, a hit rate in Hz is determined. A good channel has a hit rate between 3.2 Hz and 3,200 Hz. For more details, see the BadChannel technical note[47].

CalHit associates each hit with a detector plane and cell within the plane. It uses calibration information to convert pulse height measurements to the energy deposited in the cell, as described in Chapter 8. This process also determines the time of the hit more precisely and in units of ns.

7.3 Slicer

Data from NOνA is packaged in 500 μs readout windows[1]. This data is for the entire detector. However, physic interactions within the detector take place over much shorter time periods and are often confined to specific regions of the detector. Slicer is a module that groups hits within a readout window in time and space. The clusters of hits that are roughly contiguous in time and space are called *slices*. If Slicer worked perfectly, each slice would correlate to one physics interaction (neutrino interaction or cosmic ray). Slicer also identifies cell hits from noise and puts them into a single junk slice.

The beam spill (the length of time that we expect neutrino interactions from the beam to occur within) is only 10 μs. The 500 μs readout window allows the synchronization of the detector to the beam spill to drift slightly without losing data. Also, the time outside the beam spill allows a background determination to be made. Finally, the longer livetime allows one to determine the data quality close to the time the data signal occurred.

Physics interactions in the detector can be 10's of meters long and a few meters wide. A particle moving at the speed of light takes 10's of ns to traverse the detector. Propagating the light signal through different lengths of wavelength shifting fiber and electronics paths can add another 100 ns to the time duration of a typical interaction in the detector. The Slicer algorithm clusters contiguous hits together in time and space using the variations expected. Slices have an average duration of 130 ns, which is much finer grained than either the 10 μs beam spill or the 500 μs readout window.

To group cell hits into clusters, Slicer uses a score function to determine the distance between cell hits in space and time. The algorithm attempts to group hits that could be causally linked by particles traveling at the speed of light. Any point that does not get clustered is defined as a noise hit. For more information, see Appendix C and the technical note[48].

[1]This is a configurable parameter - for many data driven triggers, we record only a 50 μs readout window.

7.4 KalmanTrack and KalmanTrackMerge

KalmanTrack is a module which takes in clusters of hits from Slicer and groups the cell hits from either the vertical or horizontal cells into 2D tracks. KalmanTrackMerge takes the 2D tracks from each view and attempts to merge them into 3D tracks. The information from the 3D tracks and unmatched 2D tracks are written out to the file. This information includes the position of the most upstream (lowest z) part of the track, called the start position, the position of the most downstream (highest z) part of the track, called the end position, the all the hits associated with each track, and the trajectory points determined to exist along the path of the track. The tracker is based on Kalman filtering, described here[49][50]. A track is defined as a mostly continuous string of cells that would result from the energy deposited by a single, non-showering particle. This tracker uses linear, piecewise segments to describe a trajectory and assumes the scatter is consistent with a muon. For more information on what particle tracks look like in the NOνA detector, see Section 5.1. This analysis focuses primarily on 3D tracks, since 2D tracks have less information and are often more poorly reconstructed than 3D tracks.

To make 2D tracks, KalmanTrack starts with all the hits in the cluster corresponding to a single view (either XZ or YZ). Seeds are created, where a seed is a pair of hits that are less than 3 planes apart within that view. These seeds then form test tracks that are propagated to the next plane using the current value of the track position and slope. Any hits found in the next plane consistent with being on the test track are added to the track, the slope and intercept of the track are updated, and the process continues to the next plane. See Figure 7.1 for a diagram of this process. Hits are determined to be consistent with the track by requiring that the $\Delta\chi^2$ of the track fit after including this hit is less than 8.0. Tracks are allowed to continue propagation until they cross 3 consecutive planes in a view without adding a hit. If the projected track indicates that it would encounter a bad channel while crossing the plane, that plane is not counted as a missed plane. This allows the tracker to successfully cross known gaps of bad channels within the detector. After a complete track is found, the procedure is redone, starting at the other end of the found track and without trying to add new hits, to determine if any of the previously added hits should now be rejected.

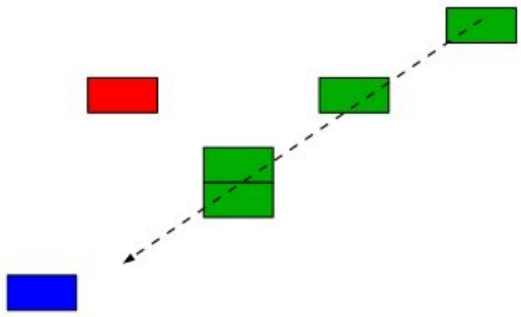


Figure 7.1: Diagram of the creation of 2D Kalman tracks. Each colored rectangle is a hit. The green hits are already included on the track. The track estimate of position and slope is shown as the dotted line. The red hit is inconsistent with the track, while the blue hit is consistent and will be added. This process is then repeated for the next plane.

A test track is promoted to being a valid track if it contains at least 4 track hits. After a valid track has been identified, the method uses the remaining seeds not included on an existing track and continues the procedure of creating more 2D tracks. The optimization process for the definition of a good track was primarily based on maximizing the efficiency of reconstructing long tracks typical of muons, while limiting the number of junk tracks created.

After all the 2D tracks have been made in each view independently, KalmanTrackMerge looks at the consistency between 2D tracks of different views and tries to create one 3D track from two 2D tracks. The score that determines the consistency is defined as:

$$score = \frac{|z_{low\ of\ xz\ track} - z_{low\ of\ yz\ track}| + |z_{high\ of\ xz\ track} - z_{high\ of\ yz\ track}|}{length\ of\ overlap\ in\ z}$$

Lowest scores are matched first. In some cases, a long 2D track in one view is matched to multiple 2D tracks from the other view. This is necessary when a scatter breaks a track into two pieces in one view. See Figure 7.2 for an example. Track merging continues until no more 2D tracks with overlapping z positions exist. The merging algorithm balances the desire to have tracks not broken into segments with the ability to reconstruct multiple tracks emerging from the same vertex. For more information on the tracking algorithms, see this internal NOνA document[51] or this thesis by Raddatz[52].

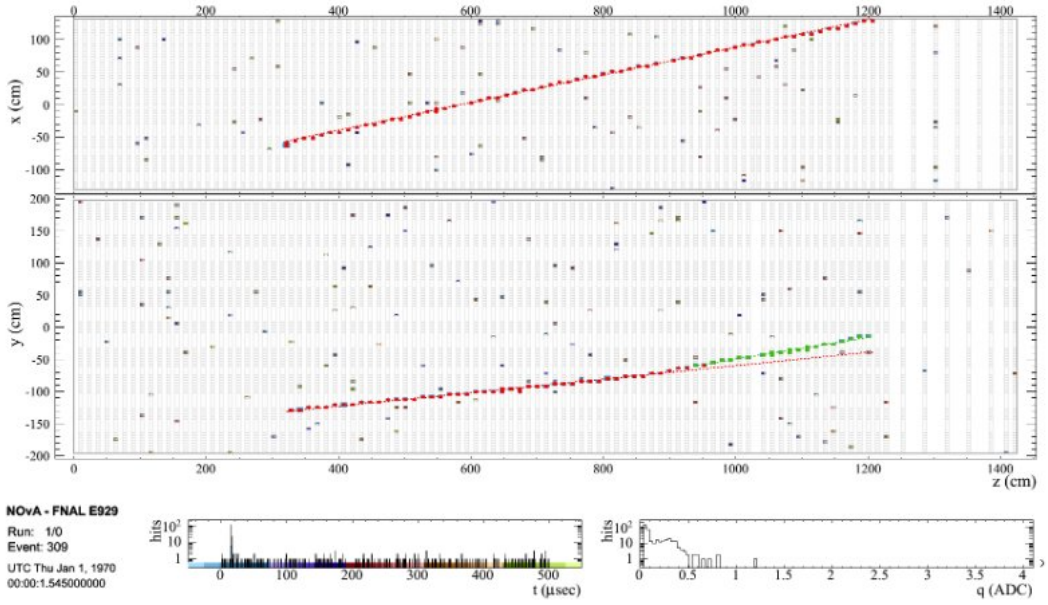


Figure 7.2: Event display showing Kalman tracks. Here, the muon scatter is large enough that KalmanTrack created 2 2D tracks in the yz view. If KalmanTrackMerge could only match individual 2D tracks to each other, the red 3D KalmanTrack represents the best match that could be accomplished. However, KalmanTrackMerge would instead merge the green 2D track with the red 3D track, resulting in a single reconstructed muon track.

7.5 CosmicTrack

CosmicTrack is an alternative method for reconstructing tracks from slice hits. It is a less complicated algorithm than KalmanTrack and is optimized for finding high energy vertical tracks. It makes, at most, one 3D Cosmic track for each slice. This track is precisely linear. CosmicTrack is a quick algorithm to run and works well for cosmic rays, which are dominated by slices with one long muon. The analysis presented in this thesis uses CosmicTrack only for a few cosmic rejection cuts, discussed in Section 10.3.

7.6 ReMId

ReMId (Reconstructed Muon Identification) is an algorithm that attempts to identify a muon track. A ν_μ charged current interaction is then identified if it has a 3D muon track, as determined by ReMId.

ReMId uses a k-Nearest Neighbors (kNN) algorithm to make its determination. For more information about kNN algorithms, see [53]. The input variables are: log-likelihoods which use the $\frac{dE}{dx}$ of the track, log-likelihoods based on the scattering observed in the track, the track length, and the fraction of planes used to created the $\frac{dE}{dx}$ log-likelihood.

To use the $\frac{dE}{dx}$ information of a track, ReMId determines the log-likelihood (LL^E) that a particle of type j created the track, where:

$$LL^E_j = \frac{1}{N_{plane}} \sum_i P^j_i$$

For LL^E, the $\frac{dE}{dx}$ of the track is measured at plane i and the probability, P^j_i, of a particle of type j to have the measured $\frac{dE}{dx}$ at this distance from the end of the track is calculated from a histogram created using simulated events under the assumption of particle type j. The number of planes for which $\frac{dE}{dx}$ is measured on the track is N_{plane} and is used to normalize the LL. The difference in the values of the LL under the pion and the muon assumptions is taken to form the final $\frac{dE}{dx}$ LL variable input into the ReMId kNN.

To use the scattering information of a track, the LL^{scat} is defined as:

$$LL_j^{scat} = \frac{1}{N_{scat}} \sum_i P_i^j$$

where P_i^j is the probability that a particle of type j has the measured scatter at a distance of i from the end of the track, as determined by a histogram created using simulated events under the assumption of particle type j. N_{plane} is used to normalize the LL. The measured scatter s is defined as:

$$s = \frac{\theta^2}{d}$$

where θ is the measured scattering angle for that trajectory point and d is the distance from the last measured scatter.

The kNN returns a value between 0 and 1, 1 being muon-like and 0 being background-like. The analysis uses a cut of 0.7 to determine if the track is deemed muonic. This cut was optimized to create the smallest statistical error on the oscillation parameters $\sin^2\theta_{23}$ and Δm^2_{32}. Figure 7.3 demonstrates the ability of this metric to differentiate between ν_μ charged current interactions and other types of neutrino interactions. For more information, see the technical note[54] or the thesis by Raddatz[52].

7.7 NumuEnergy

NumuEnergy attempts to determine the energy of the initial neutrino, under the assumption that the interaction was a ν_μ charged current interaction. It treats the 3D KalmanTrack with the highest ReMId value as the muon. All other energy in the slice is treated as hadronic energy. The muon energy is determined using a linear spline fit of the reconstructed track length of the muon to true muon energy. The hadronic energy is determined using a linear spline fit of visible hadronic energy to true neutrino energy minus reconstructed muon energy. The fits used are created using simulation files and truth information. For full details of this energy estimator, see Chapter 9.

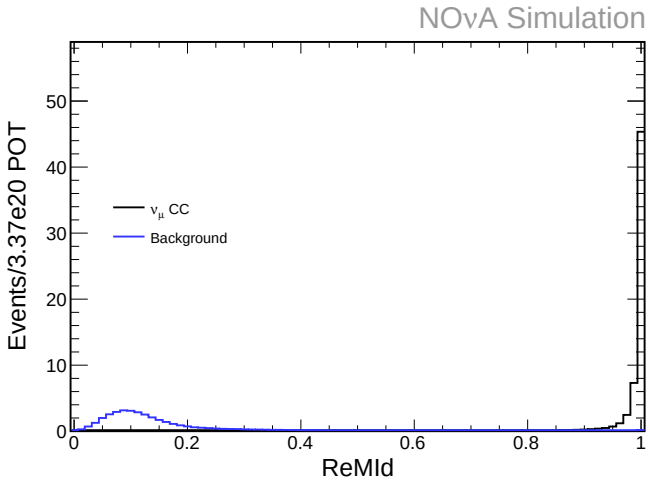


Figure 7.3: Plot of the ReMId value for the 3D Kalman track with the highest ReMId value in a slice. This is for the far detector. The distribution for ν_μ charged current events is drawn in black; the distribution for other types of neutrino interactions is drawn in blue. Candidates with values greater than 0.7 are used by the analysis. The simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $\Delta m^2_{32} = 2.4 \times 10^{-3}$ eV^2. Each population is scaled to the expected levels for 3.52×10^{20} POT. All cuts listed in Tables 10.4, 10.5, and 10.6 are applied.

7.8 QePId

QePId (Quasielastic Particle Identification) is an algorithm that attempts to identify an interaction as either quasielastic (QE) or non-quasielastic (non-QE), given the assumption that it is a ν_μ charged current interaction. In general, it reports a PID value between 0 and 1, where 1 is likely to be a QE interaction and 0 is unlikely. It only tries to identify an interaction as QE if there are one or two Kalman tracks (of which only one is required to be 3D) in the slice; if there are more tracks (either 2D or 3D), it is unlikely to be QE and is given a PID score of -1.

For slices with one or two tracks, QePId uses the variables described below as inputs to a k-Nearest Neighbors (kNN) algorithm to create the PID value. It relies on values determined by NumuEnergy and described in Chapter 9. The input variables are: the ratio of the energy in the slice not associated with a track to the sum of the energy in the slice on tracks, the relative difference in neutrino energies using the reconstructed QE neutrino energy for this analysis and the reconstructed QE neutrino energy using the angle formula, an energy difference z-test variable that takes into account the uncertainties associated with the calculation of QE neutrino energy when using the angle formula, and, in the case of two tracks, the $\frac{dE}{dx}$ ratio between the muonic and nonmuonic tracks. The analysis uses a cut of 0.4 for the one-track sample and 0.45 for the two-track sample to determine that a slice is QE. Figure 7.4 demonstrates the ability of this metric to differentiate between ν_μ QE charged current interactions and other types of neutrino interactions in the one track sample. Figure 7.5 is for the two track sample. For more information, see the technical note[55] or the thesis by Raddatz[52].

7.9 CAFMaker

It is convenient to have small files that are extremely quick to process that contain only summary information required for a final analysis. These files are called CAF files. They are easily readable by ROOT[56], ROOT macros, and code written to work in the cafe environment. CAF files do not contain information at the hit level, so can not be used for all analysis tasks or for creating event displays. For more information about CAF files and cafe, please see [57][58].

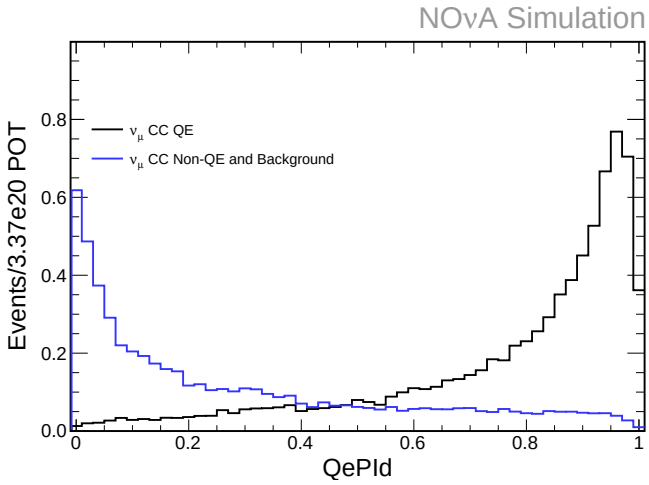


Figure 7.4: Plot of the QePId value for slices in the one track sample. This is for the far detector. The distribution for ν_μ quasielastic charged current events is drawn in black; the distribution for other types of neutrino interactions, including ν_μ non-quasielastic charged current events, is drawn in blue. Candidates with values greater than 0.4 are part of the QE population; slices with values less than 0.4 are considered part of the nonQE population. The simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $\Delta m^2_{32} = 2.4 \times 10^{-3}$ eV^2. Each population is scaled to the expected levels for 3.52×10^{20} POT. All cuts listed in Tables 10.4, 10.5, and 10.6 are applied, along with the ReMId cut listed in Table 10.8.

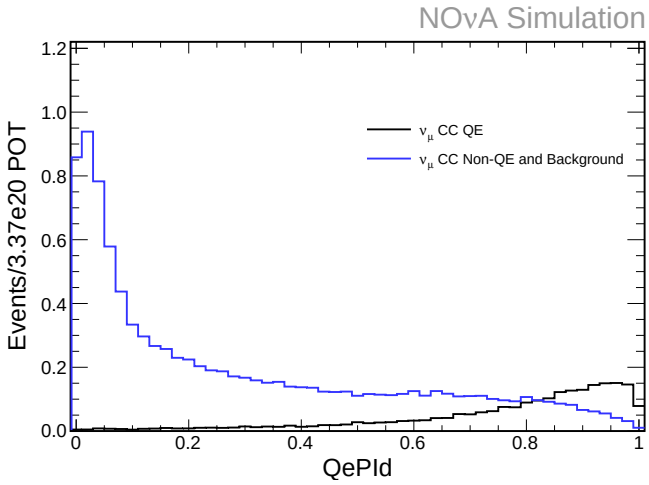


Figure 7.5: Plot of the QePId value for slices in the two track sample. This is for the far detector. The distribution for ν_μ quasielastic charged current events is drawn in black; the distribution for other types of neutrino interactions, including ν_μ non-quasielastic charged current events, is drawn in blue. Candidates with values greater than 0.45 are part of the QE population; slices with values less than 0.45 are considered part of the nonQE population. The simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $\Delta m^2_{32} = 2.4 \times 10^{-3}$ eV^2. Each population is scaled to the expected levels for 3.52×10^{20} POT. All cuts listed in Tables 10.4, 10.5, and 10.6 are applied, along with the ReMId cut listed in Table 10.8.

Chapter 8

Calibration

Calibration is necessary to understand how to translate a hit in a specific cell at a specific depth into an energy deposition. Calibration has sequential stages that allow this translation:

1. Wavelength Shifting Fiber Attenuation Correction
2. Absolute Energy Correction

Each step will be covered in its own section.

8.1 Attenuation Correction

Attenuation calibration corrects cell-to-cell differences. To do this, a collection of hits is formed. Hits are required to be on well-reconstructed muon cosmic rays. If possible, the tri-cell criteria is also required; that is, that a hit is only used if the cells above and below it in the plane were also on the track. This allows for the most precise knowledge of path length in the central cell. In cases where there isn't enough statistics to make the tri-cell criterion (such as having bad channels next to a good channel), one can instead require that the cells in the planes before and after the hit were also hit. In the worst cases, such as corner cells, there is no requirement and instead an average path length over all the directions that the track can have is used.

Given a collection of hits in a cell with known path length, one can make populations that have relatively the same path length. Next, a 2D plot of distance down the length

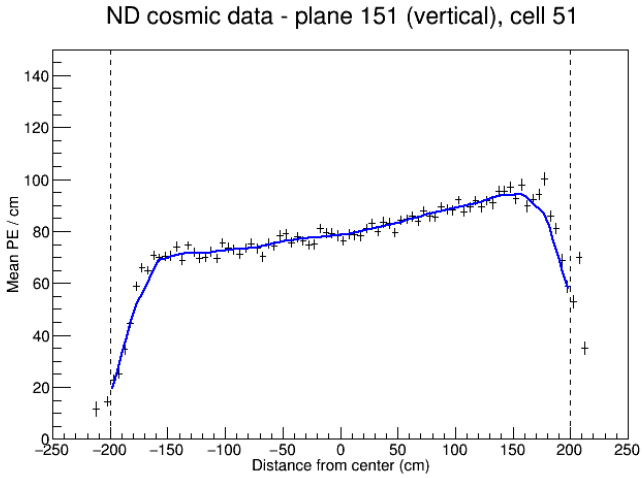


Figure 8.1: The final attenuation calibration fit for a channel in the physical near detector (plane 151, a vertical plane, and cell 51). This fit, drawn as a blue line, is considered good. The horizontal axis is distance along the cell depth, measured in cm away from the central depth. The vertical axis is average photo-electrons per path length in cm. Dashed vertical lines indicate the ends of the cell.

of the cell vs. (ADC/cm) is constructed. The profile of this plot is taken and fit to determine the attenuation correction for this cell.

For most of the length of the fiber, the shape of the attenuation correction is approximated to be that of two exponentials. Near the top and bottom of the cells, the data shows a strong roll-off effect due to the end structure of the cells. This is not seen as strongly in the Monte Carlo, since the Monte Carlo doesn't accurately simulate structure of the ends of the module properly. Also, some channels show strong deviations from the double exponential fit that are not related to roll-off effects. In some cases, this was due to extremely noisy channels located nearby in space, allowing fake cosmic tracks to be produced. However, most of the time noisy behavior is not in evidence and the source of the discrepancy is unknown. An interpolation fit is applied on top of the double exponential fit. If the interpolation fit is more than 15% different from the original fit, the channel is marked as uncalibrated. Figure 8.1 displays the attenuation fit for a typically good channel in the near detector; Figure 8.2 displays the same for a typically good channel in the far detector. For more information about the attenuation calibration, see the technical note[59].

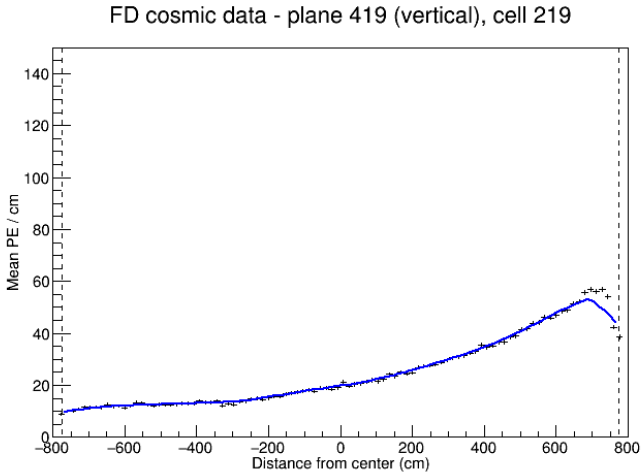


Figure 8.2: The final attenuation calibration fit for a channel in the physical far detector (plane 419, a vertical plane, and cell 219). This fit, drawn as a blue line, is considered good. The horizontal axis is distance along the cell depth, measured in cm away from the central depth. The vertical axis is average photo-electrons per path length in cm. Dashed vertical lines indicate the ends of the cell.

8.2 Absolute Energy Correction

After cell-to-cell differences are removed, one needs to set an absolute scale for the energy. This allows one to translate a hit's energy deposition into a reconstructed GeV. To set the absolute energy scale, a population of muon cosmic tracks that stop within the detector is used. Since muon energy loss in the detector is well understood, one can start at the end of the track (where the muon energy is zero) and work backwards with the Bethe-Bloch formula to determine the $\frac{dE}{dx}$ along the track. The MIP region of the track, where $\frac{dE}{dx}$ is approximately constant, is found. Figure 8.3 displays the MIP region for stopping muon tracks in the far detector. A scale factor is then determined that translates the energy units given by the attenuation calibration into GeV. For more information about the absolute energy calibration, see the technical note[60].

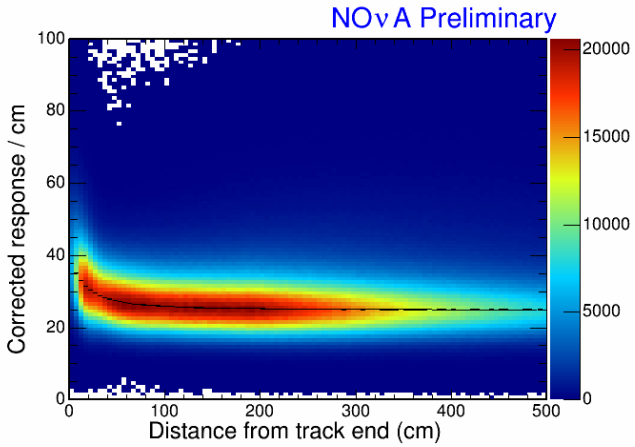


Figure 8.3: Plot used to create absolute energy calibration for the physical far detector. Each energy in the histogram is a tri-cell hit on a cosmic ray muon which stopped inside the detector. The horizontal axis is the distance of the hit from the end of the track in cm. The vertical axis is the attenuation-corrected energy per path length in cm. The black fit points show the mean of the fit to the distribution for each horizontal bin. Values between 100 and 200 cm from the end of the track are considered the MIP region and used for the absolute energy calibration.

Chapter 9

Neutrino Energy Reconstruction

The software module NumuEnergy attempts to determine the energy of the initial neutrino, under the assumption that the interaction was a ν_μ charged current interaction. It treats the 3D KalmanTrack with the highest ReMId value as the muon. All other energy in the slice is treated as hadronic energy. No attempt is made to remove noise hits from consideration as hadronic energy. Energy fits are created using unoscillated simulation files and truth information.

This module creates three primary estimates of the neutrino energy; two use the muon track length and visible hadronic energy to create a neutrino energy, while the other uses the muon track length and muon angle relative to the beam direction under the assumption of a quasielastic interaction to compute the neutrino energy. The first two energies are called TrkQEE (Track Quasielastic Energy) and TrkNonQEE (Track Non-quasielastic Energy). The third energy is referred to as AngleQEE (Angle Quasielastic Energy).

The fitting techniques used for NumuEnergy depend on generating a large statistics simulation sample and creating various 2D histograms. These 2D histograms are then turned into graphs of points that can be fit with a series of straight line segments. The fit values are then turned into functions that allow one to translate from the measured quantity to the desired reconstructed quantity. For detailed information about how the 2D histograms are turned into graphs, please see Appendix D. For information about the functional form of the spline fits used, see Appendix E. For information about the populations used to create the fit, see Appendix F.

For all three reconstructed neutrino energies, the muon energy is determined from the track length of the 3D Kalman track with the highest ReMId value. This track length is converted into a reconstructed muon energy using a spline fit. The spline fit was determined using simulation files to compared the reconstructed muon track length to the true muon energy. This is described further in Section 9.1 for the far detector and Section 9.4 for the near detector.

Although the muon energy does not depend on any assumption as to whether the event is considered QE or non-QE, the hadronic energy does. Visible hadronic energy is defined as the sum of the energy deposited by all hits not on the muon track as well as the energy on the muon track in the vertex region that exceeds minimum ionizing values. For information about the hadronic model used by the simulation, see Chapter 6. For more information about determining hadronic energy contamination on the muon track, see the technical note[61]. No attempt is made to remove noise hits contained within the slice from consideration as hadronic energy. The visible hadronic energy is then converted into a reconstructed hadronic energy using a spline fit. The spline fit was determined using simulation files to compare the visible hadronic energy to the true neutrino energy minus the reconstructed muon energy. This is described further in Section 9.2 for the far detector and Section 9.5 for the near detector.

For TrkQEE and TrkNonQEE, the final neutrino energy is a sum of the reconstructed muon and hadronic energies. This is described further in Section 9.3 for the far detector and Section 9.6 for the near detector. For AngleQEE, the neutrino energy is calculated using a formula described in Section 9.7.

9.1 Muon Energy in the Far Detector

Muons created in neutrino interactions travel through the NOνA detectors, losing energy until they stop. These muons are primarily minimum ionizing particles. See Figure 9.1 for a plot of muon momentum to average energy loss. Minimum ionizing particles have a roughly flat average energy loss as they travel through a medium. This means that over a wide range of momenta, encountered as the muon slows down, the muon still loses the same amount of energy per path length in the detector. At the lowest values of momentum, the muon loses much more energy per path length in the detector.

When one considers the final range of such particles, it is a mostly linear relationship between track length and initial muon energy. For very low energy muons, however, this relationship deviates from linear. Low energy muons proportionally have less track length in the minimum ionizing regime and more track length in lower momentum regimes. To account for these deviations from linear, we use multiple linear splines. A more proper treatment would use an curve that integrates the relationship shown in Figure 9.1. See Chapter 16 for a discussion about future improvements to muon energy reconstruction.

NumuEnergy uses a four spline fit is to describe the relationship between muon track length to true muon energy, determined using simulation files. For more information on the spline fit functional form, see Appendix E. Figure 9.2 is the 2D histogram that relates reconstructed muon track length in cm to true muon energy in GeV. For more information on the population used to fill this histogram, see Appendix F. This is then transformed into a graph, using the method described in Appendix D. Figures 9.3 and 9.4 show the graph overlaid with the 2D histogram. Figure 9.4 has a logarithmic color scale.

I chose four splines because it markedly improved the χ^2/NDF. The three spline fit had a $\chi^2/NDF = 3368.8 \ / \ 141 = 23.9$, while the four spline fit had a $\chi^2/NDF = 1169.98 \ / \ 139 = 8.42$. Further improvement could be made by using either more splines or a better functional form. See Chapter 16 for more discussion of improvements. See Table 9.1 for the final fit parameters. Figure 9.5 shows the fit splines overlaid on the graph points used to find the fit. Figure 9.6 shows the fit splines overlaid on the original 2D histogram. Figure 9.7 is the same as Figure 9.6, except that it has a logarithmic color axis.

Having obtained a fit which allows translation from reconstructed track length to a reconstructed muon energy, one can now consider the muon energy resolutions. All of the following plots were produced using a statistically independent sample from the one used to find the fit. Figure 9.8 is a 2D histogram which compares reconstructed muon energy to true muon energy. Then one can consider the relative energy resolution:

$$\frac{Reconstructed\,Energy - True\,Energy}{True\,Energy}. \tag{9.1}$$

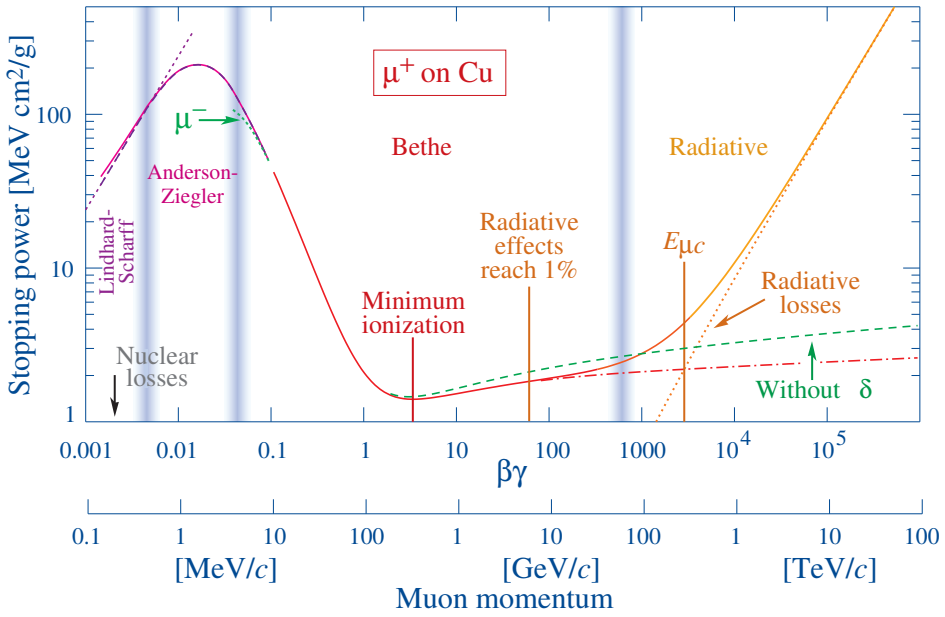


Figure 9.1: This figure was created by the Particle Data Group[1]. The logarithmic horizontal axis displays the muon momentum. The stopping power for a muon, otherwise known as the average loss of energy per path length ($\langle -dE/dx \rangle$), is shown on the logarithmic vertical axis. A muon in the NOνA detectors typically has a momentum in the minimum ionizing range, although as the muon stops, the momentum lowers. When a particle is in the minimum ionizing regime, it looses approximately a constant amount of energy per amount of detector traveled.

Parameter	**Value**	**Units**
Offset	0.1503	GeV
Slope 1	0.001910	GeV/cm
Slope 2	0.001987	GeV/cm
Slope 3	0.002039	GeV/cm
Slope 4	0.002159	GeV/cm
Stitch 1	334	cm
Stitch 2	539	cm
Stitch 3	1064	cm

Table 9.1: Fit parameters used to determine the far detector muon reconstructed energy.

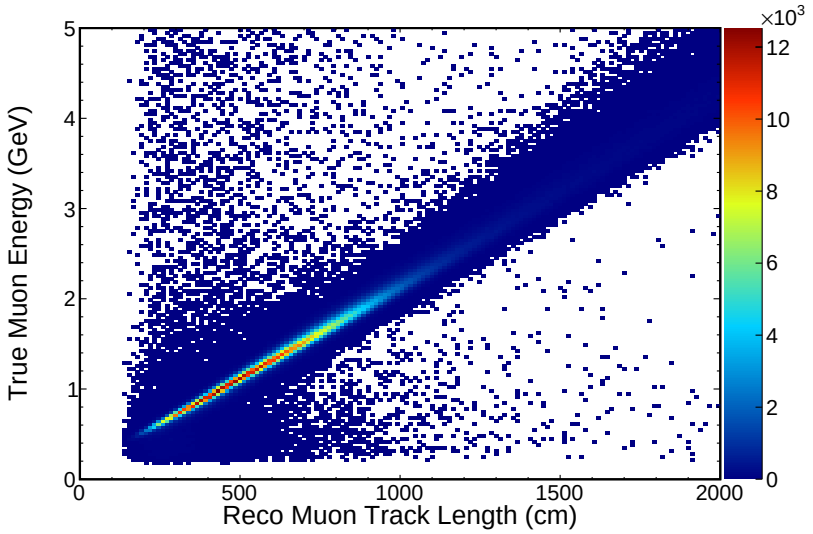


Figure 9.2: Reconstructed muon track length in cm vs. true muon energy in GeV. The color axis is arbitrary number of interactions. This histogram was created using simulated events in the far detector.

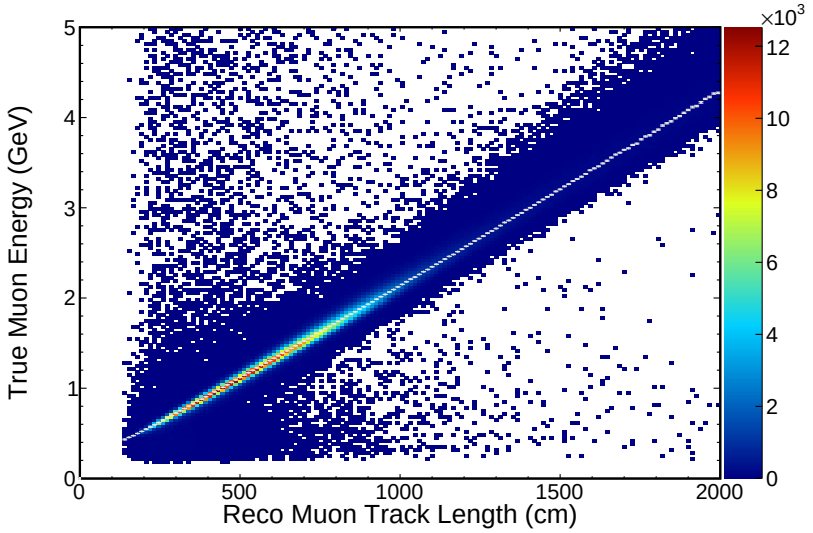


Figure 9.3: Graph points overlaid on 2D histogram of reconstructed muon track length in cm vs. true muon energy in GeV. The color axis is arbitrary number of interactions. This plot was created using simulated events in the far detector.

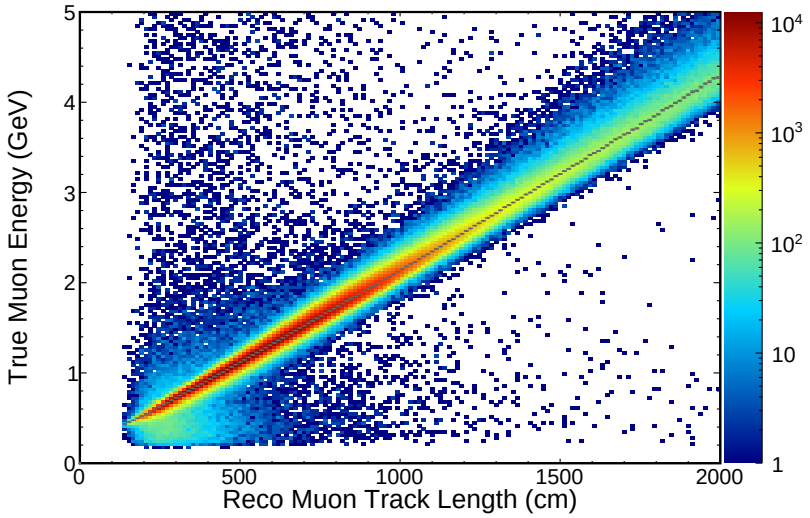


Figure 9.4: Graph points overlaid on 2D histogram of reconstructed muon track length in cm vs. true muon energy in GeV. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the far detector.

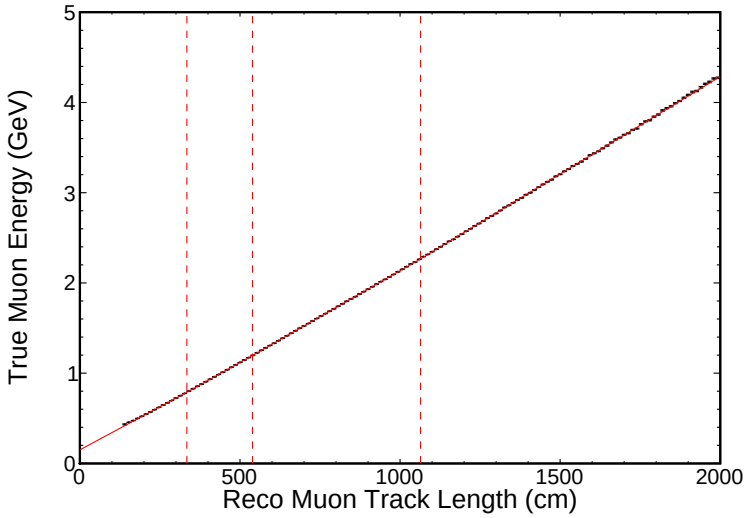


Figure 9.5: Best fit line overlaid on graph points used to make the fit. The fit relates reconstructed muon track length in cm to true muon energy in GeV. The dashed vertical lines indicate the stitch locations of the splines. This plot was created using simulated events in the far detector.

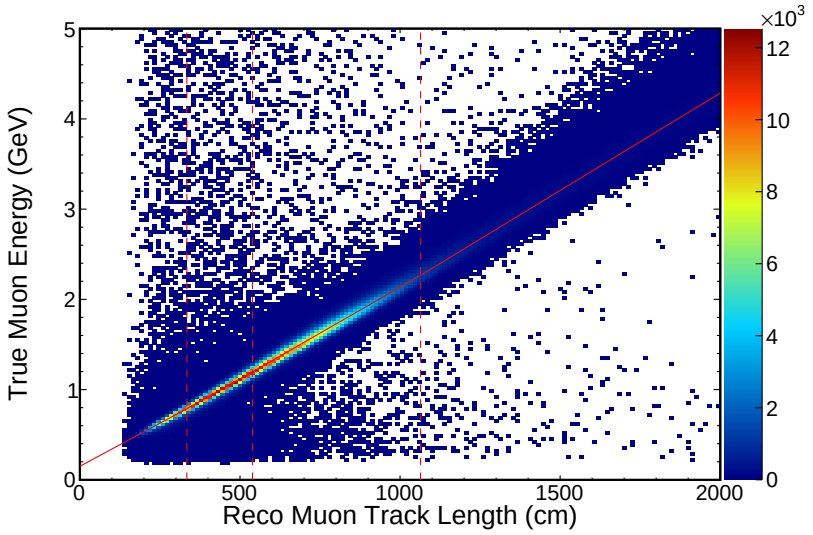


Figure 9.6: Best fit line overlaid on original 2D histogram. The fit relates reconstructed muon track length in cm to true muon energy in GeV. The dashed vertical lines indicate the stitch locations of the splines. The color axis is arbitrary number of interactions. This plot was created using simulated events in the far detector.

Figure 9.9 plots muon track length in cm vs. the relative energy resolution. One can also consider a 1D histogram of the relative energy resolution, Figure 9.10.

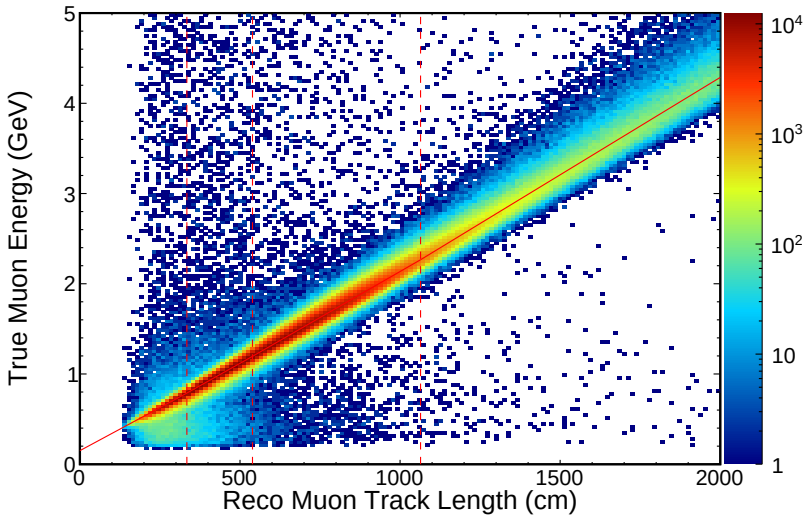


Figure 9.7: Best fit line overlaid on original 2D histogram with a logarithmic color axis. The fit relates reconstructed muon track length in cm to true muon energy in GeV. The dashed vertical lines indicate the stitch locations of the splines. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the far detector.

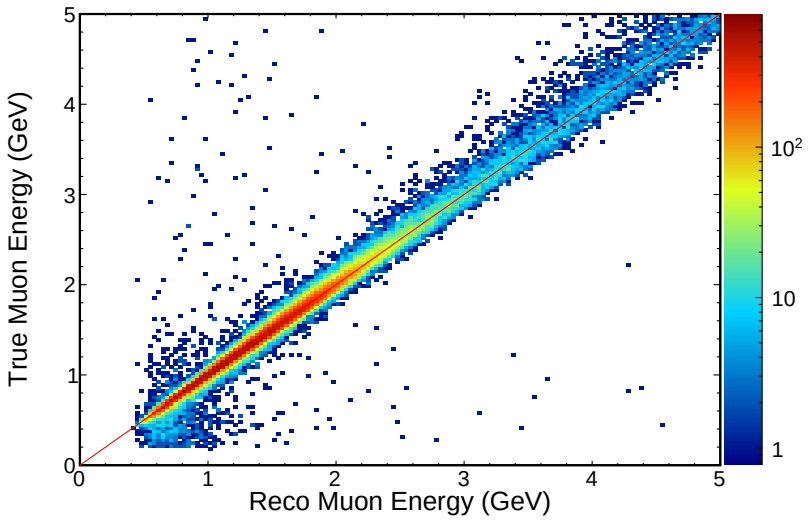


Figure 9.8: Histogram of reconstructed muon energy in GeV vs. true muon energy in GeV. Red line is 45 degrees - if we did a perfect job, everything would lay on this line. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the far detector.

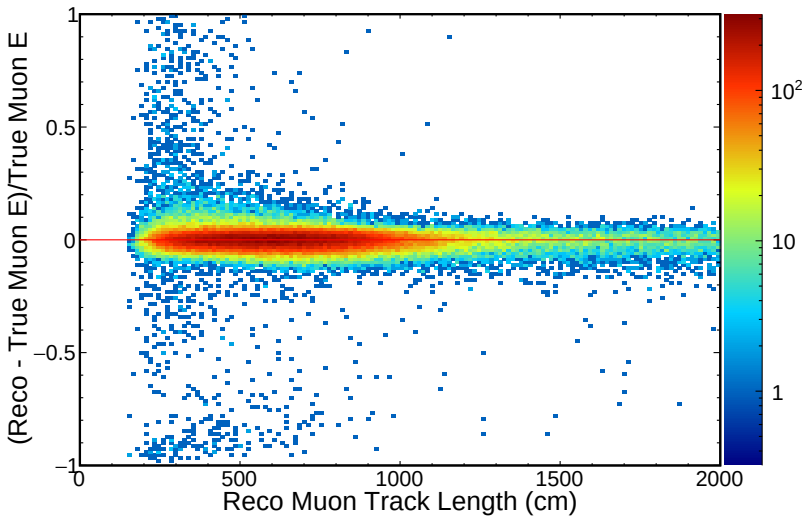


Figure 9.9: Histogram of reconstructed muon track length in cm vs. the relative energy resolution. Red line is flat at zero - if we did a perfect job, everything would lay on this line. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the far detector.

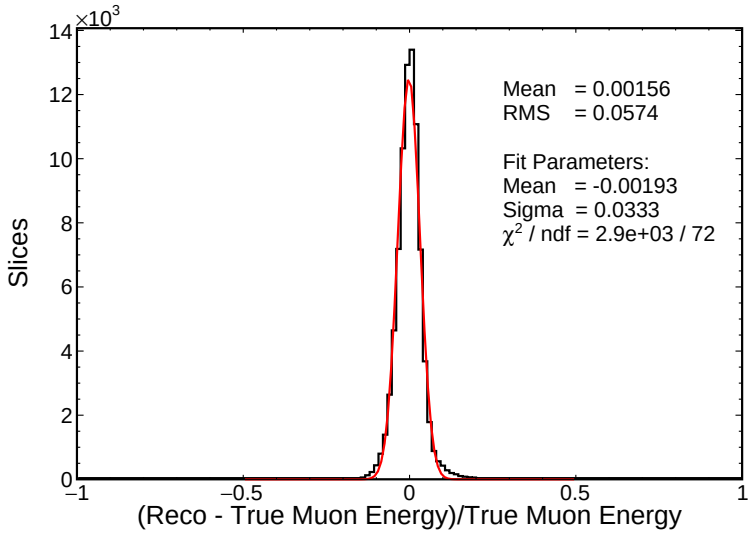


Figure 9.10: Relative energy resolution for reconstructed muon energy. Red line is a Gaussian fit to the peak. This plot only includes events with true neutrino energy less than 5 GeV. This plot was created using simulated events in the far detector.

9.2 Hadronic Energy in the Far Detector

After one has the far detector muon spline fit, one moves on to considering far detector hadronic energy spline fits. Visible hadronic energy is defined as the sum of the energy deposited by all hits not on the muon track as well as the energy on the muon track in the vertex region that exceeds minimum ionizing values. For information about the hadronic model used by the simulation, see Chapter 6. For more information about determining hadronic energy contamination on the muon track, see the technical note[61]. No attempt is made to remove noise hits contained within the slice from consideration as hadronic energy.

Both the QE and non-QE populations in the far detector were fit with a four spline fit. For more information on the spline fit functional form, see Appendix E. For both populations, variable binning was employed for the underlying 2D histograms of visible hadronic energy in GeV vs. true neutrino energy minus reconstructed muon energy in GeV. For more information on the population used to fill these histograms, see Appendix F. The 2D histograms are then transformed into a graph, using the method described in Appendix D.

9.2.1 QE Events

The results for the QE population will be presented first. Figure 9.11 is the 2D histogram that relates visible hadronic energy in GeV to true neutrino energy minus reconstructed muon energy in GeV. This is then transformed into a graph, using the method described in Appendix D. Figures 9.12 and 9.13 show the graph to be fit overlaid with the 2D histogram. Figure 9.13 has a logarithmic color scale.

I looked at the effect of using various numbers of splines in the fit. The χ^2/NDF values for each fit is reported in Table 9.2. Visually, only the four spline fit captured most of the structure. The four spline fit was chosen for the final fit; in the future, using more splines could be investigated. See Table 9.3 for the final fit parameters. Figure 9.14 shows the fit line overlaid on the graph points used to find the fit. Figure 9.15 shows the fit line overlaid on the original 2D histogram. Figure 9.16 is the same as Figure 9.15, except that it has a logarithmic color axis.

Having obtained a fit which allows translation from visible hadronic energy to a

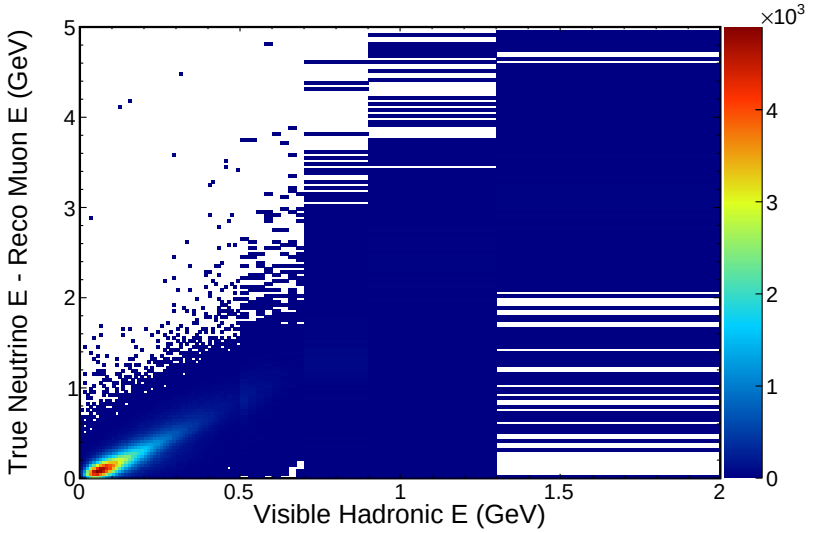


Figure 9.11: Visible hadronic energy in GeV vs. true neutrino energy minus reconstructed muon energy in GeV. This plot is for the QE population. The horizontal axis has variable binning. The color axis is arbitrary number of interactions. This plot was created using simulated events in the far detector.

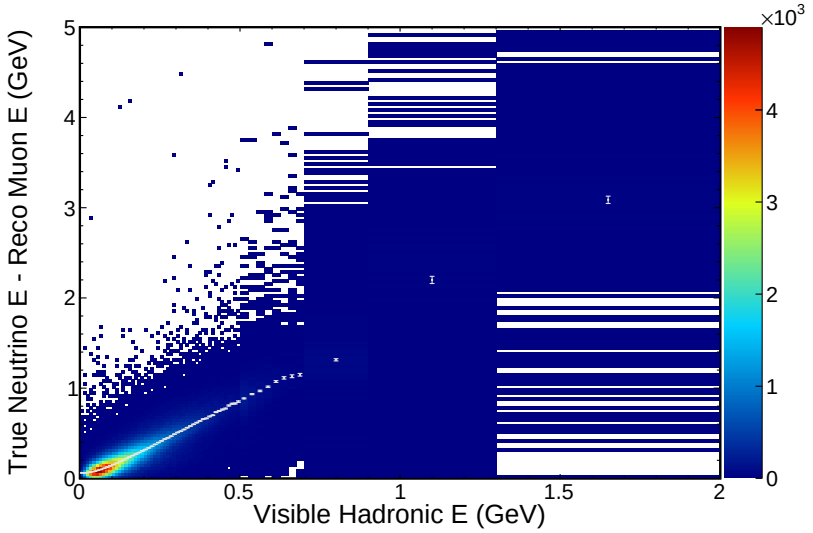


Figure 9.12: Graph points overlaid on 2D histogram of visible hadronic energy in GeV vs. true neutrino energy minus reconstructed muon energy in GeV. This plot is for the QE population. The horizontal axis has variable binning. The color axis is arbitrary number of interactions. This plot was created using simulated events in the far detector.

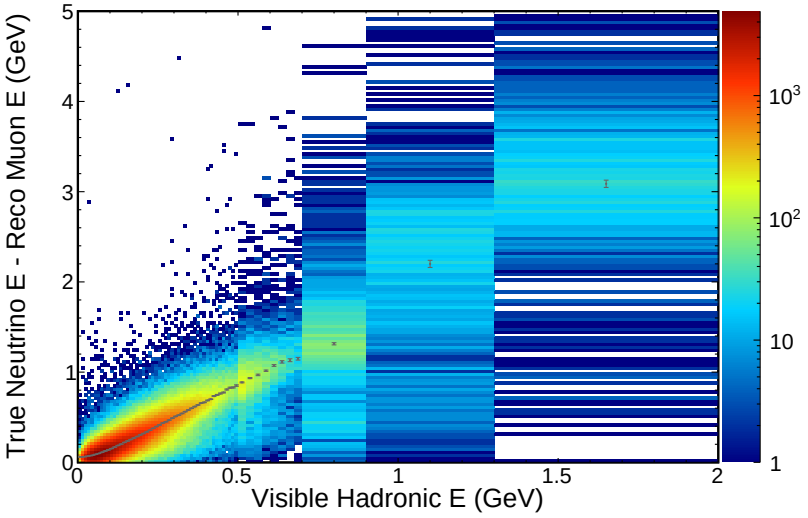


Figure 9.13: Graph points overlaid on 2D histogram of visible hadronic energy in GeV vs. true neutrino energy minus reconstructed muon energy in GeV. This plot is for the QE population. The horizontal axis has variable binning. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the far detector.

No. of Splines	χ^2/NDF
One	10,600 / 75 = 141
Two	1,340 / 73 = 18.3
Three	440 / 71 = 6.19
Four	298 / 69 = 4.31

Table 9.2: χ^2/NDF values for differing numbers of splines used to fit the far detector QE hadronic reconstructed energy.

Parameter	Value	Units
Offset	0.0515	GeV
Slope 1	0.621	unitless
Slope 2	1.47	unitless
Slope 3	1.81	unitless
Slope 4	2.06	unitless
Stitch 1	0.0597	GeV
Stitch 2	0.139	GeV
Stitch 3	0.800	GeV

Table 9.3: Fit parameters used to determine the far detector QE hadronic reconstructed energy.

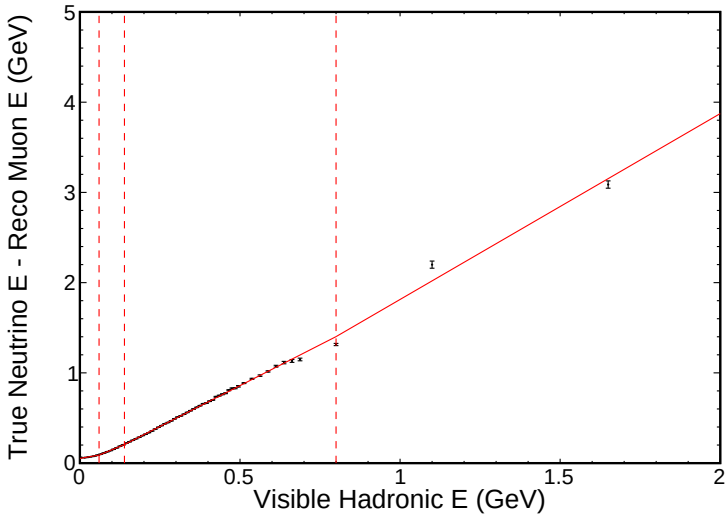


Figure 9.14: Best fit line overlaid on graph points used to make the fit. The fit relates visible hadronic energy in GeV to true neutrino energy minus reconstructed muon energy in GeV. The dashed vertical lines indicate the stitch locations of the splines. This plot is for the QE population. The horizontal axis has variable binning. This plot was created using simulated events in the far detector.

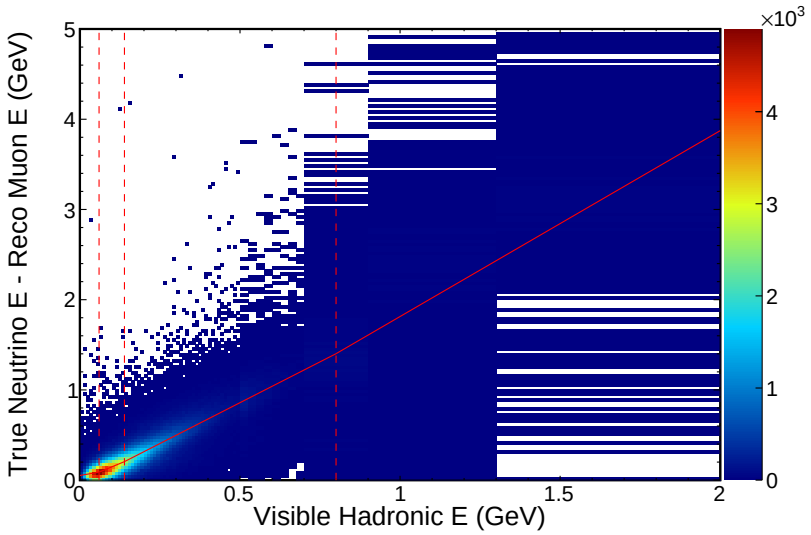


Figure 9.15: Best fit line overlaid on original 2D histogram. The fit relates visible hadronic energy in GeV to true neutrino energy minus reconstructed muon energy in GeV. The dashed vertical lines indicate the stitch locations of the splines. This plot is for the QE population. The horizontal axis has variable binning. The color axis is arbitrary number of interactions. This plot was created using simulated events in the far detector.

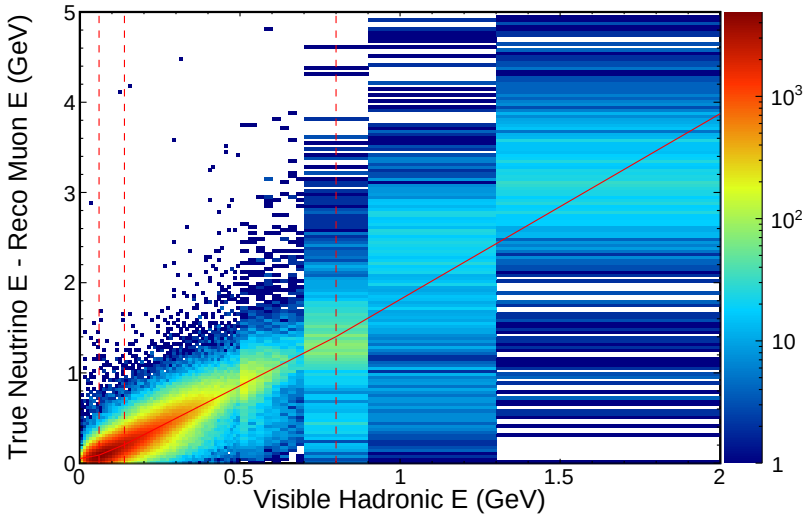


Figure 9.16: Best fit line overlaid on original 2D histogram. The fit relates visible hadronic energy in GeV to true neutrino energy minus reconstructed muon energy in GeV. The dashed vertical lines indicate the stitch locations of the splines. This plot is for the QE population. The horizontal axis has variable binning. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the far detector.

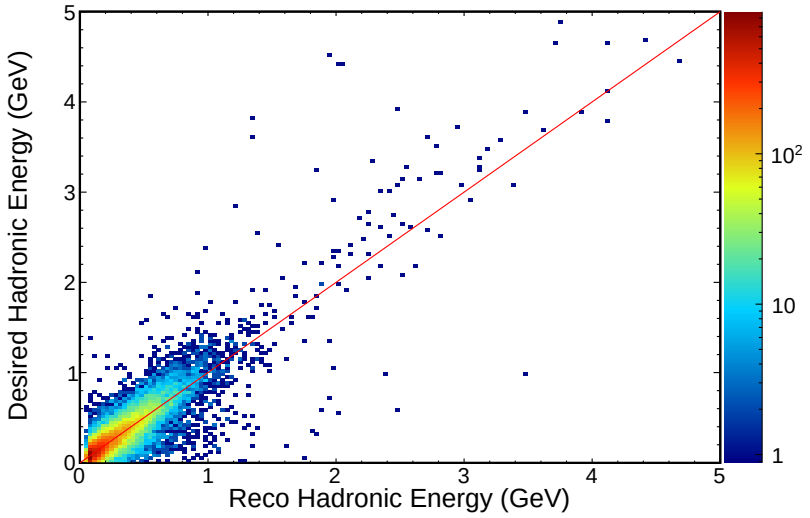


Figure 9.17: Histogram of reconstructed hadronic energy in GeV vs. true neutrino energy minus reconstructed muon energy in GeV, notated as "desired hadronic energy." Red line is 45 degrees - if we did a perfect job, everything would lay on this line. This plot is for the QE population. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the far detector.

reconstructed hadronic energy, one can now consider the hadronic energy resolutions. All of the following plots were produced using a statistically independent sample from the one used to find the fit. Figure 9.17 is a 2D histogram which compares reconstructed hadronic energy to desired hadronic energy, defined as true neutrino energy minus reconstructed muon energy. Figure 9.18 plots reconstructed hadronic energy vs. the relative energy resolution. One can also consider a 1D histogram of the relative energy resolution, Figure 9.19.

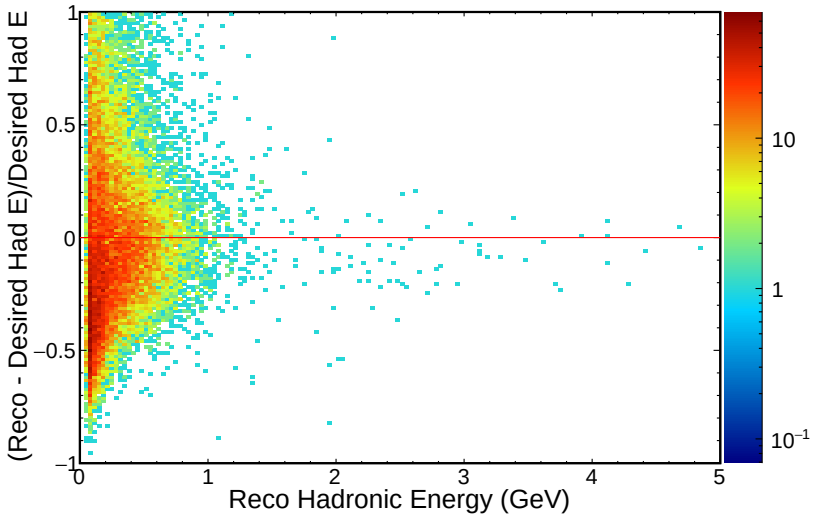


Figure 9.18: Histogram of visible hadronic energy in GeV vs. the relative energy resolution. The desired hadronic energy is defined as true neutrino energy minus reconstructed muon energy in GeV. Red line is flat at zero - if we did a perfect job, everything would lay on this line. This plot is for the QE population. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the far detector.

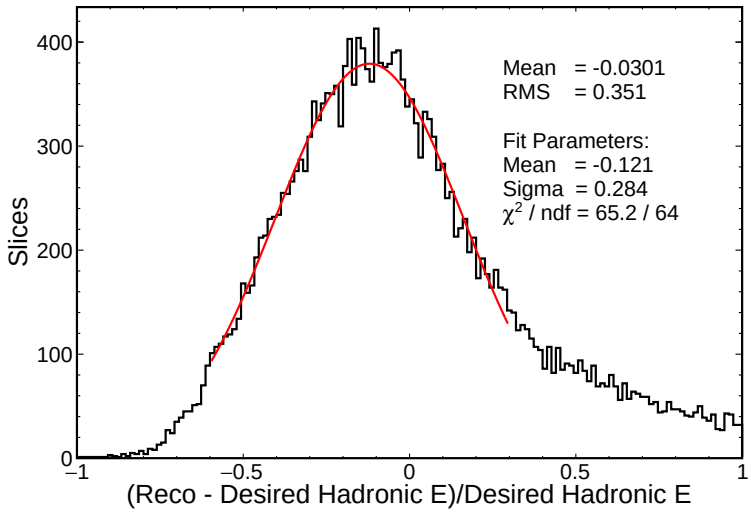


Figure 9.19: Relative energy resolution for reconstructed hadronic energy. The desired hadronic energy is defined as true neutrino energy minus reconstructed muon energy in GeV. This plot is for the QE population. This plot only includes events with true neutrino energy less than 5 GeV. This plot was created using simulated events in the far detector. Red line is a Gaussian fit to the peak.

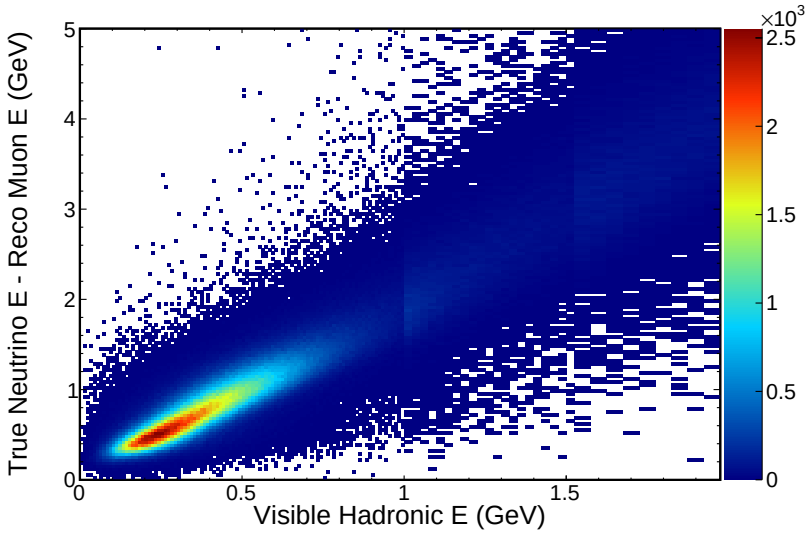


Figure 9.20: Visible hadronic energy in GeV vs. true neutrino energy minus reconstructed muon energy in GeV. This plot is for the non-QE population. The horizontal axis has variable binning. The color axis is arbitrary number of interactions. This plot was created using simulated events in the far detector.

9.2.2 Non-QE Events

For the non-QE population, Figure 9.20 is the 2D histogram that relates visible hadronic energy in GeV to true neutrino energy minus reconstructed muon energy in GeV. For more information on the population used to fill these histograms, see Appendix F. This is then transformed into a graph, using the method described in Appendix D. Figures 9.21 and 9.22 show the graph to be fit overlaid with the 2D histogram. Figure 9.22 has a logarithmic color scale.

For this population,the four spline fit did a good job visually. It had a lower χ^2/NDF, which was 818.403 / 121 = 6.76, than fits using less splines did. See Table 9.4 for the final fit parameters. In the future, using more splines could be investigated. Figure 9.23 shows the fit line overlaid on the graph points used to find the fit. Figure 9.24 shows the fit line overlaid on the original 2D histogram. Figure 9.25 is the same as Figure 9.24, except that it has a logarithmic color axis.

Having now obtained a fit which allows translation from visible hadronic energy

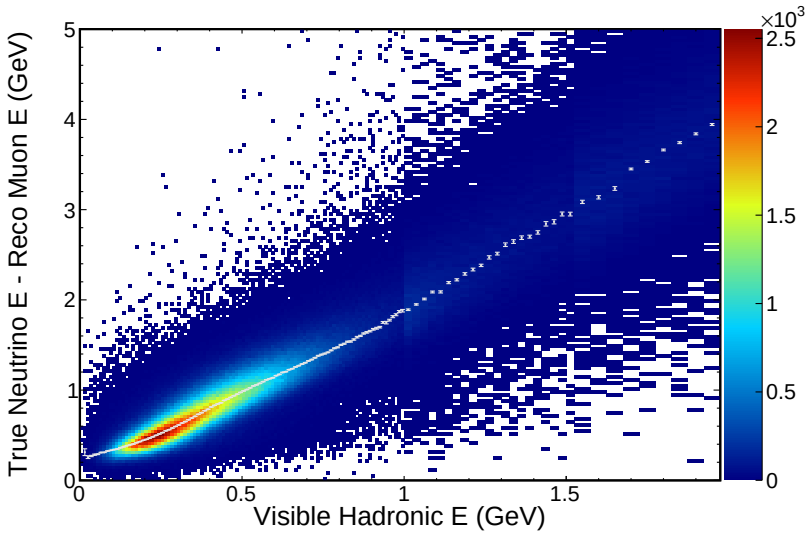


Figure 9.21: Graph points overlaid on 2D histogram of visible hadronic energy in GeV vs. true neutrino energy minus reconstructed muon energy in GeV. This plot is for the non-QE population. The horizontal axis has variable binning. The color axis is arbitrary number of interactions. This plot was created using simulated events in the far detector.

Parameter	**Value**	**Units**
Offset	0.241	GeV
Slope 1	0.94	unitless
Slope 2	1.21	unitless
Slope 3	1.78	unitless
Slope 4	2.20	unitless
Stitch 1	0.132	GeV
Stitch 2	0.223	GeV
Stitch 3	0.964	GeV

Table 9.4: Fit parameters used to determine the far detector non-QE hadronic reconstructed energy.

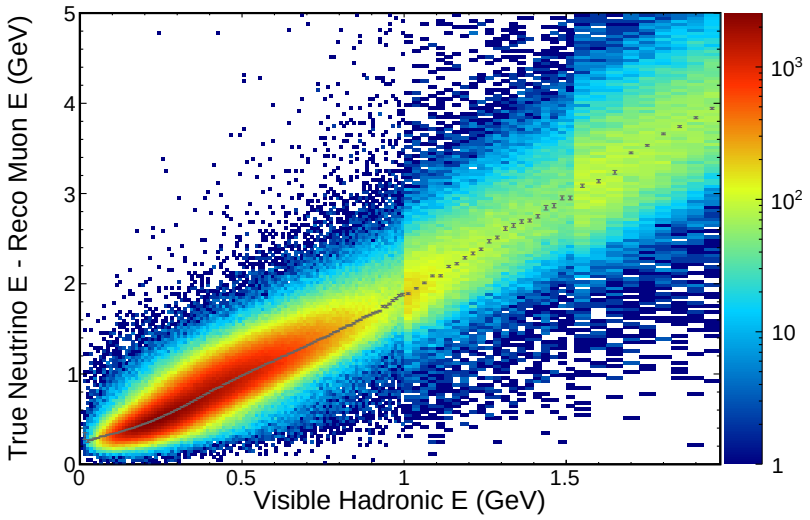


Figure 9.22: Graph points overlaid on 2D histogram of visible hadronic energy in GeV vs. true neutrino energy minus reconstructed muon energy in GeV. This plot is for the non-QE population. The horizontal axis has variable binning. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the far detector.

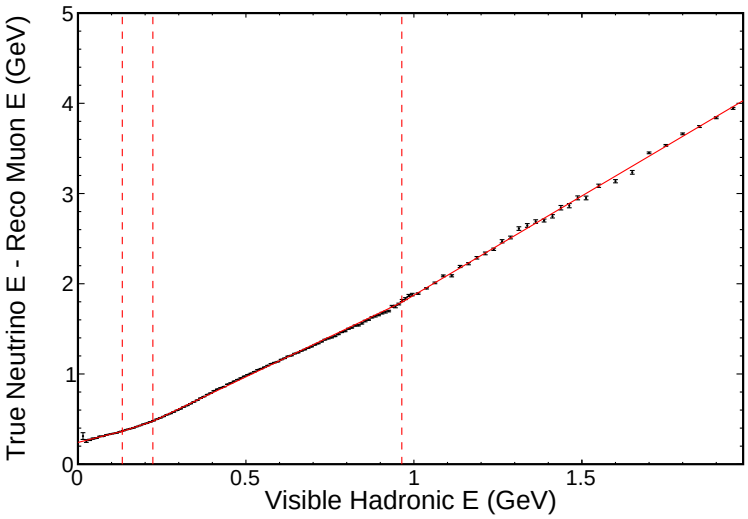


Figure 9.23: Best fit line overlaid on graph points used to make the fit. The fit relates visible hadronic energy in GeV to true neutrino energy minus reconstructed muon energy in GeV. The dashed vertical lines indicate the stitch locations of the splines. This plot is for the non-QE population. The horizontal axis has variable binning. This plot was created using simulated events in the far detector.

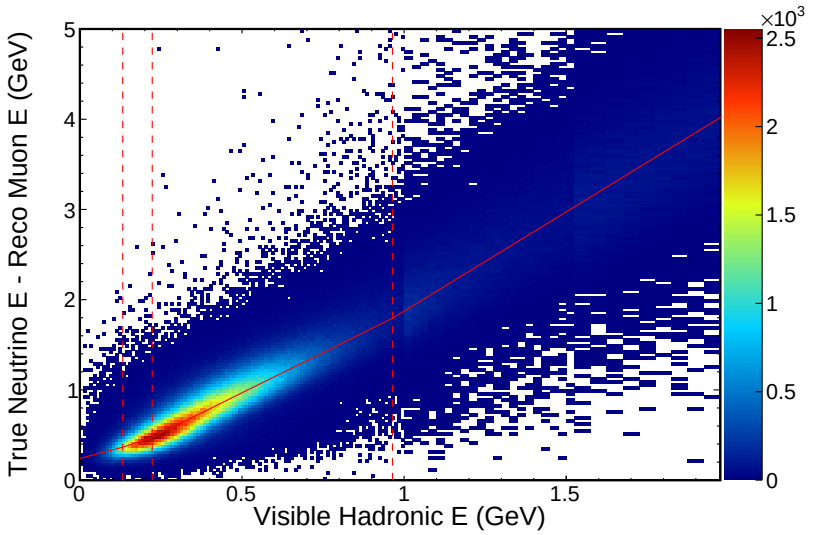


Figure 9.24: Best fit line overlaid on original 2D histogram. The fit relates visible hadronic energy in GeV to true neutrino energy minus reconstructed muon energy in GeV. The dashed vertical lines indicate the stitch locations of the splines. This plot is for the non-QE population. The horizontal axis has variable binning. The color axis is arbitrary number of interactions. This plot was created using simulated events in the far detector.

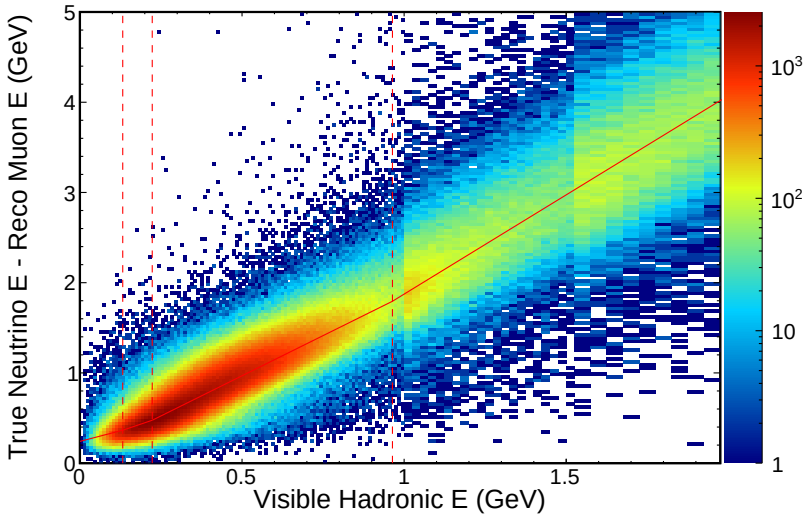


Figure 9.25: Best fit line overlaid on original 2D histogram with a logarithmic color axis. The fit relates visible hadronic energy in GeV to true neutrino energy minus reconstructed muon energy in GeV. The dashed vertical lines indicate the stitch locations of the splines. This plot is for the non-QE population. The horizontal axis has variable binning. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the far detector.

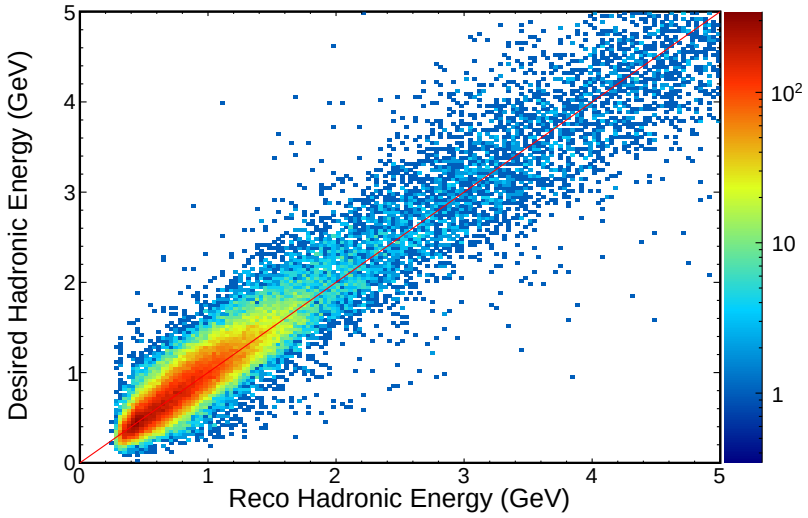


Figure 9.26: Histogram of visible hadronic energy in GeV vs. true neutrino energy minus reconstructed muon energy in GeV, notated as "desired hadronic energy." Red line is 45 degrees - if we did a perfect job, everything would lay on this line. This plot is for the non-QE population. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the far detector.

to a reconstructed hadronic energy, one can now consider the hadronic energy resolutions. All of the following plots were produced using a statistically independent sample from the one used to find the fit. Figure 9.26 is a 2D histogram which compares reconstructed hadronic energy to true hadronic energy, defined as true neutrino energy minus reconstructed muon energy. Figure 9.27 plots reconstructed hadronic energy vs. the relative energy resolution. One can also consider a 1D histogram of the relative energy resolution, Figure 9.28.

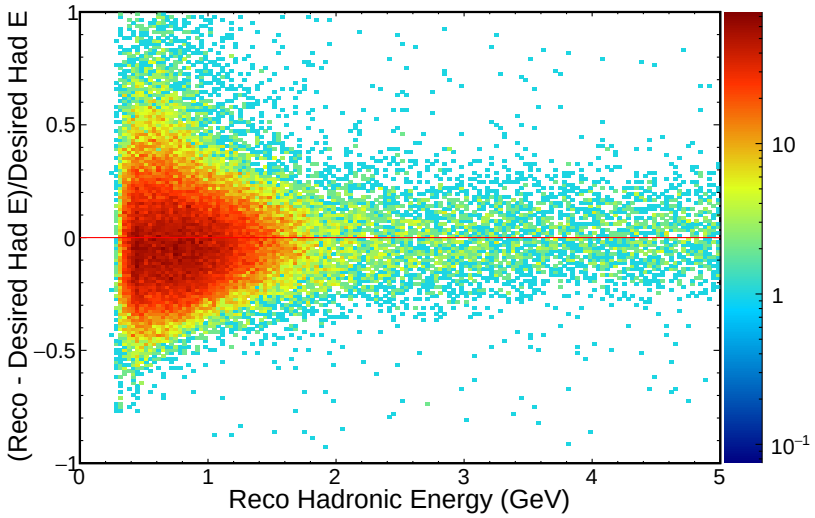


Figure 9.27: Histogram of visible hadronic energy in GeV vs. the relative energy resolution. The desired hadronic energy is defined as true neutrino energy minus reconstructed muon energy in GeV. Red line is flat at zero - if we did a perfect job, everything would lay on this line. This plot is for the non-QE population. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the far detector.

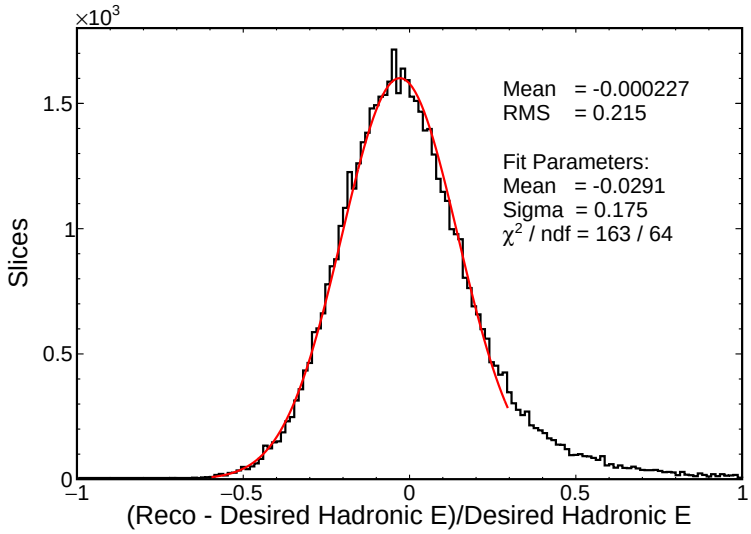


Figure 9.28: Relative energy resolution for reconstructed hadronic energy. The desired hadronic energy is defined as true neutrino energy minus reconstructed muon energy in GeV. This plot is for the non-QE population. This plot only includes events with true neutrino energy less than 5 GeV. This plot was created using simulated events in the far detector. Red line is a Gaussian fit to the peak.

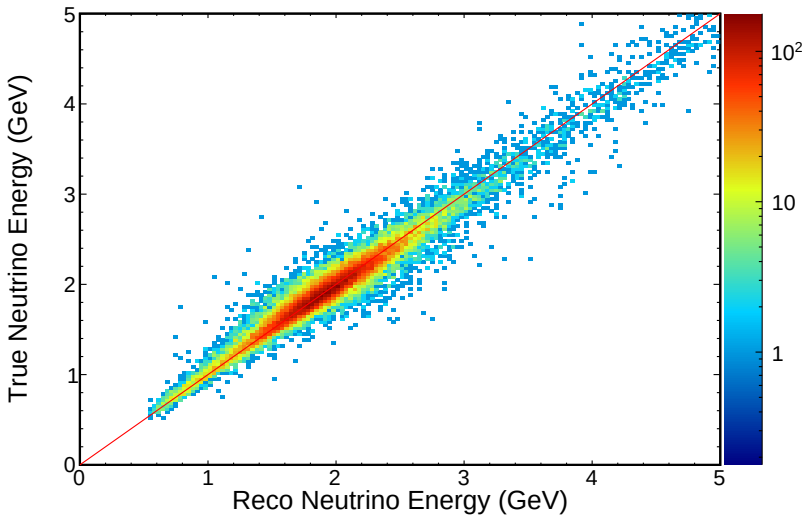


Figure 9.29: Histogram of reconstructed neutrino energy in GeV vs. true neutrino energy in GeV for the QE population. Red line is 45 degrees - if we did a perfect job, everything would lay on this line. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the far detector.

9.3 Neutrino Energy in the Far Detector

For each population (QE or non-QE), one can now look at neutrino energies and their resolutions. The reconstructed neutrino energy TrkQEE or TrkNonQEE is just defined as the sum of the reconstructed muon energy and the reconstructed hadronic energy. All of the following plots were produced using a statistically independent sample from the one used to find the fits.

For the QE population, Figure 9.29 is a 2D histogram which compares reconstructed neutrino energy to true neutrino energy. Figure 9.30 plots reconstructed neutrino energy vs. the relative energy resolution. Figure 9.31 is a 1D histogram of the relative energy resolution.

Figure 9.32 is a 2D histogram of the nonQE population which compares reconstructed neutrino energy to true neutrino energy. Figure 9.33 plots reconstructed neutrino energy vs. the relative energy resolution. One can also consider a 1D histogram of the relative energy resolution, Figure 9.34.

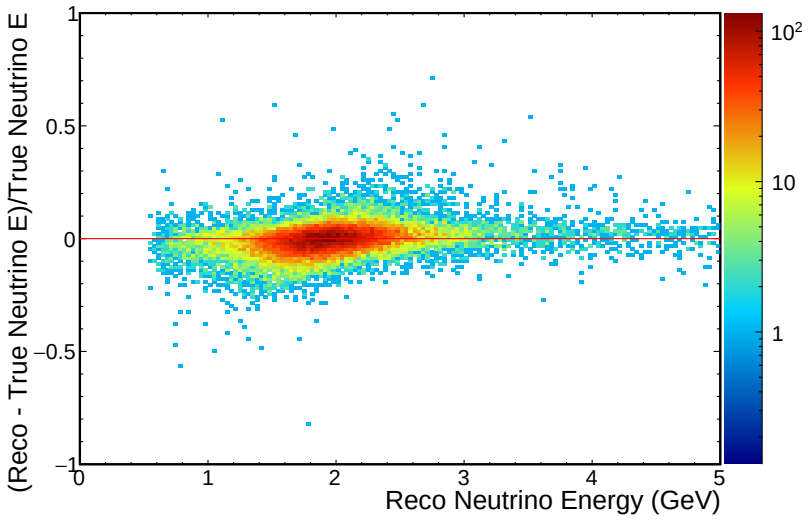


Figure 9.30: Histogram of reconstructed neutrino energy in GeV vs. the relative energy resolution for the QE population. Red line is flat at zero - if we did a perfect job, everything would lay on this line. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the far detector.

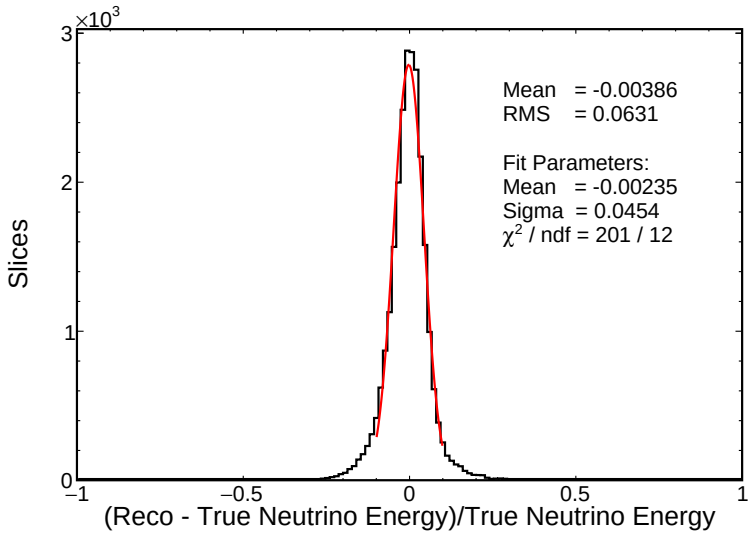


Figure 9.31: Relative energy resolution for reconstructed neutrino energy for the QE population. This plot only includes events with true neutrino energy less than 5 GeV. This plot was created using simulated events in the far detector. Red line is a Gaussian fit to the peak.

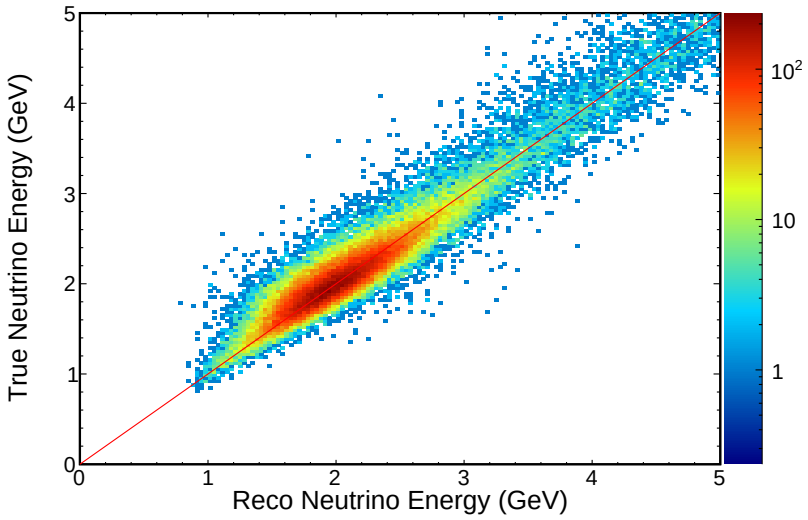


Figure 9.32: Histogram of reconstructed neutrino energy in GeV vs. true neutrino energy in GeV for the non-QE population. Red line is 45 degrees - if we did a perfect job, everything would lay on this line. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the far detector.

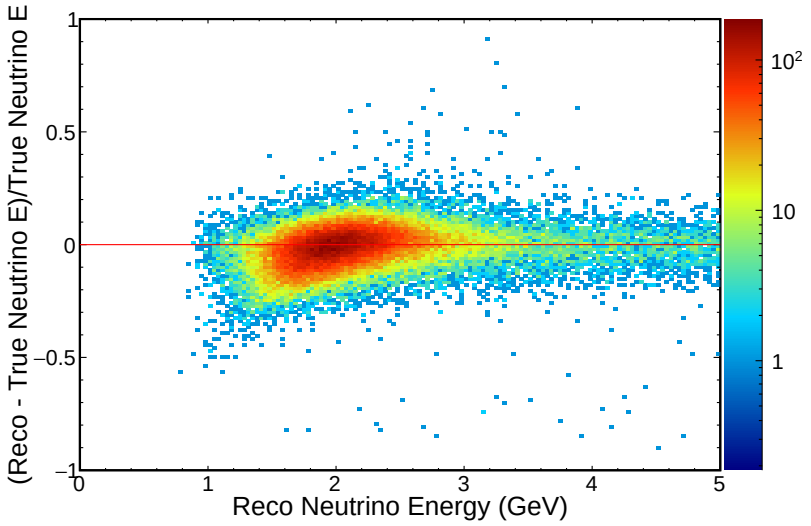


Figure 9.33: Histogram of reconstructed neutrino energy in GeV vs. the relative energy resolution for the non-QE population. Red line is flat at zero - if we did a perfect job, everything would lay on this line. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the far detector.

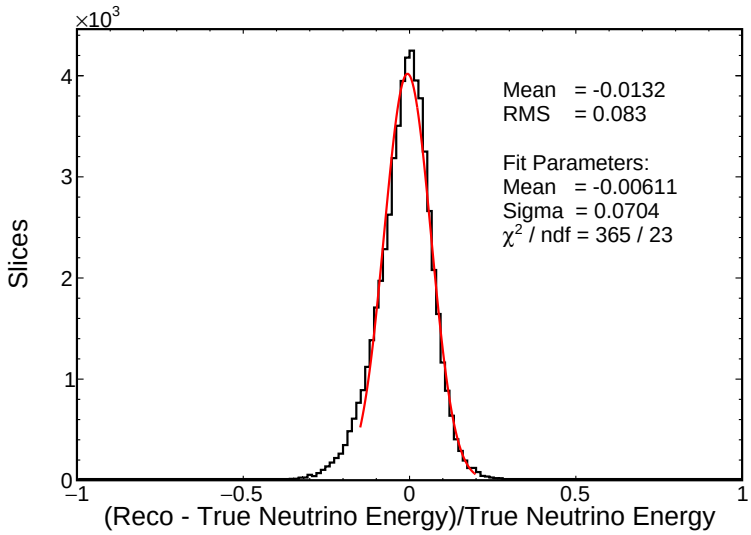


Figure 9.34: Relative energy resolution for reconstructed neutrino energy for the non-QE population. This plot only includes events with true neutrino energy less than 5 GeV. This plot was created using simulated events in the far detector. Red line is a Gaussian fit to the peak.

9.4 Muon Energy in the Near Detector

The near detector is composed of a fully active region and a muon catcher. The fully active region is composed of the same cells as the far detector; the muon catcher interspersed steel planes with active cells. See Section 5.1 for more information about the near detector and its composition.

For muons contained in the fully active region, the same functional relationship used in the far detector to convert reconstructed track length in cm into reconstructed muon energy in GeV should be valid. Section 9.4.1 evaluates the efficacy of applying the far detector muon fit to the near detector fully active population. For more information on the population used to fill these histograms, see Appendix F.

Another useful population for analysis consists of muons which start in the fully active region and end in the muon catcher. The rate of energy loss is radically different for the muon catcher portion of the detector, compared to the fully active region. For these muons, the reconstructed track length is divided into the length in the fully active region and the length in muon catcher. A linear fit is created from a 2D plot of reconstructed muon track length in the muon catcher vs. the true muon energy at the point where the muon entered the muon catcher. For more information on the population used to fill this histogram, see Appendix F. The 2D histogram is transformed into a graph, using the method described in Appendix D.

One can then use the linear fit to determine the muon reconstructed energy as it enters the muon catcher. By inverting the far detector muon spline fit, one can determine the effective track length of the muon in the muon catcher in units of track length in a fully active detector. This is added to the track length in the fully active region for a total track length. This final track length is converted into reconstructed muon energy using the far detector muon spline fit. See Section 9.4.2 for more information.

Finally, these two populations of muons are combined for a total muon population in the near detector. The total muon resolution is presented in Section 9.4.3.

9.4.1 Fully Active Population

For muons contained in the fully active region, the same functional relationship used in the far detector to convert reconstructed track length in cm into reconstructed muon

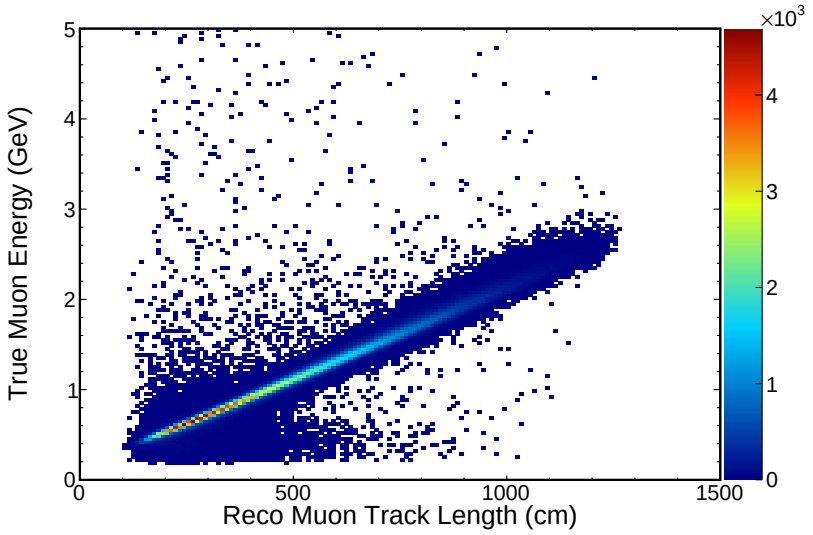


Figure 9.35: Reconstructed muon track length in cm vs. true muon energy in GeV for the fully active population. The color axis is arbitrary number of interactions. This histogram was created using simulated events in the near detector.

energy in GeV should be valid. For more information on the fit used for far detector muons, see Section 9.1.

A four spline fit is used to describe the relationship between muon track length to true muon energy, determined using simulation files of the far detector. For more information on the spline fit functional form, see Appendix E. Figure 9.35 is the 2D histogram that relates reconstructed muon track length in cm to true muon energy in GeV for the near detector fully active population. For more information on the population used to fill this histogram, see Appendix F. This is then transformed into a graph, using the method described in Appendix D. Figures 9.36 and 9.37 show the graph overlaid with the 2D histogram. Figure 9.37 has a logarithmic color scale.

Table 9.1 lists the far detector muon fit parameters. Figure 9.38 shows the far detector fit splines overlaid on the near detector graph points. Figure 9.39 shows the far detector fit splines overlaid on the near detector 2D histogram. Figure 9.40 is the same as Figure 9.39, except that it has a logarithmic color axis.

One can now consider the muon energy resolutions for the fully active population. Figure 9.41 is a 2D histogram which compares reconstructed muon energy to true muon

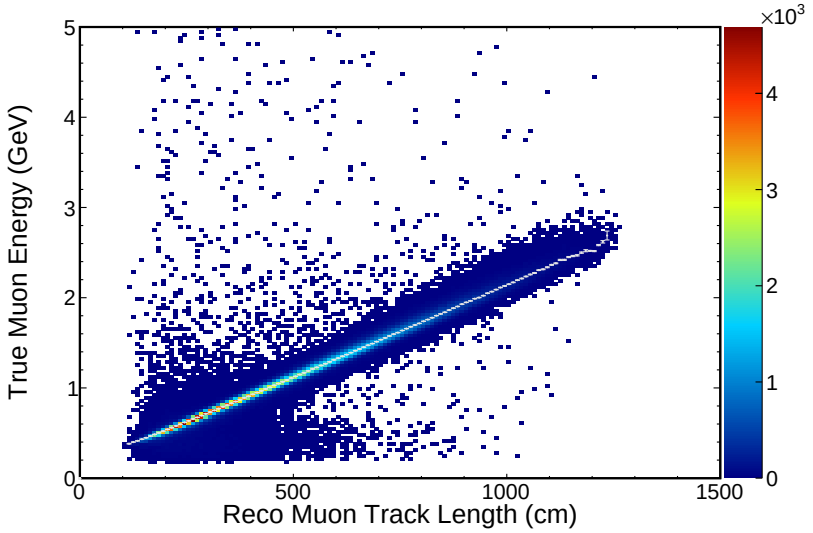


Figure 9.36: Graph points overlaid on 2D histogram of reconstructed muon track length in cm vs. true muon energy in GeV for the fully active population. The color axis is arbitrary number of interactions. This plot was created using simulated events in the near detector.

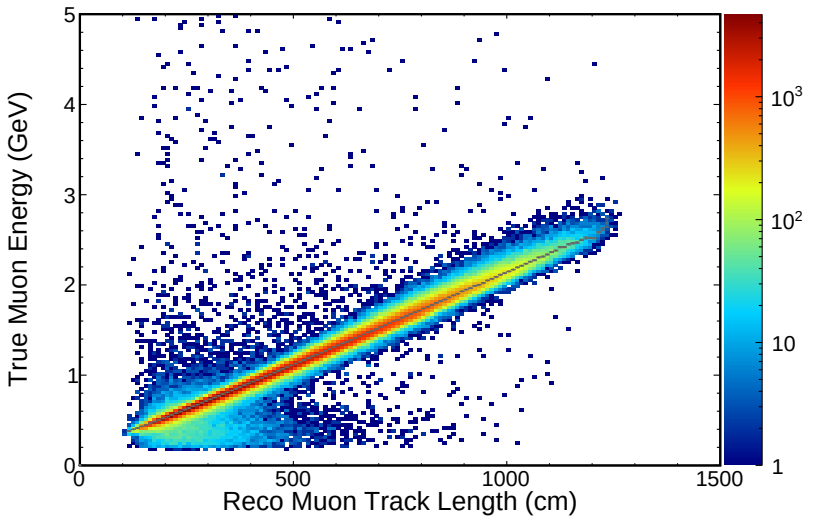


Figure 9.37: Graph points overlaid on 2D histogram of reconstructed muon track length in cm vs. true muon energy in GeV for the fully active population. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the near detector.

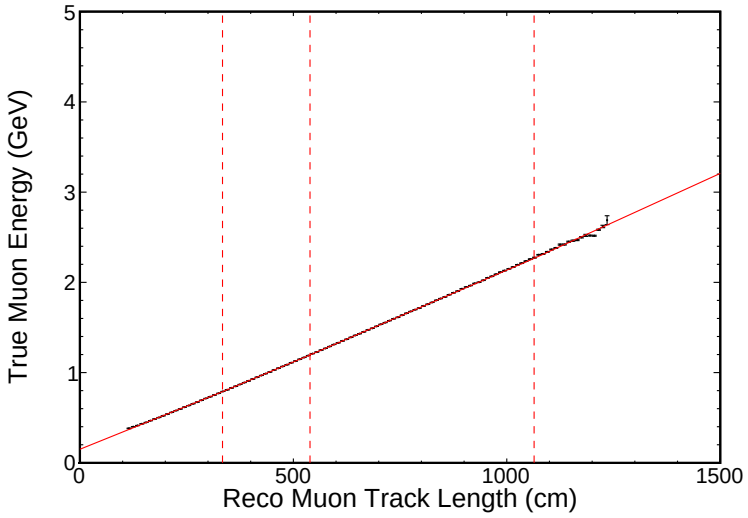


Figure 9.38: Far detector best fit line overlaid on graph points for the near detector fully active population. The fit relates reconstructed muon track length in cm to true muon energy in GeV. The dashed vertical lines indicate the stitch locations of the splines. This plot was created using simulated events in the near detector.

energy. Figure 9.42 plots muon track length in cm vs. the relative energy resolution. One can also consider a 1D histogram of the relative energy resolution, Figure 9.43.

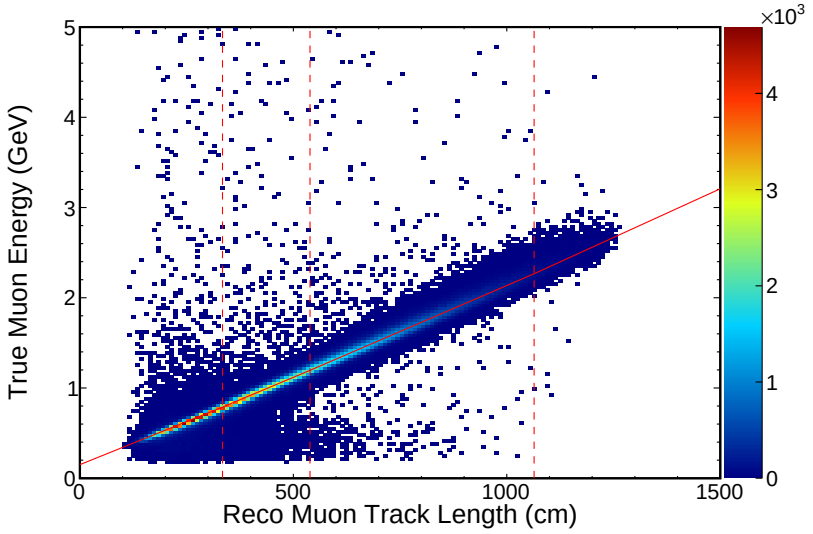


Figure 9.39: Far detector best fit line overlaid on 2D histogram for the near detector fully active population. The fit relates reconstructed muon track length in cm to true muon energy in GeV. The dashed vertical lines indicate the stitch locations of the splines. The color axis is arbitrary number of interactions. This plot was created using simulated events in the near detector.

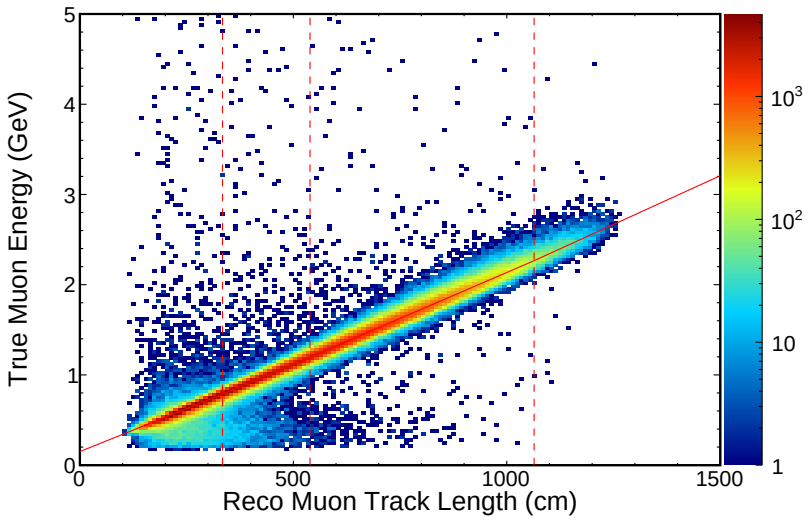


Figure 9.40: Far detector best fit line overlaid on 2D histogram for the near detector fully active population with a logarithmic color axis. The fit relates reconstructed muon track length in cm to true muon energy in GeV. The dashed vertical lines indicate the stitch locations of the splines. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the near detector.

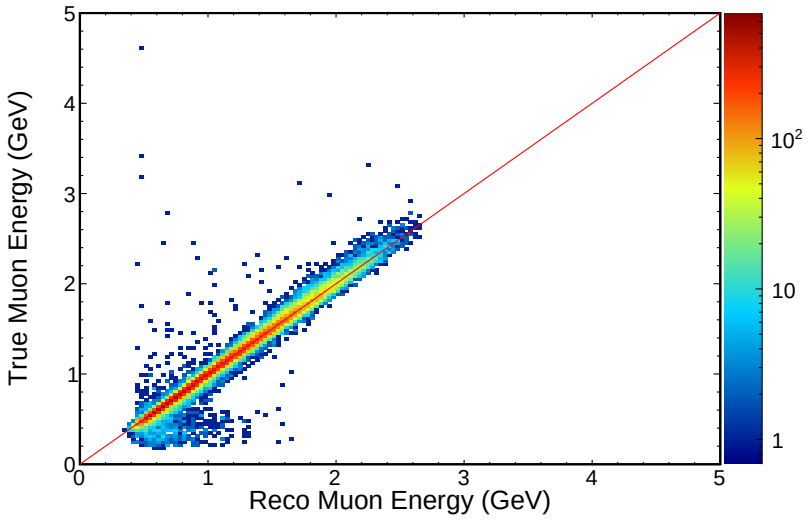


Figure 9.41: Histogram of reconstructed muon energy in GeV vs. true muon energy in GeV for the fully active population. Red line is 45 degrees - if we did a perfect job, everything would lay on this line. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the near detector.

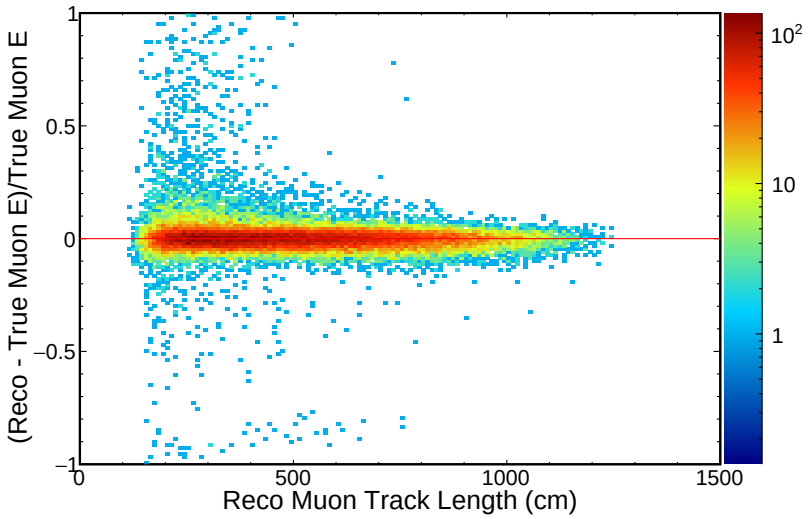


Figure 9.42: Histogram of reconstructed muon track length in cm vs. the relative energy resolution for the fully active population. Red line is flat at zero - if we did a perfect job, everything would lay on this line. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the near detector.

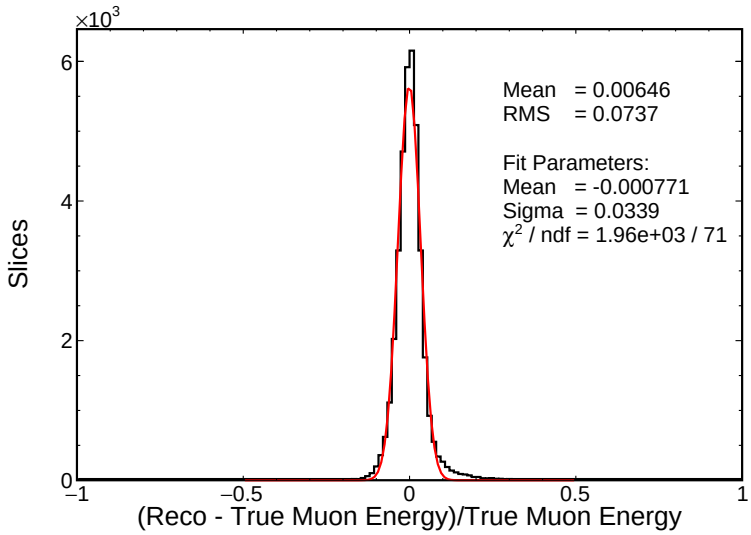


Figure 9.43: Relative energy resolution for reconstructed muon energy for the fully active population. Red line is a Gaussian fit to the peak. This plot only includes events with true neutrino energy less than 5 GeV. This plot was created using simulated events in the near detector.

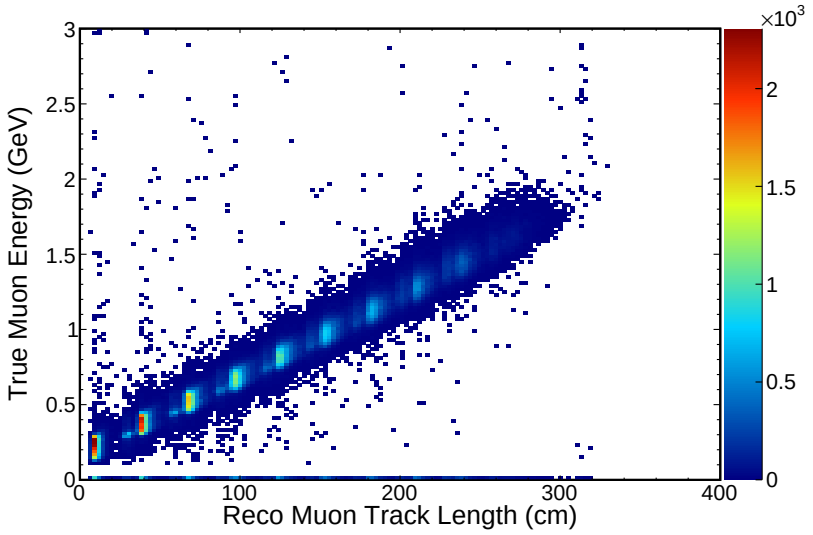


Figure 9.44: Reconstructed muon track length in the muon catcher in cm vs. true muon energy when entering the muon catcher in GeV for the active and catcher population. The color axis is arbitrary number of interactions. This histogram was created using simulated events in the near detector.

9.4.2 Active and Catcher Population

Another useful population for analysis consists of muons which start in the fully active region and end in the muon catcher. The rate of energy loss is radically different for the muon catcher portion of the detector, compared to the fully active region. For these muons, the reconstructed track length is divided into the length in the fully active region and the length in muon catcher. A linear fit is created from a 2D plot of reconstructed muon track length in the muon catcher vs. the true muon energy at the point where the muon entered the muon catcher, Figure 9.44. For more information on the population used to fill this histogram, see Appendix F. The 2D histogram is transformed into a graph, using the method described in Appendix D. Figures 9.45 and 9.46 show the graph overlaid with the 2D histogram. Figure 9.46 has a logarithmic color scale.

Table 9.5 lists the muon catcher fit parameters. Figure 9.47 shows the muon catcher fit overlaid on the graph points. Figure 9.48 shows the muon catcher fit overlaid on the 2D histogram. Figure 9.49 is the same as Figure 9.48, except that it has a logarithmic color axis.

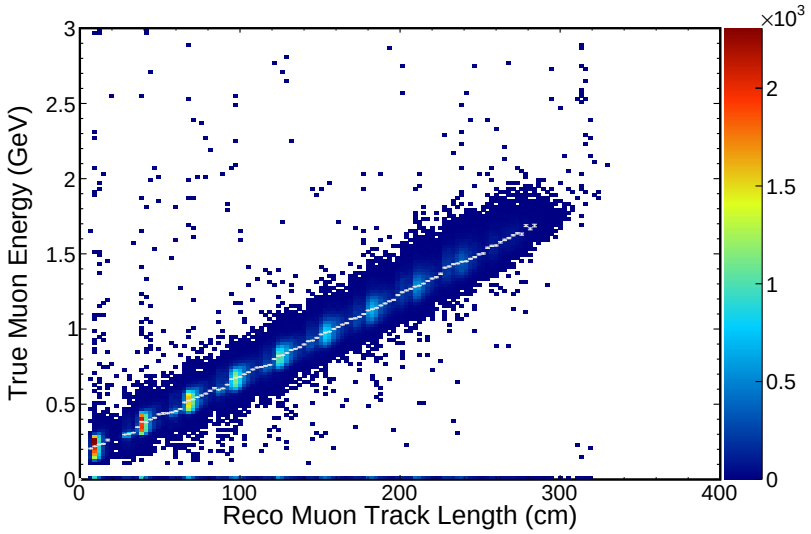


Figure 9.45: Graph points overlaid on 2D histogram of reconstructed muon track length in the muon catcher in cm vs. true muon energy when entering the muon catcher in GeV for the active and catcher population. The color axis is arbitrary number of interactions. This plot was created using simulated events in the near detector.

Parameter	**Value**	**Units**
Offset	0.152	GeV
Slope	0.00536	GeV/cm

Table 9.5: Fit parameters used to determine the near detector muon reconstructed energy for track length in the muon catcher.

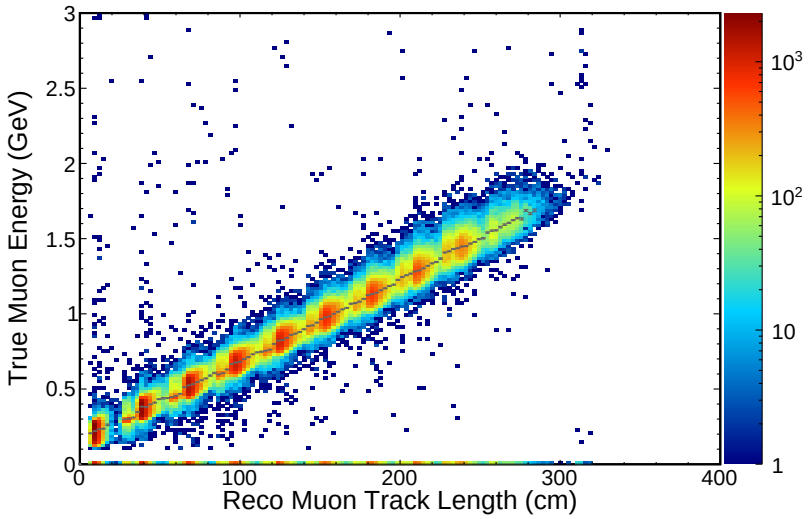


Figure 9.46: Graph points overlaid on 2D histogram of reconstructed muon track length in the muon catcher in cm vs. true muon energy when entering the muon catcher in GeV for the active and catcher population. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the near detector.

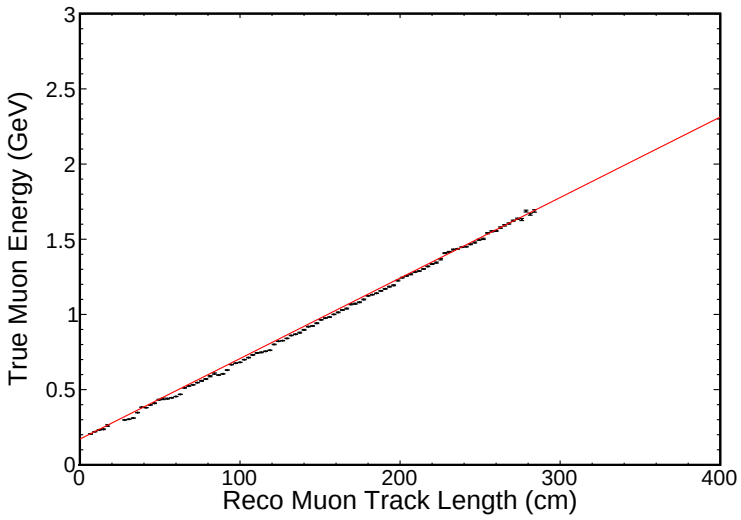


Figure 9.47: Best fit line overlaid on graph points for the active and catcher population. The fit relates reconstructed muon track length in the muon catcher in cm to true muon energy when entering the muon catcher in GeV. This plot was created using simulated events in the near detector.

One can now consider the muon energy resolutions in the muon catcher for the active and catcher population. All of the following plots were produced using a statistically independent sample from the one used to find the fit. Figure 9.50 is a 2D histogram which compares reconstructed muon energy in the muon catcher to true muon energy in the muon catcher. Figure 9.51 plots muon track length in the muon catcher in cm vs. the relative energy resolution. One can also consider a 1D histogram of the relative energy resolution, Figure 9.52.

One can then use the linear fit to determine the muon reconstructed energy as it enters the muon catcher. By inverting the far detector muon spline fit, one can determine the effective track length for the muon catcher in units of track length in a fully active detector. This is added to the track length in the fully active region for a total track length. This final track length is converted into reconstructed muon energy using the far detector muon spline fit.

Total muon energy resolutions for the active and catcher population can be plotted. All of the following plots were produced using a statistically independent sample from

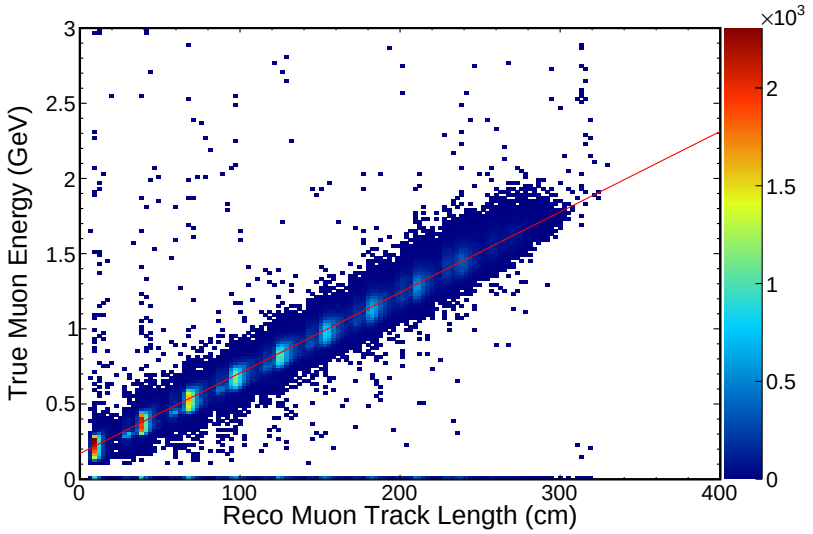


Figure 9.48: Best fit line overlaid on 2D histogram for the near detector active and catcher population. The fit relates reconstructed muon track length in the muon catcher in cm to true muon energy when entering the muon catcher in GeV. The color axis is arbitrary number of interactions. This plot was created using simulated events in the near detector.

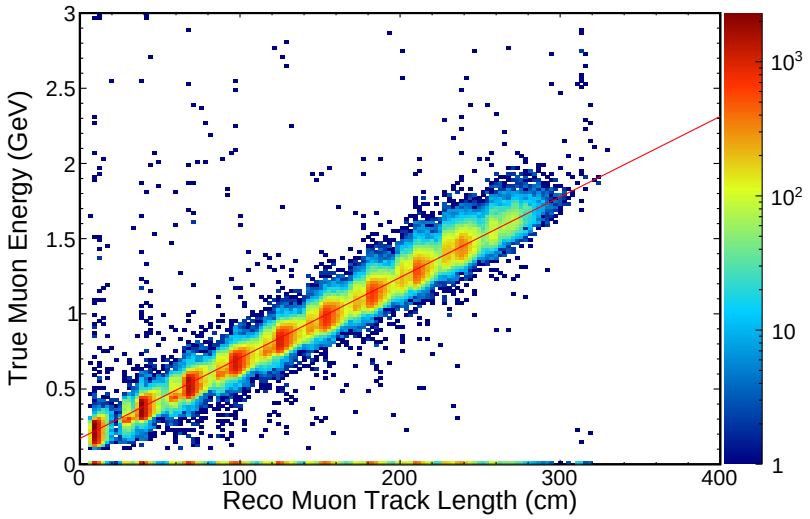


Figure 9.49: Best fit line overlaid on 2D histogram for the near detector active and catcher population with a logarithmic color axis. The fit relates reconstructed muon track length in the muon catcher in cm to true muon energy when entering the muon catcher in GeV. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the near detector.

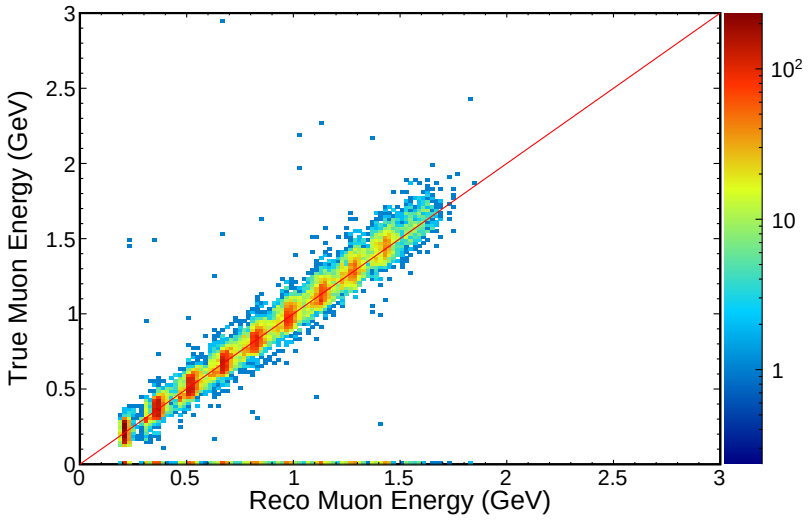


Figure 9.50: Histogram of reconstructed muon energy in the muon catcher in GeV vs. true muon energy in the muon catcher in GeV for the active and catcher population. Red line is 45 degrees - if we did a perfect job, everything would lay on this line. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the near detector.

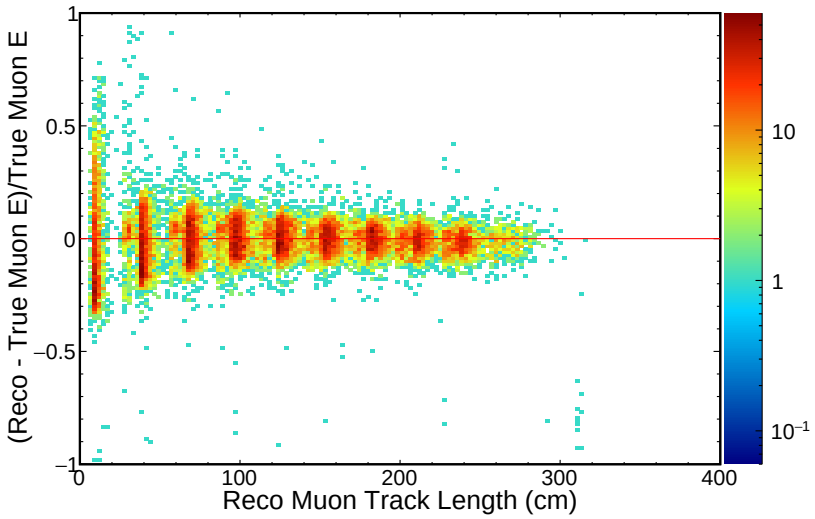


Figure 9.51: Histogram of reconstructed muon track length in the muon catcher in cm vs. the relative energy resolution for the active and catcher population. Red line is flat at zero - if we did a perfect job, everything would lay on this line. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the near detector.

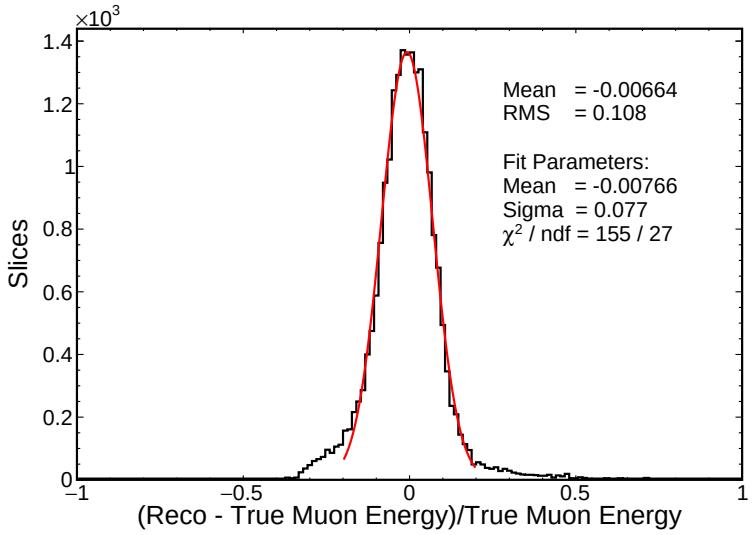


Figure 9.52: Relative energy resolution for reconstructed muon energy in the muon catcher for the active and catcher population. Red line is a Gaussian fit to the peak. This plot only includes events with true neutrino energy less than 5 GeV. This plot was created using simulated events in the near detector.

the one used to find the fit. Figure 9.53 is a 2D histogram which compares the total reconstructed muon energy to the total true muon energy. Figure 9.54 plots total effective muon track length in cm vs. the relative energy resolution. One can also consider a 1D histogram of the relative energy resolution, Figure 9.55.

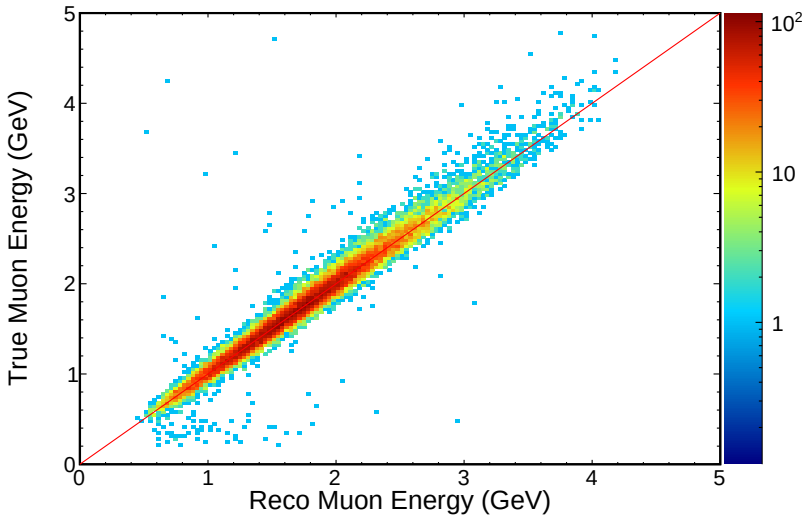


Figure 9.53: Histogram of the total reconstructed muon energy in GeV vs. the total true muon energy in GeV for the active and catcher population. Red line is 45 degrees - if we did a perfect job, everything would lay on this line. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the near detector.

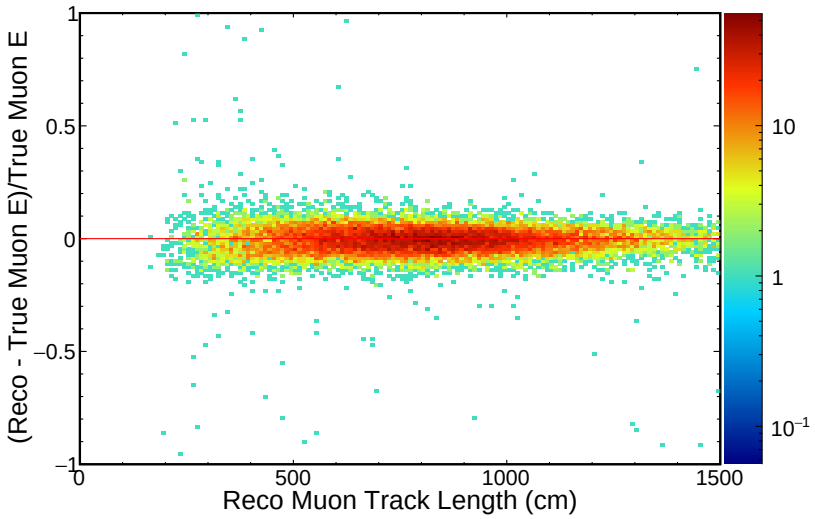


Figure 9.54: Histogram of the total reconstructed muon track length in cm vs. the relative energy resolution for the active and catcher population. Red line is flat at zero - if we did a perfect job, everything would lay on this line. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the near detector.

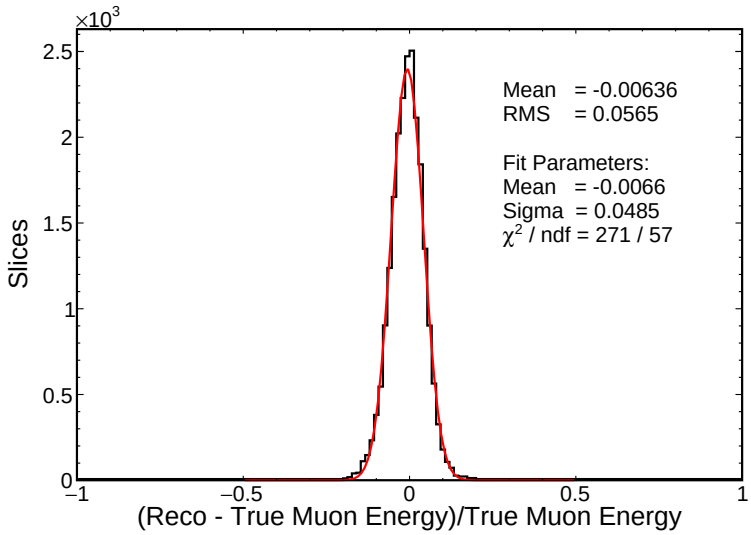


Figure 9.55: Relative energy resolution for the total reconstructed muon energy for the active and catcher population. Red line is a Gaussian fit to the peak. This plot only includes events with true neutrino energy less than 5 GeV. This plot was created using simulated events in the near detector.

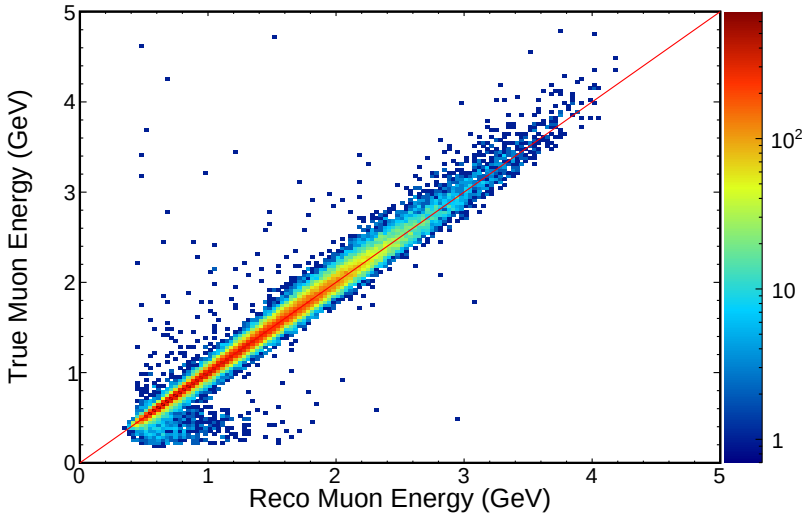


Figure 9.56: Histogram of reconstructed muon energy in GeV vs. true muon energy in GeV. Red line is 45 degrees - if we did a perfect job, everything would lay on this line. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the near detector.

9.4.3 Total Muon Near Detector Population

Finally, one can combine the two near detector populations of muons and evaluate the total muon resolution. All of the following plots were produced using a statistically independent sample from the one used to determine fits. Figure 9.56 is a 2D histogram which compares the reconstructed muon energy to the true muon energy. Figure 9.57 plots muon track length in cm vs. the relative energy resolution. One can also consider a 1D histogram of the relative energy resolution, Figure 9.58.

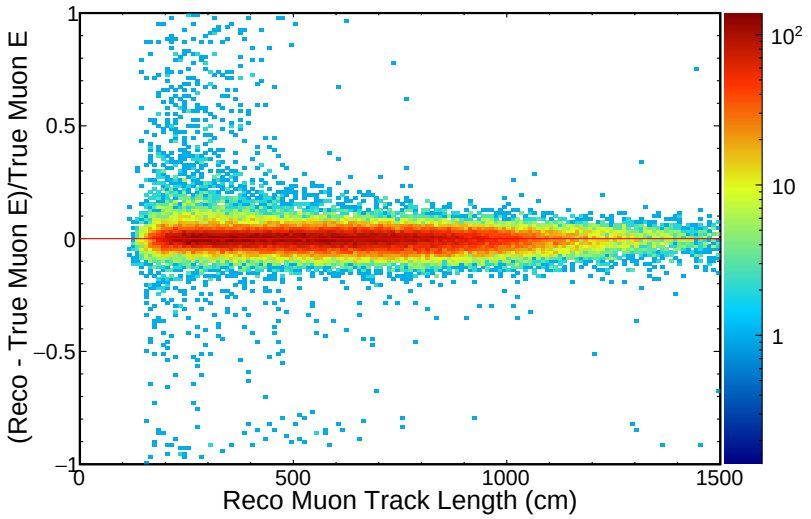


Figure 9.57: Histogram of the reconstructed muon track length in cm vs. the relative energy resolution. Red line is flat at zero - if we did a perfect job, everything would lay on this line. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the near detector.

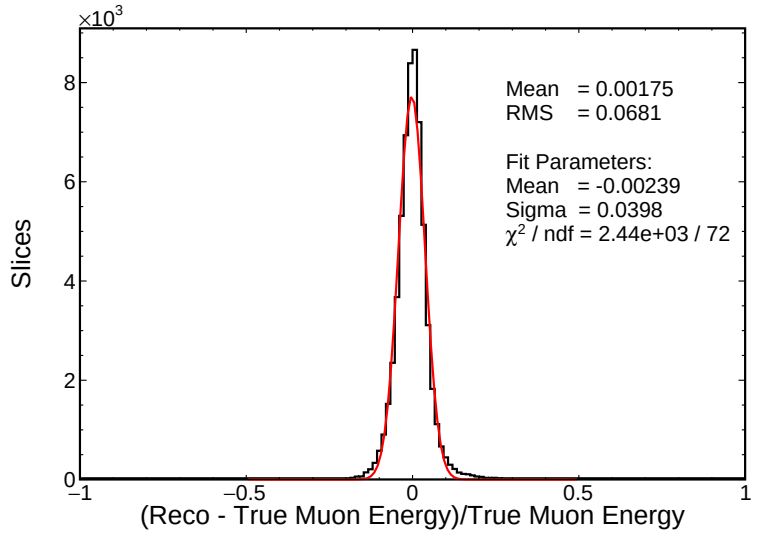


Figure 9.58: Relative energy resolution for reconstructed muon energy. Red line is a Gaussian fit to the peak. This plot only includes events with true neutrino energy less than 5 GeV. This plot was created using simulated events in the near detector.

9.5 Hadronic Energy in the Near Detector

After one has the fits to create reconstructed muon energy in the near detector, one creates near detector hadronic energy spline fits. Visible hadronic energy is defined as the sum of the energy deposited by all hits not on the muon track as well as the energy on the muon track in the vertex region that exceeds minimum ionizing values. For information about the hadronic model used by the simulation, see Chapter 6. For more information about determining hadronic energy contamination on the muon track, see the technical note[61]. No attempt is made to remove noise hits contained within the slice from consideration as hadronic energy.

The far and near detectors need to have different hadronic energy fits. This is driven primarily by the small size of the near detector; it is much easier for hadronic energy to escape and not be visible in the near detector. The small size also means that all hits in the near detector are relatively close to the electronics readout and less likely to fall below threshold than in the far detector. Finally, only lower energy events can be contained in the near detector. Because of the limited energy range, only two splines have been used to capture the structure in the visible hadronic energy. In the far detector, four were used.

If one wanted to account for large amounts of hadronic energy in the muon catcher, special care and extra fits would be required. This has not been done at this time. Instead, the containment criteria require that very little visible hadronic energy is present in the muon catcher. This allows events with a few noise hits in the muon catcher to still be included in the sample.

Both the QE and non-QE populations in the near detector were fit with a two spline fit. For more information on the spline fit functional form, see Appendix E. For both populations, variable binning was employed for the underlying 2D histograms of visible hadronic energy in GeV vs. true neutrino energy minus reconstructed muon energy in GeV. For more information on the population used to fill these histograms, see Appendix F. The 2D histograms are then transformed into a graph, using the method described in Appendix D.

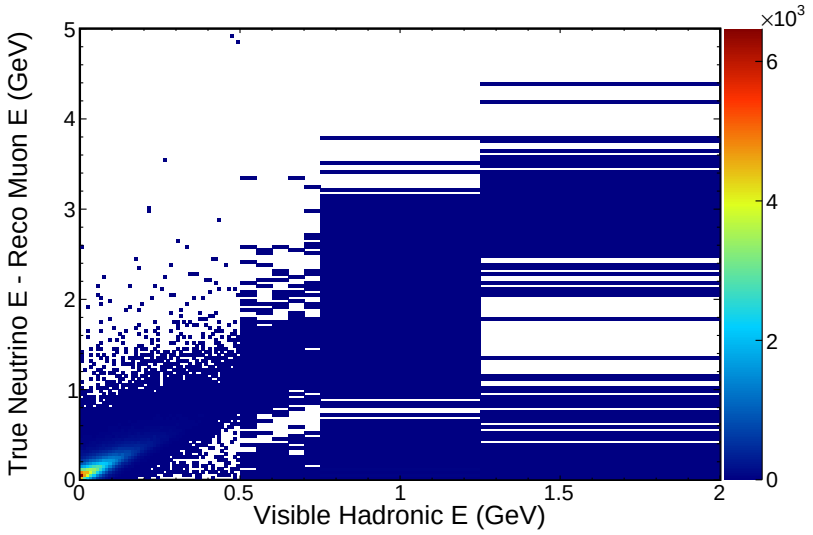


Figure 9.59: Visible hadronic energy in GeV vs. true neutrino energy minus reconstructed muon energy in GeV. This plot is for the QE population. The horizontal axis has variable binning. The color axis is arbitrary number of interactions. This plot was created using simulated events in the near detector.

9.5.1 QE Events

The results for the QE population will be presented first. Figure 9.59 is the 2D histogram that relates visible hadronic energy in GeV to true neutrino energy minus reconstructed muon energy in GeV. This is then transformed into a graph, using the method described in Appendix D. Figures 9.60 and 9.61 show the graph to be fit overlaid with the 2D histogram. Figure 9.61 has a logarithmic color scale.

Table 9.6 displays the final fit parameters. Figure 9.62 shows the fit line overlaid on the graph points used to find the fit. Figure 9.63 shows the fit line overlaid on the original 2D histogram. Figure 9.64 is the same as Figure 9.63, except that it has a logarithmic color axis.

Having obtained a fit which allows translation from visible hadronic energy to a reconstructed hadronic energy, one can now consider the hadronic energy resolutions. All of the following plots were produced using a statistically independent sample from the one used to find the fit. Figure 9.65 is a 2D histogram which compares reconstructed hadronic energy to desired hadronic energy, defined as true neutrino energy

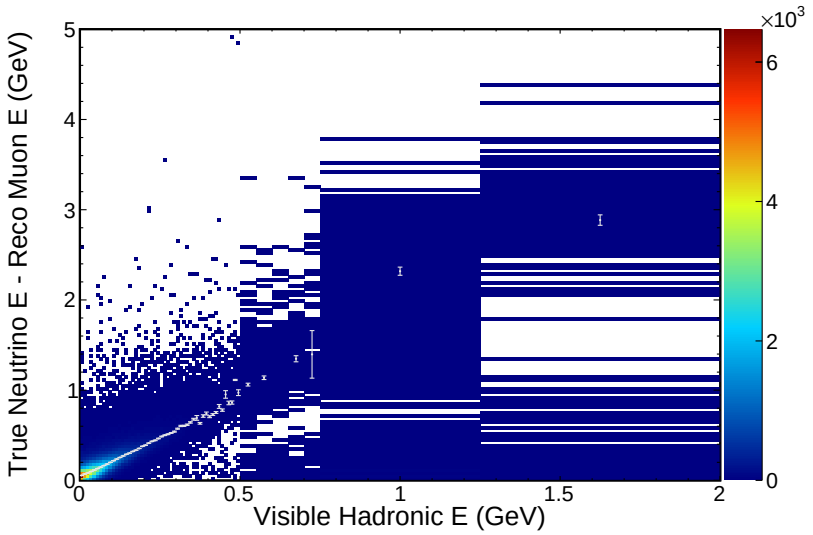


Figure 9.60: Graph points overlaid on 2D histogram of visible hadronic energy in GeV vs. true neutrino energy minus reconstructed muon energy in GeV. This plot is for the QE population. The horizontal axis has variable binning. The color axis is arbitrary number of interactions. This plot was created using simulated events in the near detector.

Parameter	**Value**	**Units**
Offset	0.0450	GeV
Slope 1	1.567	unitless
Slope 2	1.822	unitless
Stitch	0.0820	GeV

Table 9.6: Fit parameters used to determine the near detector QE hadronic reconstructed energy.

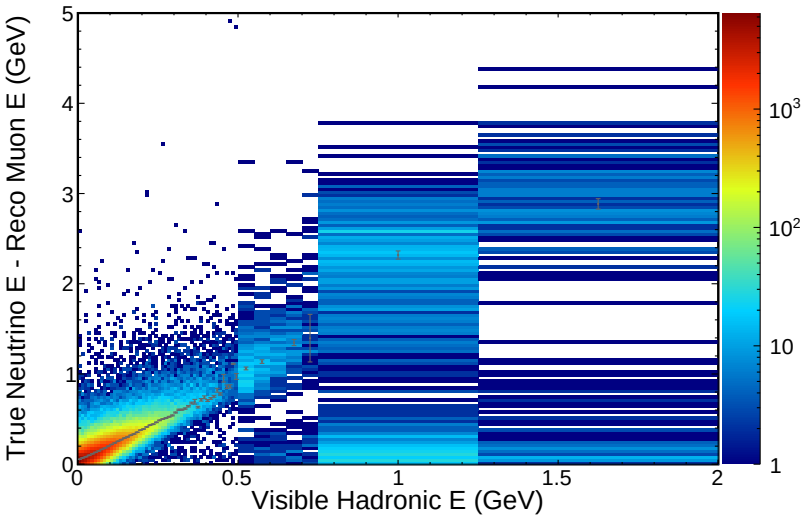


Figure 9.61: Graph points overlaid on 2D histogram of visible hadronic energy in GeV vs. true neutrino energy minus reconstructed muon energy in GeV. This plot is for the QE population. The horizontal axis has variable binning. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the near detector.

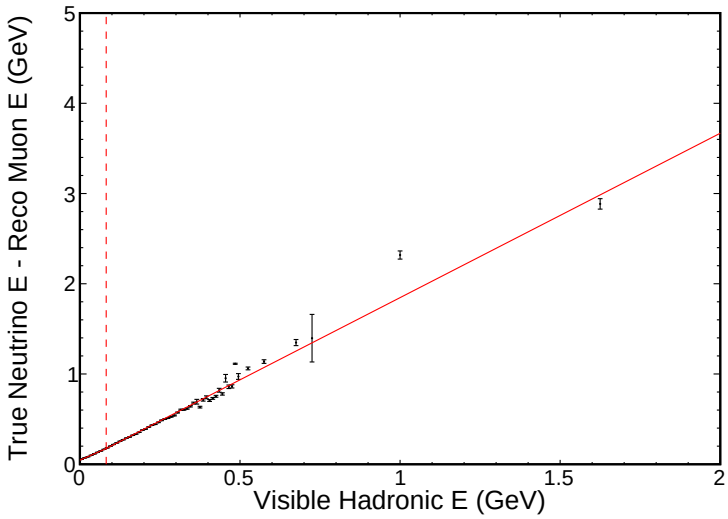


Figure 9.62: Best fit line overlaid on graph points used to make the fit. The fit relates visible hadronic energy in GeV to true neutrino energy minus reconstructed muon energy in GeV. The dashed vertical line indicates the stitch location of the splines. This plot is for the QE population. The horizontal axis has variable binning. This plot was created using simulated events in the near detector.

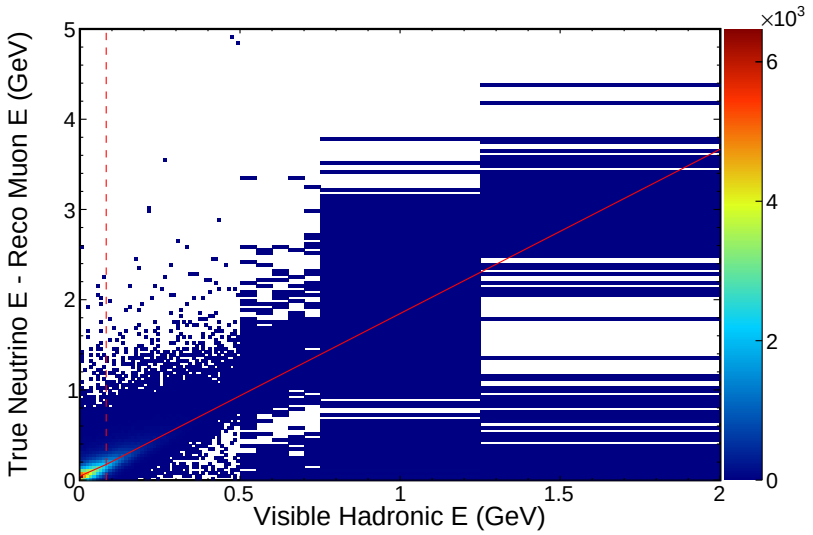


Figure 9.63: Best fit line overlaid on original 2D histogram. The fit relates visible hadronic energy in GeV to true neutrino energy minus reconstructed muon energy in GeV. The dashed vertical line indicates the stitch location of the splines. This plot is for the QE population. The horizontal axis has variable binning. The color axis is arbitrary number of interactions. This plot was created using simulated events in the near detector.

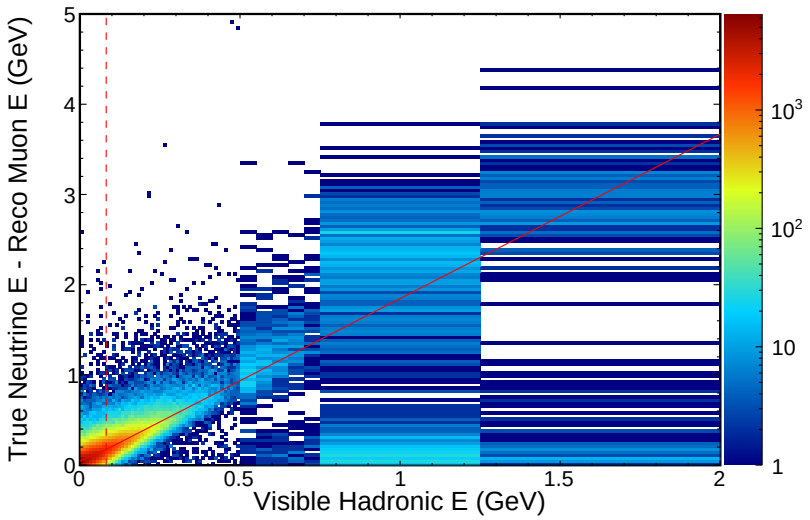


Figure 9.64: Best fit line overlaid on original 2D histogram. The fit relates visible hadronic energy in GeV to true neutrino energy minus reconstructed muon energy in GeV. The dashed vertical line indicates the stitch locations of the splines. This plot is for the QE population. The horizontal axis has variable binning. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the near detector.

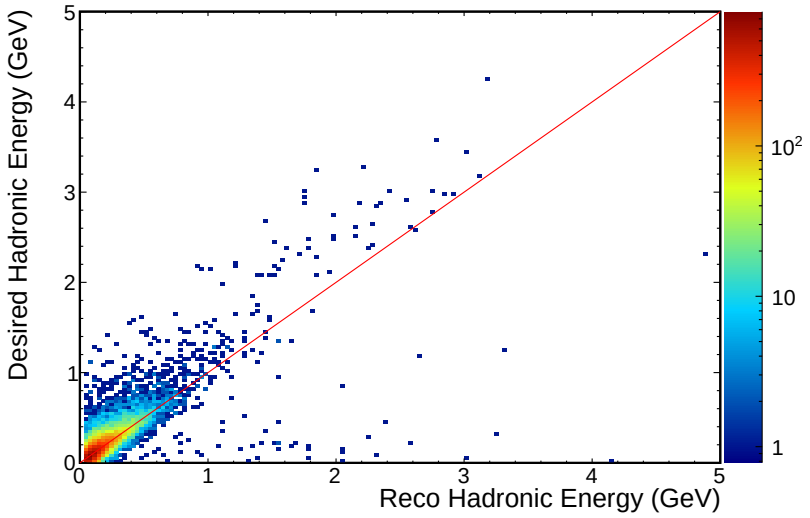


Figure 9.65: Histogram of reconstructed hadronic energy in GeV vs. true neutrino energy minus reconstructed muon energy in GeV, notated as "desired hadronic energy." Red line is 45 degrees - if we did a perfect job, everything would lay on this line. This plot is for the QE population. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the near detector.

minus reconstructed muon energy. Figure 9.66 plots reconstructed hadronic energy vs. the relative energy resolution. One can also consider a 1D histogram of the relative energy resolution, Figure 9.67.

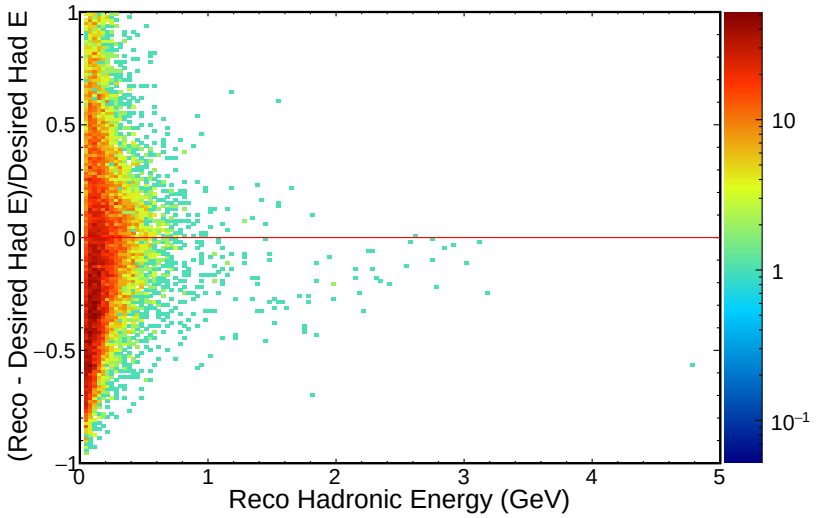


Figure 9.66: Histogram of visible hadronic energy in GeV vs. the relative energy resolution. The desired hadronic energy is defined as true neutrino energy minus reconstructed muon energy in GeV. Red line is flat at zero - if we did a perfect job, everything would lay on this line. This plot is for the QE population. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the near detector.

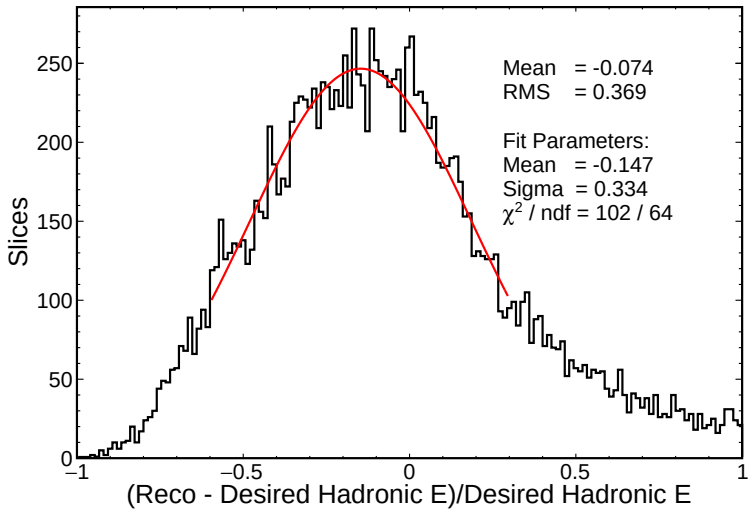


Figure 9.67: Relative energy resolution for reconstructed hadronic energy. The desired hadronic energy is defined as true neutrino energy minus reconstructed muon energy in GeV. This plot is for the QE population. This plot only includes events with true neutrino energy less than 5 GeV. This plot was created using simulated events in the near detector. Red line is a Gaussian fit to the peak.

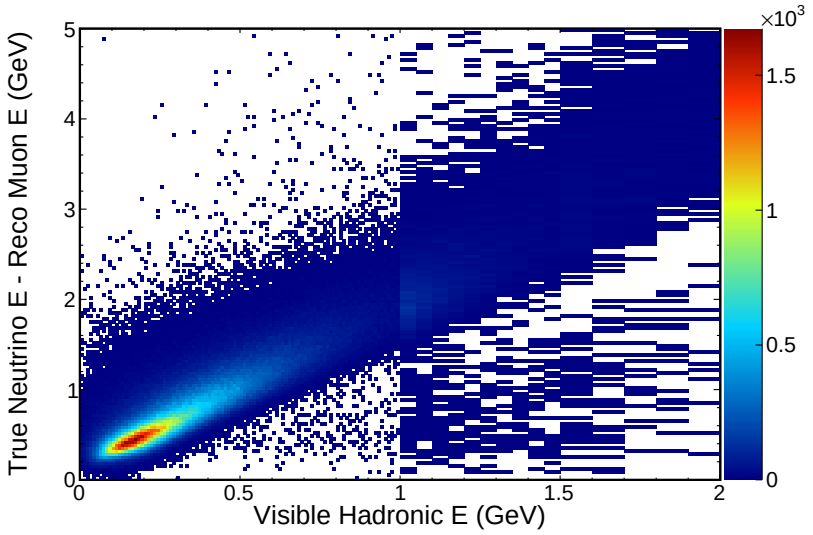


Figure 9.68: Visible hadronic energy in GeV vs. true neutrino energy minus reconstructed muon energy in GeV. This plot is for the non-QE population. The horizontal axis has variable binning. The color axis is arbitrary number of interactions. This plot was created using simulated events in the near detector.

9.5.2 Non-QE Events

For the non-QE population, Figure 9.68 is the 2D histogram that relates visible hadronic energy in GeV to true neutrino energy minus reconstructed muon energy in GeV. For more information on the population used to fill these histograms, see Appendix F. This is then transformed into a graph, using the method described in Appendix D. Figures 9.69 and 9.70 show the graph to be fit overlaid with the 2D histogram. Figure 9.70 has a logarithmic color scale.

Table 9.7 displays the final fit parameters. Figure 9.71 shows the fit line overlaid on the graph points used to find the fit. Figure 9.72 shows the fit line overlaid on the original 2D histogram. Figure 9.73 is the same as Figure 9.72, except that it has a logarithmic color axis.

Having now obtained a fit which allows translation from visible hadronic energy to a reconstructed hadronic energy, one can now consider the hadronic energy resolutions. All of the following plots were produced using a statistically independent sample

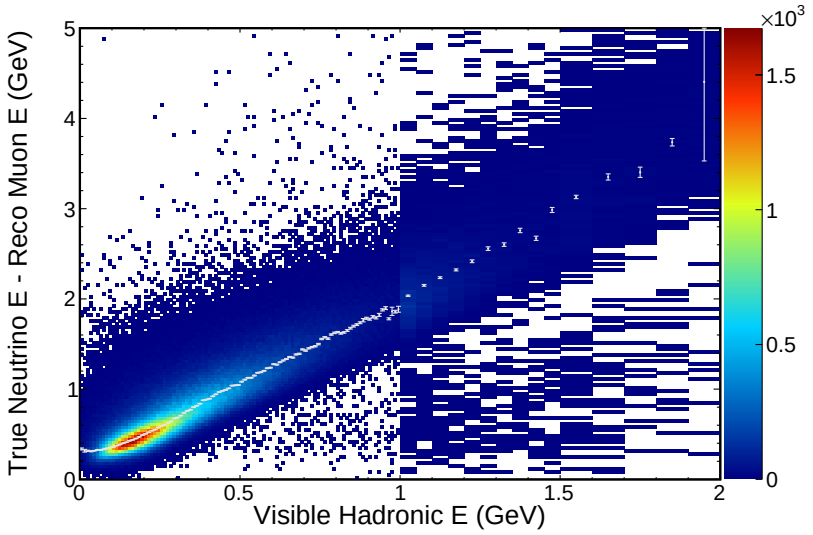


Figure 9.69: Graph points overlaid on 2D histogram of visible hadronic energy in GeV vs. true neutrino energy minus reconstructed muon energy in GeV. This plot is for the non-QE population. The horizontal axis has variable binning. The color axis is arbitrary number of interactions. This plot was created using simulated events in the near detector.

Parameter	**Value**	**Units**
Offset	0.254	GeV
Slope 1	1.080	unitless
Slope 2	1.889	unitless
Stitch	0.169	GeV

Table 9.7: Fit parameters used to determine the near detector non-QE hadronic reconstructed energy.

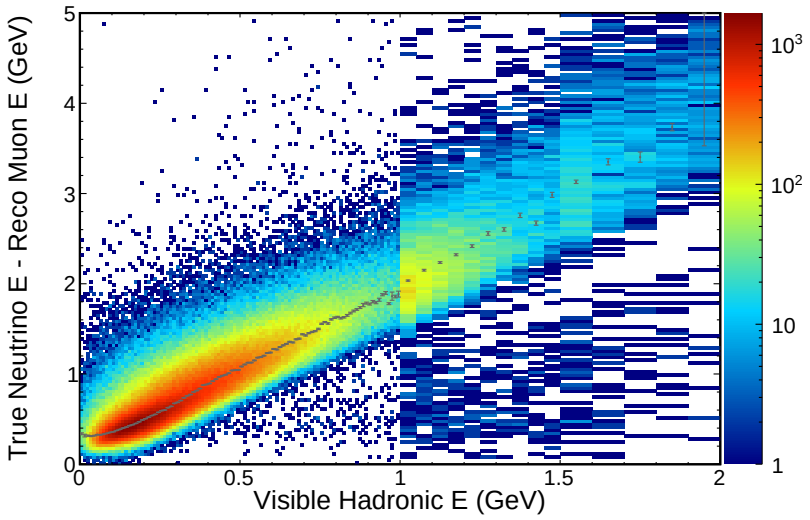


Figure 9.70: Graph points overlaid on 2D histogram of visible hadronic energy in GeV vs. true neutrino energy minus reconstructed muon energy in GeV. This plot is for the non-QE population. The horizontal axis has variable binning. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the near detector.

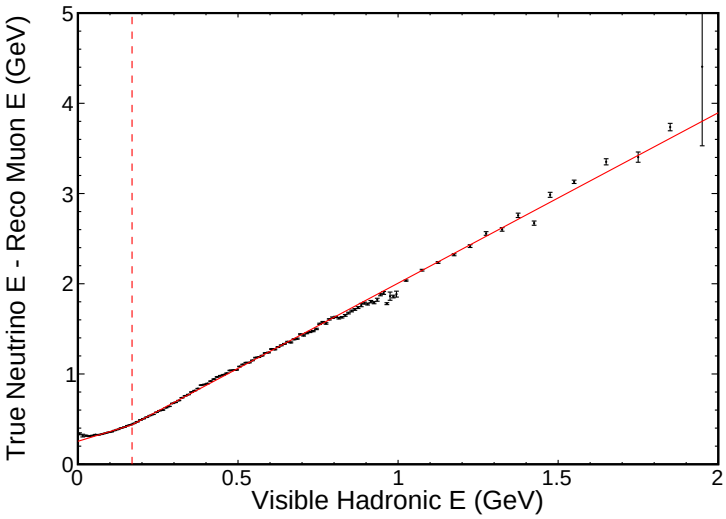


Figure 9.71: Best fit line overlaid on graph points used to make the fit. The fit relates visible hadronic energy in GeV to true neutrino energy minus reconstructed muon energy in GeV. The dashed vertical line indicates the stitch location of the splines. This plot is for the non-QE population. The horizontal axis has variable binning. This plot was created using simulated events in the near detector.

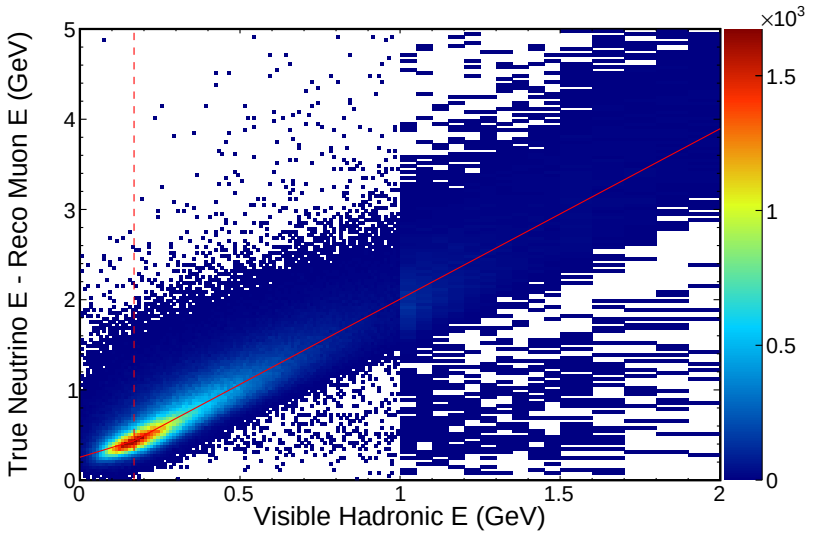


Figure 9.72: Best fit line overlaid on original 2D histogram. The fit relates visible hadronic energy in GeV to true neutrino energy minus reconstructed muon energy in GeV. The dashed vertical line indicates the stitch location of the splines. This plot is for the non-QE population. The horizontal axis has variable binning. The color axis is arbitrary number of interactions. This plot was created using simulated events in the near detector.

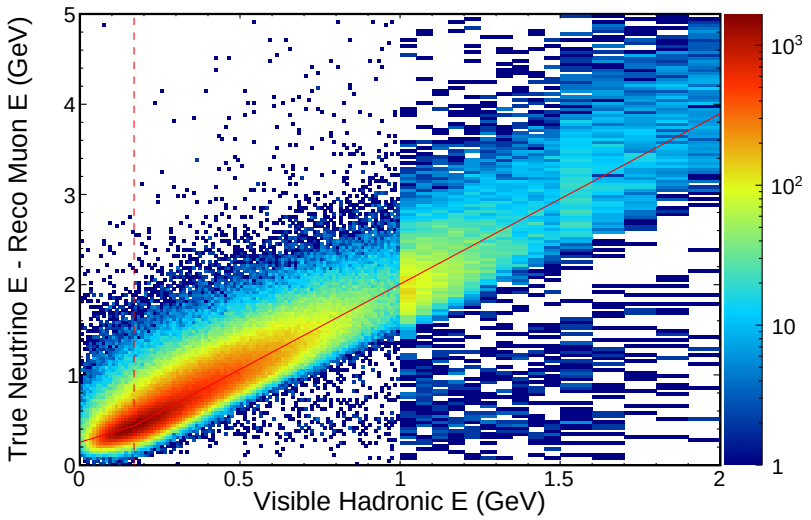


Figure 9.73: Best fit line overlaid on original 2D histogram with a logarithmic color axis. The fit relates visible hadronic energy in GeV to true neutrino energy minus reconstructed muon energy in GeV. The dashed vertical line indicates the stitch location of the splines. This plot is for the non-QE population. The horizontal axis has variable binning. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the near detector.

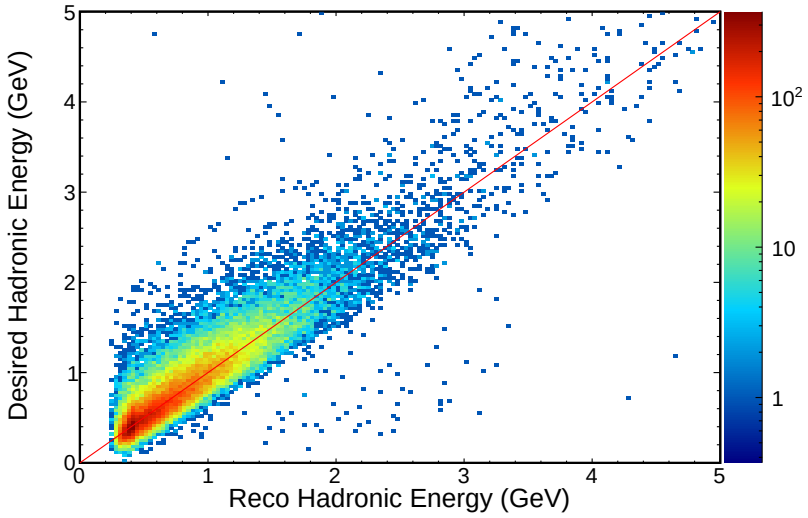


Figure 9.74: Histogram of visible hadronic energy in GeV vs. true neutrino energy minus reconstructed muon energy in GeV, notated as "desired hadronic energy." Red line is 45 degrees - if we did a perfect job, everything would lay on this line. This plot is for the non-QE population. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the near detector.

from the one used to find the fit. Figure 9.74 is a 2D histogram which compares reconstructed hadronic energy to true hadronic energy, defined as true neutrino energy minus reconstructed muon energy. Figure 9.75 plots reconstructed hadronic energy vs. the relative energy resolution. One can also consider a 1D histogram of the relative energy resolution, Figure 9.76.

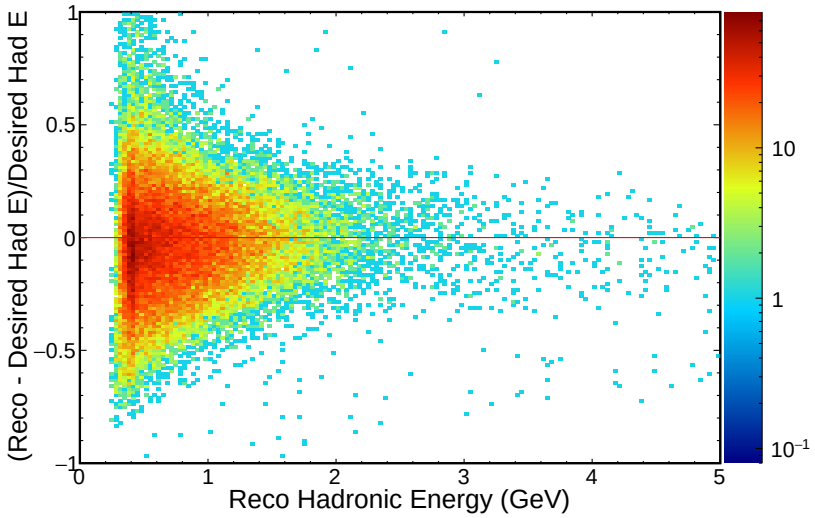


Figure 9.75: Histogram of visible hadronic energy in GeV vs. the relative energy resolution. The desired hadronic energy is defined as true neutrino energy minus reconstructed muon energy in GeV. Red line is flat at zero - if we did a perfect job, everything would lay on this line. This plot is for the non-QE population. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the near detector.

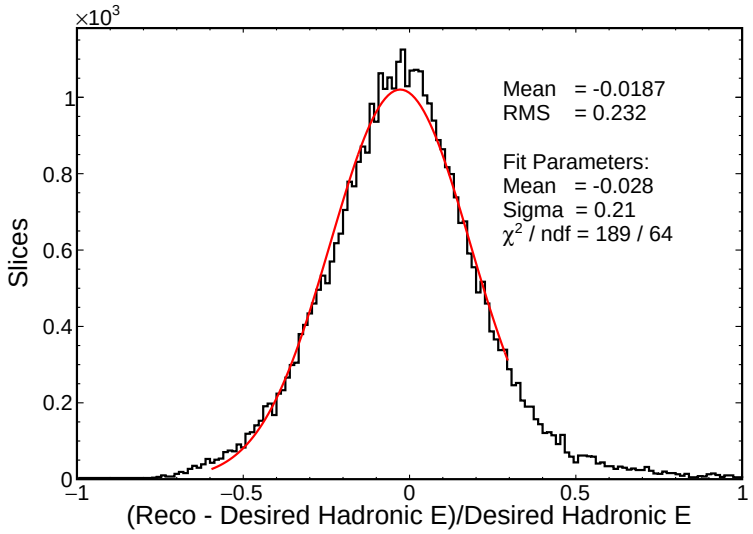


Figure 9.76: Relative energy resolution for reconstructed hadronic energy. The desired hadronic energy is defined as true neutrino energy minus reconstructed muon energy in GeV. This plot is for the non-QE population. This plot only includes events with true neutrino energy less than 5 GeV. This plot was created using simulated events in the near detector. Red line is a Gaussian fit to the peak.

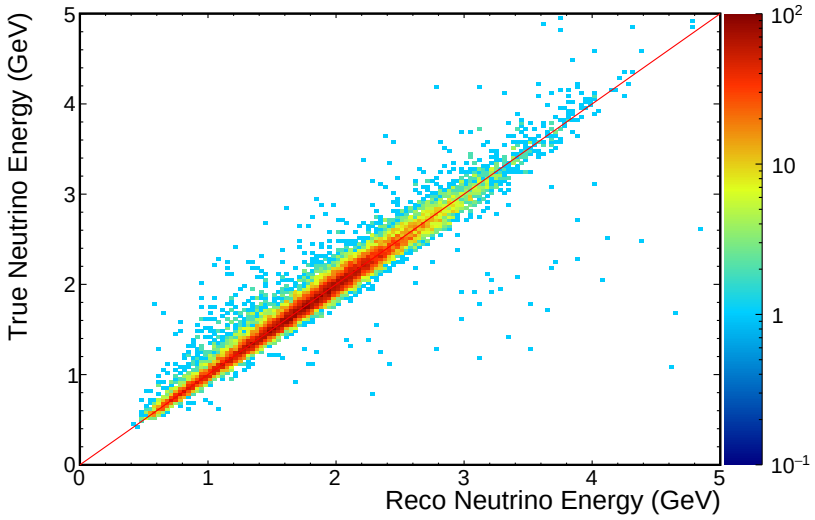


Figure 9.77: Histogram of reconstructed neutrino energy in GeV vs. true neutrino energy in GeV for the QE population. Red line is 45 degrees - if we did a perfect job, everything would lay on this line. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the near detector.

9.6 Neutrino Energy in the Near Detector

For each population (QE or non-QE), one can now look at neutrino energies and their resolutions. The reconstructed neutrino energy TrkQEE or TrkNonQEE is just defined as the sum of the reconstructed muon energy and the reconstructed hadronic energy. All of the following plots were produced using a statistically independent sample from the one used to find the fits.

For the QE population, Figure 9.77 is a 2D histogram which compares reconstructed neutrino energy to true neutrino energy. Figure 9.78 plots reconstructed neutrino energy vs. the relative energy resolution. Figure 9.79 is a 1D histogram of the relative energy resolution.

Figure 9.80 is a 2D histogram of the nonQE population which compares reconstructed neutrino energy to true neutrino energy. Figure 9.81 plots reconstructed neutrino energy vs. the relative energy resolution. One can also consider a 1D histogram of the relative energy resolution, Figure 9.82.

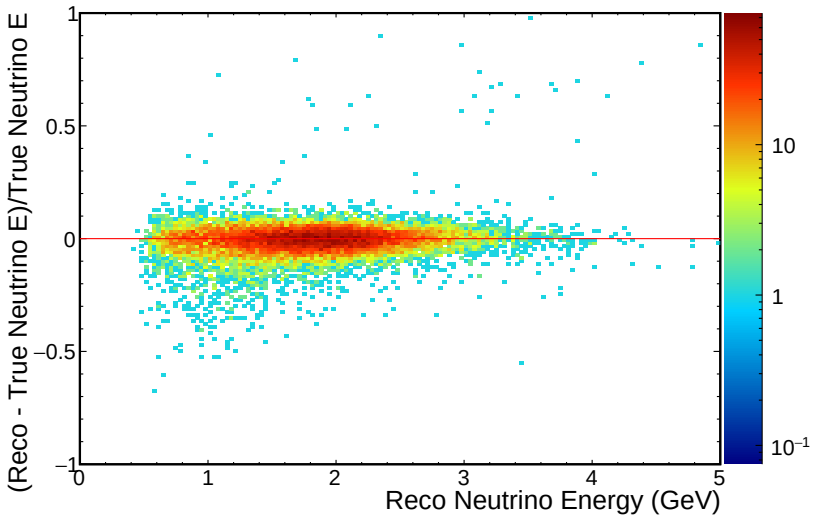


Figure 9.78: Histogram of reconstructed neutrino energy in GeV vs. the relative energy resolution for the QE population. Red line is flat at zero - if we did a perfect job, everything would lay on this line. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the near detector.

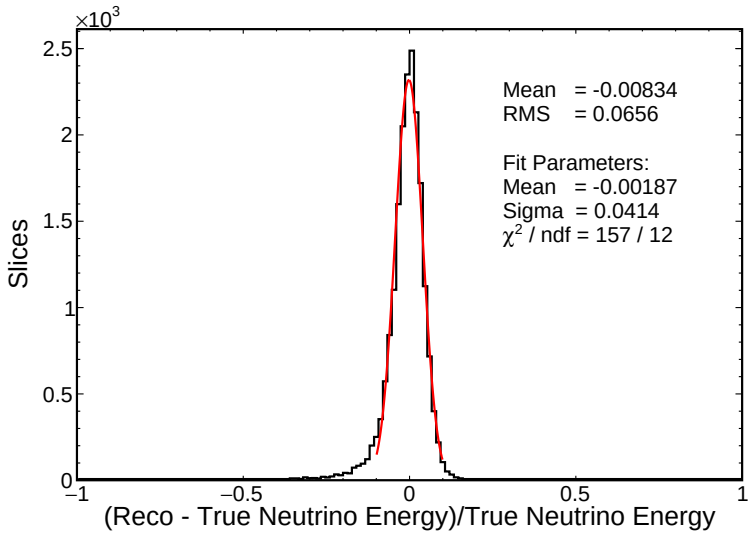


Figure 9.79: Relative energy resolution for reconstructed neutrino energy for the QE population. This plot only includes events with true neutrino energy less than 5 GeV. This plot was created using simulated events in the near detector. Red line is a Gaussian fit to the peak.

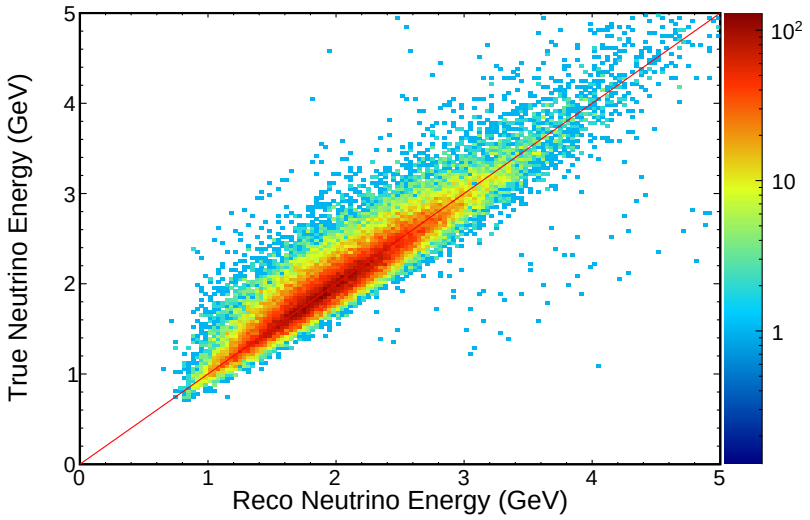


Figure 9.80: Histogram of reconstructed neutrino energy in GeV vs. true neutrino energy in GeV for the non-QE population. Red line is 45 degrees - if we did a perfect job, everything would lay on this line. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the near detector.

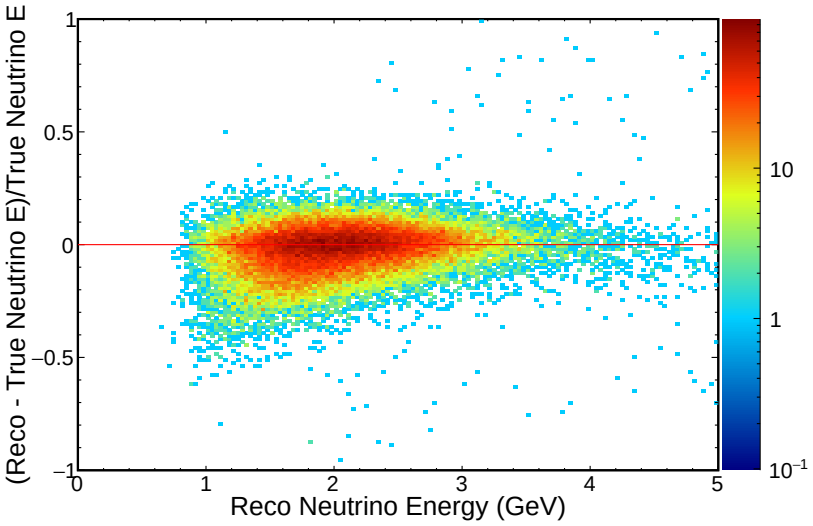


Figure 9.81: Histogram of reconstructed neutrino energy in GeV vs. the relative energy resolution for the non-QE population. Red line is flat at zero - if we did a perfect job, everything would lay on this line. The color axis is arbitrary number of interactions and is display logarithmically. This plot was created using simulated events in the near detector.

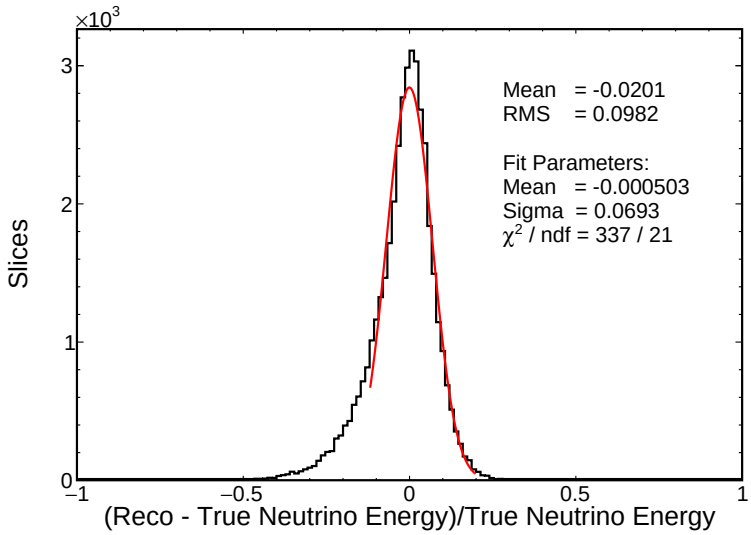


Figure 9.82: Relative energy resolution for reconstructed neutrino energy for the non-QE population. This plot only includes events with true neutrino energy less than 5 GeV. This plot was created using simulated events in the near detector. Red line is a Gaussian fit to the peak.

9.7 QE Energy using the Angle Formula

AngleQEE should only be used for QE interactions. It uses Equation 9.2 to relate the muon energy and the angle between the muon and the incoming neutrino to calculate the neutrino energy:

$$E_\nu = \frac{2M'_n - \left(M'^{\,2}_n + M^2_\mu + M^2_p\right)}{2\left(M'_n - E_\mu + \sqrt{E^2_\mu - M^2_\mu}\cos\theta\right)}, \tag{9.2}$$

where E_ν is the reconstructed neutrino energy, M'_n is the modified mass of the neutron ($M'_n = M_n - E_{binding}$), $E_{binding}$ is the binding energy of nucleus, M_μ is the mass of the muon, M_p is the mass of the proton, E_μ is the energy of the muon, and θ is the angle between the muon and the incoming neutrino. The muon track is chosen to be the 3D Kalman track with the highest ReMId value and its energy is reconstructed in the same manner as in TrkQEE (discussed in Section 9.1). The nuclear binding energy is set to be 25 MeV. This is consistent with what was used in the NOνA ν_μ QE cross-section analysis[62] and what GENIE uses in the simulation.

The AngleQEE energy is not as robust to reconstruction failures as TrkQEE. Some phase space of track length and angle can cause crazy energy values, including negative ones. Also, its resolution is limited by the Fermi momentum in the nucleus.

Chapter 10

Event Selection

Selection criteria were developed to increase sensitivity to the measurement of the oscillation parameters $\sin^2\theta_{23}$ and $|\Delta m^2_{32}|$. Basic quality checks were imposed to ensure the detector was performing normally and that the event was able to be reconstructed. Then, criteria designed to maximize the amount of signal (ν_μ CC interactions) and minimize the amount of background were applied. Two main sources of background exist: beam neutrino interactions that are not ν_μ CC interactions interactions and cosmic ray interactions. The beam neutrino interactions that are not ν_μ CC interactions are primarily NC interactions (Section 4.2). These are primarily removed by requiring a reconstructed track in the event to be identified as a muon. The cosmic ray interactions are a large background for the far detector, which is relatively close to the surface of the earth, but not the near detector, which is deep underground (Chapter 5). By requiring that the event occurred at the same time that the NuMI beam was on, we reduce the cosmic ray background by a factor of 10^5. Criteria applied only to the far detector events that require containment and alignment with the NuMI beam direction reduce this background by a further factor of 10^7. The final reconstructed neutrino energy spectra and event counts for the expected signal and background in the far detector is given in Chapter 11.

Good neutrino energy resolution is also necessary to make precision measurements of the oscillation parameters. This will be discussed further in Chapter 11. To ensure that all the energy of the neutrino interaction is visible, we applied containment criteria to both the near and far detector samples. Since the near detector has a muon catcher with

coarse energy resolution, criteria restricting the type of energy deposition in the muon catcher were developed. Finally, we divided the analysis population into two categories: a population of QE-like events and a population of nonQE-like events. QE events have better neutrino energy resolution than nonQE events. By dividing into two populations, we don't dilute the precise energy information but can still take advantage of the larger statistics of the nonQE population. This chapter gives a detailed description of all of the criteria used to select events.

10.1 File Selection

Data for the measurement of $\sin^2\theta_{23}$ and $|\Delta m^2_{32}|$ presented in this thesis corresponds to Run Periods 1, 2a and 2b[1]. Run Period 1 corresponded to when the far detector was in the process of being built. Far detector data was taken with partial detector configurations. For the first 9 months of this period, the near detector had not yet been constructed. For the final month of the period, the near detector was operational. The NuMI beam was shut down for three months for upgrades from the end of Run Period 1 to the start of Run Period 2. For Run Period 2, the far detector data mostly corresponded to the full detector configuration. Run Period 2a corresponded to 5 months. The near detector data from Run Period 2b, a period of 2 months, was not used due to time constraints.

For a data subrun file to be used, it must first pass basic criteria. It must have successfully been reconstructed through the CAF stage[2]. It must contain at least one neutrino beam spill. For the far detector, the amount of live detector must equal 4 diblocks or larger. This allows for containment of ν_μ CC interactions and reduces the ratio of cosmic background to contained signal events. Figure 10.1 contains diagrams showing which types of variable detector configurations were used. The far detector was

[1]Run Period 1 is from before the beam shutdown in fall 2014, between October 25th, 2013 to August 5th, 2014. For the near detector, the run period started July 18th, 2014. Run Period 2a occurred after the beam shutdown, from October 24th, 2014 until March 13th, 2015. Run Period 2b is the data from March 13th, 2015 to May 15th, 2015. Run Period 1 corresponds to far detector runs 11496-17078 and near detector runs 10377-10407. Run Period 2a corresponds to far detector runs 17967-19096 and near detector runs 10496-10824. Only far detector runs from Run Period 2b were used in this analysis; they corresponded to runs 19097-19586.

[2]Files can fail to be reconstructed for various reasons. Some events can exceed the allowed memory usage of the grid. Some files can fail to have proper information in the database.

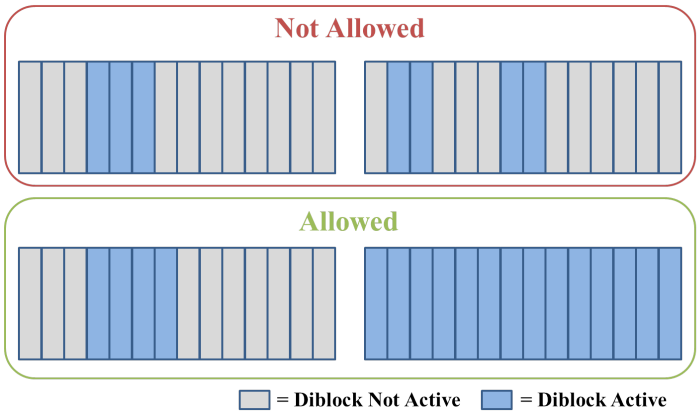


Figure 10.1: Diagram of different far detector configurations. A diblock is 1/14th of the total detector. In the diagram, blue indicates a live diblock and grey indicates a diblock not included in the readout. To be used for analysis, the detector must have had 4 or more contiguous diblocks live. Therefore, if only 3 diblocks are live (top left) or only two contiguous diblocks are live (top right), the data was not used. If 4 diblocks are live (bottom left) or, in the best case, the entire detector is live (bottom right), the data was used. If the detector had 4 or more contiguous diblocks live as well as another, not contiguous set of diblocks live, only the data from the largest contiguous portion would be used.

being constructed during the data taking period, so much of the data was taken with only a portion of the detector active. For the near detector, the entire detector must be live to be used for analysis. The near detector is too small to make partial detector configurations useful. The subrun must have passed the criteria applied to remove bad running conditions. For details on the data quality cuts applied, see [63],[64].

10.2 Spill Cuts

Once the subrun file is selected, one then selects beam spills which have good data.

The first category of far detector spill cuts are those made to ensure data quality. These are listed in Table 10.1. The first cut requires that every data concentrator module (DCM, for more information see Section 5.1) in the active detector has at least one hit during the time window. When DCMs receive too much information in too short of a time period, they can become overwhelmed and fail to send their data to the buffer nodes (Section 5.1) . This results in missing data that can cause the DCM to

have no hits. Secondly, the mask applied to create the active detector by Live Geometry must be valid[3]. The mask is used to define the boundaries of the active, usable portion of the far detector. These are needed since far detector data corresponds to variable detector configurations. The third cut ensures that the DCMs are synchronized relative to each other. To detect DCM synchronization failures, the DCM edge metric software examines all the cosmic tracks (reconstructed with the Cosmic Tracker, Section 7.5) within the readout time window. It determines when these tracks are near boundaries in the detector that are readout out by different DCMs and determines the fraction of tracks that successfully cross the DCM boundary. If all DCMs are not synchronized relative to one another, this metric has a low value.

These metrics are not all applied to the near detector. Live Geometry doesn't apply a mask to the near detector, so we can not check the status of the mask. We only use near detector data that corresponds to the full detector, so a mask is not needed. The DCM edge metric can not be applied to the near detector, because it relies on a high flux of cosmic rays. Since the near detector is underground, it doesn't have a high enough flux.

Instead, a different set of spill cuts is applied to the near detector data. A different software algorithm from the one used in the far detector is used to determine if any of the DCMs in the near detector are missing. Before the 2014 beam shutdown, the near detector was sometimes run in with the detector hall lights on. This allowed light to leak into the electronics and create many noise hits, which could confuse reconstruction attempts or obscure physics interactions. During the beam shutdown, extra light protection was added to the electronics to prevent this failure mode. A metric was developed to remove data with light pollution. A study was done to find cells affected by the light pollution. It looked for channels with high noise rates correlated with light-on running. The light pollution metric creates a ratio between the number of hits in affected cells not associated with the beam spill and the number of hits in the horizontal cells associated with the beam spill. For more information on the light pollution metric, see the technical note[65].

[3]Mask status must equal 1. If not, this could indicate that database problems caused an incorrect mask.

Table 10.1: Far Detector Data Quality Spill Cuts

Cut	Description
Number of Missing DCMs = 0	Require that all DCMs have at least one hit for the read-out time window
Good Mask = 1	Require Live Geometry have a valid diblock mask
DCM Edge Metric > 0.2	The DCMs are synchronized relative to each other

Table 10.2: Near Detector Data Quality Spill Cuts

Cut	Description
Number of Missing DCMs = 0	Require that all DCMs have at least one hit for the time window
Light Pollution Metric > 0.45	The electronic readout didn't have excess light pollution

The second type of spill cuts that can be made relate to the quality of the beam. These are listed in Table 10.3. First, the trigger time recorded by the detector must be within 0.5 sec of a NuMI beam trigger time listed in the Intensity Frontier beam DataBase (IFDB).[4] Since the NuMI beam typically has a spill roughly every second, this cut ensures that we have properly matched the detector data to the beam spill. The protons on target (POT) for the spill must be greater than 2×10^{12}. The current inside the focusing horns[5] must be within the acceptable range of current, given in Table 10.3. The range was determined by looking at values of the current corresponding to normal running. All of the normal beam running used in the study was within the cut range. Also, the beam position[6] must at the expected location and with the expected width. Again, the cut values listed in Table 10.3 were determined by examining normal beam running and placing cuts to encompass the bulk of the normal running period. For more information, see this document[66].

[4]This database is filled by Fermilab every time the accelerator sends POT to the NuMI beam.
[5]Calculated using the sum of 4 calibrated stripline values: NSLINA, NSLINB, NSLINC, and NSLIND
[6]As measured by the beam position monitors

Table 10.3: Beam Quality Spill Cuts

Cut	Description
$\lvert\Delta SpillTime\rvert < 0.5$ sec	Time of trigger recorded on detector to be within 0.5 sec of a trigger time in IFDB
POT of spill $> 2 \times 10^{12}$	Protons on target for the spill must be greater than 2×10^{12}
$-202 < Horn\ Current < -198$	The current within the focusing horns must be between -202 and -198 kAmps
$0.02 < X\ Position < 2.00$	The horizontal position of the beam must be between 0.02 and 2.00 mm
$0.02 < Y\ Position < 2.00$	The vertical position of the beam must be between 0.02 and 2.00 mm
$0.57 < X\ Width < 1.58$	The horizontal width of the beam must be between 0.57 and 1.58 mm
$0.57 < Y\ Width < 1.58$	The vertical height of the beam must be between 0.57 and 1.58 mm

10.3 Slice Cuts

The next level of cuts is on the slice level. Slices correspond to individual physics interactions; for more detail, see Section 7.3. These cuts have three main purposes: first, to ensure that all the necessary reconstruction, particle identification, and energy information exists in the CAF file and is valid[7]; second, to ensure that the interaction is contained within the detector; and third, to reject cosmic ray backgrounds. For an analysis that uses an uncontained sample, see Raddatz's thesis[52]. These cuts vary somewhat between the far and near detectors; for instance, the cosmic ray background in the near detector is so low that no cosmic rejection cuts need to be applied. Section 10.3.1 discusses the slice cuts for the far detector population; Section 10.3.2 discusses the slice cuts for the near detector.

[7]For instance, a low energy event could fail to have any tracks reconstructed. Then we would not be able to identify if a track was a muon or create a reconstructed neutrino energy.

10.3.1 Far Detector Slice Cuts

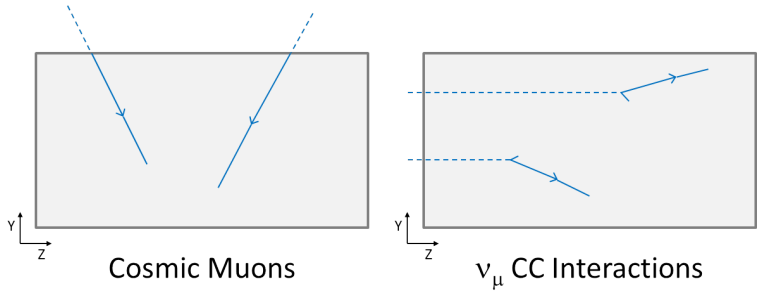
This analysis defines two populations: those that have downgoing primary Kalman tracks (called the *negative population*) and those that have upgoing primary Kalman tracks (called the *positive population*). For more information about Kalman tracks, see Section 7.4. The primary Kalman track for each slice is defined as the 3D Kalman track with the highest ReMId value, that is, the track most likely to be a muon. For more information about ReMId, see Section 7.6. Two populations are created to more effectively cut the cosmic muon background. See Figure 10.2 for diagrams that define these populations and Figure 10.3 for a plot of the populations.

Cosmic ray muons primarily start at the top of the detector and travel downwards. Muons created by NuMI neutrino interactions, instead, tend to start at the front of the detector and travel to the back of the detector. For this reason, KalmanTrack always defines the most upstream Z position as the start of the track and the most downstream Z position as the end of the track. For cosmic ray muons, however, this definition of start and end of the track is wrong roughly half of the time. For downward-going cosmic ray muons, the negative population corresponds to cases where the KalmanTrack definitions of start and end of the track are probably correct. The positive population corresponds to cases where the background muons probably have start and end of the tracks opposite to that which KalmanTrack defines. If one, for instance, were to cut on the true start position of the cosmic background, one would cut on the start Z position of the negative population and the end Z position of the positive population. See Figure 10.2 for diagrams that define these populations.

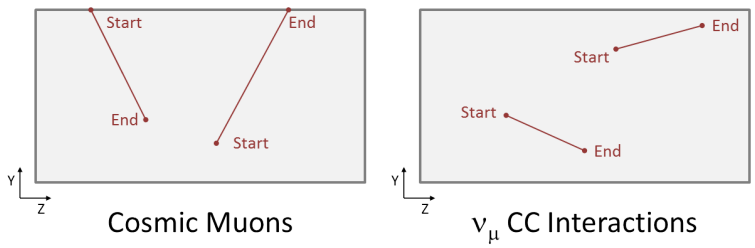
Slice cuts which are applied to both populations are listed in Table 10.4. Cuts which are applied only to the negative population are listed in Table 10.5, while those applied only to the positive population are listed in Table 10.6.

To take advantage of the fact that we know when the NuMI beam occurs relative to the triggered time window, we cut on the mean time of the slice to be within the window of the beam spill. The readout window is typically 500 μs long and the beam spill is 10 μs long.[8] For the early running period, a malfunction in the timing system caused the beam spill to have the possibility of being shifted $+64\mu s$ relative to the readout

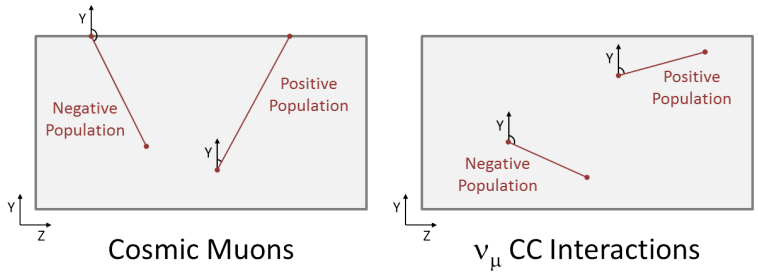
[8] We add 2 μs of padding to the beam spill cut to ensure it is fully contained within our cut.



(a) Common topologies for cosmic ray muons, on the left, and ν_μ CC interactions, on the right.



(b) Red lines represent reconstructed Kalman tracks for the topologies in Figure 10.2a. The start and end positions defined by KalmanTrack are labeled.



(c) The Kalman tracks from Figure 10.2b are sorted into negative and positive populations, based on the angle between the Y detector axis and the track.

Figure 10.2: Diagram of common topologies. The left side of Figure 10.2a shows two cosmic ray muons, entering from the top of the detector. The right side of Figure 10.2a shows two cases of ν_μ CC interactions. Figure 10.2b displays the reconstructed tracks KalmanTrack would create, given the underlying topologies. KalmanTrack always defines the start of the track as the side with the lowest Z position. The end is defined as the side with the largest Z position. Note that the choice of start and end is incorrect for the cosmic ray muon on the right. Figure 10.2c displays the population definitions for each track.

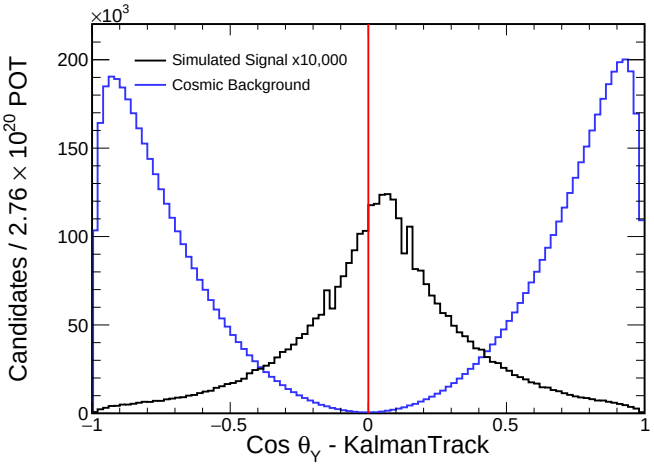


Figure 10.3: Plot that defines the negative and positive populations. The horizontal axis displays $\cos\theta_Y$, where θ_Y is the angle between the primary 3D Kalman track and the vertical detector axis. A red line is drawn at $\cos\theta_Y = 0$. Candidates with $\cos\theta_Y > 0$ are defined as the positive population; otherwise, they are part of the negative population. Simulated signal is shown in black; cosmic background data is shown in blue. The simulated signal is unoscillated. Each population is scaled to the expected levels for 2.76×10^{20} POT. Note that the simulated signal is multiplied by 10,000 to make it visible on this scale. Basic quality cuts requiring a 3D Kalman track, at least one valid ReMId object, a valid neutrino energy, and a valid Live Geometry mask are applied. Data quality cuts requiring no missing DCMs and that the detector is synced are applied. The primary Kalman track is required to have at least 5 hits and have a ReMId value greater than 0.5.

window. Therefore, for these runs, we have two beam spill windows. This doubles the cosmic ray background for this period. For more information on the malfunction, see the summary by Messier[67].

For basic quality cuts, we required that the slice has at least one 3D Kalman track, at least one Cosmic track, at least one valid ReMId object[9], and a valid neutrino energy[10]. These are all required because later cuts make use of their presence. To ensure that the reconstructed track is not an artifact, we require that the primary Kalman track have more than five hits.

[9]Vector index isn't equal to 999, a default value used when no ReMId objects are found. This might happen, for instance, if no 3D Kalman track existed.

[10]Neutrino energy isn't equal to -5, a default value used when no neutrino energy could successfully be calculated. This might happen, for instance, if no 3D Kalman track existed.

Table 10.4: Far Detector Slice Cuts

Cut	Description
218.1 <Slice Time < 230.1	Require slice time, in μs relative to the start of the readout window, coincident with beam spill[11]
Number of 3D Kalman Tracks $\neq 0$	Require at least one 3D Kalman track
Number of Cosmic Tracks $\neq 0$	Require at least one Cosmic track
Number of ReMId Objects $\neq 0$	Require at least one ReMId value
Best ReMId Index $\neq 999$	Require the vector index of highest ReMId track is valid
Neutrino Energy $\neq$ -5	Require a valid neutrino energy value
No. Track Hits > 5	Track has more than 5 hits
$\frac{No.\ Slice\ Hits}{No.\ Track\ Planes} < 3.0$	The ratio of the number of hits in the slice to the number of planes that the track crosses is less than 3.0
$\lvert StartX\rvert < 740$	Track x start position absolute value less than 740 cm
$StartY > -750$	Track y start position greater than -750 cm
$\lvert EndX\rvert < 740$	Track x end position absolute value less than 740 cm
$EndY > -750$	Track y end position greater than 750 cm
$\lvert StartOtherX\rvert < 740$	For all other Kalman tracks within the slice, track x start position absolute value less than 740 cm
$-750 < StartOtherY < 725$	For all other Kalman tracks within the slice, track y start position greater than -750 cm and less than 725 cm
$StartOtherZ > Det.\ Front + 20$	For all other Kalman tracks within the slice, track z start position greater than 20 cm from the front face of the active detector
$\lvert EndOtherX\rvert < 740$	For all other Kalman tracks within the slice, track x end position absolute value less than 740 cm

Continued on next page

[11] For runs numbers less than 17500, slice times between 282.1 and 294.1 μs are also allowed

Table 10.4 – continued from previous page

Cut	Description
$-750 < EndOtherY < 725$	For all other Kalman tracks within the slice, track y end position greater than -750 cm and less than 725 cm
$EndOtherZ > Det.\ Back - 20$	For all other Kalman tracks within the slice, track z end position greater than 20 cm from the front face of the active detector
Transverse Momentum Frac. < 0.65	The fraction of the vector of the mean position of the slice relative to the start position of the track transverse to the NuMI beam direction
Cosmic Track Proj. Back Dist. > 50	Projected backwards distance of the track from CosmicTrack is more than 50 cm from the detector edge

Table 10.5: Far Detector Negative Population Slice Cuts

Cut	Description
$DirY < 0$	Track initial direction has a negative cosine with respect to the y axis
$StartY < 650$	Track y start position less than 650 cm
$StartZ > Det.\ Front + 40$	Track z start position greater than 40 cm from the front face of the active detector
$EndY < 725$	Track y end position less than 725 cm
$EndZ < Det.\ Back - 20$	Track z end position less than 20 cm from the back face of the active detector
Cosmic Track Proj. Back Dist. and Comic Track $DirBeam$ not within certain values	The values of the projected backwards distance of the track from CosmicTrack and the initial track direction cosine with respect to the NuMI beam direction doesn't fall within a forbidden region.[12]

[12] In the 2D space of the CosmicTrack parameters of $BeamDir$ and backwards projected distance, make a rectangle cut: $0.2 < BeamDir < 0.65$ and backwards projected distance < 725 cm. If the track has values within the box, then consider parameter y, where $y = 0.65 - (9.0/14500.0) *$projected backwards distance. If $y > BeamDir$, then the track falls into the forbidden region. Another forbidden region is defined by the box cut of $-0.3 < BeamDir < 0.2$ and backwards projected distance < 1500 cm

Table 10.6: Far Detector Positive Population Slice Cuts

Cut	Description
$DirY \geq 0$	Track initial direction has a positive cosine with respect to the y axis
$StartY < 725$	Track y start position less than 725 cm
$StartZ > Det.\ Front + 20$	Track z start position greater than 20 cm from the front face of the active detector
$EndY < 650$	Track y end position less than 650 cm
$EndZ < Det.\ Back - 40$	Track z end position less than 40 cm from the back face of the active detector
Cosmic Track Proj. Back Dist. and Comic Track $DirBeam$ not within certain values	The values of the projected backwards distance of the track from CosmicTrack and the initial track direction cosine with respect to the NuMI beam direction doesn't fall within a forbidden region.[13]

A cut is introduced to remove background due to front end board (FEB; for more information, see Section 5.1) ringing. When a very high-energy particle travels through the detector, it can saturate the avalanche photodiode (APD; for more information, see Section 5.1). As the APD recovers from this saturation, false hits on other pixels can be created. The time scale of these false hits is 10's of μsec. These false hits can create slices that do not include the initial high-energy particle. Slices of false hits often look contained and often have many hits on each APD. These can sometimes mimic nonQE interactions. To remove these slices, one calculates the ratio of the number of slice hits to the number of planes crossed by the primary Kalman track. This ratio is defined as the *flasher metric*. If the ratio is greater than or equal to 3, it can be a case of FEB ringing. Figure 10.4 shows the effect of this cut.

[13] In the 2D space of the CosmicTrack parameters of $BeamDir$ and backwards projected distance, make a rectangle cut: $-0.85 < BeamDir < -0.2$ and backwards projected distance < 725 cm. If the track has values within the box, then consider parameter y, where $y = -0.85 - (13.0/14500.0) *$projected backwards distance. If $y < BeamDir$, then the track falls into the forbidden region. Another forbidden region is defined by the box cut of $-0.2 < BeamDir < 0.2$ and backwards projected distance < 1500 cm

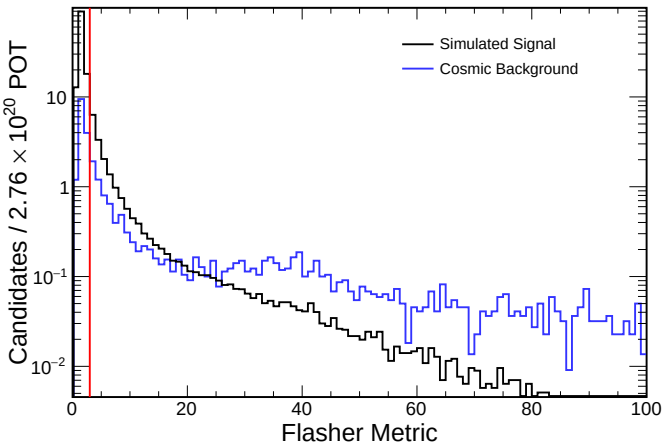
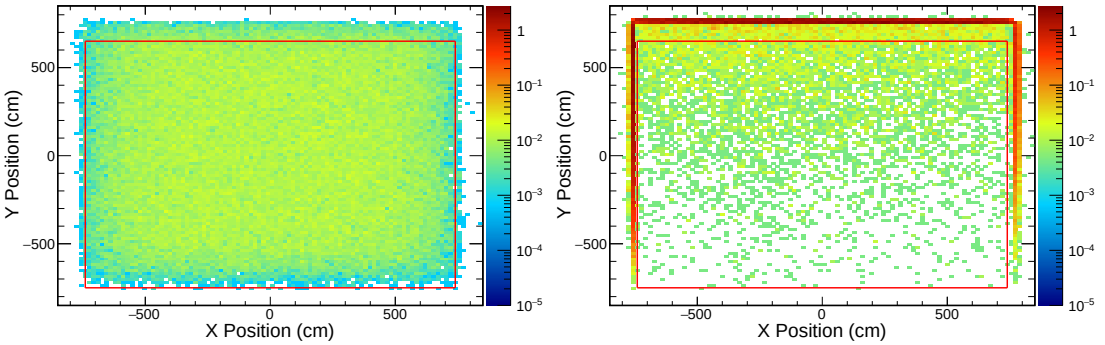


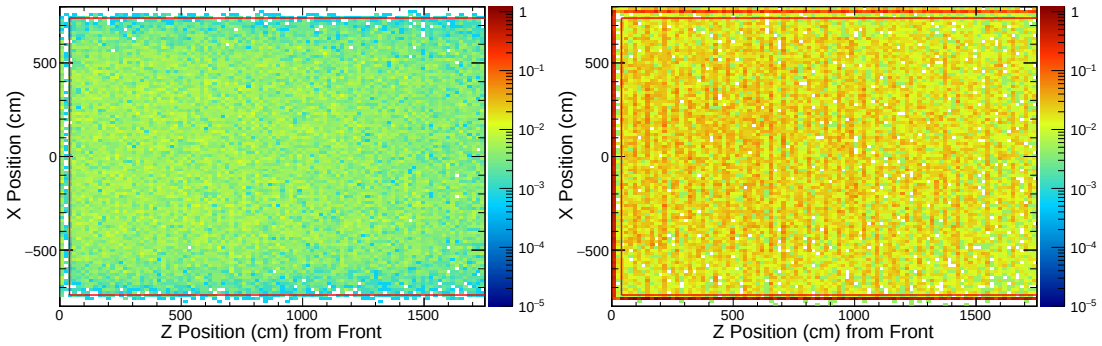
Figure 10.4: Plot of the number of hits in the slice divided by the number of planes that the primary Kalman track crossed. Note the logarithmic vertical axis. A red line is drawn at $\frac{No.\ Slice\ Hits}{No.\ Track\ Planes} = 3$. Candidates with values greater than 3 are rejected. Simulated signal is shown in black; cosmic background data is shown in blue. The simulated signal is unoscillated. Each population is scaled to the expected levels for 2.76×10^{20} POT. Data quality cuts requiring no missing DCMs and that the detector is synced are applied. The primary Kalman track is required to have a ReMId value greater than 0.5. All cuts listed in Tables 10.4, 10.5, and 10.6, except for the cut on $\frac{No.\ Slice\ Hits}{No.\ Track\ Planes}$, are applied.



(a) Simulated Signal (b) Cosmic Background

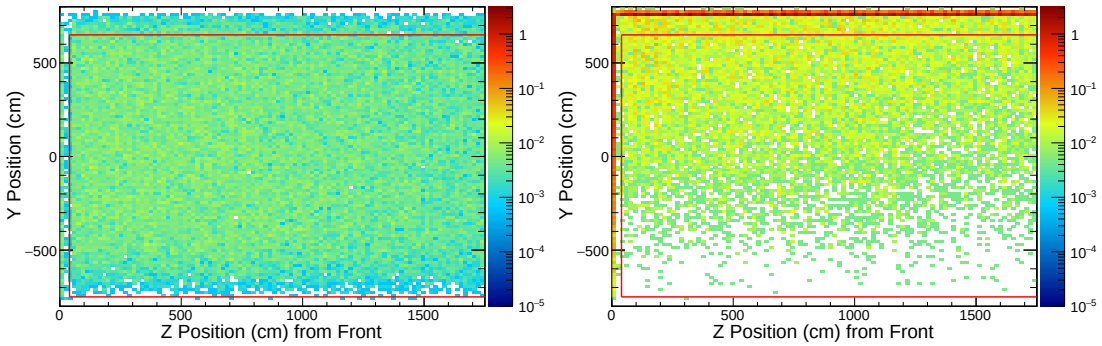
Figure 10.5: Plot of the primary Kalman track start position in detector X and Y coordinates for the negative population. Note the logarithmic color axis. Red lines are drawn at the cut location for the start position; candidates outside the red box are rejected. Simulated signal is shown on the left; cosmic background data is shown on the right. The simulated signal is unoscillated. Each population is scaled to the expected levels for 2.76×10^{20} POT. Data quality cuts requiring no missing DCMs and that the detector is synced are applied. The primary Kalman track is required to have a ReMId value greater than 0.5. All cuts listed in Tables 10.4 and 10.5, except for cuts on the primary Kalman track start position, are applied.

To eliminate background from cosmic rays and ensure that the interaction is well contained within the detector, a series of cuts are applied to the start and end position of the primary Kalman track. The y and z cuts differ somewhat between the two populations, for the reasons described earlier in this section. Figures 10.5, 10.6, and 10.7 show the effect of cutting on primary Kalman track start position for the negative population. Figures 10.8, 10.9, and 10.10 display the same information, but for the positive population. Figures 10.11, 10.12, and 10.13 show the effect of cutting on the primary Kalman track end position for the negative population. Figures 10.14, 10.15, and 10.16 display the same information, but for the positive population.



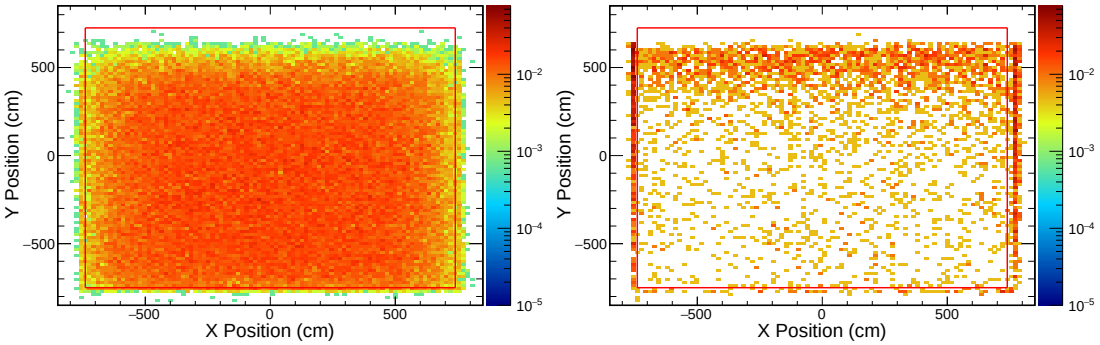
(a) Simulated Signal (b) Cosmic Background

Figure 10.6: Plot of the primary Kalman track start position for the negative population. The horizontal axis is distance, in cm, from the front of the active detector. The vertical axis is the detector X coordinate. Note the logarithmic color axis. Red lines are drawn at the cut location for the start position; candidates outside the red lines are rejected. Simulated signal is shown on the left; cosmic background data is shown on the right. The simulated signal is unoscillated. Each population is scaled to the expected levels for 2.76×10^{20} POT. Data quality cuts requiring no missing DCMs and that the detector is synced are applied. The primary Kalman track is required to have a ReMId value greater than 0.5. All cuts listed in Tables 10.4 and 10.5, except for cuts on the primary Kalman track start position, are applied.



(a) Simulated Signal (b) Cosmic Background

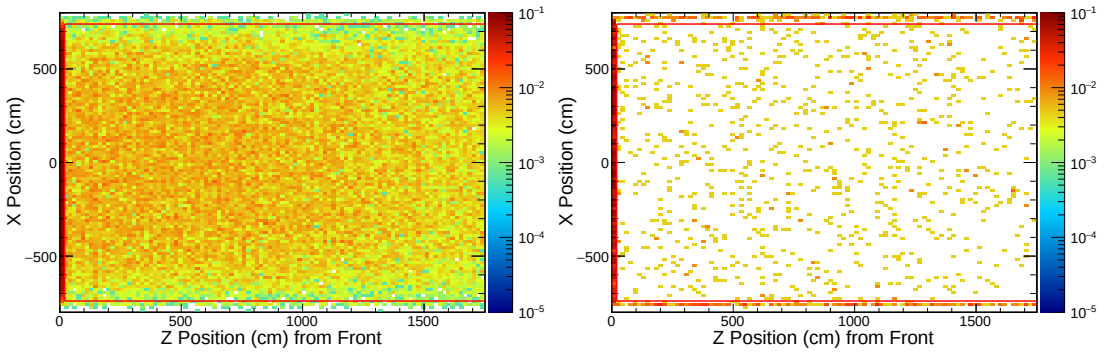
Figure 10.7: Plot of the primary Kalman track start position for the negative population. The horizontal axis is distance, in cm, from the front of the active detector. The vertical axis is the detector Y coordinate. Note the logarithmic color axis. Red lines are drawn at the cut location for the start position; candidates outside the red lines are rejected. Simulated signal is shown on the left; cosmic background data is shown on the right. The simulated signal is unoscillated. Each population is scaled to the expected levels for 2.76×10^{20} POT. Data quality cuts requiring no missing DCMs and that the detector is synced are applied. The primary Kalman track is required to have a ReMId value greater than 0.5. All cuts listed in Tables 10.4 and 10.5, except for cuts on the primary Kalman track start position, are applied.



(a) Simulated Signal

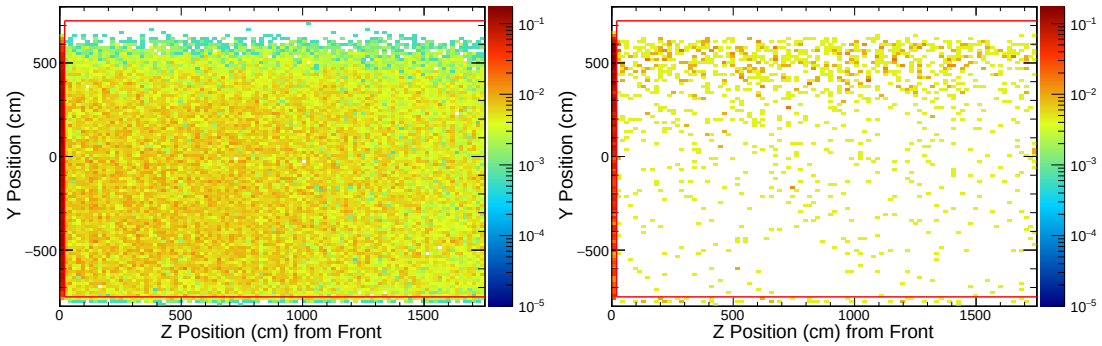
(b) Cosmic Background

Figure 10.8: Plot of the primary Kalman track start position in detector X and Y coordinates for the positive population. Note the logarithmic color axis. Red lines are drawn at the cut location for the start position; candidates outside the red box are rejected. Simulated signal is shown on the left; cosmic background data is shown on the right. The simulated signal is unoscillated. Each population is scaled to the expected levels for 2.76×10^{20} POT. Data quality cuts requiring no missing DCMs and that the detector is synced are applied. The primary Kalman track is required to have a ReMId value greater than 0.5. All cuts listed in Tables 10.4 and 10.6, except for cuts on the primary Kalman track start position, are applied.



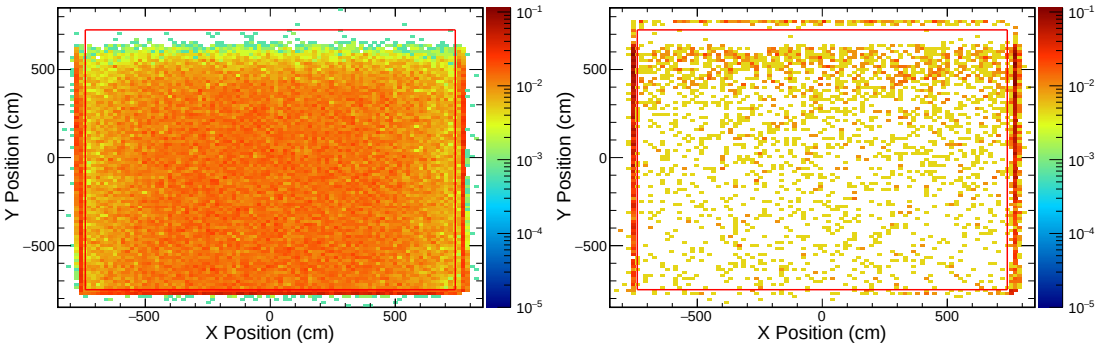
(a) Simulated Signal (b) Cosmic Background

Figure 10.9: Plot of the primary Kalman track start position for the positive population. The horizontal axis is distance, in cm, from the front of the active detector. The vertical axis is the detector X coordinate. Note the logarithmic color axis. Red lines are drawn at the cut location for the start position; candidates outside the red lines are rejected. Simulated signal is shown on the left; cosmic background data is shown on the right. The simulated signal is unoscillated. Each population is scaled to the expected levels for 2.76×10^{20} POT. Data quality cuts requiring no missing DCMs and that the detector is synced are applied. The primary Kalman track is required to have a ReMId value greater than 0.5. All cuts listed in Tables 10.4 and 10.6, except for cuts on the primary Kalman track start position, are applied.



(a) Simulated Signal (b) Cosmic Background

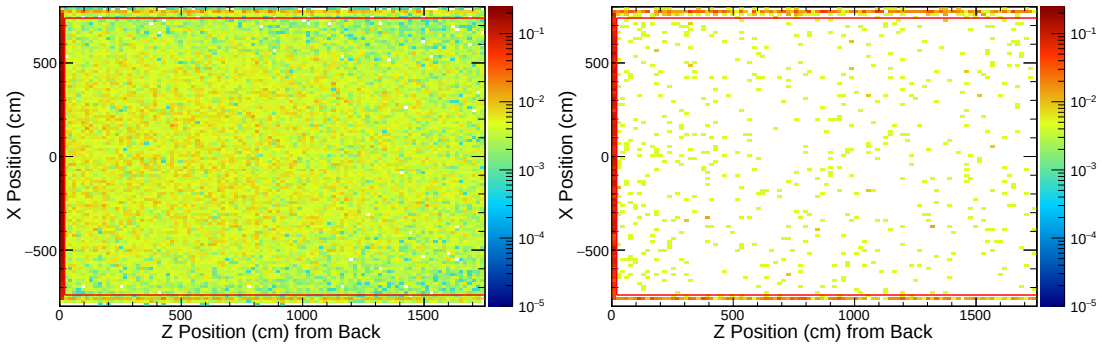
Figure 10.10: Plot of the primary Kalman track start position for the positive population. The horizontal axis is distance, in cm, from the front of the active detector. The vertical axis is the detector Y coordinate. Note the logarithmic color axis. Red lines are drawn at the cut location for the start position; candidates outside the red lines are rejected. Simulated signal is shown on the left; cosmic background data is shown on the right. The simulated signal is unoscillated. Each population is scaled to the expected levels for 2.76×10^{20} POT. Data quality cuts requiring no missing DCMs and that the detector is synced are applied. The primary Kalman track is required to have a ReMId value greater than 0.5. All cuts listed in Tables 10.4 and 10.6, except for cuts on the primary Kalman track start position, are applied.



(a) Simulated Signal

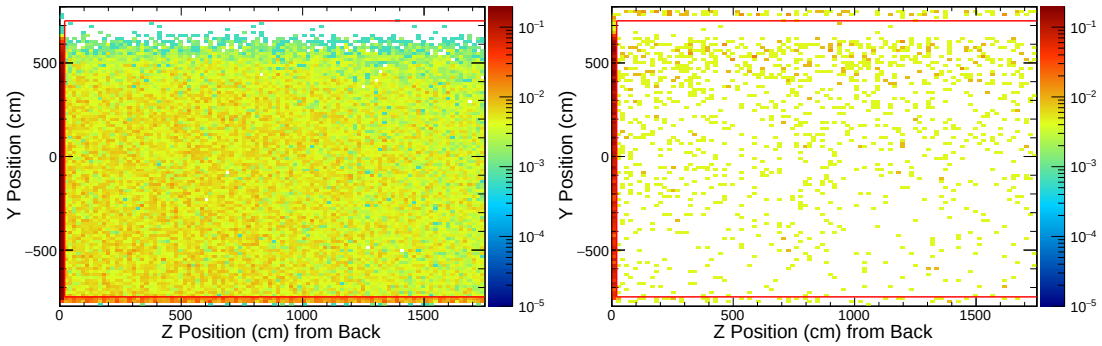
(b) Cosmic Background

Figure 10.11: Plot of the primary Kalman track end position in detector X and Y coordinates for the negative population. Note the logarithmic color axis. Red lines are drawn at the cut location for the end position; candidates outside the red box are rejected. Simulated signal is shown on the left; cosmic background data is shown on the right. The simulated signal is unoscillated. Each population is scaled to the expected levels for 2.76×10^{20} POT. Data quality cuts requiring no missing DCMs and that the detector is synced are applied. The primary Kalman track is required to have a ReMId value greater than 0.5. All cuts listed in Tables 10.4 and 10.5, except for cuts on the primary Kalman track end position, are applied.



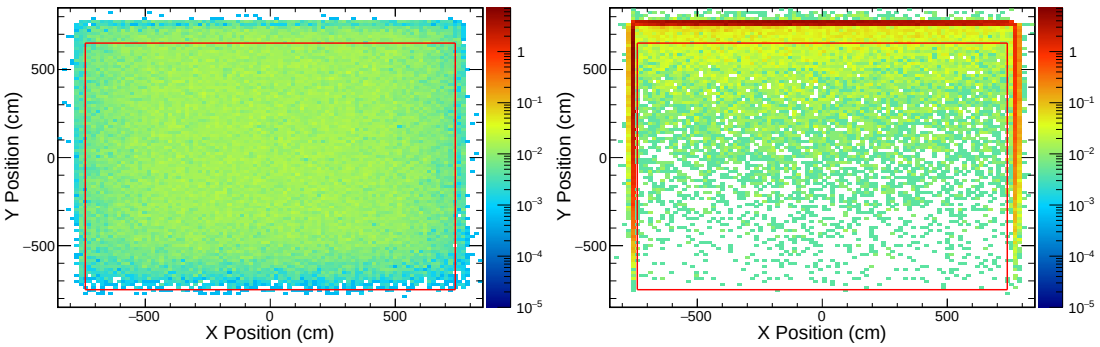
(a) Simulated Signal (b) Cosmic Background

Figure 10.12: Plot of the primary Kalman track end position for the negative population. The horizontal axis is distance, in cm, from the back of the active detector. The vertical axis is the detector X coordinate. Note the logarithmic color axis. Red lines are drawn at the cut location for the end position; candidates outside the red lines are rejected. Simulated signal is shown on the left; cosmic background data is shown on the right. The simulated signal is unoscillated. Each population is scaled to the expected levels for 2.76×10^{20} POT. Data quality cuts requiring no missing DCMs and that the detector is synced are applied. The primary Kalman track is required to have a ReMId value greater than 0.5. All cuts listed in Tables 10.4 and 10.5, except for cuts on the primary Kalman track end position, are applied.



(a) Simulated Signal (b) Cosmic Background

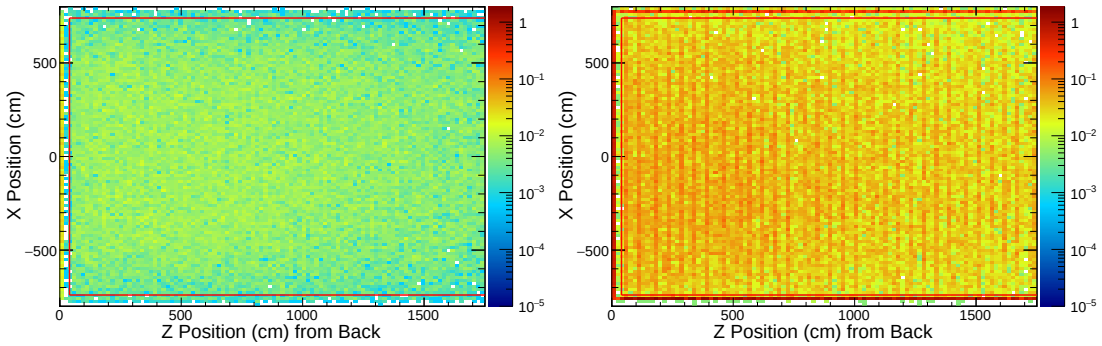
Figure 10.13: Plot of the primary Kalman track end position for the negative population. The horizontal axis is distance, in cm, from the back of the active detector. The vertical axis is the detector Y coordinate. Note the logarithmic color axis. Red lines are drawn at the cut location for the end position; candidates outside the red lines are rejected. Simulated signal is shown on the left; cosmic background data is shown on the right. The simulated signal is unoscillated. Each population is scaled to the expected levels for 2.76×10^{20} POT. Data quality cuts requiring no missing DCMs and that the detector is synced are applied. The primary Kalman track is required to have a ReMId value greater than 0.5. All cuts listed in Tables 10.4 and 10.5, except for cuts on the primary Kalman track end position, are applied.



(a) Simulated Signal

(b) Cosmic Background

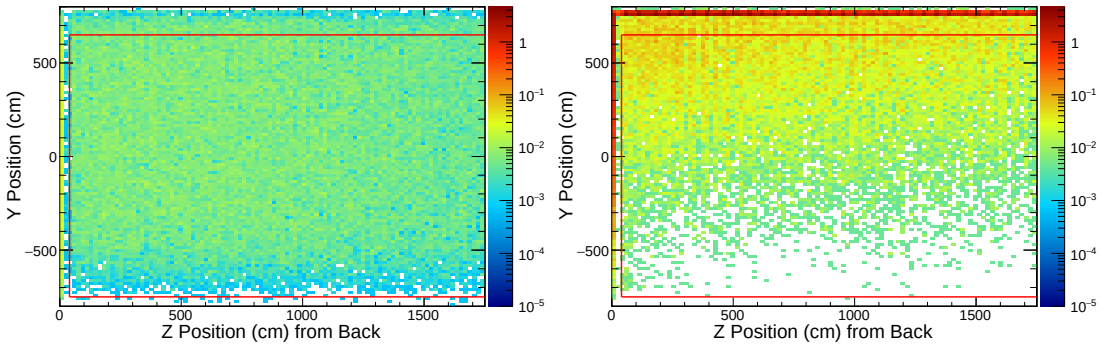
Figure 10.14: Plot of the primary Kalman track end position in detector X and Y coordinates for the positive population. Note the logarithmic color axis. Red lines are drawn at the cut location for the end position; candidates outside the red box are rejected. Simulated signal is shown on the left; cosmic background data is shown on the right. The simulated signal is unoscillated. Each population is scaled to the expected levels for 2.76×10^{20} POT. Data quality cuts requiring no missing DCMs and that the detector is synced are applied. The primary Kalman track is required to have a ReMId value greater than 0.5. All cuts listed in Tables 10.4 and 10.6, except for cuts on the primary Kalman track end position, are applied.



(a) Simulated Signal

(b) Cosmic Background

Figure 10.15: Plot of the primary Kalman track end position for the positive population. The horizontal axis is distance, in cm, from the back of the active detector. The vertical axis is the detector X coordinate. Note the logarithmic color axis. Red lines are drawn at the cut location for the end position; candidates outside the red lines are rejected. Simulated signal is shown on the left; cosmic background data is shown on the right. The simulated signal is unoscillated. Each population is scaled to the expected levels for 2.76×10^{20} POT. Data quality cuts requiring no missing DCMs and that the detector is synced are applied. The primary Kalman track is required to have a ReMId value greater than 0.5. All cuts listed in Tables 10.4 and 10.6, except for cuts on the primary Kalman track end position, are applied.



(a) Simulated Signal (b) Cosmic Background

Figure 10.16: Plot of the primary Kalman track end position for the positive population. The horizontal axis is distance, in cm, from the back of the active detector. The vertical axis is the detector Y coordinate. Note the logarithmic color axis. Red lines are drawn at the cut location for the end position; candidates outside the red lines are rejected. Simulated signal is shown on the left; cosmic background data is shown on the right. The simulated signal is unoscillated. Each population is scaled to the expected levels for 2.76×10^{20} POT. Data quality cuts requiring no missing DCMs and that the detector is synced are applied. The primary Kalman track is required to have a ReMId value greater than 0.5. All cuts listed in Tables 10.4 and 10.6, except for cuts on the primary Kalman track end position, are applied.

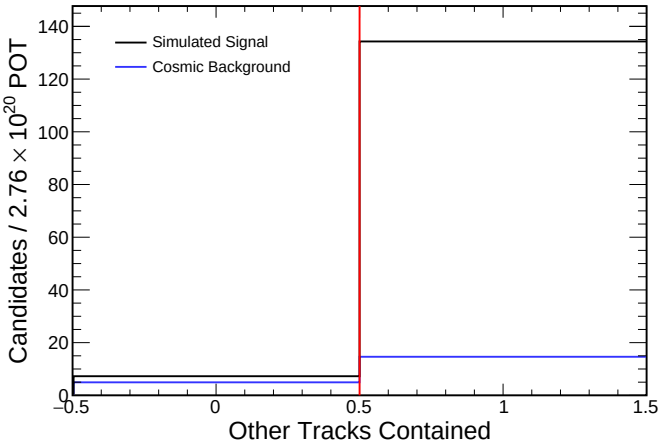


Figure 10.17: Plot of the number of candidates that have uncontained secondary tracks (plotted with value 0) and have no uncontained secondary tracks (plotted with value 1). A red line is drawn at between the two values; those candidates with value 0 are rejected. Simulated signal is shown in black; cosmic background data is shown in blue. The simulated signal is unoscillated. Each population is scaled to the expected levels for 2.76×10^{20} POT. Data quality cuts requiring no missing DCMs and that the detector is synced are applied. The primary Kalman track is required to have a ReMId value greater than 0.5. All cuts listed in Tables 10.4, 10.5 and 10.6, except for the cut requiring secondary tracks are contained, are applied.

For well-contained neutrino interactions, no Kalman tracks from the interaction should touch the detector edge. Therefore, we require that all 2D and 3D Kalman tracks associated with the slice pass containment cuts on their start and end positions. This also helps to remove a particular type of reconstruction failure for cosmic muons. Sometimes a cosmic muon is reconstructed as multiple tracks. If an interior section has the highest ReMId value, it can look contained. However, by requiring that all the Kalman tracks within the slice are contained, this failure mode can be rejected. Figure 10.17 displays the power of this cut.

ν_μ CC interactions are usually well-aligned with the beam direction. Cosmic ray background is typically not well-aligned with the beam direction. The fraction of transverse momentum of a ν_μ CC interaction relative to the beam should not be large. To calculate the fraction of transverse momentum, we first make a unit vector that corresponds to the difference between the average position of the slice and the start position

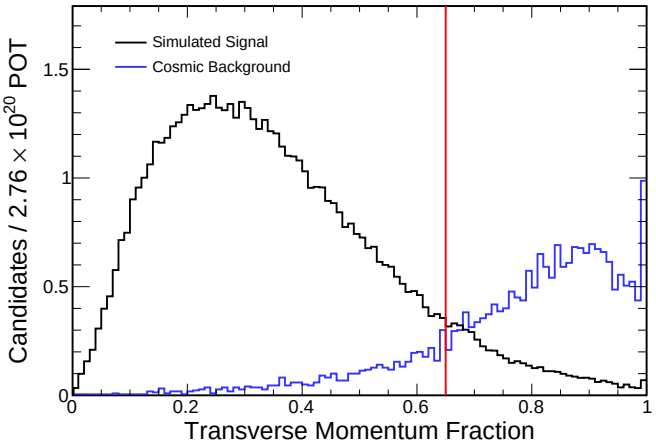


Figure 10.18: Plot of the transverse momentum fraction for the negative population. A red line is drawn at 0.65. Candidates with values more than 0.65 are rejected. Simulated signal is shown in black; cosmic background data is shown in blue. The simulated signal is unoscillated. Each population is scaled to the expected levels for 2.76×10^{20} POT. Data quality cuts requiring no missing DCMs and that the detector is synced are applied. The primary Kalman track is required to have a ReMId value greater than 0.5. All cuts listed in Tables 10.4 and 10.5, except for the cut on the transverse momentum fraction, are applied.

of the primary Kalman track. The sine of this vector relative to the beam direction represents the fraction of the momentum of the visible interaction which is transverse to the beam. We require that this fraction is less than 0.65. Figure 10.18 shows the effect of this cut for the negative population; Figure 10.19 shows the effect of this cut for the positive population.

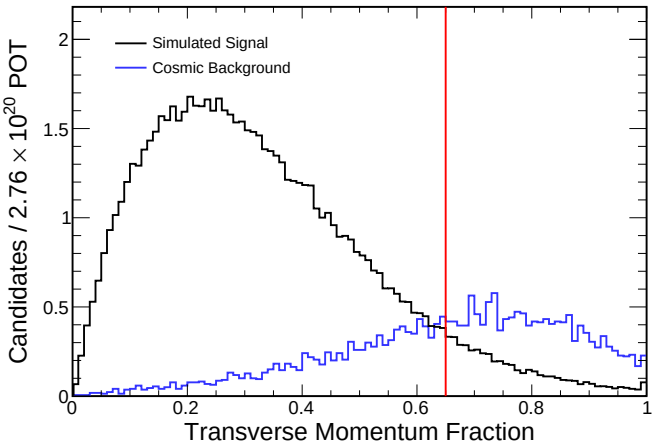
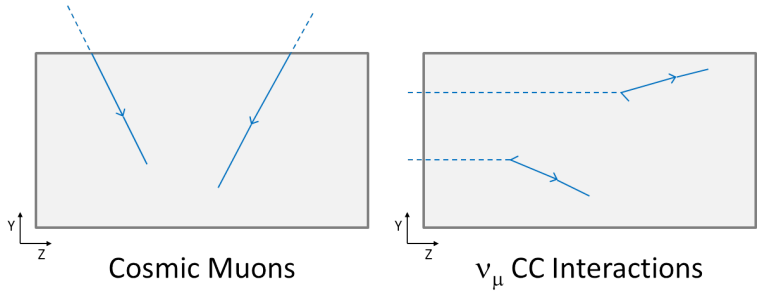
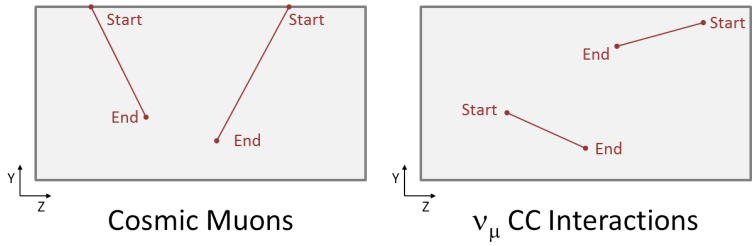


Figure 10.19: Plot of the transverse momentum fraction for the positive population. A red line is drawn at 0.65. Candidates with values more than 0.65 are rejected. Simulated signal is shown in black; cosmic background data is shown in blue. The simulated signal is unoscillated. Each population is scaled to the expected levels for 2.76×10^{20} POT. Data quality cuts requiring no missing DCMs and that the detector is synced are applied. The primary Kalman track is required to have a ReMId value greater than 0.5. All cuts listed in Tables 10.4 and 10.6, except for the cut on the transverse momentum fraction, are applied.

Cosmic rejection can be improved by using information from a second, independent tracking algorithm. Any tracking algorithm can have failure modes; by using two different algorithms which have different strengths and weaknesses, one is more robust to the failure modes of any particular tracker. For this analysis, we use information from CosmicTrack (Section 7.5) as well as KalmanTrack (Section 7.4). CosmicTrack creates, at most, one 3D track per slice. The track is entirely linear and does not follow curves. The end closest to the top of the detector is defined as the start of the track; the end farthest from the top of the detector is defined as the end of the track. For cosmic ray tracks, these definitions of start and end of the track are usually correct. Figure 10.20 diagrams how CosmicTrack defines track start and end positions.



(a) Common topologies for cosmic ray muons, on the left, and ν_μ CC interactions, on the right.



(b) Red lines represent reconstructed Cosmic tracks for the topologies in Figure 10.20a. The start and end positions defined by CosmicTrack are labeled.

Figure 10.20: Diagram of common topologies. The left side of Figure 10.20a shows two cosmic ray muons, entering from the top of the detector. The right side of Figure 10.20a shows two cases of ν_μ CC interactions. Figure 10.20b displays the reconstructed tracks CosmicTrack would create, given the underlying topologies. CosmicTrack always defines the start of the track as the side with the largest Y position. The end is defined as the side with the smallest Y position. Note that the choice of start and end is incorrect for the ν_μ CC interaction on the right.

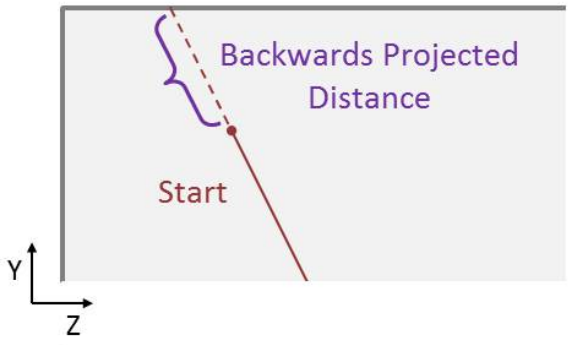


Figure 10.21: The track direction is projected backwards from the start of the track until it intersects a detector edge. The distance between the edge and the start of the track is called the backwards projected distance.

A useful cut can be made on the backwards projected distance of the Cosmic track. The backwards projected distance is calculated by projecting the track direction backwards from the start position to the edge of the detector. See Figure 10.21 for a diagram of this variable. We require that this distance is more than 50 cm. Figure 10.22 displays the effect of this cut on the negative population; Figure 10.23 shows the effect of this cut for the positive population.

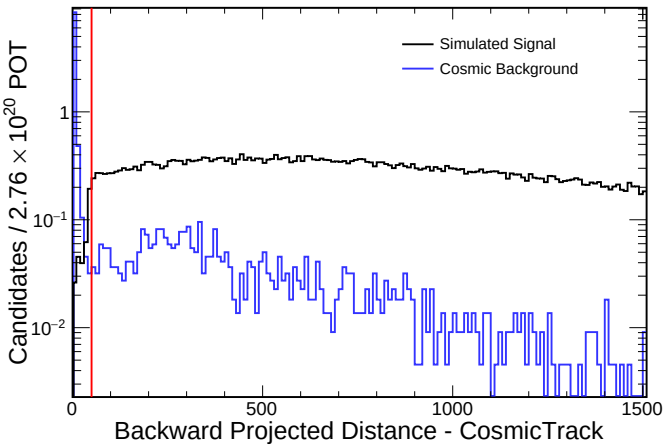


Figure 10.22: Plot of the backwards projected distance to the detector edge for the CosmicTrack. This is plot only includes the negative population. Note the logarithmic vertical axis. A red line is drawn at backwards projected distance equal to 50 cm. Candidates with values less than 50 cm are rejected. Simulated signal is shown in black; cosmic background data is shown in blue. The simulated signal is unoscillated. Each population is scaled to the expected levels for 2.76×10^{20} POT. Data quality cuts requiring no missing DCMs and that the detector is synced are applied. The primary Kalman track is required to have a ReMId value greater than 0.5. All cuts listed in Tables 10.4 and 10.5, except for the cut on backwards projected distance, are applied.

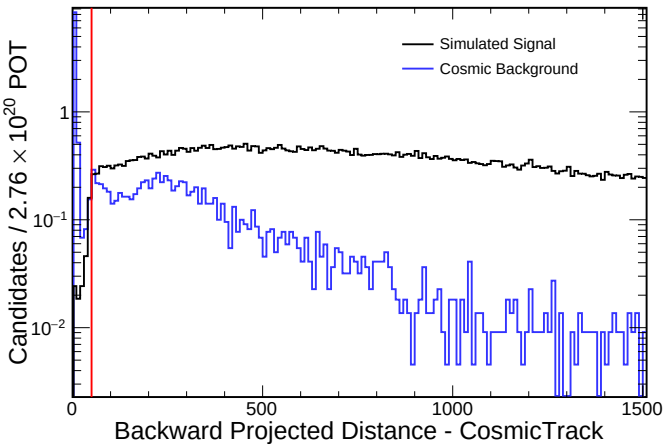
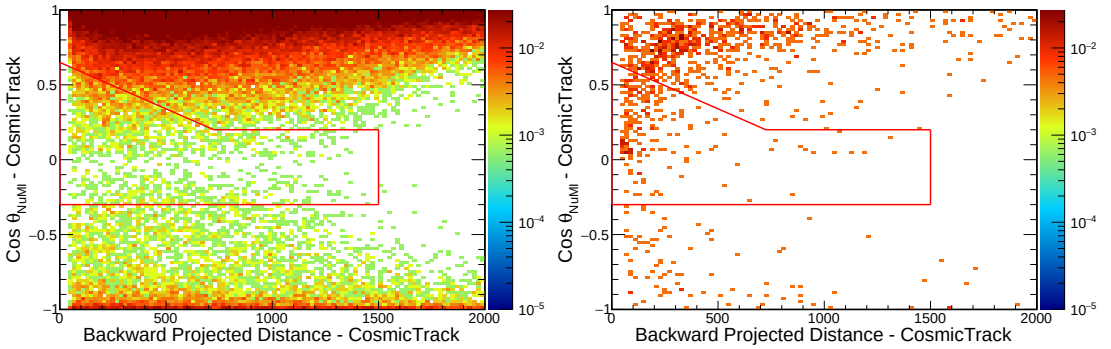


Figure 10.23: Plot of the backwards projected distance to the detector edge for the CosmicTrack. This is plot only includes the positive population. Note the logarithmic vertical axis. A red line is drawn at backwards projected distance equal to 50 cm. Candidates with values less than 50 cm are rejected. Simulated signal is shown in black; cosmic background data is shown in blue. The simulated signal is unoscillated. Each population is scaled to the expected levels for 2.76×10^{20} POT. Data quality cuts requiring no missing DCMs and that the detector is synced are applied. The primary Kalman track is required to have a ReMId value greater than 0.5. All cuts listed in Tables 10.4 and 10.6, except for the cut on backwards projected distance, are applied.

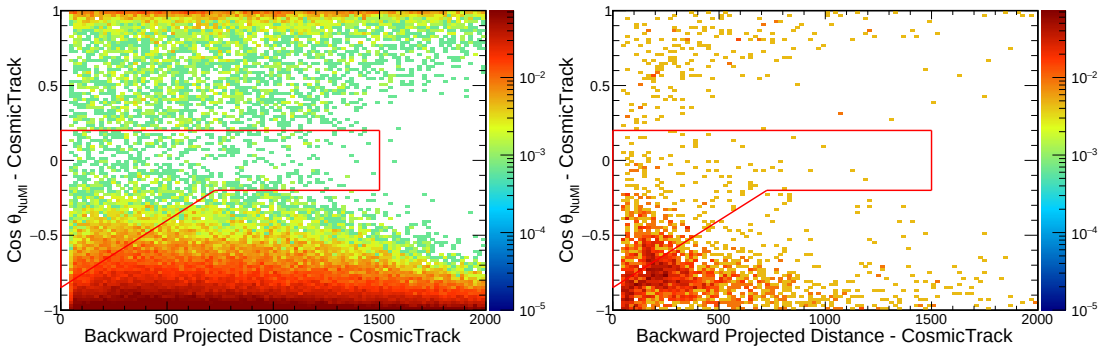


(a) Simulated Signal

(b) Cosmic Background

Figure 10.24: The horizontal axis is the backwards projected distance in cm of the Cosmic track. The vertical axis is the $\cos\theta_{NuMI}$ for the Cosmic track, where θ_{NuMI} is the angle between the Cosmic track and the direction of the neutrino beam. This plot is only for the negative population. Note the logarithmic color axis. Red lines are drawn to define the cut region; candidates inside the red lines are rejected. Simulated signal is shown on the left; cosmic background data is shown on the right. The simulated signal is unoscillated. Each population is scaled to the expected levels for 2.76×10^{20} POT. Data quality cuts requiring no missing DCMs and that the detector is synced are applied. The primary Kalman track is required to have a ReMId value greater than 0.5. All cuts listed in Tables 10.4 and 10.5, except for cuts on the 2D space of the backwards projected distance of the Cosmic track and the cosine between the Cosmic track and the beam direction, are applied.

Many cosmic background candidates can be rejected when one cuts in the 2D space of the backwards projected distance of the Cosmic track and the direction cosine between the Cosmic track and the beam direction. The exact values of box and triangle cuts vary between the positive and negative populations. See Figure 10.24 for a pictorial representation of the cut for the negative population; see Figure 10.25 for the positive population.



(a) Simulated Signal

(b) Cosmic Background

Figure 10.25: The horizontal axis is the backwards projected distance in cm of the Cosmic track. The vertical axis is the $\cos\theta_{NuMI}$ for the Cosmic track, where θ_{NuMI} is the angle between the Cosmic track and the direction of the neutrino beam. This plot is only for the positive population. Note the logarithmic color axis. Red lines are drawn to define the cut region; candidates inside the red lines are rejected. Simulated signal is shown on the left; cosmic background data is shown on the right. The simulated signal is unoscillated. Each population is scaled to the expected levels for 2.76×10^{20} POT. Data quality cuts requiring no missing DCMs and that the detector is synced are applied. The primary Kalman track is required to have a ReMId value greater than 0.5. All cuts listed in Tables 10.4 and 10.6, except for cuts on the 2D space of the backwards projected distance of the Cosmic track and the cosine between the Cosmic track and the beam direction, are applied.

10.3.2 Near Detector Slice Cuts

The near detector event selection doesn't need to reject cosmic background; the flux that reaches the underground near detector is small. However, the near detector does have a more complicated geometry than the far detector. The near detector is much smaller than the far detector; this makes proportionally more of the detector near an edge. Also, the near detector has a muon catcher with steel plates. The muon catcher has coarser resolution than the active portion of the detector. Finally, the muon catcher is shorter in height than the active region. We avoid events which travel through the uninstrumented notch. For more information on the near and far detectors, see Section 5.1. These differences must be taken into account when creating containment cuts to optimize neutrino energy resolution. Slice cuts which are applied to the near detector population are listed in Table 10.7.

For basic quality cuts, we required that the slice has at least one 3D Kalman track, at least one Cosmic track, at least one valid ReMId object[14], and a valid neutrino energy[15]. These are all required for later cuts, except the Cosmic track. This is simply required to mirror the reconstruction requirements for the far detector population. We do not require that the mask created by Live Geometry is valid; this masking is only done for the far detector. For the near detector, we only use events where the entire detector is active.

Table 10.7: Near Detector Slice Cuts

Cut	Description
Number of 3D Kalman Tracks $\neq 0$	Require at least one 3D Kalman track
Number of Cosmic Tracks $\neq 0$	Require at least one Cosmic track

Continued on next page

[14]Vector index isn't equal to 999, a default value used when no ReMId objects are found. This might happen, for instance, if no 3D Kalman track existed.

[15]Neutrino energy isn't equal to -5, a default value used when no neutrino energy could successfully be calculated. This might happen, for instance, if no 3D Kalman track existed.

Table 10.7 – continued from previous page

Cut	Description
Number of ReMId Objects $\neq 0$	Require at least one ReMId value
Best ReMId Index $\neq 999$	Require the vector index of highest ReMId track is valid
Neutrino Energy $\neq$ -5	Require a valid neutrino energy value
No. Slice Hits > 20	Slice has more than 20 hits
No. Slice Cont. Planes > 4	Slice has more than 4 contiguous planes
No. Cells From Edge > 2 for Slice	Slice is more than 2 cells away from XY edge of detector
First Slice Plane > 2	First plane of slice is not plane 0, 1 or 2
Last Slice Plane < 211	Last plane of slice is not plane 211, 212 or 213, last active planes of detector
$StartZ < 1150$	Track z start position less than 1150 cm
$EndZ < 1275$ or $TranY < 55$	Track z end position less than 1275 cm or has a y position less than 55 cm when crossing the transition plane
Proj. Forward Cell > 5	Projected forward number of cells to edge of detector from end of track is more than 5
Proj. Backwards Cell > 8	Projected backwards number of cells to edge of detector from start of track is more than 8
Off-track Energy in Muon Catcher < 0.03 GeV	Visible energy in transition plane and muon catcher not associated with the muon track is less than 0.03 GeV

To ensure the slice isn't junk, we require that the slice has more than 20 hits. To remove primarily vertical events, which might be cosmic background, we require that the slice has hits in a least 5 contiguous planes.

For containment, we want to make sure that the slice doesn't have hits in the outside layers of the detector. We require that the slice hit closest to a detector X or Y edge is more than 2 cells away from the edge. We require that the plane number of the lowest detector Z slice hit is not 0, 1 or 2, the first three planes in the detector. We require that the plane number of the highest detector Z slice hit is not 211, 212 or 213, the last

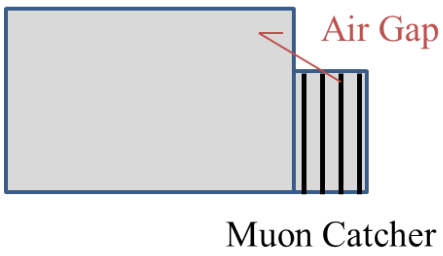


Figure 10.26: Diagram of the notch between the active region and the muon catcher region of the near detector. We require that muon tracks do not travel through the air gap of the notch.

three active planes of the detector. These planes are in the muon catcher. The start position of the 3D Kalman track with the highest ReMId value must have a detector Z value of less than 1,150 cm. This keeps the vertex of the event in the fully active region of the detector.

The fully active region of the detector is taller than the muon catcher and we do not want the muon track to leave the detector at this transition. See Figure 10.26 for a diagram of the shorter muon catcher in the near detector. We require that either the 3D Kalman track with the highest ReMId value has a stop detector Z position less than 1,275 cm or that the detector Y position in the transition plane is less than 55 cm. The transition plane is defined as the last full-height plane in the fully active region before the muon catcher.

Finally, we require that the start and end position of the 3D Kalman track with the highest ReMId value are contained. To do this, we project the track forwards from the end of the track along the direction of the track. We count the number of cells crossed by this projection before it intersects the edge of the detector. The number of cells crossed is required to be more than 5. We also project the track direction backwards from the start of the track and count the number of cells crossed before the edge of the detector. This number of cells must be less than 8.

The neutrino energy reconstruction does not attempt to account for hadronic energy loss in the muon catcher. We make a cut to ensure that the hadronic energy is mostly contained in the fully active region. We sum the visible energy of the slice not associated with the muon track in the transition plane and muon catcher. This energy must be less

than 0.03 GeV. This allows for a few noise hits in the muon catcher but not significant hadronic energy.

10.4 Analysis Cuts

Now that one has a population of contained slices, one can make the final analysis cuts, listed in Table 10.8. To be considered a ν_μ CC candidate, we require the primary Kalman track to have a ReMId value greater than 0.7. This ensures that the event has a muon-like track. For more information on Kalman tracks, see Section 7.4. For more information on ReMId, see Section 7.6. Figure 7.3 demonstrates the power of this cut.

Within the contained, ν_μ CC population, we make one further division. Using QePId, which identifies quasielastic(QE)-like events, we separate candidates into a QE and nonQE population. For more information on QePId, see Section 7.8. To be considered QE, the slice must have either one or two Kalman tracks. One of these Kalman tracks must be 3D; the other can be either 2D or 3D. For the one track sample, we require a QePId value greater than 0.45. For the two track sample, the QePId value must be greater than 0.4. Figure 7.4 shows the power of this cut for the one track sample; Figure 7.5 is for the two track sample.

Table 10.8: Analysis Cuts

Cut	Description
Highest ReMId Value > 0.7	The track with the highest ReMId value is consistent with a muon hypothesis
QE Population	If the slice is consistent with a QE hypothesis, use in QE population. Otherwise, use in nonQE population. [16]

[16] The QE population consists of two samples. For slices with precisely one 3D Kalman track and no other Kalman tracks, require the QePId value to be > 0.45. Otherwise, for slices with one 3D Kalman track and precisely one other Kalman track (either 2D or 3D), require the QePId value to be > 0.4.

Chapter 11

Expected Signal and Background

Having defined an event selection procedure in the previous chapter, we can now evaluate how many signal and background events we expect. The far detector data corresponds to 3.45×10^{20} POT with variable detector configurations. Our predicted results are scaled to this exposure.

11.1 Signal Events

Maximal mixing is defined as $\theta_{23} = \pi/4$. If θ_{23} is slightly larger or smaller in value, we would expect slightly more neutrino interactions, hence the name "maximal". Maximal mixing is consistent with the world's current understanding of the oscillation parameters[1]. See Section 4.1.2 for more information about the current measurements of the oscillation parameters. We will use it as a convenient approximation of what we expect to see in our data, along with setting $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2. If no neutrino oscillations occurred, we would expect many events in each analysis population. However, in the case of maximal mixing, we would expect fewer events.

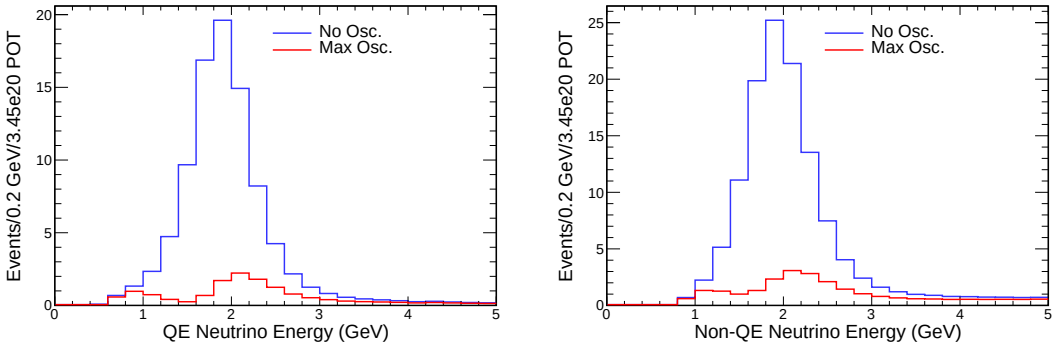
Table 11.1 lists the predicted event counts for each population for no neutrino oscillations and maximal mixing. To be counted, the event must have a reconstructed neutrino energy between 0 and 5 GeV. This analysis can easily determine if neutrino oscillations occur by simply noting if our total event count is less than that expected with no oscillation, roughly 200 events.

Figure 11.1 displays the different energy spectra expected for the assumptions of no

Table 11.1: Predicted and measured event counts for the far detector. The total POT for the far detector is 3.45×10^{20} POT and this corresponds to variable detector configurations. The counts for each population only include events with reconstructed neutrino energy between 0 and 5 GeV. The predicted total event count for each population under the hypothesis of no neutrino oscillations is given first. The next line correspond to the hypothesis of maximal mixing, setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2 as well as using the oscillation parameters listed in Table 4.2.

	QE	NonQE
Total Events (no osc.)	89.2	122.3
Total Events (max. mix)	14.3	24.2

neutrino oscillations and maximal mixing. The measurement of the dip in the energy spectra seen around 1.5 GeV allows us to determine θ_{23} and $|\Delta m^2_{32}|$. The depth of the dip tells us the value of θ_{23} and the location of the dip along the horizontal axis tells us the value of $|\Delta m^2_{32}|$. Having good neutrino energy resolution and a precise understanding of the absolute energy scale are key to measuring these parameters.



(a) QE Population (b) NonQE Population

Figure 11.1: Plot of the reconstructed neutrino energy in GeV. Figure 11.1a is for the QE population in the far detector; Figure 11.1b is for the nonQE population. The blue line represents the predicted spectrum from simulation if no neutrino oscillations occurred. The red line represents the predicted spectrum from simulation under the assumption of maximal mixing. For maximal mixing, the simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2. The simulation is scaled down to match the data POT, 3.45×10^{20} POT.

11.2 Cosmic Ray Background

Cosmic ray interactions at the far detector are an important source of background for this analysis. By requiring that an event is in time with a NuMI beam spill, we reduce this background by a factor of 10^5. With our selection cuts (Section 10.3.1), especially the containment cuts and cuts requiring that the event is aligned with the NuMI beam direction, we reduce this background by a further factor of 10^7.

This section discusses how the remaining cosmic ray background is modeled. The cosmic ray background that we examined which passes all the analysis cuts is relatively small. If we studied enough cosmic background events, the predicted cosmic background shape for the reconstructed neutrino energy should be relatively smooth. We did not want the analysis to use a spiked cosmic background shape determined by the accidents of low statistics. Instead of looking at an even larger statistics sample of cosmic ray background, we fit a smooth shape to the reconstructed neutrino energy for the cosmic ray background seen. We then determined a normalization for this shape and used that as our predicted cosmic ray background.

This analysis has two source of data which represent the cosmic background. First, it has data from cosmic trigger files. These readout windows are written at a rate of 10 or 40 Hz, depending on the running period. Initially, 40 Hz was used to rapidly acquire statistics useful for the calibration procedure. Later, 10 Hz was used to reduce the amount of data requiring transfer to and storage at Fermilab. A filter is used to ensure that these readout windows do not include beam spills.

The second source is from the out-of-time data in the NuMI trigger files. These readout windows correspond to NuMI beam spills and contain the neutrino interaction data for this analysis. However, the beam spill is only 10 μs long, while the readout window is 500 μs long. Using appropriate cuts to remove the beam spill, one can treat the out-of-time data as a cosmic background estimate.

The cosmic trigger files have the advantage of much larger statistics - the NuMI triggers occur at 1.33 Hz. Large statistics is crucial for understanding the shape and characteristics of the cosmic ray background. However, the out-of-time NuMI data is useful because it exactly replicates the running conditions of the NuMI data. For instance, by having a mix of 40 Hz and 10 Hz running in the cosmic trigger files, some

detector configurations are oversampled compared to others. Also, the NuMI beam might have only sent beam spills for part of a run, but we would have cosmic trigger files for the whole run. With effort, these differences could be mitigated, but in the interest of time, they were not.

A large statistics sample of cosmic trigger files was processed. For every run used in the analysis, the cosmic trigger files corresponding to every 4th subrun was used. More subruns could have been processed for a larger sample, but were not due to time constraints. The cosmic ray plots presented in Chapter 10 used this dataset.

Normalizing to the expected livetime of the analysis, reconstructed energy spectra for the cosmic ray background for each population were made using the cosmic trigger data. Although the statistics of the initial sample was large, the number of cosmic ray events which pass the analysis event selection was relatively small. Therefore, a fit was made to the shape of the energy spectrum for each population. This fit was then used to create a histogram for each population representing the shape of the expected cosmic ray background. The fit used a Landau distribution with a constant offset. Figure 11.2 displays the cosmic trigger data for the QE population, along with the fit and the histogram created using the fit. Figure 11.3 is for the nonQE population.

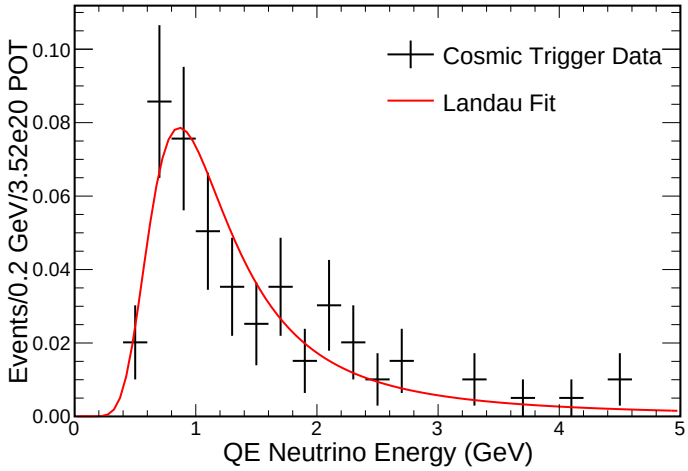
Having determined the shape of the cosmic ray background for each population, we must determine a method of normalizing the shape. First, we used the relative size of the events in the QE population and nonQE population seen in the cosmic trigger data to normalize the shape histograms. Looking at all the cosmic ray background events which passed our analysis cuts, 20% were classified as part of the QE population and 80% were classified as part of the nonQE population. Therefore, the QE histogram was normalized to have an area of 0.2, while the nonQE histogram was normalized to have an area of 0.8.

The cosmic trigger data could be used to determine an absolute normalization, but it suffers from not exactly matching the running conditions seen in the analysis. Instead, the out-of-time NuMI trigger data was used. This sample had 80 total background events which passed the analysis cuts. It was decided that 80 was a large-enough number to have a small-enough statistical error for the absolute normalization; however, it wasn't large enough to use for the relative normalizations between samples.

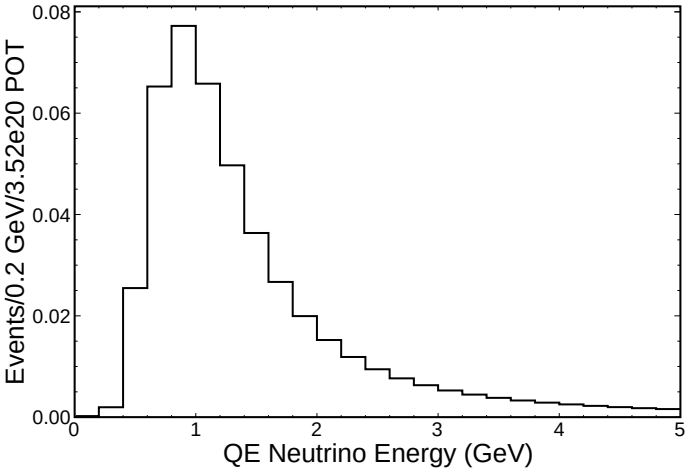
The 80 events in the out-of-time NuMI trigger data was multiplied by a scale factor

that accounted for the roughly 20-to-1 longer livetime seen in the out-of-time data compared to the amount of livetime used for the in-time sample.[1] This resulted in an expectation of 3.4 cosmic ray background events for the analysis. For the QE sample, the expected cosmic ray background is 0.7 events. For the nonQE sample, the expected cosmic ray background is 2.7 events.

[1]The readout window is 500 μs long and the beam spill is 10 μs long. We use a 12 μs cut to define the beam spill to ensure the beam spill is fully captured. The first and last 52 μs of the readout window is not used because physics interactions might have only been partially captured by the readout window. 9 μs on either side of the beam spill are not used for either the in-time or out-of-time sample to ensure there is no overlap. For far detector data taken before October 2014, a second beam timing window was also used. After all these effects are taken into account, the out-of-time exposure is roughly 20 times longer than the in-time data.

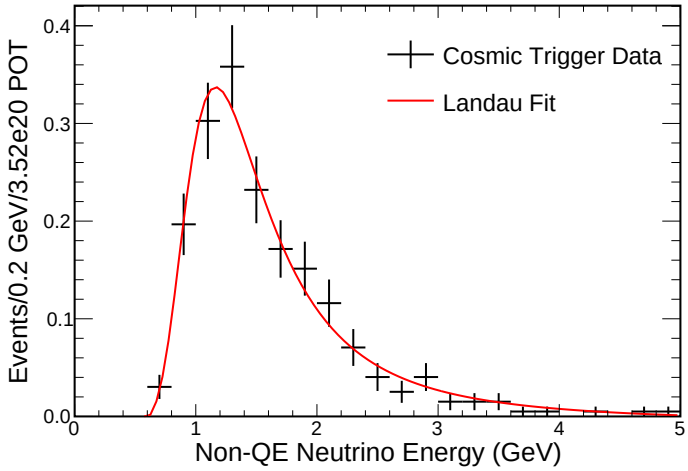


(a) Plot of the expected cosmic ray background, as determined using the cosmic trigger files. A fit was made of this data, using a Landau distribution and a constant offset. The fit is drawn in red.

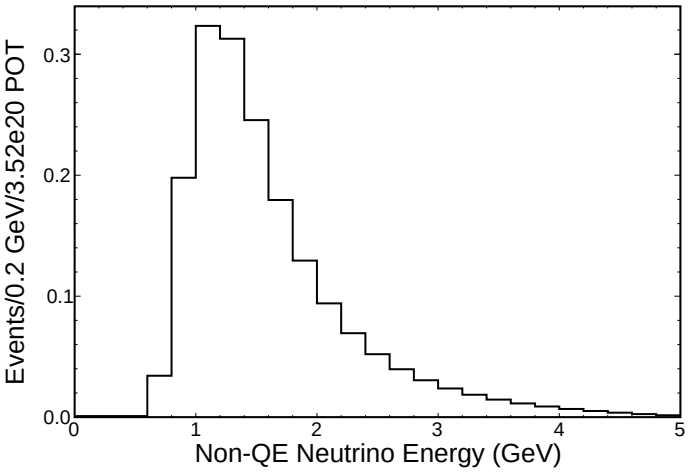


(b) Histogram created using the fit from Figure 11.2a.

Figure 11.2: Plot of the expected cosmic ray background for the QE population in the far detector. Figure 11.2a displays the data from the cosmic trigger files and the fit to the data. Figure 11.2b displays the histogram created from the fit.



(a) Plot of the expected cosmic ray background, as determined using the cosmic trigger files. A fit was made of this data, using a Landau distribution and a constant offset. The fit is drawn in red.



(b) Histogram created using the fit from Figure 11.2a.

Figure 11.3: Plot of the expected cosmic ray background for the nonQE population in the far detector. Figure 11.2a displays the data from the cosmic trigger files and the fit to the data. Figure 11.2b displays the histogram created from the fit.

11.3 Signal and Background Events

We have now discussed the expected signal and cosmic ray background event counts for this analysis. We can also count the expected number of events in our analysis from neutrino interactions that are not ν_μ CC interactions. These are primarily NC interactions. This count is done using the extrapolated far detector prediction, discussed in Chapter 13. Table 11.2 lists the total expected events under the assumptions of no neutrino oscillations as well as maximal mixing. It also lists the expected number of background events for each sample under the assumption of maximal mixing. For the QE population, we expect about 14 total events, with 1 background event. For the nonQE population, we expect about 24 events with 5 background events.

Table 11.2: Predicted and measured event counts for the far detector. The total POT for the far detector is 3.45×10^{20} POT and this corresponds to variable detector configurations. The counts for each population only include events with reconstructed neutrino energy between 0 and 5 GeV. The predicted total event count for each population under the hypothesis of no neutrino oscillations is given first. The rest of the lines correspond to the hypothesis of maximal mixing, setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2 as well as using the oscillation parameters listed in Table 4.2. The expected background event counts are given for each population, with neutrino interactions that are not ν_μ CC interactions and cosmic ray backgrounds separated.

	QE	NonQE
Total Events (no osc.)	89.2	122.3
Total Events (max. mix)	14.3	24.2
Neutrino Background	0.2	2.1
Cosmic Ray Background	0.7	2.7

Chapter 12

Near Detector Comparisons

Having defined a set of cuts to select events for analysis (in Chapter 10), we can now compare the data in the near detector to our simulation for the near detector. The near detector simulation models multiple beam neutrino interactions per spill, including interactions occurring in the surrounding rock that enter the detector. The simulation uses two different noise models; a more complicated one for electronic channels with energy depositions from physics interactions and a simple, histogram-based one for electronic channels without energy from physics interactions. There is no cosmic ray background included in the near detector simulation, since the background rate is so low. For more information on the simulation, see Chapter 6.

If the distributions comparing the near detector data and simulation align well, we can have some confidence in the framework and process we are using. Distributions that disagree can point to areas of future work and investigation.

12.1 Contained Slices

The first set of plots looks at characteristics of the slices which pass our selection criteria. The criteria applied include file selection, detailed in Section 10.1, spill cuts, listed in Section 10.2, and near detector slice cuts, from Section 10.3.2. The analysis level cuts, Section 10.4, have not been applied at this stage. Primarily, this means that the ReMId cut, used to identify slices with a reconstructed muon, is not present. For more information on slices, see Section 7.3.

By looking at lower-level plots before our final analysis cuts, we can understand if the detector and simulation are working on the most basic level. We can evaluate if the geometry of the detector is simulated correctly. Also, by working from lower-level plots to higher-level analysis plots, we have a better chance of understanding any discrepancies seen in the final plots. We will understand what lower-level quantities agree and what could be the source of the problem.

Figure 12.1 plots the number of hits for each slice. This agrees fairly well, both in shape and overall normalization. The next three plots, Figures 12.2, 12.3 and 12.4, show the minimum X, Y and Z detector hit positions for the slice. Figures 12.5, 12.6 and 12.7 show the maximum X, Y and Z detector hit positions for the slice. These show that we understand the distribution of activity in our detector fairly well, although we consistently see differences at the X and Y the edges of the detector. This might point to more compact activity in our detector than our simulation suggests.

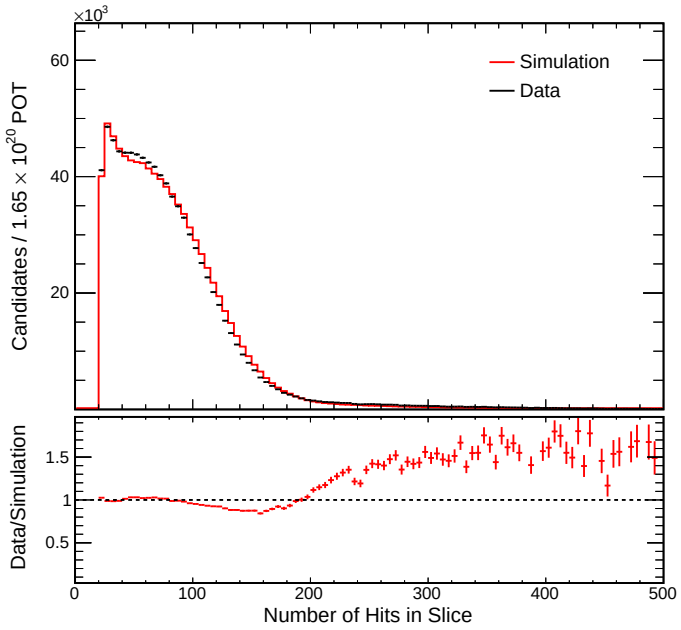


Figure 12.1: Plot of the number of hits in a slice for the near detector. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

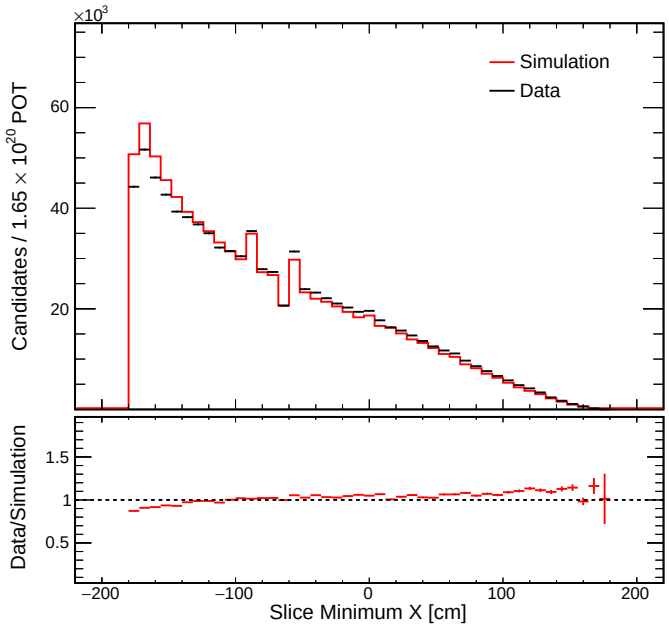


Figure 12.2: Plot of the minimum position in the detector X coordinate in cm for hits in a slice for the near detector. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

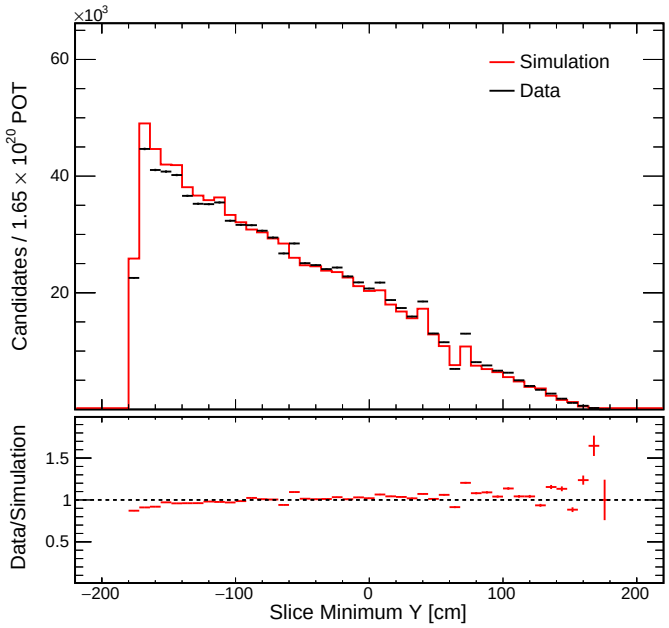


Figure 12.3: Plot of the minimum position in the detector Y coordinate in cm for hits in a slice for the near detector. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

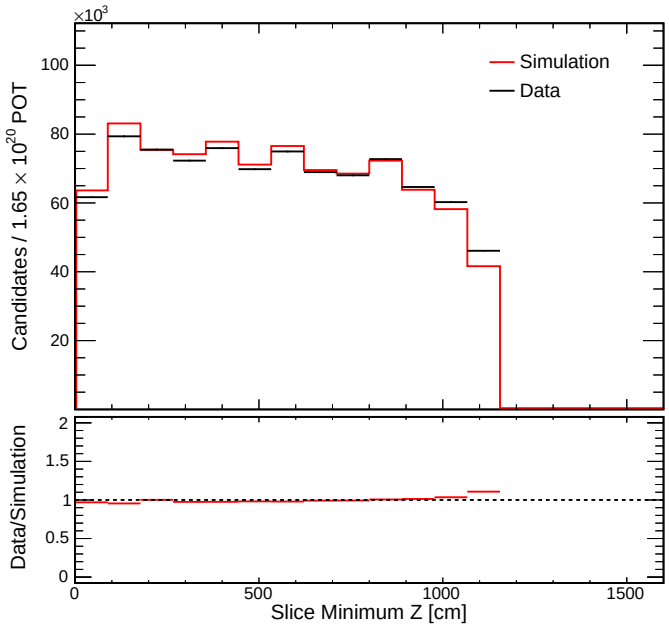


Figure 12.4: Plot of the minimum position in the detector Z coordinate in cm for hits in a slice for the near detector. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

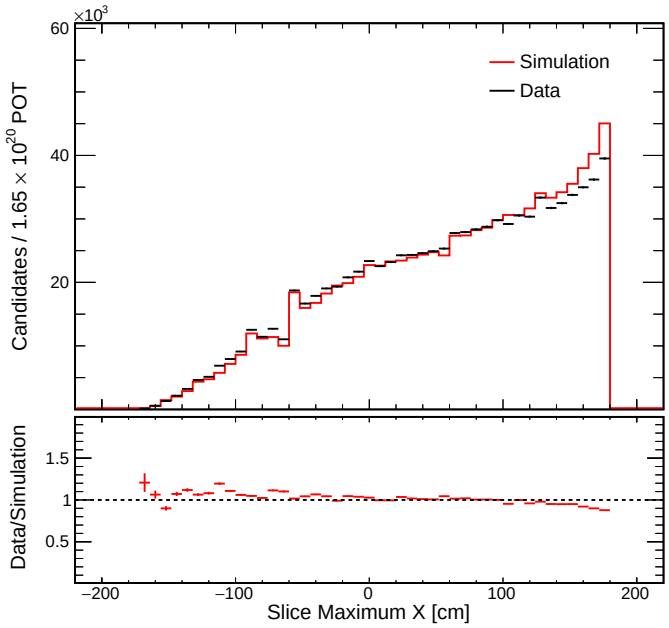


Figure 12.5: Plot of the maximum position in the detector X coordinate in cm for hits in a slice for the near detector. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

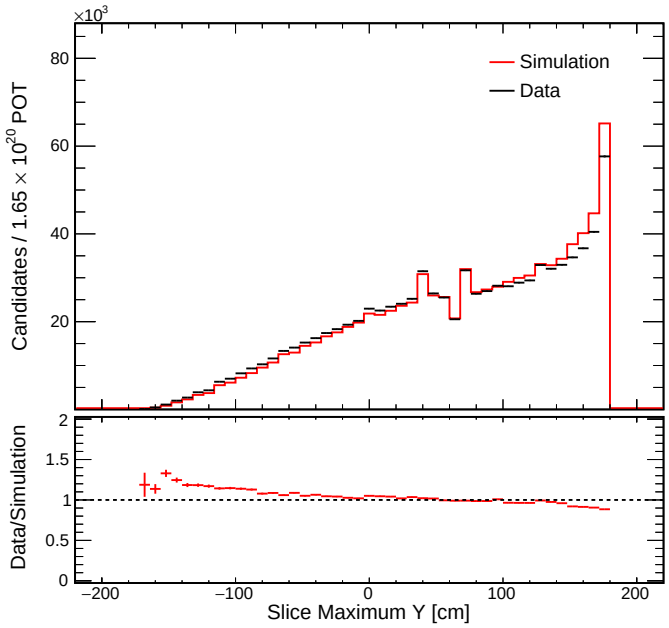


Figure 12.6: Plot of the maximum position in the detector Y coordinate in cm for hits in a slice for the near detector. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

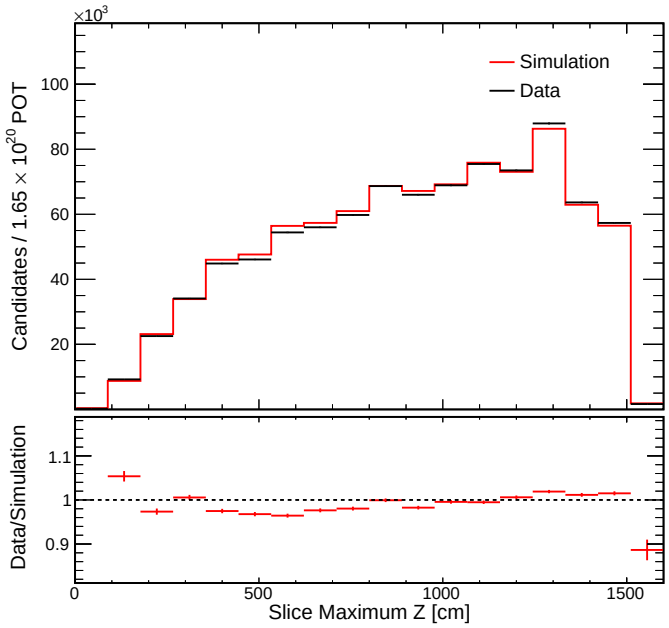


Figure 12.7: Plot of the maximum position in the detector Z coordinate in cm for hits in a slice for the near detector. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

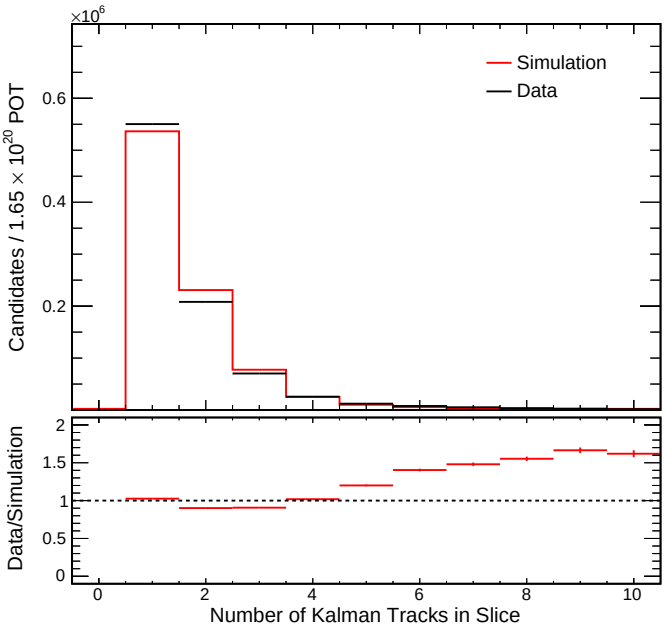


Figure 12.8: Plot of the number of 3D Kalman tracks in a slice for the near detector. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

12.2 Contained Tracks

The next set of plots looks at characteristics of the 3D Kalman tracks which pass our selection criteria. The criteria applied include file selection, detailed in Section 10.1, spill cuts, listed in Section 10.2, and near detector slice cuts, from Section 10.3.2. The analysis level cuts, Section 10.4, have not been applied at this stage. Primarily, this means that the ReMId cut, used to identify slices with a reconstructed muon, is not present. For more information on Kalman tracks, see Section 7.4. For more information on ReMId, see Section 7.6.

By understanding the tracks that we will apply a ReMId cut to, we can better understand any features seen in the ReMId selection. First, we can look at the multiplicity of the tracks. Figure 12.8 shows the number of reconstructed 3D Kalman tracks for each

slice. We see that the data has more single track events (which tend to be defined as QE) and slightly fewer 2 and 3 track events (which tend to be defined as nonQE). This should be kept in mind when we later make QE and nonQE selections.

The rest of the contained track plots have an entry only for the primary 3D Kalman track, defined as the track with the highest ReMId value. Figure 12.9 shows the number of hits on the primary track; Figure 12.10 shows the track length distribution for the primary tracks. We see that, before selecting for muon tracks, the data has fewer short primary tracks than the simulation predicts. Figures 12.11, 12.12 and 12.13, show the X, Y and Z detector coordinate positions for start of the primary track. Figures 12.14, 12.15 and 12.16, show the X, Y and Z detector coordinate positions for end of the primary track. These plots show that we are modeling the track geometry fairly well. There are a few differences between the data and the simulation that could probably be improved with a detector geometry simulation that matched the physical detector even better than the one we are using. We can also look at the initial track direction. Figures 12.17, 12.18 and 12.19 plot the cosine of the angle between the initial track direction and the detector X, Y and Z axes. Figure 12.20 is the cosine of the angle between the initial track direction and the NuMI beam direction. In general, these show good agreement and give us confidence that we are modeling the track angular distributions well.

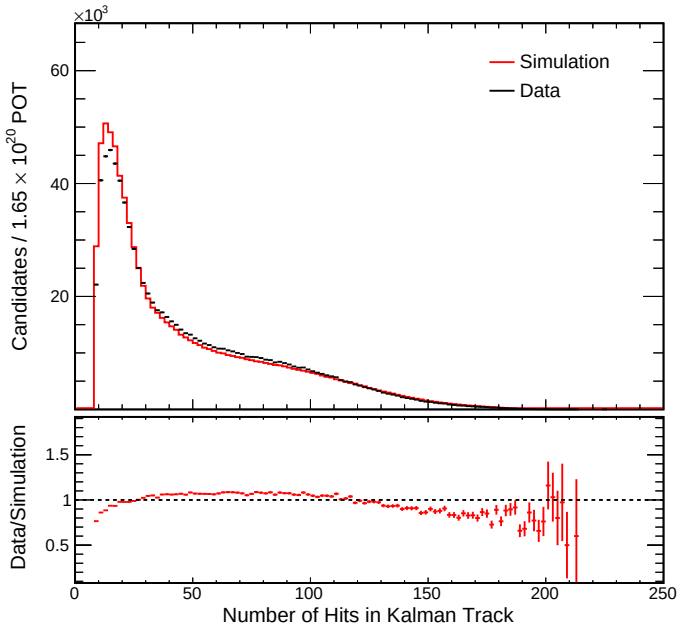


Figure 12.9: Plot of the number of hits on the 3D Kalman track with the highest ReMId value in a slice for the near detector. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

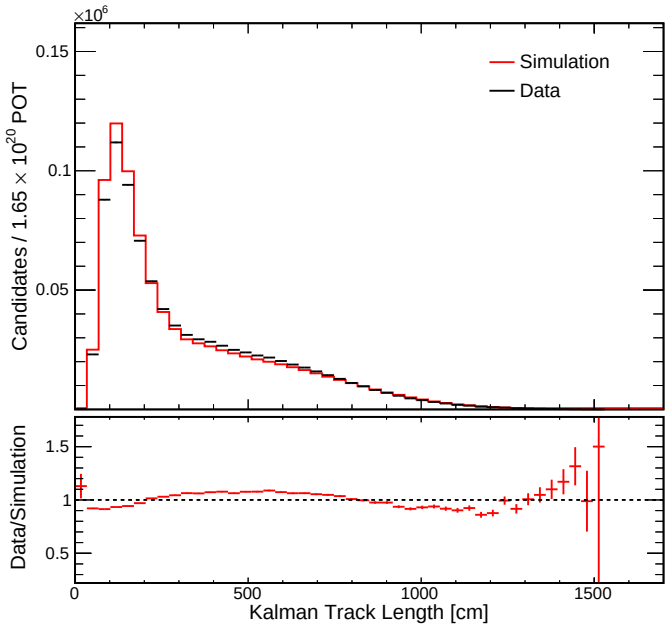


Figure 12.10: Plot of the length, in cm, of the 3D Kalman track with the highest ReMId value in a slice for the near detector. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

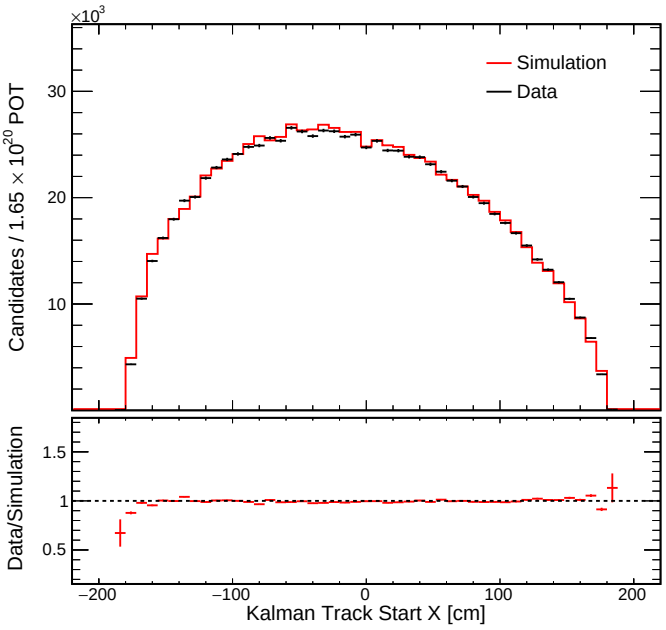


Figure 12.11: Plot of the start position in the detector X coordinate in cm for the 3D Kalman track with the highest ReMId value in a slice for the near detector. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

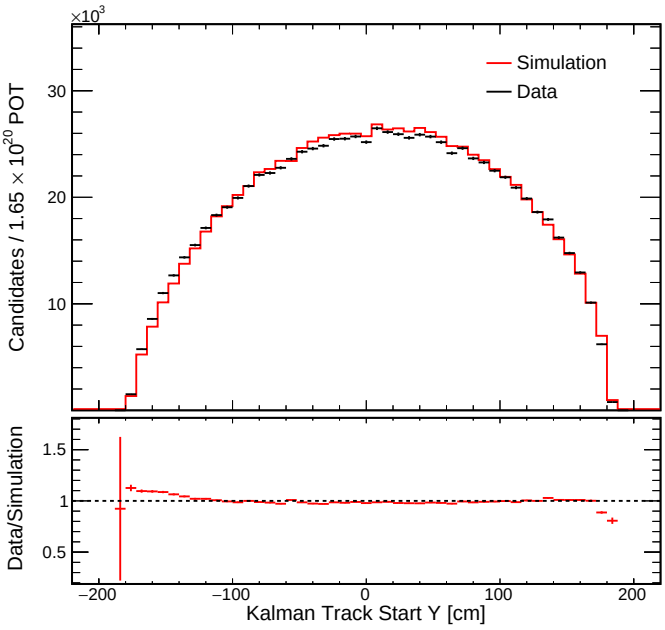


Figure 12.12: Plot of the start position in the detector Y coordinate in cm for the 3D Kalman track with the highest ReMId value in a slice for the near detector. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

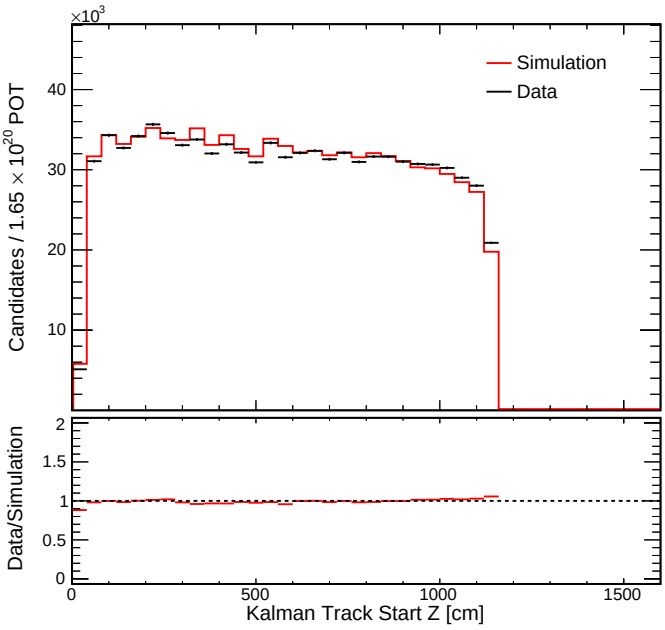


Figure 12.13: Plot of the start position in the detector Z coordinate in cm for the 3D Kalman track with the highest ReMId value in a slice for the near detector. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

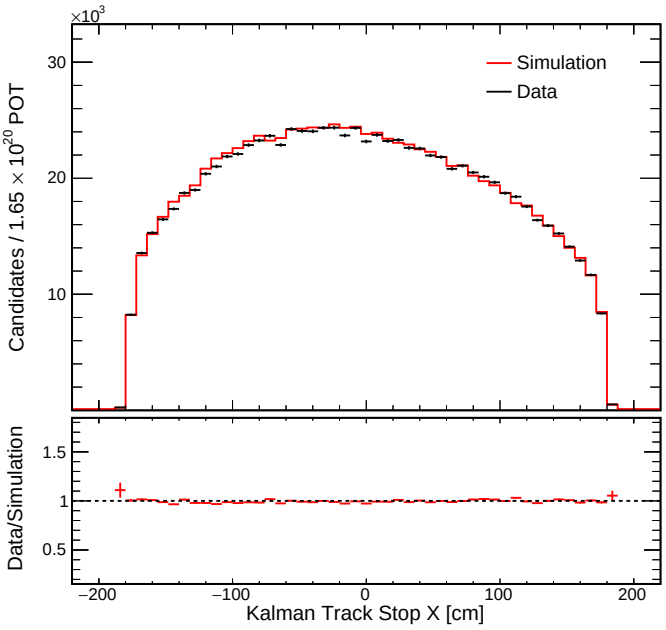


Figure 12.14: Plot of the end position in the detector X coordinate in cm for the 3D Kalman track with the highest ReMId value in a slice for the near detector. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

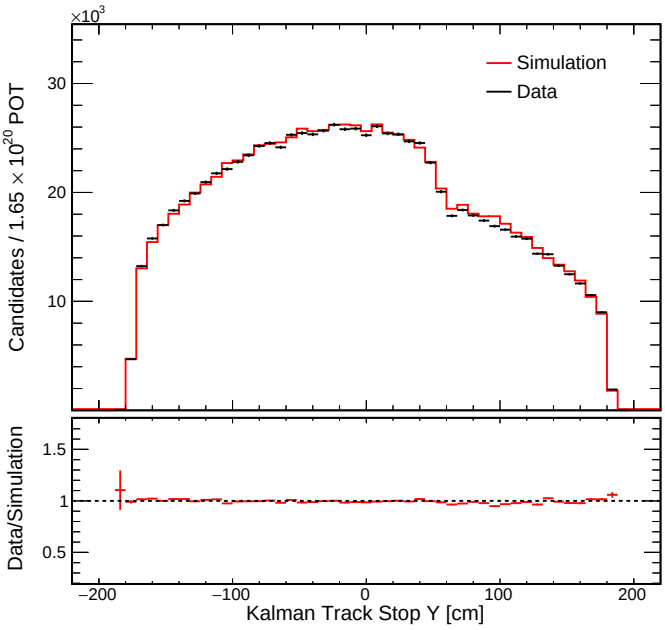


Figure 12.15: Plot of the end position in the detector Y coordinate in cm for the 3D Kalman track with the highest ReMId value in a slice for the near detector. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

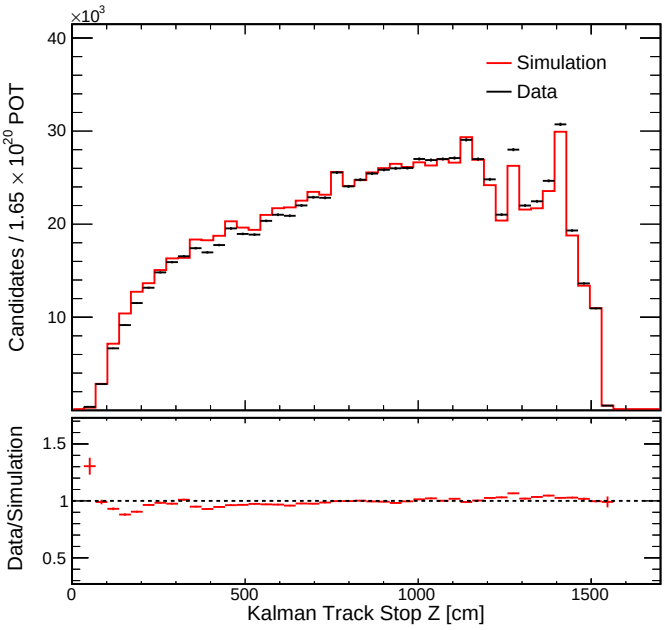


Figure 12.16: Plot of the end position in the detector Z coordinate in cm for the 3D Kalman track with the highest ReMId value in a slice for the near detector. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

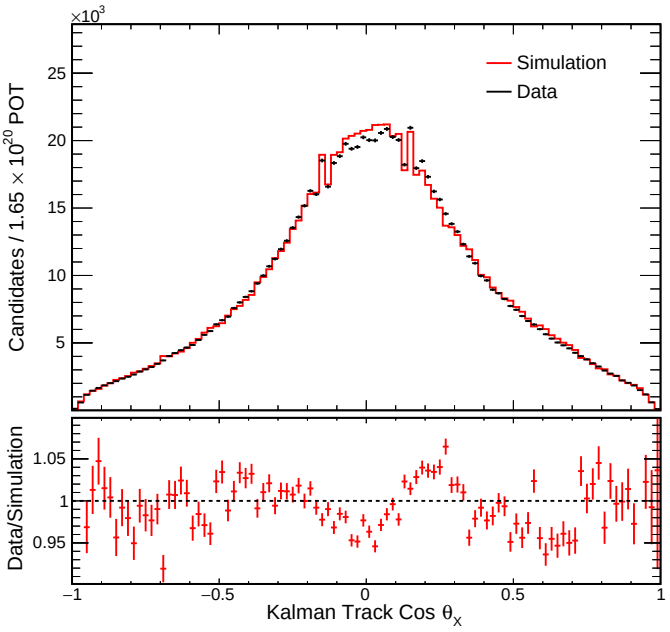


Figure 12.17: Plot of $\cos\theta_X$, where θ_X is the angle between the 3D Kalman track with the highest ReMId value in a slice and the detector X axis. This plot is for the near detector populations. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

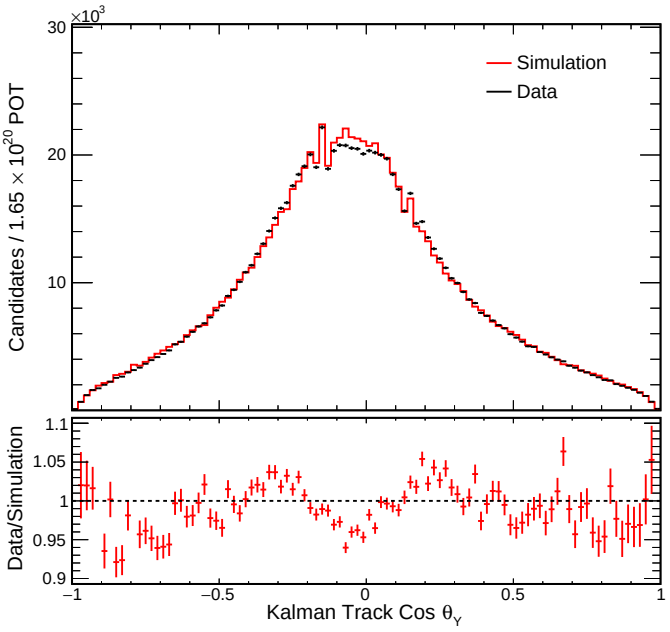


Figure 12.18: Plot of $\cos\theta_Y$, where θ_Y is the angle between the 3D Kalman track with the highest ReMId value in a slice and the detector Y axis. This plot is for the near detector populations. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

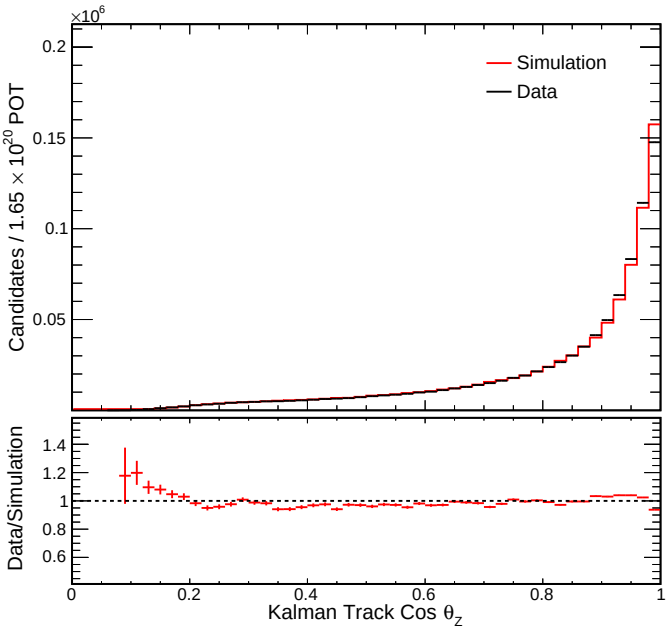


Figure 12.19: Plot of $\cos\theta_Z$, where θ_Z is the angle between the 3D Kalman track with the highest ReMId value in a slice and the detector Z axis. This plot is for the near detector populations. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

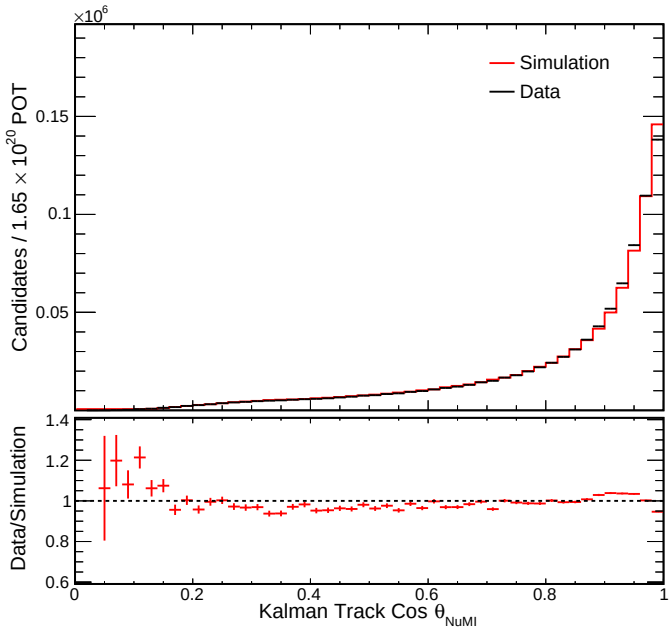


Figure 12.20: Plot of $\cos\theta_{NuMI}$, where θ_{NuMI} is the angle between the 3D Kalman track with the highest ReMId value in a slice and the NuMI beam direction. This plot is for the near detector populations. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

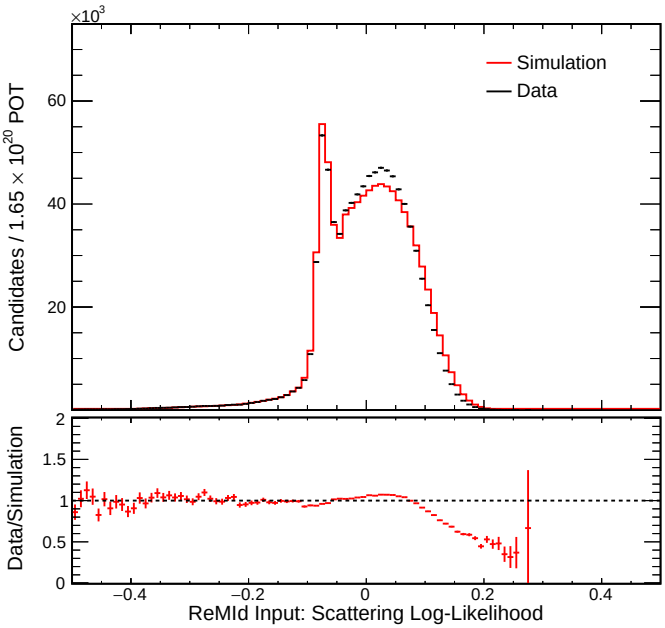


Figure 12.21: Plot of the scattering log-likelihood for the the 3D Kalman track with the highest ReMId value in a slice. This plot is for the near detector populations. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

12.3 ReMId Inputs

To make our analysis sample, we need to apply a ReMId cut. First, let us understand the inputs to ReMId. The criteria applied to create our populations include file selection, detailed in Section 10.1, spill cuts, listed in Section 10.2, and near detector slice cuts, from Section 10.3.2. The analysis level cuts, Section 10.4, have not been applied at this stage. Primarily, this means that the ReMId cut, used to identify slices with a reconstructed muon, is not present. For more information on ReMId and the inputs to ReMId, see Section 7.6. Each plot has one entry per slice, corresponding to the metric for the 3D Kalman track with the highest ReMId value in the slice.

By looking at the inputs to ReMId, we will be able to better understand our final

selected population. Figure 12.10, the track length of the 3D Kalman track with the highest ReMId value, shown in the previous section, is one input. We already noted that this plot showed some difference between the data and the simulation. Figures 12.21, 12.22 and 12.23 display the other three inputs to ReMId. We can see that these distributions do not match perfectly between the data and the simulation. For the scattering log-likelihood, we see the the peak structure is somewhat different. It is possible that an improved geometry simulation would help this distribution to agree better. For the dE/dx log-likelihood, the data shows a different peak and tail shape from the simulation. Having the absolute calorimetric energy scale set exactly the same in the data and simulation is key for this distribution to show close agreement. Finally, the non-hadronic plane fraction shows that the data sees less hadronic contamination than the simulation. This could point to less hadronic energy or more compact hadronic energy in the data.

Finally, Figure 12.24 shows the ReMId distribution. The shape of the data and simulation agree well, except at the very lowest values of ReMId (which correspond to things that look the least like a muon). However, there is 10% more data than simulation candidates for most of the ReMId distribution. After making a ReMId cut, the data population will be larger than the simulation population. A cut at 0.7 is made on this distribution to select slices with a reconstructed muon for our analysis. The rest of the plots in this chapter will have this cut applied.

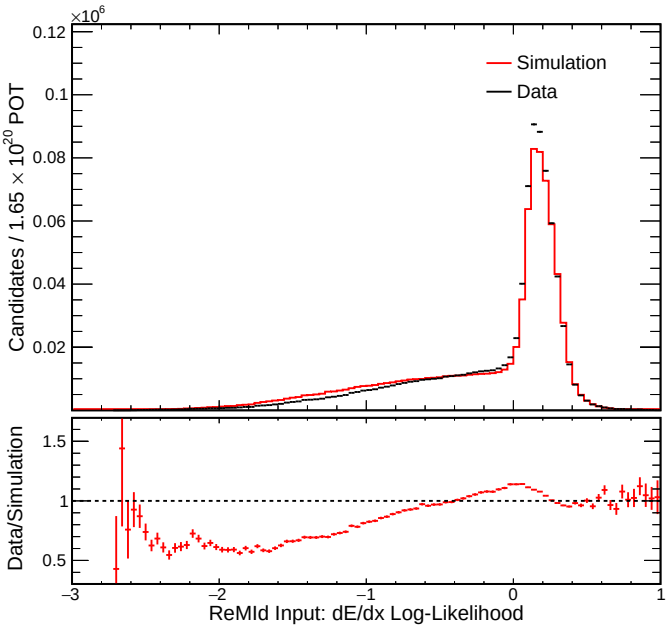


Figure 12.22: Plot of the dE/dx log-likelihood for the the 3D Kalman track with the highest ReMId value in a slice. This plot is for the near detector populations. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

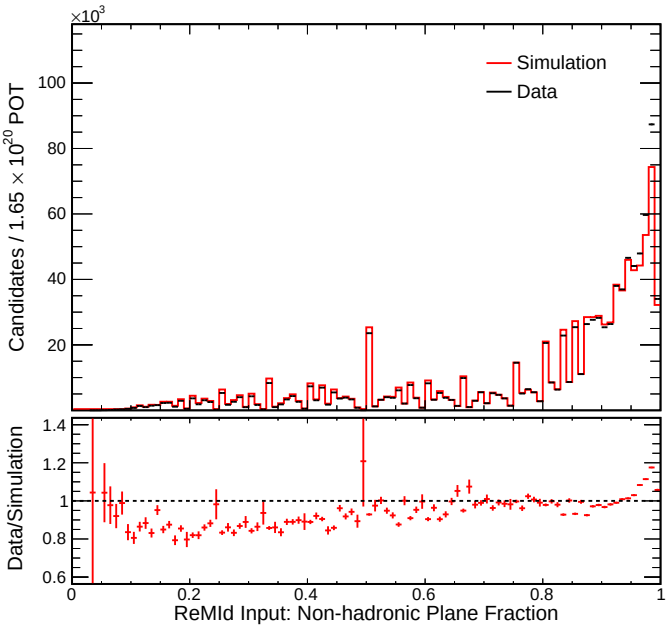


Figure 12.23: Plot of the non-hadronic plane fraction for the the 3D Kalman track with the highest ReMId value in a slice. This plot is for the near detector populations. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

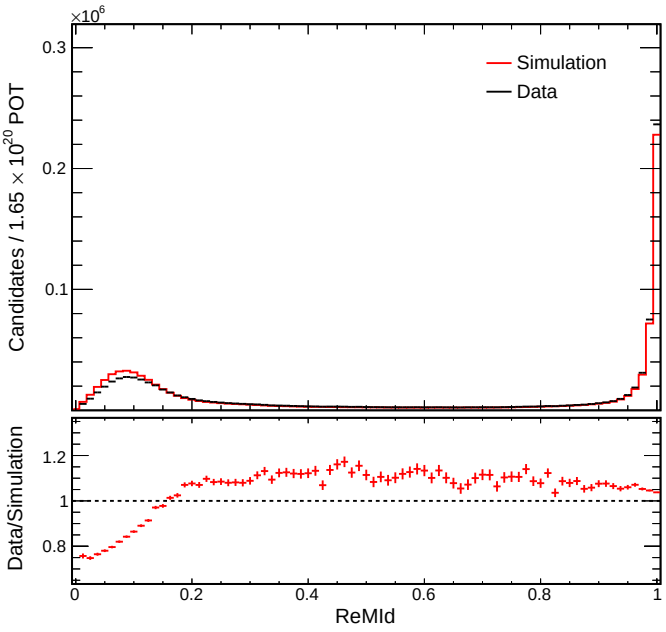


Figure 12.24: Plot of the ReMId value for the the 3D Kalman track with the highest ReMId value in a slice. This plot is for the near detector populations. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

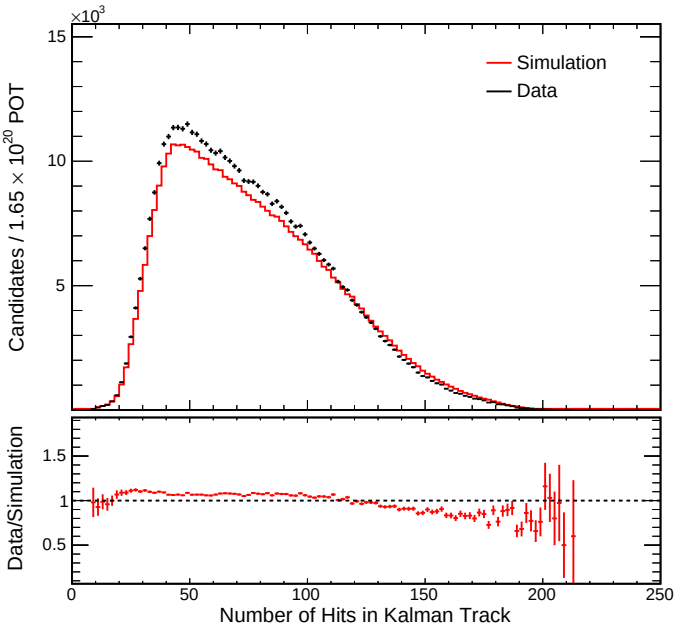


Figure 12.25: Plot of the number of hits on the 3D Kalman track with the highest ReMId value in a slice for the near detector. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

12.4 ν_μ CC Sample

Having made a ReMId cut, we now have a sample which is primarily composed of contained ν_μ CC interactions. We can look again at the properties of the primary tracks in this sample and inspect the hadronic component of the slice. The criteria applied to create our populations include file selection, detailed in Section 10.1, spill cuts, listed in Section 10.2, and near detector slice cuts, from Section 10.3.2. The analysis level cuts, Section 10.4, have not been fully applied at this stage. A ReMId cut requiring a value of greater than 0.7 has been applied, but the samples have not been divided into QE and non-QE populations. For more information on ReMId, see Section 7.6. Each plot has one entry per slice. When track properties are plotted, only the 3D Kalman track with the highest ReMId value, defined as the reconstructed muon track, is used.

Figure 12.25 shows the number of hits on the primary track; Figure 12.26 shows the track length distribution for the primary tracks. These shapes disagree somewhat; we see that the excess data events tend to have short to medium length tracks. For the longer tracks, the data actually has slightly fewer tracks. Next, we can look at the properties of the hadronic part of the slice. Figure 12.27 plots the number of hits in the slice not on the primary track. These hits are defined as the hadronic hits. This plot shows one of the largest data and simulation discrepancies in the near detector. The discrepancy and its implications will be discussed further in Section 14.12. We can also look at the visible calorimetric energy of the hadronic sector. Figure 12.28 plots the visible energy, in GeV, for the hadronic hits. Figure 12.29 plots the visible energy, in GeV, associated with hadronic energy in the vertex region of the muon track. These hadronic energies have not had the fits, discussed in Chapter 9, applied. For more information on the visible hadronic energy, see Section 9.5. Here, the trend seen in the number of hadronic hits continues. The data has significantly less hadronic energy than the simulation.

Average energy per hit can tell us if the energy per hit is modeled correctly. Figure 12.30 displays, for each slice, the total visible energy in GeV divided by the total number of hits in the slice. Each slice has one entry in the plot. We can see the overall excess in the data, as well as a slight shift to lower energy per hit. Figure 12.31 plots the total visible energy in GeV for hits on the muon track divided by the total number of hits associated with the muon track. Again, each slice has one entry in the plot. We again see a slight shift to lower energies in the data. This could indicate an offset between than data and simulation absolute calorimetric energy scale. The hadronic energy sector has more structure; Figure 12.32 plots the total visible energy in GeV for hits in the slice not associated with the muon track, divided by the total number of hadronic hits. Noise hits, which can be reconstructed as part of the hadronic cluster, tend to have low energy values. Much of the disagreement in the distributions below 0.005 GeV is probably due to problems in the noise simulation. However, because these hits have very little energy, the overall effect on the analysis is limited.

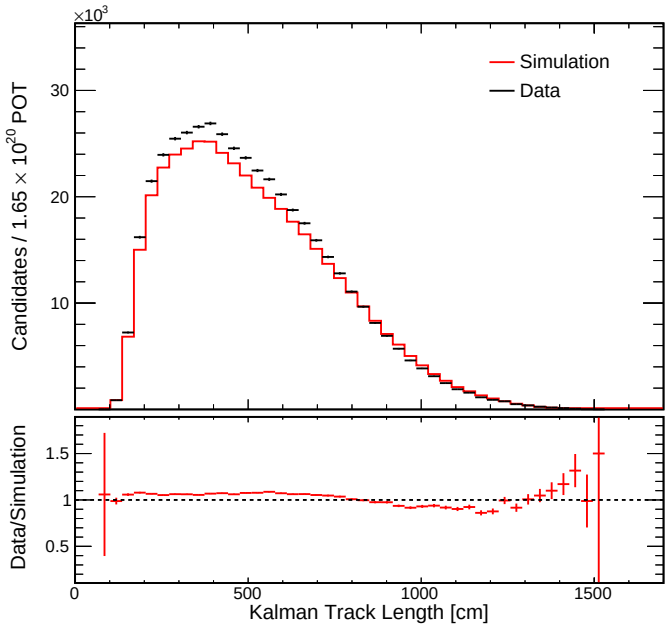


Figure 12.26: Plot of the length, in cm, of the 3D Kalman track with the highest ReMId value in a slice for the near detector. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

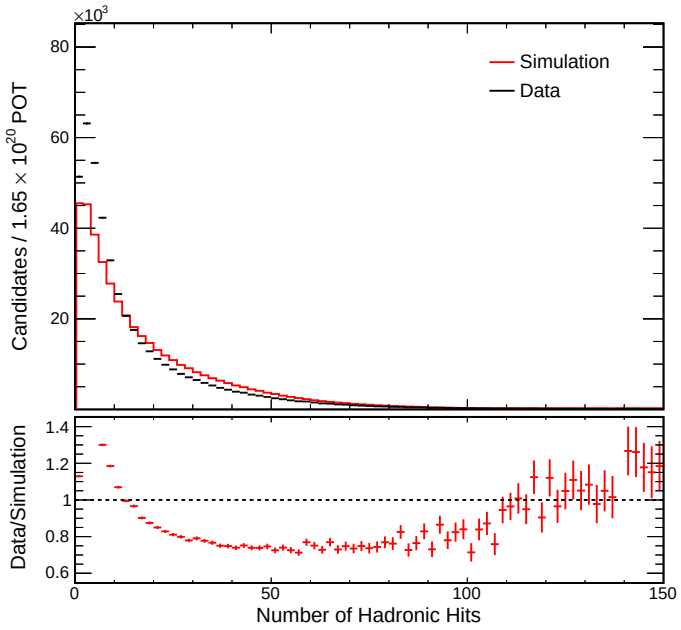


Figure 12.27: Plot of the number of hits in the slice not on the 3D Kalman track with the highest ReMId value for the near detector. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

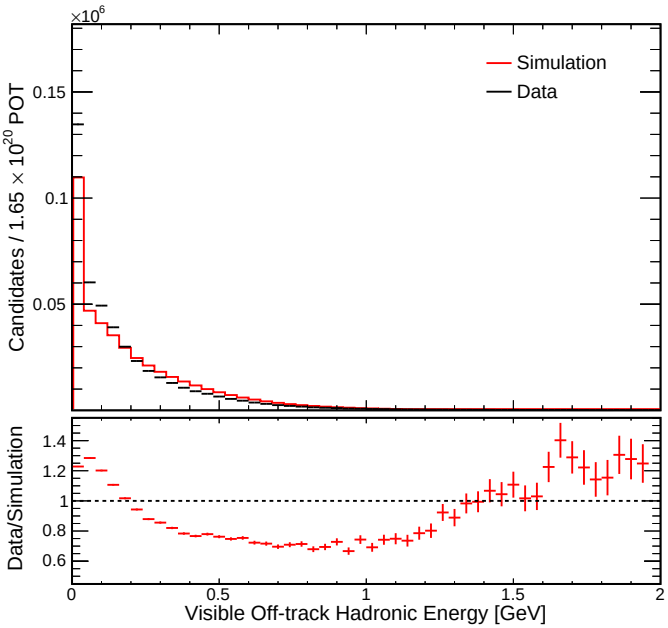


Figure 12.28: Plot of the sum of the visible energy (in GeV) of hits in the slice not on the 3D Kalman track with the highest ReMId value. This plot is for the near detector populations. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

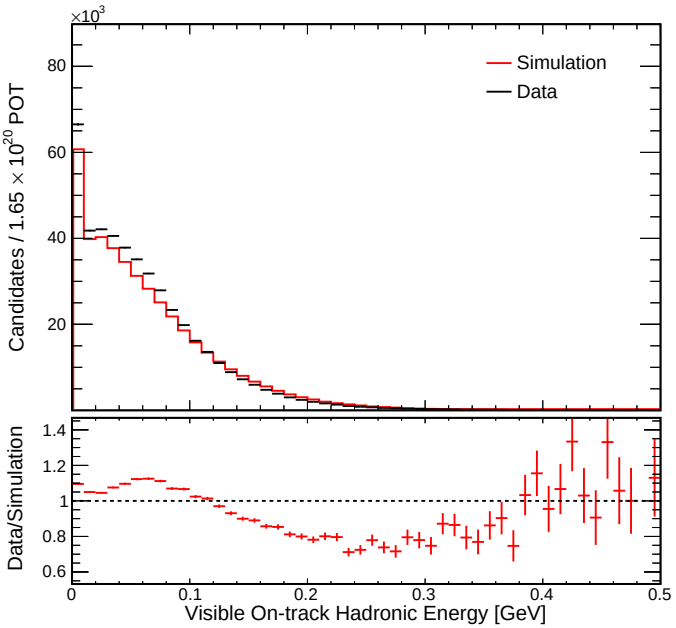


Figure 12.29: Plot of the sum of the visible energy (in GeV) associated with hadronic energy in the vertex region of the 3D Kalman track with the highest ReMId value. This plot is for the near detector populations. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

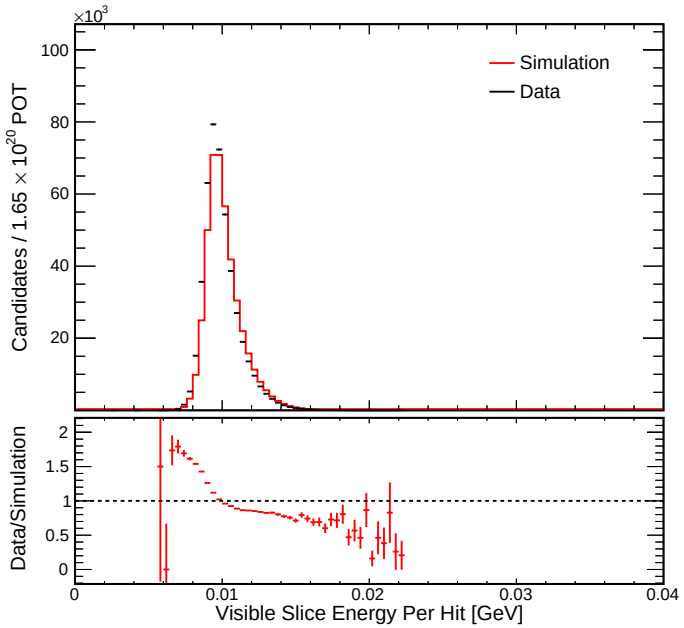


Figure 12.30: Plot of the sum of the visible energy (in GeV) for the slice hits divided by the total number of hits in the slice. Each slice is one entry in the histogram. This plot is for the near detector populations. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

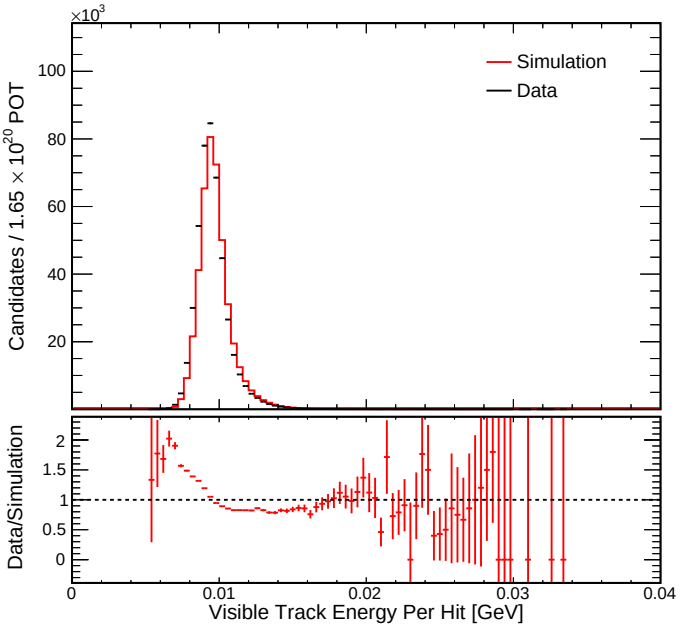


Figure 12.31: Plot of the sum of the visible energy (in GeV) for the hits associated the 3D Kalman track with the highest ReMId value divided by the number of hits associated with the primary track. Each primary track is one entry in the histogram. This plot is for the near detector populations. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

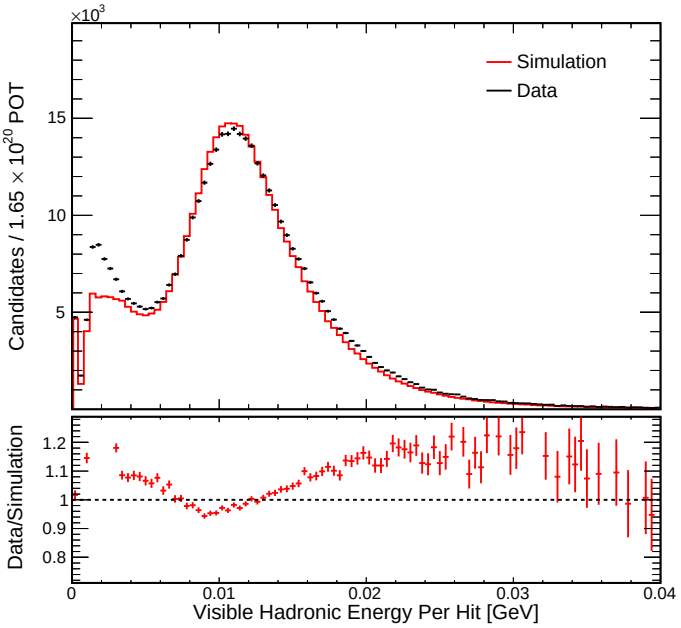


Figure 12.32: Plot of the sum of the visible energy (in GeV) for the slice hits not associated the 3D Kalman track with the highest ReMId value divided by the number of hits in the slice not associated with the primary track. Each slice is one entry in the histogram. This plot is for the near detector populations. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

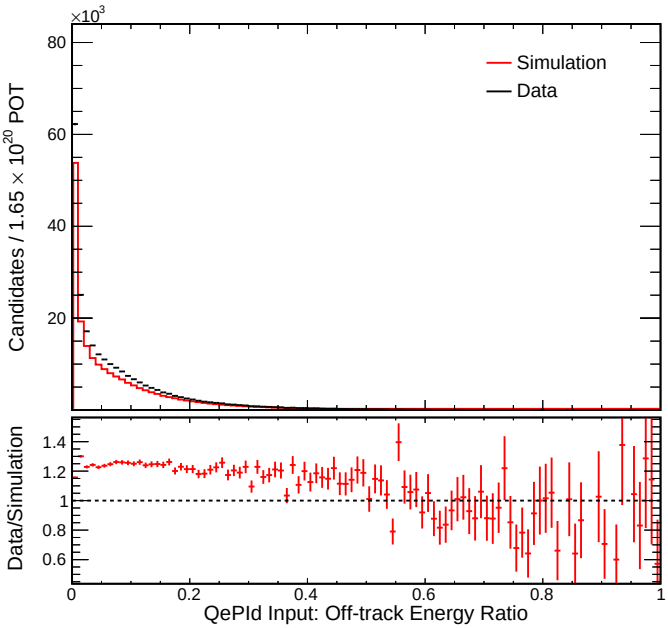


Figure 12.33: Plot of off-track energy ratio for a slice in the near detector. This is for the one track sample. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

12.5 QePId Inputs

To make our final analysis sample, we need to apply a QePId cut to create a QE and a nonQE population. First, let us understand the inputs to QePId. The criteria applied to create our populations include file selection, detailed in Section 10.1, spill cuts, listed in Section 10.2, and near detector slice cuts, from Section 10.3.2. A ReMId cut requiring a value of greater than 0.7 has been applied, but the samples have not been divided into QE and non-QE populations. For more information on ReMId, see Section 7.6. For more information on QePId, see Section 7.8. Each plot has one entry per slice. Plots are either for the one track or two track sample. For more information on these samples, see Section 7.8.

By looking at the inputs to QePId, we will better understand any differences in our

final analysis populations. Figure 12.33, the off-track energy ratio for the one track sample, is one input to QePId. Figure 12.34 displays the off-track energy ratio, but for the two track sample. We see that the data simply has more candidates for the one track sample than the simulation predicts. This makes sense given the general trend of less hadronic energy seen overall. For the two track sample, we see that the data has an excess at low values of non-muon energy. Again, this is consistent with our previous observations. Figure 12.35 shows the fractional energy difference for the one track sample and Figure 12.36 is for the two track sample. Again, we note the overall excess in data for the one track sample. The fractional energy difference Z-test for the one track sample is plotted in Figure 12.37; the two track sample is shown in Figure 12.38. The shapes agree somewhat for these distributions. Figure 12.39 plots the dE/dx ratio, only applicable to the two track sample. Here we see some shape differences in the lower region, although the disagreement isn't catastrophic.

Finally, Figure 12.40 shows the QePId distribution for the one track sample and Figure 12.41 shows the QePId distribution for the two track sample. For the one track sample, a cut is made at 0.45. Slices with higher values are considered part of the QE sample; slices with lower values are part of the nonQE sample. Again, we note the overall excess in data of one track candidates. For the two track sample, the cut is made at 0.4. We see a clear trend in the data towards events being more QE-like. The rest of the plots in this chapter will include either the QE or the nonQE population.

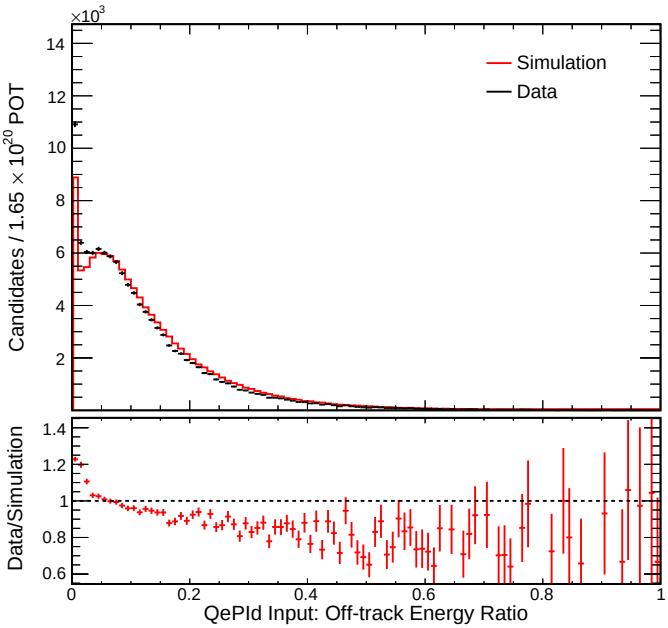


Figure 12.34: Plot of off-track energy ratio for a slice in the near detector. This is for the two track sample. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. When the ratio is too large for the scale, the point and its error bars are not drawn. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

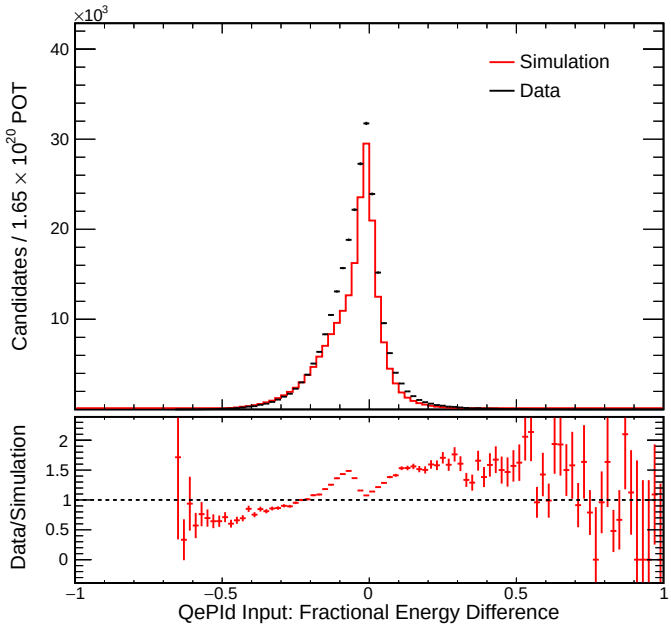


Figure 12.35: Plot of the fractional energy difference for a slice in the near detector. This is for the one track sample. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

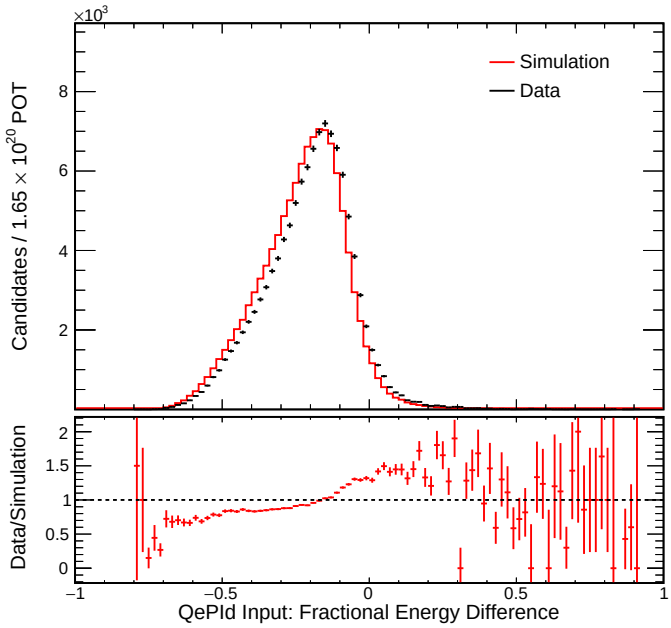


Figure 12.36: Plot of the fractional energy difference for a slice in the near detector. This is for the two track sample. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

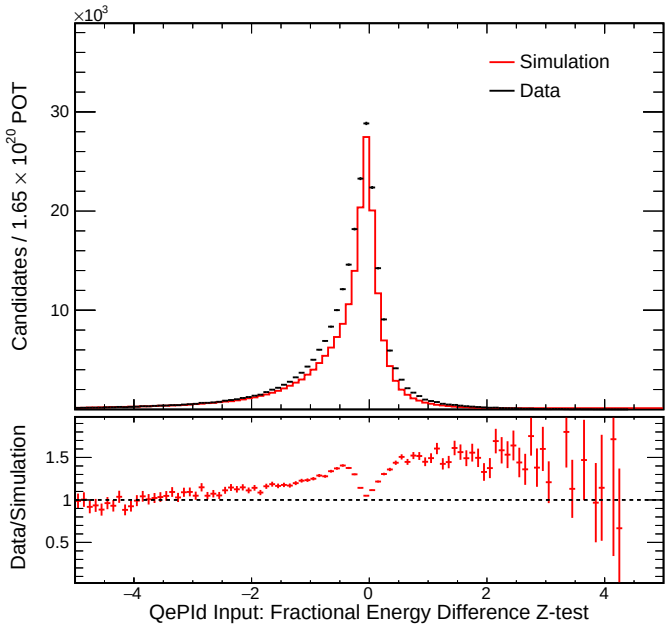


Figure 12.37: Plot of the fractional energy difference Z-test for a slice in the near detector. This is for the one track sample. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

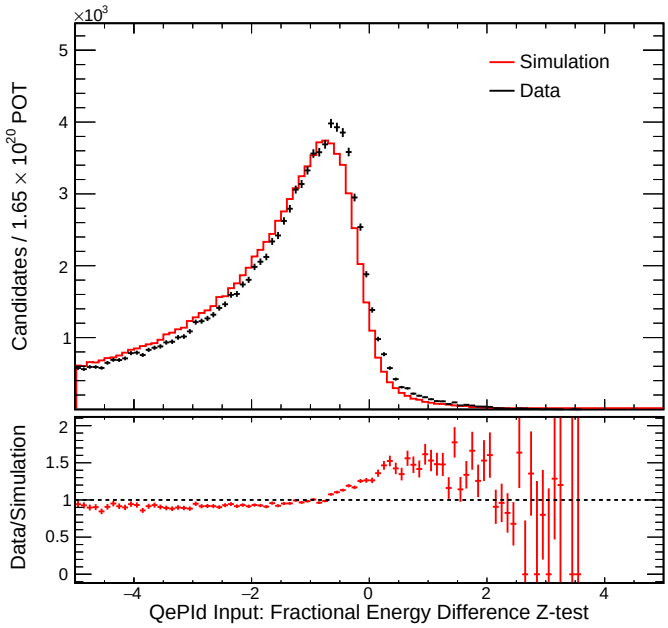


Figure 12.38: Plot of the fractional energy difference Z-test for a slice in the near detector. This is for the two track sample. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

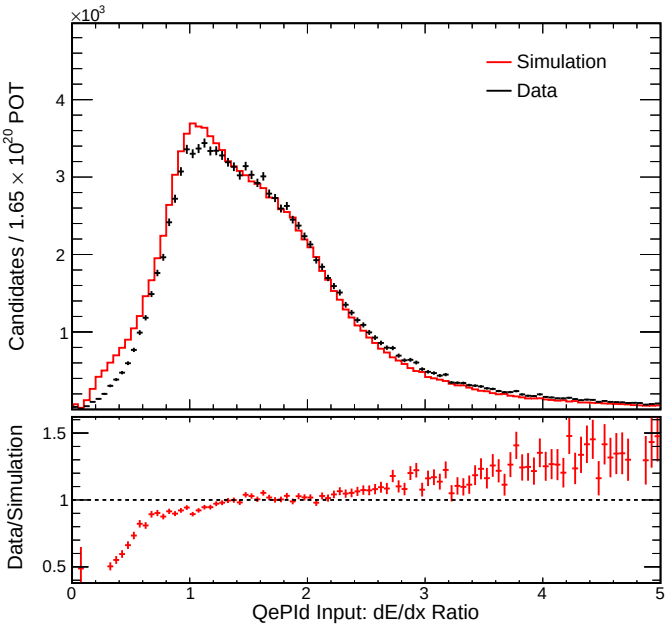


Figure 12.39: Plot of the dE/dx ratio for a slice in the near detector. This is for the two track sample. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

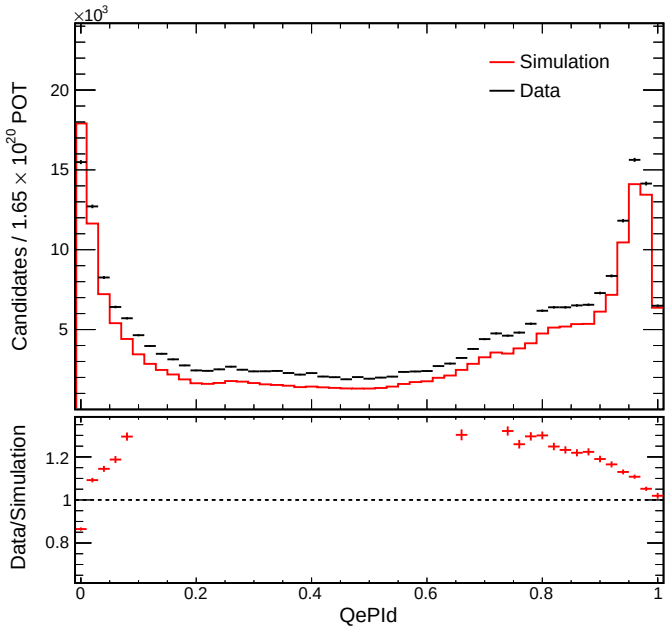


Figure 12.40: Plot of QePId for a slice in the near detector. This is for the one track sample. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

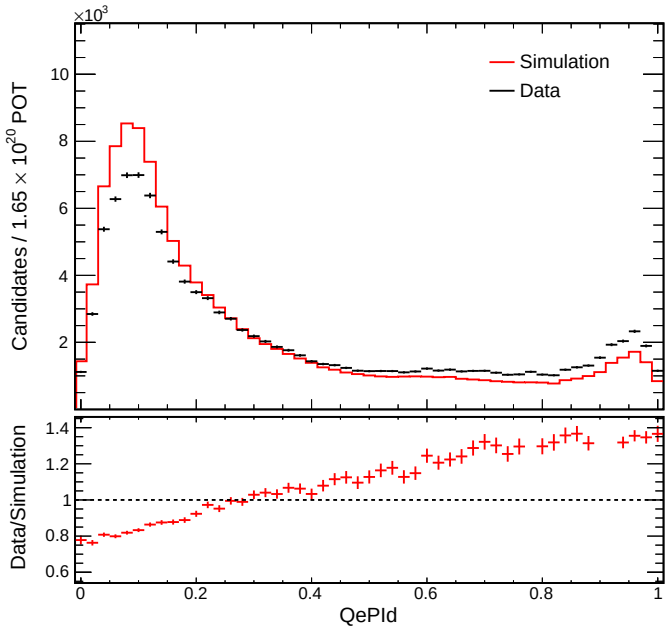


Figure 12.41: Plot of QePId for a slice in the near detector. This is for the two track sample. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

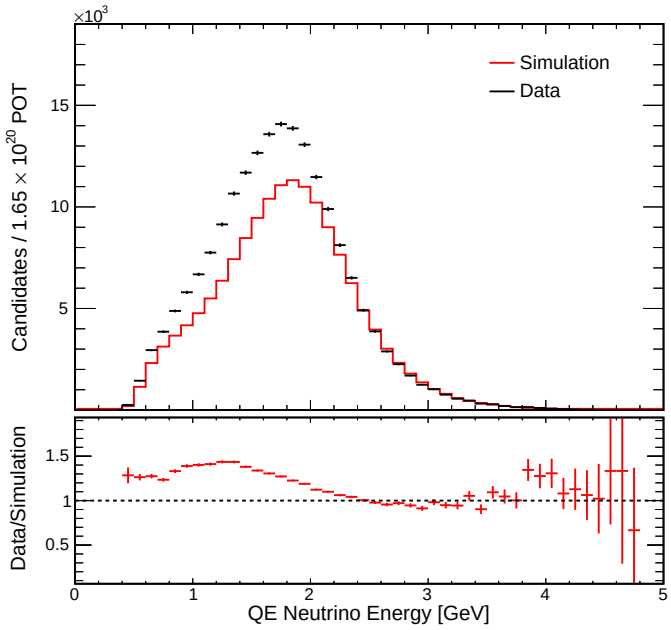


Figure 12.42: Plot of reconstructed neutrino energy in GeV for a slice in the near detector. This is for the QE sample. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2, 10.3.2 and 10.4 are applied.

12.6 Reconstructed Neutrino Energy

Having made our final analysis populations, we can now look at the reconstructed neutrino energy for each population. The reconstructed neutrino energy distributions for the far detector will be fit to make our measurement of the oscillation parameters. Figure 12.42 displays the reconstructed neutrino energy for the QE sample. We can see that the data has an excess compared to the simulated prediction. Also, the data is shifted towards lower neutrino energies. Figure 12.43 displays the reconstructed neutrino energy for the nonQE sample. Again, we see the shift towards lower neutrino energies. This is primarily due to the differences in the number of hadronic hits seen between the data and the simulation. The differences in these samples will be addressed further in Section 14.12.

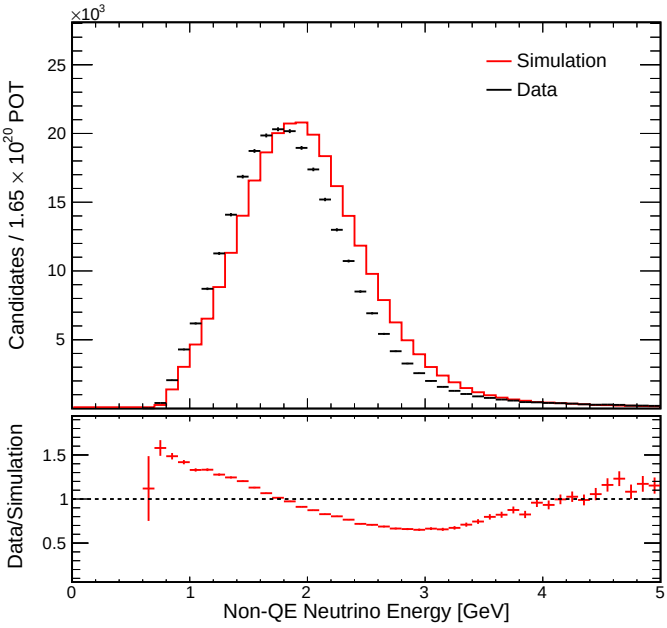


Figure 12.43: Plot of reconstructed neutrino energy in GeV for a slice in the near detector. This is for the nonQE sample. The simulation distribution is displayed as a red line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2, 10.3.2 and 10.4 are applied.

Chapter 13

Extrapolation

NOνA consists of two detectors. We can compare the near detector data and simulation. Then, the differences between the data and simulation in the near detector can be used to alter the far detector simulated prediction. This procedure is called *extrapolation*. Figure 13.1 is a diagram of the extrapolation procedure. This process reduces the need for perfect simulation and results in lower systematic errors. By using our near detector data to "fix" our far detector simulation prediction, we reduce our dependence on modeling things like the beam flux and neutrino interactions.

We perform the extrapolation procedure individually on different samples. Dividing our analysis sample into parts to be extrapolated is the process of *decomposition*. For this analysis, we extrapolate the QE and nonQE analysis samples separately. We decompose each of these samples into a ν_μ CC sample, a $\bar{\nu}_\mu$ CC sample, and a NC sample. Section 13.1 details the decomposition.

For each sample which we extrapolate, we compare the reconstructed neutrino energy spectrum between the near detector data and simulation. We alter the near detector simulation to match the data. Then, using a 2D histogram created from the simulation that relates reconstructed neutrino energy to true neutrino energy, we can create an altered true neutrino energy spectrum for the near detector. Taking into account the differences between the far and near detector, like different neutrino beam fluxes and detector sizes, we create an extrapolated true neutrino energy spectrum for the far detector. Finally, using another 2D histogram created from the far detector simulation that relates true neutrino energy and reconstructed neutrino energy, we create an

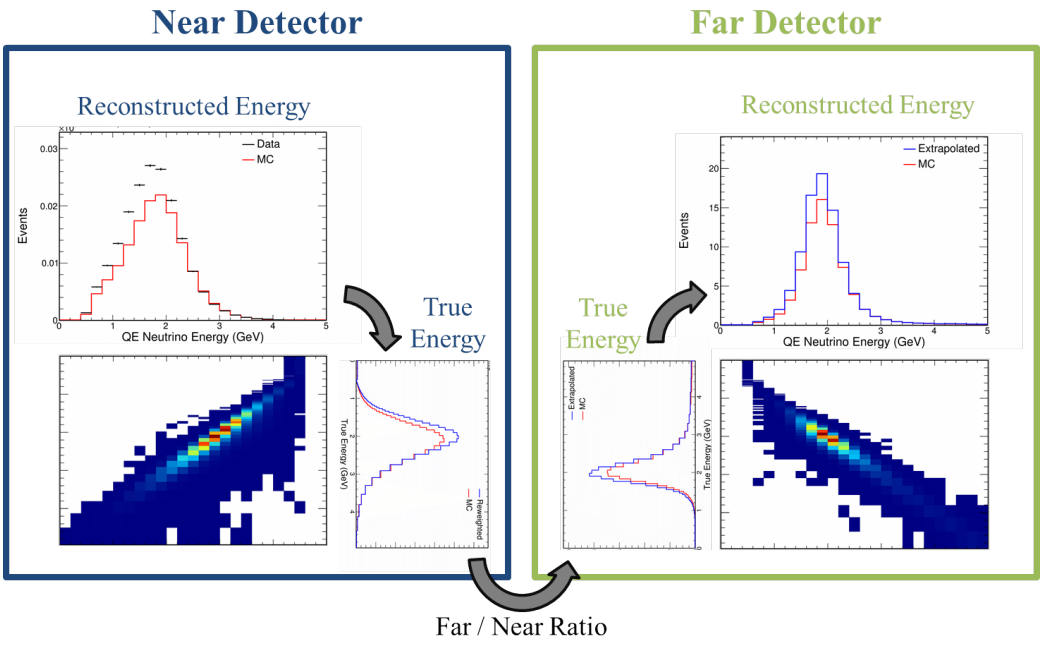


Figure 13.1: Diagram of the extrapolation procedure. For each sample, compare the near detector data and simulation reconstructed neutrino energy spectra. Alter the near detector reconstructed neutrino energy spectrum to match the data. Using a 2D histogram from the simulation which relates reconstructed neutrino energy and true neutrino energy, determine the reweighted true neutrino energy spectrum for the near detector. Taking into account the differences in the predicted true neutrino spectrum of the far and near detectors (see different solid angle of the beam, different detector sizes, etc), use the reweighted true neutrino energy spectrum at the near detector to create an extrapolated true neutrino energy spectrum at the far detector. Finally, because we measure reconstructed neutrino energy and not true neutrino energy, we use a 2D histogram from the far detector simulation that relates true neutrino energy and reconstructed neutrino energy to create a final, extrapolated far detector predicted neutrino energy spectrum. The figures used correspond to the ν_μ CC interactions for the QE population.

extrapolated far detector reconstructed neutrino energy spectrum. Section 13.2 gives more details about the extrapolation. Section 13.3 contains the plots used to create an altered far detector prediction spectrum for the ν_μ CC interactions. Section 13.4 contains the plots for the $\bar{\nu}_\mu$ CC interactions; section 13.5 contains the plots for NC interactions.

13.1 Near Detector Decomposition

Decomposition is the procedure we use to divide our sample into different parts that are extrapolated individually. Decomposition allows us to create populations of different types of neutrino interactions and extrapolate them using methods best suited for each type. The event selection described in Chapter 10 for the near detector is applied to both the data and the simulation. NonQE and QE populations are created. These populations are decomposed separately. The decomposition is done in bins of reconstructed neutrino energy.

We do a proportional decomposition for true ν_μ CC interactions, true $\bar{\nu}_\mu$ CC interactions, and true NC interactions. We do not decompose or extrapolate true ν_e CC interactions, true $\bar{\nu}_e$ CC interactions, true ν_τ CC interactions, or true $\bar{\nu}_\tau$ CC interactions. These are ignored because they represent such a small background to the analysis that simply using the simulation far detector prediction is a negligible systematic error.

For each type of interaction for which we will make a decomposition, we make a spectrum of reconstructed neutrino energy with simulation events of the correct type that pass the event selection cuts. For data, we do not know what type of true interaction each event corresponds to. Instead, for each bin in reconstructed neutrino energy, we know the proportion of the total sample in simulation that corresponds to the correct interaction type. We apply this proportion bin-by-bin to the near detector data spectrum to create a near detector interaction-type data spectrum. If, for instance, one bin of reconstructed neutrino energy in the simulation was composed of 80% true ν_μ CC interactions, then the data spectrum would be scaled by 0.8 in that bin.

Having reconstructed neutrino energy spectra for the near detector data and simulation corresponding to a single interaction type, we can now perform the extrapolation.

13.2 Extrapolation

For true ν_μ CC interactions and true $\bar{\nu}_\mu$ CC interactions, the altered reconstructed neutrino energy spectrum is used to construct an altered true neutrino energy spectrum. From the unaltered simulation, a 2D histogram is created of reconstructed neutrino energy vs. true neutrino energy. The simulated reconstructed neutrino energy spectrum is changed to match the data spectrum. This change is propagated, using the 2D histogram, to create a reweighted true neutrino energy spectrum. If a bin of reconstructed neutrino energy in the simulation is reduced by 10%, then this is also applied to each true neutrino energy bin in the 2D histogram which corresponds to that bin of reconstructed neutrino energy. By collapsing the altered 2D histogram onto the true neutrino energy axis, we can create a spectrum of altered true neutrino energy simulation for the near detector.

The altered true neutrino energy spectrum for the near detector is used to alter the predicted true neutrino energy spectrum for the far detector. This is then applied to a 2D histogram of reconstructed neutrino energy vs. true neutrino energy for the far detector simulation. This can then be collapsed to create an altered reconstructed neutrino energy prediction spectrum for far detector.

For true NC interactions, there is not a strong correlation between the reconstructed neutrino energy spectrum and the true neutrino energy spectrum. Therefore, the extra step of translation to true neutrino energy spectrum is not taken. Instead, the altered reconstructed neutrino energy spectrum for the near detector is used to alter the predicted reconstructed energy spectrum for the far detector.

For more details, see the extrapolation technical note[68].

13.3 ν_μ CC Interactions

This section contains the plots which illustrate how the far detector simulated prediction for ν_μ CC interactions is altered due to the near detector data and simulation differences. The true ν_μ CC population is created from the analysis population by performing a proportional decomposition. This decomposition uses the simulation to determine the proportion of true ν_μ CC interactions within our analysis sample. See Section 13.1 for more details. This is done independently for the QE and nonQE populations; Section

13.3.1 contains the plots for the QE population and section 13.3.2 contains the plots for the nonQE population.

13.3.1 QE Population

This section presents the plots for the QE population (as defined by the event selection in Chapter 10, specifically Section 10.4). The first step is to compare the reconstructed neutrino energy spectrum for ν_μ CC interactions in the near detector data and simulation. Figure 13.2 plots this comparison; Figure 13.3 has a logarithmic vertical axis. A 2D histogram of reconstructed neutrino energy vs. true neutrino energy for the near detector is created from the simulation. Figure 13.4 is the unaltered 2D histogram. Figure 13.5 is the same as Figure 13.4, except that it has a logarithmic color axis. The 2D histogram is then altered to make the reconstructed neutrino energy spectrum in the simulation match the near detector data. Figure 13.6 is the adjusted 2D histogram. Figure 13.7 is the same as Figure 13.6, except that it has a logarithmic color axis. The adjusted 2D histogram is projected onto the true neutrino energy axis to create an altered spectrum. Figure 13.8 compares the altered true neutrino energy spectrum and the original true neutrino energy spectrum from simulation. Figure 13.9 is the same as Figure 13.8, except that the vertical axis is displayed logarithmically.

Having an altered true neutrino energy spectrum in the simulation for the near detector, we now use it to change our far detector prediction. Figure 13.10 compares the original true neutrino energy spectrum for the far detector with one altered to match the extrapolated near detector data. Figure 13.11 is the same, except that it features a logarithmic vertical axis. A 2D histogram of reconstructed neutrino energy vs. true neutrino energy for the far detector is created from the simulation. Figure 13.12 is the unaltered 2D histogram. Figure 13.13 is the same as Figure 13.12, except that it has a logarithmic color axis. The 2D histogram is then altered to make the true neutrino energy spectrum in the simulation match the extrapolated near detector data. Figure 13.14 is the adjusted 2D histogram. Figure 13.15 is the same as Figure 13.14, except that it has a logarithmic color axis. The altered 2D histogram is then projected onto the reconstructed neutrino energy axis. This is the unoscillated prediction used for the analysis. It is compared to the unaltered simulation prediction in Figure 13.16. Figure 13.17 is the same comparison, but featuring a logarithmic vertical axis.

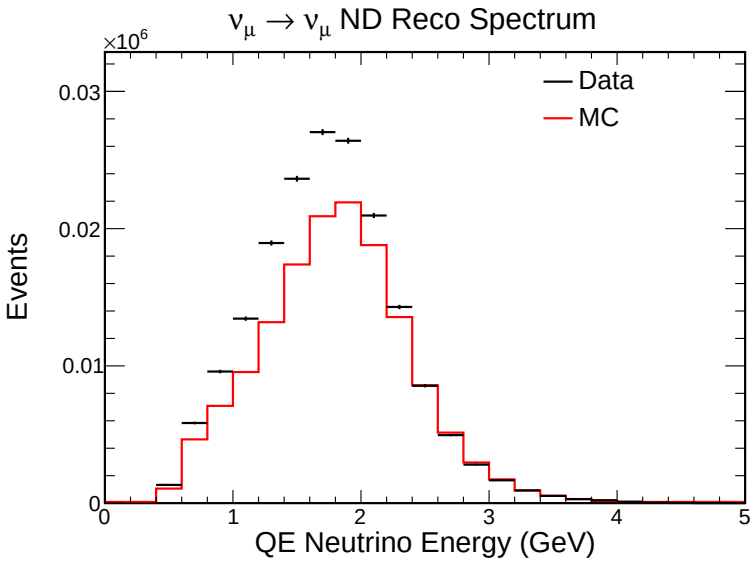


Figure 13.2: Plot comparing data and simulation reconstructed neutrino energies in GeV in the near detector for ν_μ CC interactions. The solid red histogram is simulation; the black points are data. The error bars displayed on the data are statistical only. The simulation is scaled down to match the POT of the data. This is for the QE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

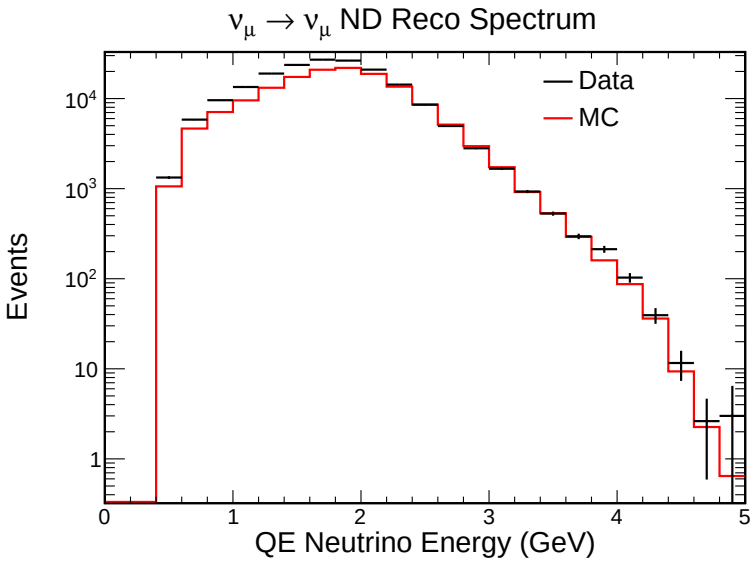


Figure 13.3: Plot comparing data and simulation reconstructed neutrino energies in GeV in the near detector for ν_μ CC interactions. The vertical axis is plotted logarithmically. The solid red histogram is simulation; the black points are data. The error bars displayed on the data are statistical only. The simulation is scaled down to match the POT of the data. This is for the QE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

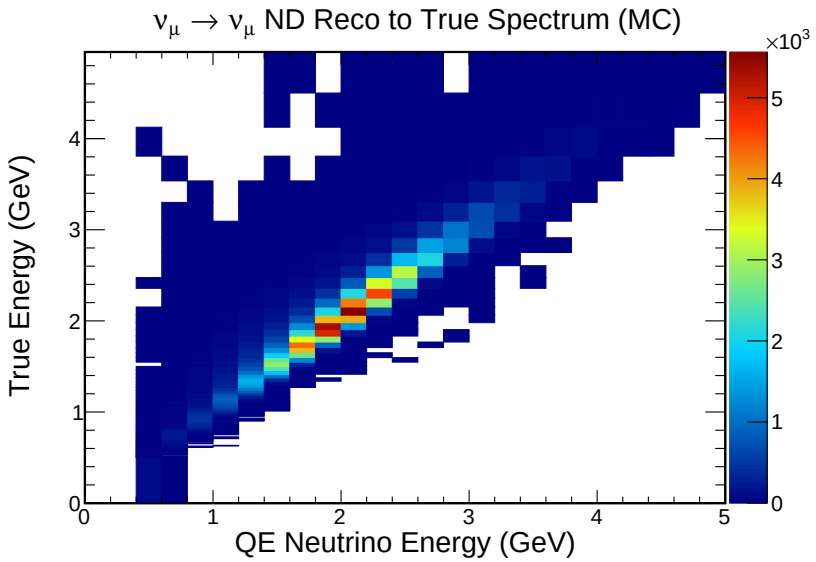


Figure 13.4: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the near detector for ν_μ CC interactions. This plot has not been altered to match near detector data. The color axis is number of interactions. The simulation is scaled down to match the POT of the data. This is for the QE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

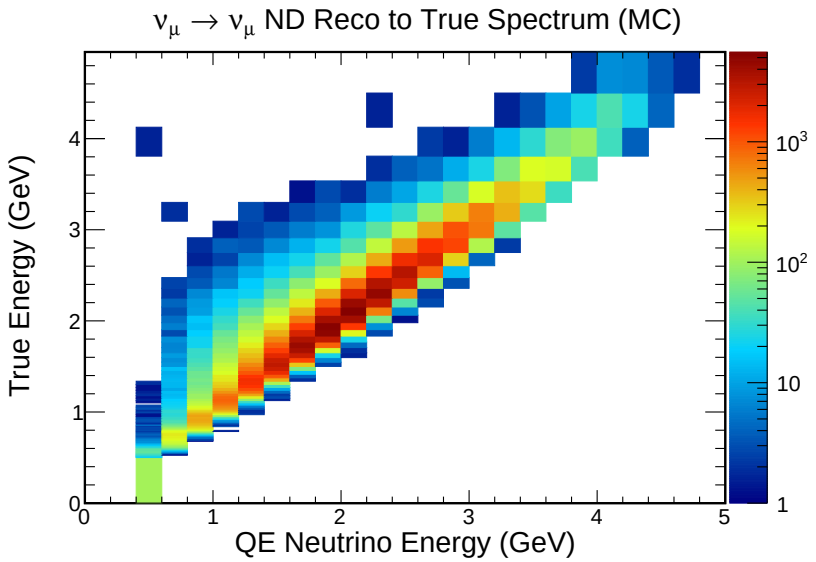


Figure 13.5: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the near detector for ν_μ CC interactions. This plot has not been altered to match near detector data. The color axis is number of interactions and is plotted logarithmically. The simulation is scaled down to match the POT of the data. This is for the QE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

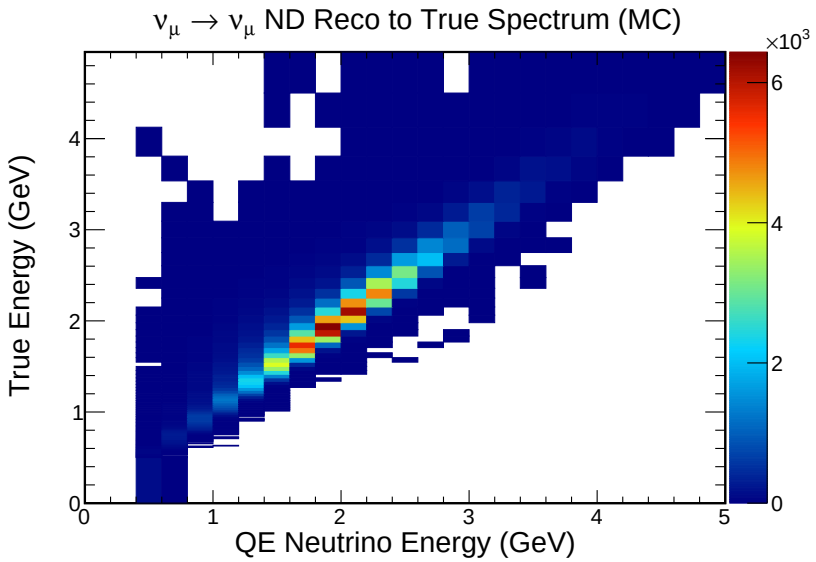


Figure 13.6: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the near detector for ν_μ CC interactions. This plot has been altered to match near detector data. The color axis is number of interactions. The simulation is scaled down to match the POT of the data. This is for the QE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

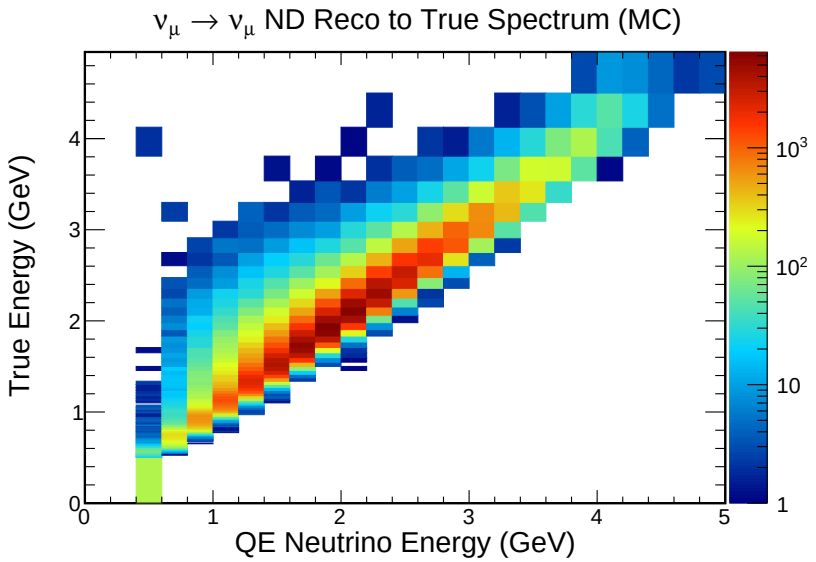


Figure 13.7: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the near detector for ν_μ CC interactions. This plot has been altered to match near detector data. The color axis is number of interactions and is plotted logarithmically. The simulation is scaled down to match the POT of the data. This is for the QE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

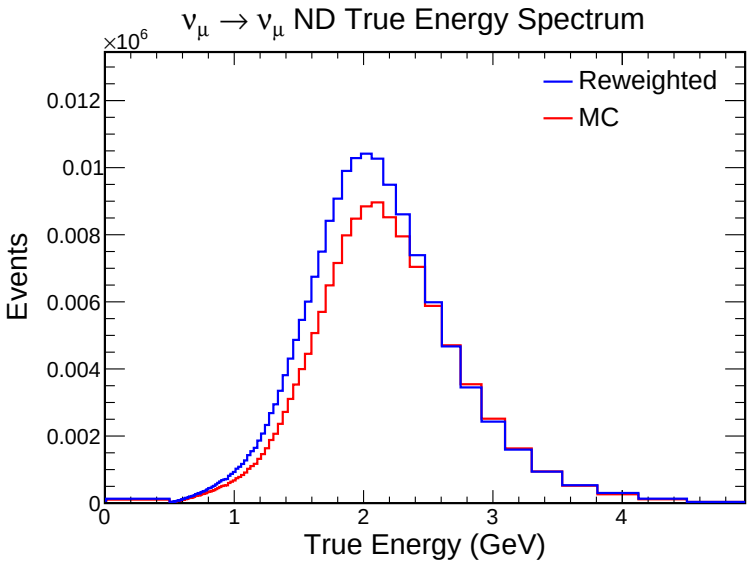


Figure 13.8: Plot comparing unaltered and reweighted-by-data true neutrino energies in GeV using simulated events in the near detector for ν_μ CC interactions. The solid red histogram is unaltered simulation; the solid blue histogram is the simulation spectrum after reweighting by data. The simulation is scaled down to match the POT of the data. This is for the QE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

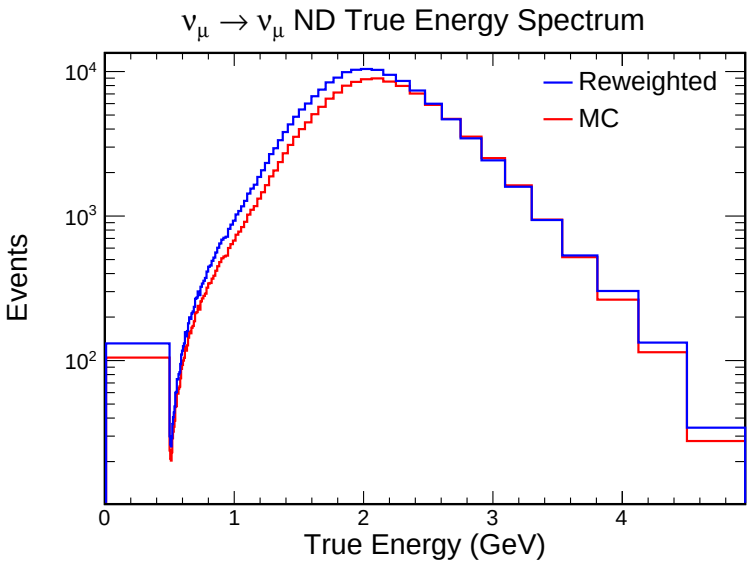


Figure 13.9: Plot comparing unaltered and reweighted-by-data true neutrino energies in GeV using simulated events in the near detector for ν_μ CC interactions. The vertical axis is plotted logarithmically. The solid red histogram is unaltered simulation; the solid blue histogram is the simulation spectrum after reweighting by data. The simulation is scaled down to match the POT of the data. This is for the QE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

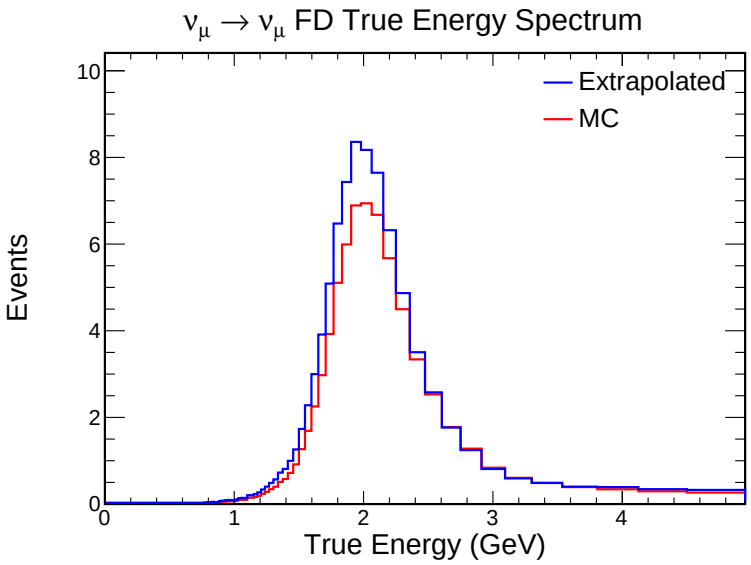


Figure 13.10: Plot comparing unaltered and reweighted-by-data true neutrino energies in GeV using simulated events in the far detector for ν_μ CC interactions. The solid red histogram is unaltered simulation; the solid blue histogram is the simulation spectrum after reweighting by the extrapolation of near detector data. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the QE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

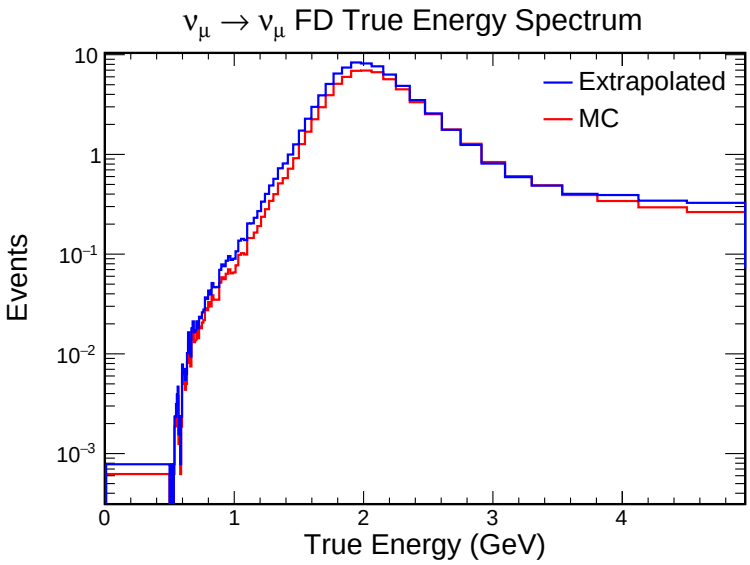


Figure 13.11: Plot comparing unaltered and reweighted-by-data true neutrino energies in GeV using simulated events in the far detector for ν_μ CC interactions. The vertical axis is plotted logarithmically. The solid red histogram is unaltered simulation; the solid blue histogram is the simulation spectrum after reweighting by the extrapolation of near detector data. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the QE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

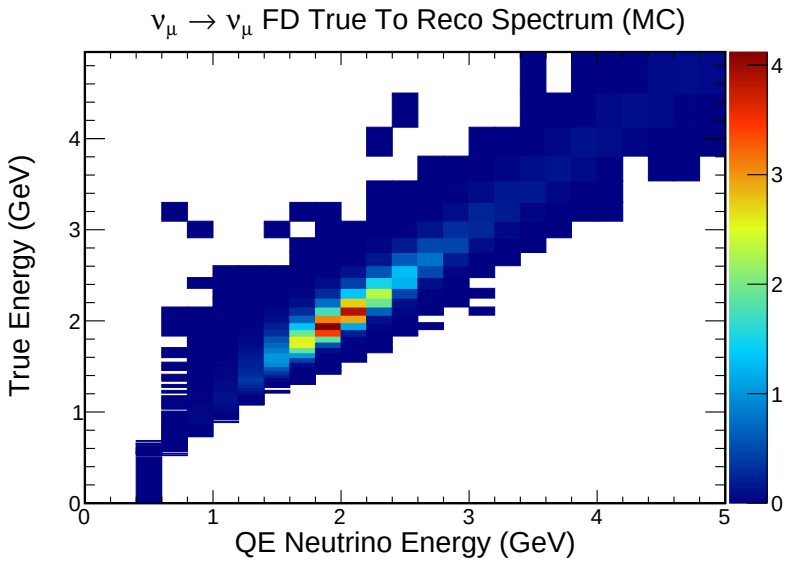


Figure 13.12: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the far detector for ν_μ CC interactions. This plot has not been altered to match the extrapolation of near detector data. The color axis is number of interactions. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the QE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

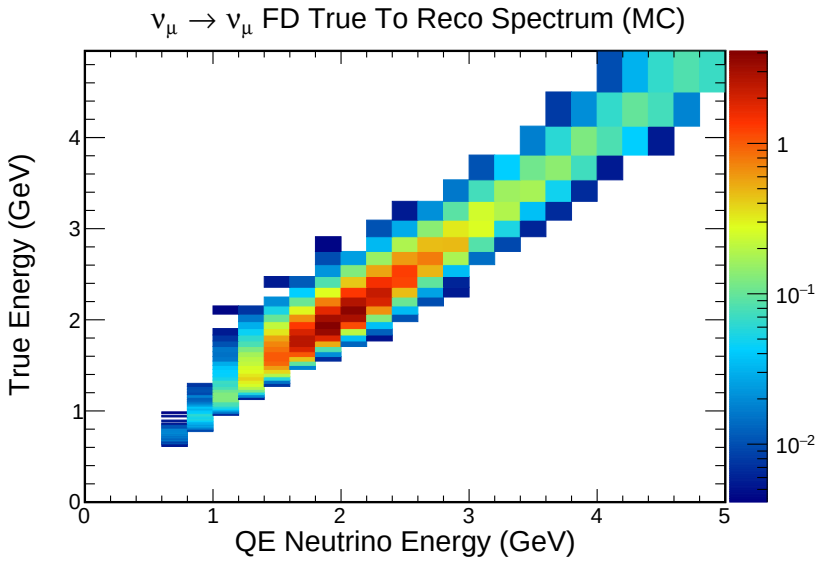


Figure 13.13: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the far detector for ν_μ CC interactions. This plot has not been altered to match the extrapolation of near detector data. The color axis is number of interactions and is plotted logarithmically. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the QE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

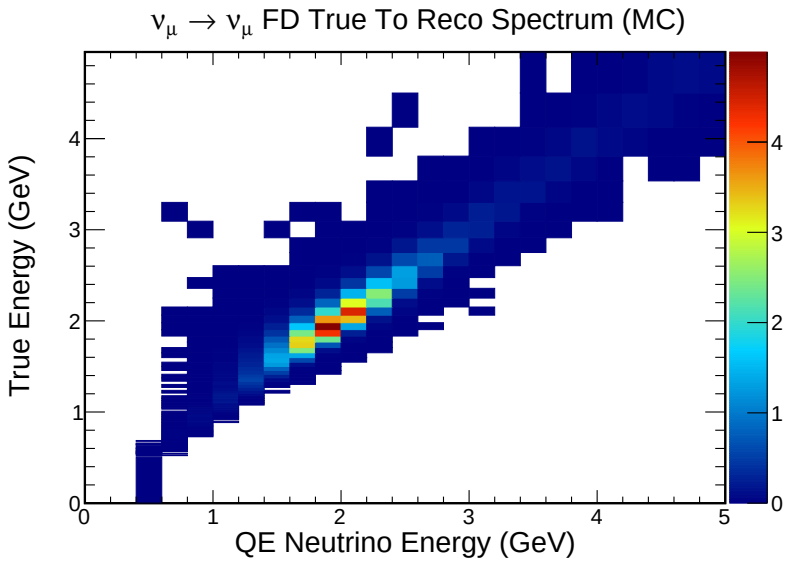


Figure 13.14: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the far detector for ν_μ CC interactions. This plot has been altered to match the extrapolation of near detector data. The color axis is number of interactions. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the QE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

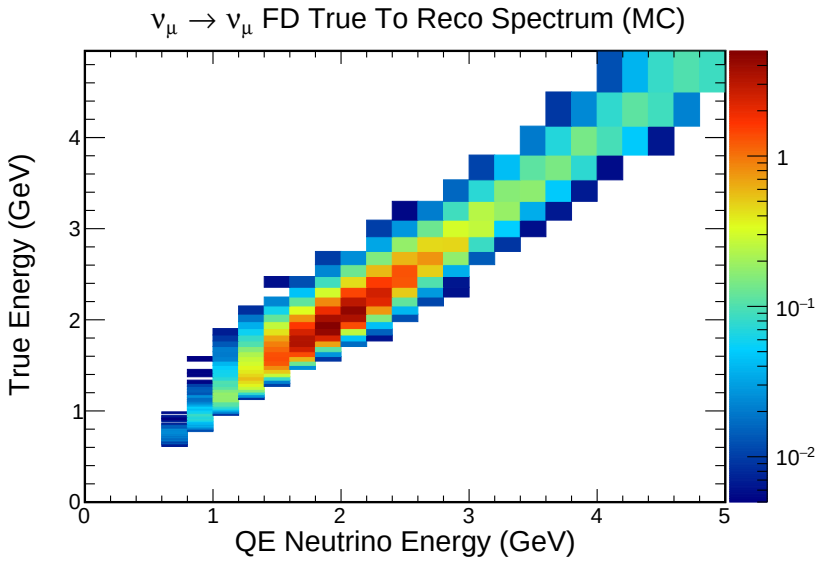


Figure 13.15: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the far detector for ν_μ CC interactions. This plot has been altered to match the extrapolation of near detector data. The color axis is number of interactions and is plotted logarithmically. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the QE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

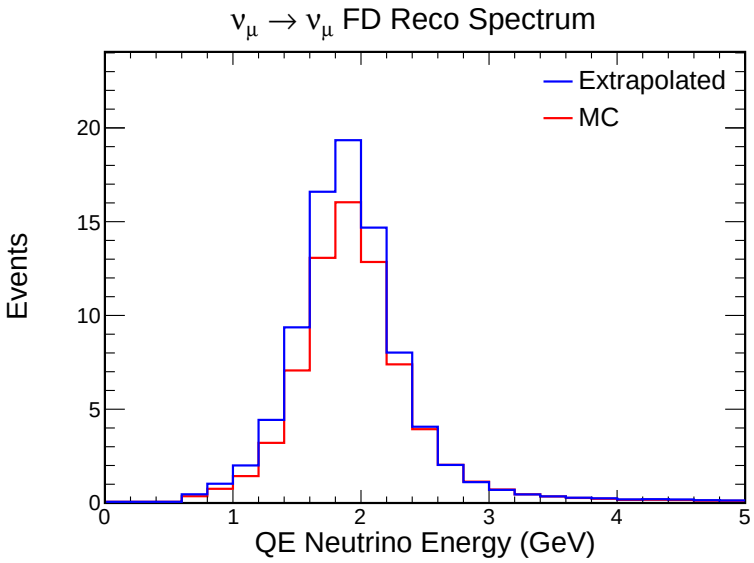


Figure 13.16: Plot comparing unaltered and reweighted-by-data reconstructed neutrino energies in GeV using simulated events in the far detector for ν_μ CC interactions. The solid red histogram is unaltered simulation; the solid blue histogram is the simulation spectrum after reweighting by the extrapolation of near detector data. The extrapolated spectrum is used by the analysis as its unoscillated prediction. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the QE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

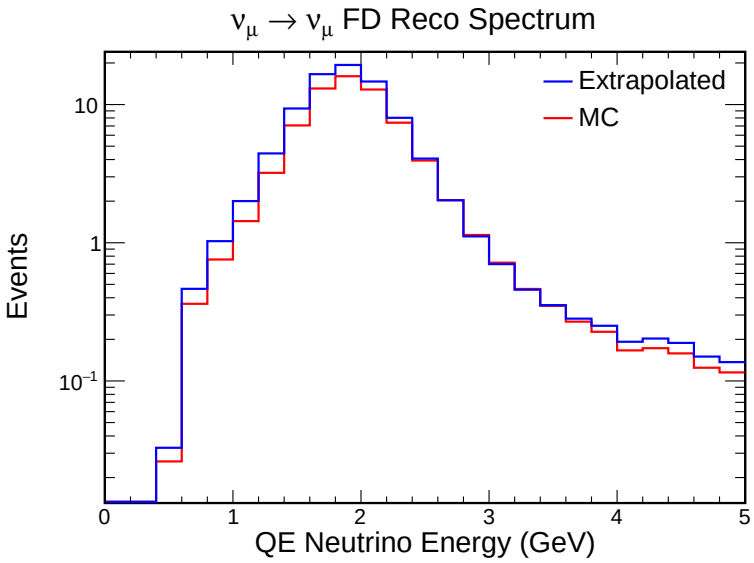


Figure 13.17: Plot comparing unaltered and reweighted-by-data reconstructed neutrino energies in GeV using simulated events in the far detector for ν_μ CC interactions. The vertical axis is plotted logarithmically. The solid red histogram is unaltered simulation; the solid blue histogram is the simulation spectrum after reweighting by the extrapolation of near detector data. The extrapolated spectrum is used by the analysis as its unoscillated prediction. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the QE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

13.3.2 NonQE Population

This section presents the plots for the nonQE population (as defined by the event selection in Chapter 10, specifically Section 10.4). The first step is to compare the reconstructed neutrino energy spectrum for ν_μ CC interactions in the near detector data and simulation. Figure 13.18 plots this comparison; Figure 13.19 has a logarithmic vertical axis. A 2D histogram of reconstructed neutrino energy vs. true neutrino energy for the near detector is created from the simulation. Figure 13.20 is the unaltered 2D histogram. Figure 13.21 is the same as Figure 13.20, except that it has a logarithmic color axis. The 2D histogram is then altered to make the reconstructed neutrino energy spectrum in the simulation match the near detector data. Figure 13.22 is the adjusted 2D histogram. Figure 13.23 is the same as Figure 13.22, except that it has a logarithmic color axis. The adjusted 2D histogram is projected onto the true neutrino energy axis to create an altered spectrum. Figure 13.24 compares the altered true neutrino energy spectrum and the original true neutrino energy spectrum from simulation. Figure 13.25 is the same as Figure 13.24, except that the vertical axis is displayed logarithmically.

Having an altered true neutrino energy spectrum for the near detector, we now use it to change our far detector prediction. Figure 13.26 compares the original true neutrino energy spectrum for the far detector with one altered to match the extrapolated near detector data. Figure 13.27 is the same, except that it features a logarithmic vertical axis. A 2D histogram of reconstructed neutrino energy vs. true neutrino energy for the far detector is created from the simulation. Figure 13.28 is the unaltered 2D histogram. Figure 13.29 is the same as Figure 13.28, except that it has a logarithmic color axis. The 2D histogram is then altered to make the true neutrino energy spectrum in the simulation match the extrapolated near detector data. Figure 13.30 is the adjusted 2D histogram. Figure 13.31 is the same as Figure 13.30, except that it has a logarithmic color axis. The altered 2D histogram is then projected onto the reconstructed neutrino energy axis. This is the unoscillated prediction used for the analysis. It is compared to the unaltered simulation prediction in Figure 13.32. Figure 13.33 is the same comparison, but featuring a logarithmic vertical axis.

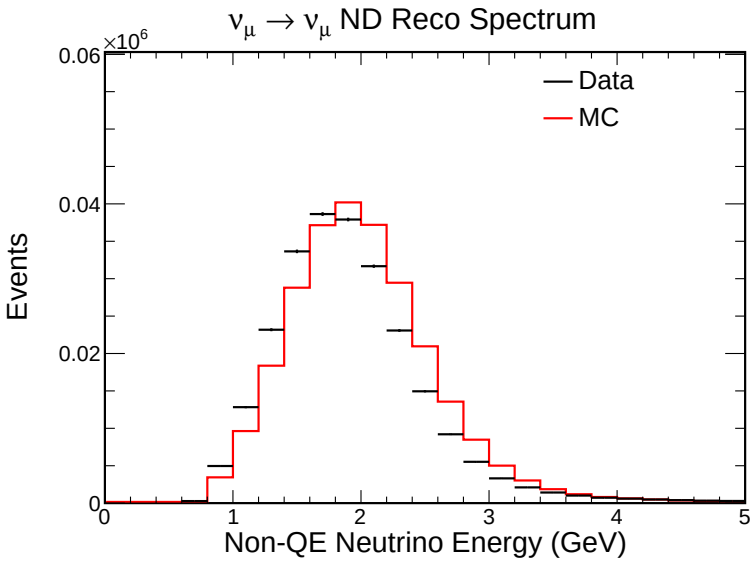


Figure 13.18: Plot comparing data and simulation reconstructed neutrino energies in GeV in the near detector for ν_μ CC interactions. The solid red histogram is simulation; the black points are data. The error bars displayed on the data are statistical only. The simulation is scaled down to match the POT of the data. This is for the nonQE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

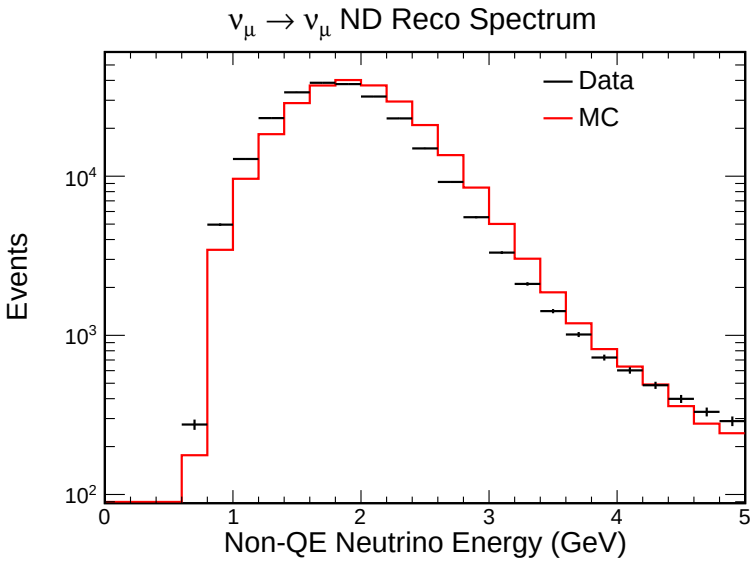


Figure 13.19: Plot comparing data and simulation reconstructed neutrino energies in GeV in the near detector for ν_μ CC interactions. The vertical axis is plotted logarithmically. The solid red histogram is simulation; the black points are data. The error bars displayed on the data are statistical only. The simulation is scaled down to match the POT of the data. This is for the nonQE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

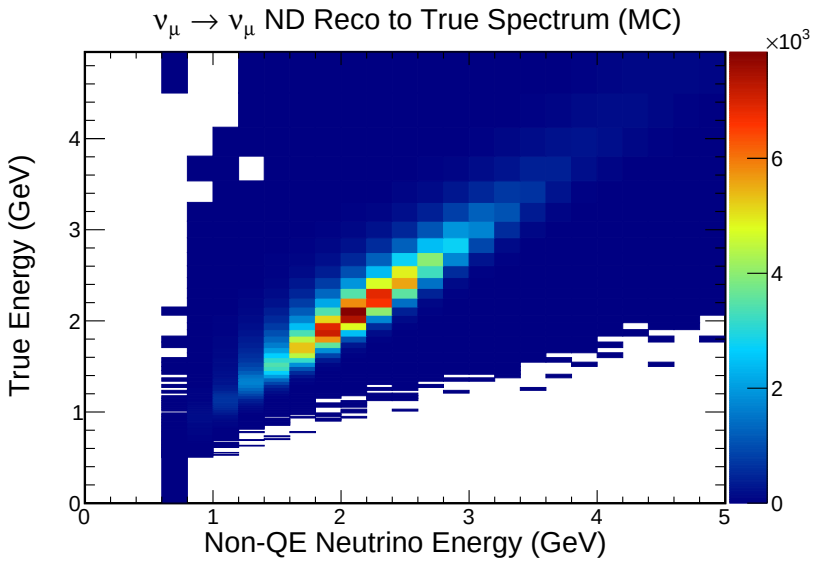


Figure 13.20: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the near detector for ν_μ CC interactions. This plot has not been altered to match near detector data. The color axis is number of interactions. The simulation is scaled down to match the POT of the data. This is for the nonQE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

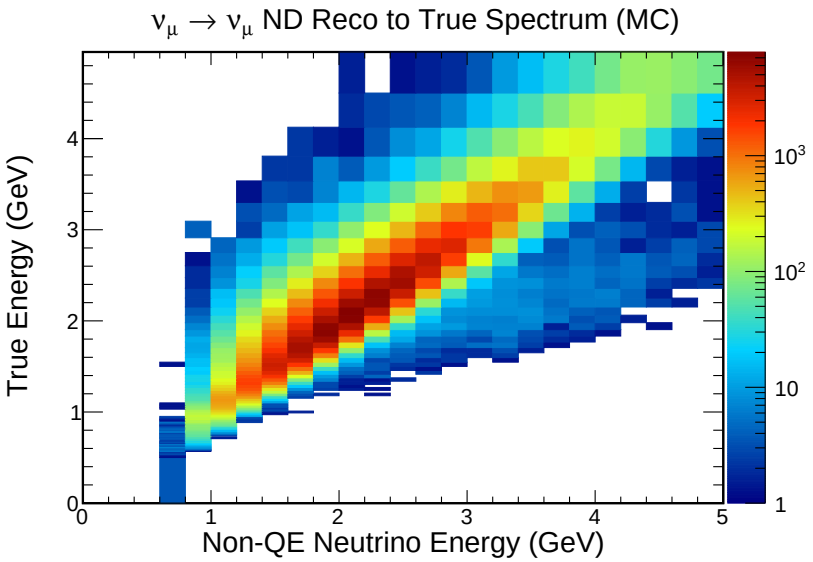


Figure 13.21: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the near detector for ν_μ CC interactions. This plot has not been altered to match near detector data. The color axis is number of interactions and is plotted logarithmically. The simulation is scaled down to match the POT of the data. This is for the nonQE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

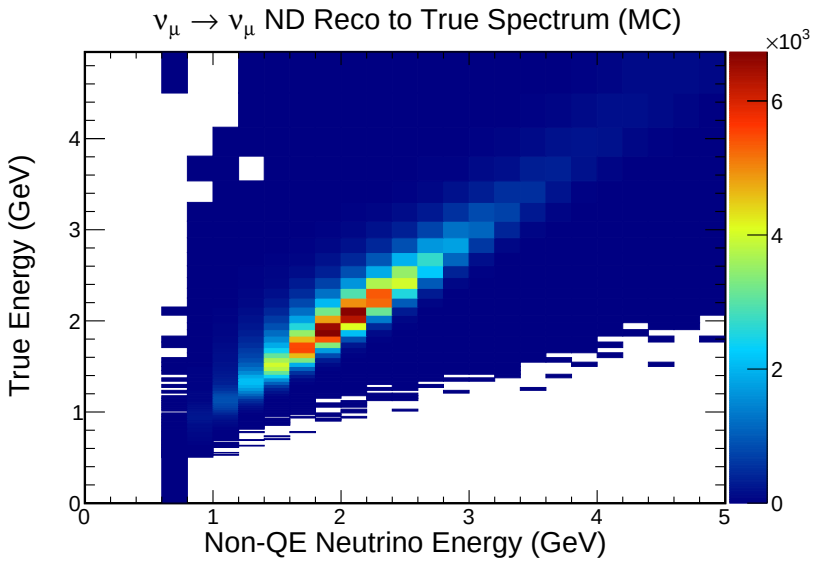


Figure 13.22: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the near detector for ν_μ CC interactions. This plot has been altered to match near detector data. The color axis is number of interactions. The simulation is scaled down to match the POT of the data. This is for the nonQE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

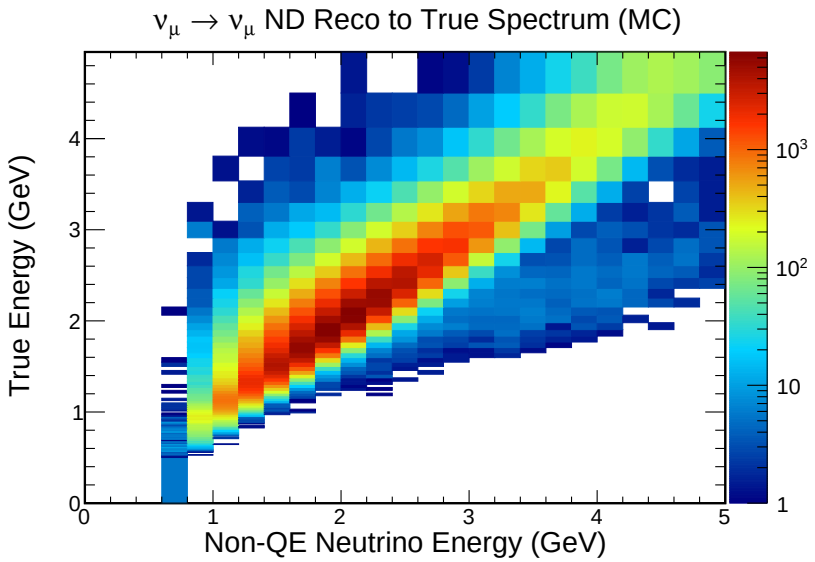


Figure 13.23: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the near detector for ν_μ CC interactions. This plot has been altered to match near detector data. The color axis is number of interactions and is plotted logarithmically. The simulation is scaled down to match the POT of the data. This is for the nonQE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

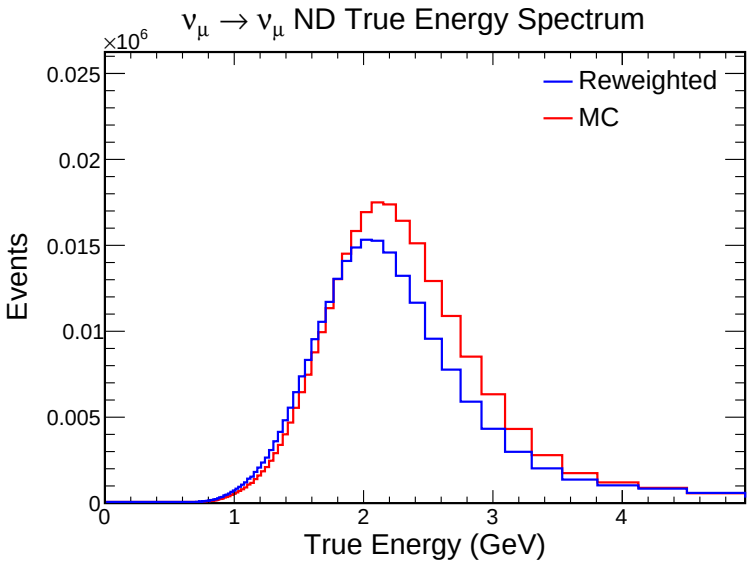


Figure 13.24: Plot comparing unaltered and reweighted-by-data true neutrino energies in GeV using simulated events in the near detector for ν_μ CC interactions. The solid red histogram is unaltered simulation; the solid blue histogram is the simulation spectrum after reweighting by data. The simulation is scaled down to match the POT of the data. This is for the nonQE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

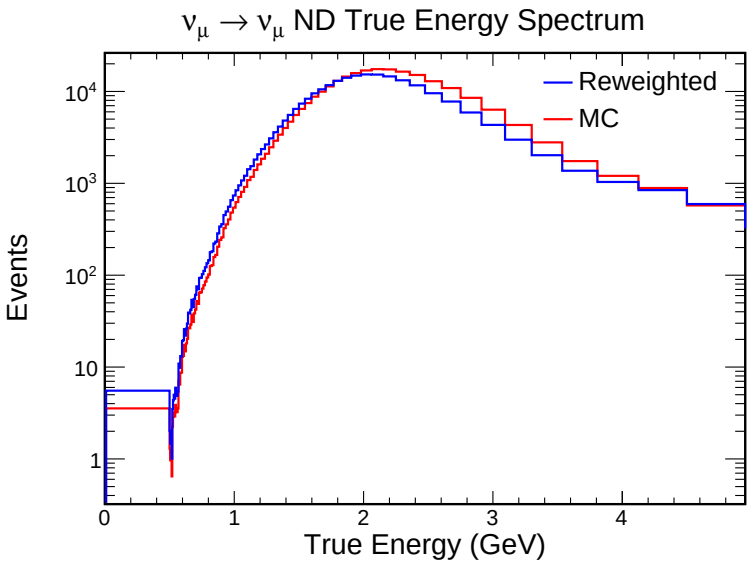


Figure 13.25: Plot comparing unaltered and reweighted-by-data true neutrino energies in GeV using simulated events in the near detector for ν_μ CC interactions. The vertical axis is plotted logarithmically. The solid red histogram is unaltered simulation; the solid blue histogram is the simulation spectrum after reweighting by data. The simulation is scaled down to match the POT of the data. This is for the nonQE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

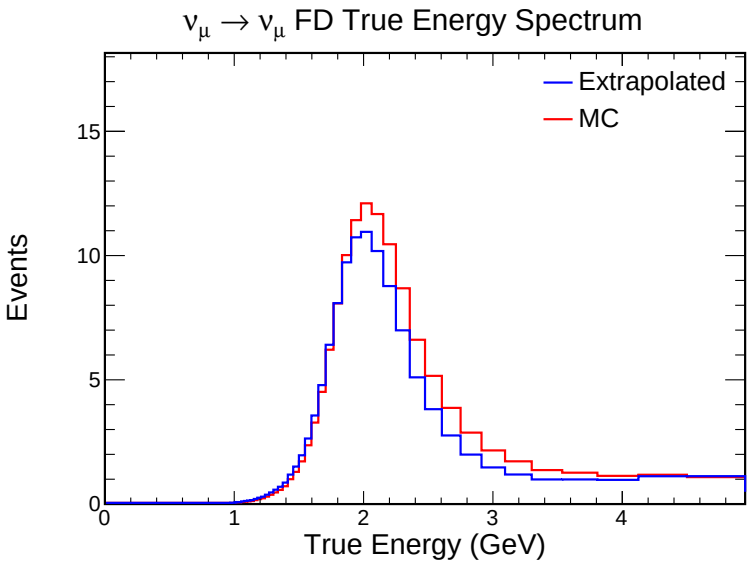


Figure 13.26: Plot comparing unaltered and reweighted-by-data true neutrino energies in GeV using simulated events in the far detector for ν_μ CC interactions. The solid red histogram is unaltered simulation; the solid blue histogram is the simulation spectrum after reweighting by the extrapolation of near detector data. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the nonQE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

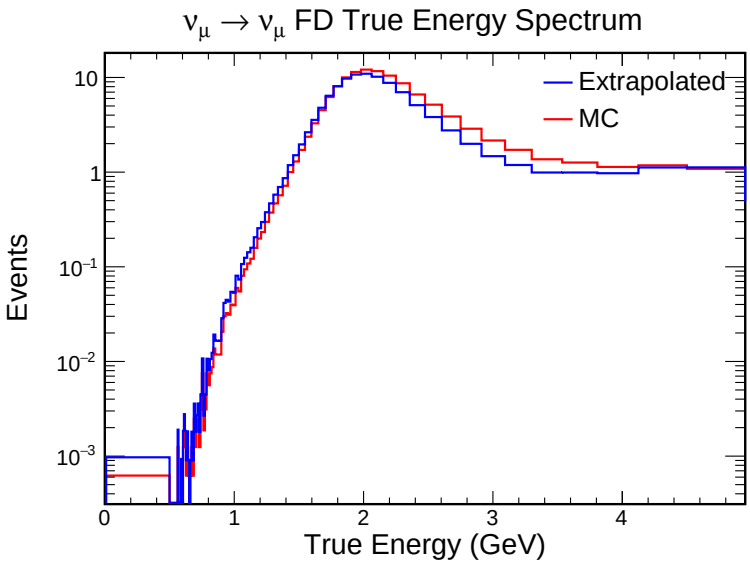


Figure 13.27: Plot comparing unaltered and reweighted-by-data true neutrino energies in GeV using simulated events in the far detector for ν_μ CC interactions. The vertical axis is plotted logarithmically. The solid red histogram is unaltered simulation; the solid blue histogram is the simulation spectrum after reweighting by the extrapolation of near detector data. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the nonQE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

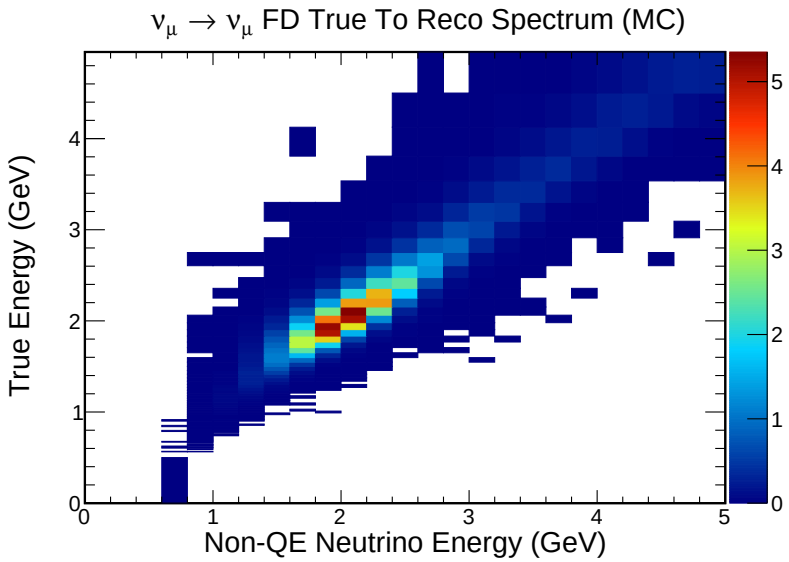


Figure 13.28: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the far detector for ν_μ CC interactions. This plot has not been altered to match the extrapolation of near detector data. The color axis is number of interactions. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the nonQE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

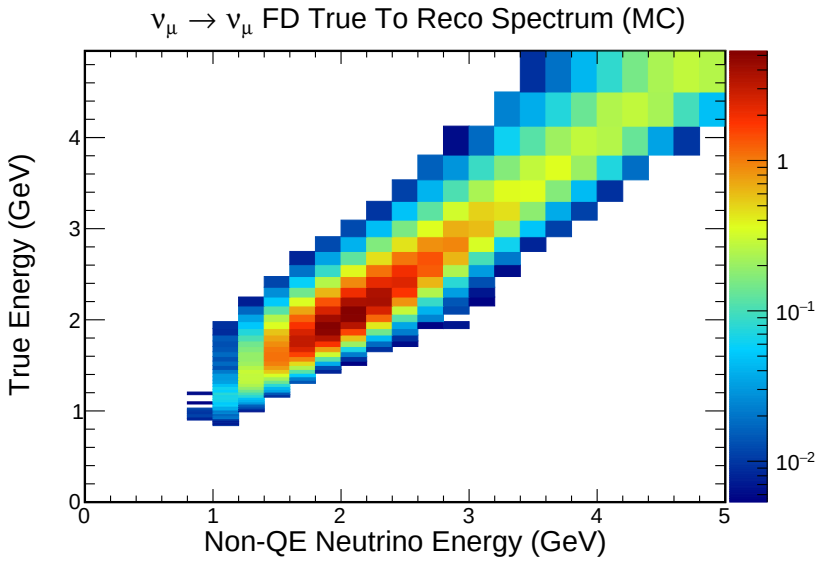


Figure 13.29: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the far detector for ν_μ CC interactions. This plot has not been altered to match the extrapolation of near detector data. The color axis is number of interactions and is plotted logarithmically. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the nonQE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

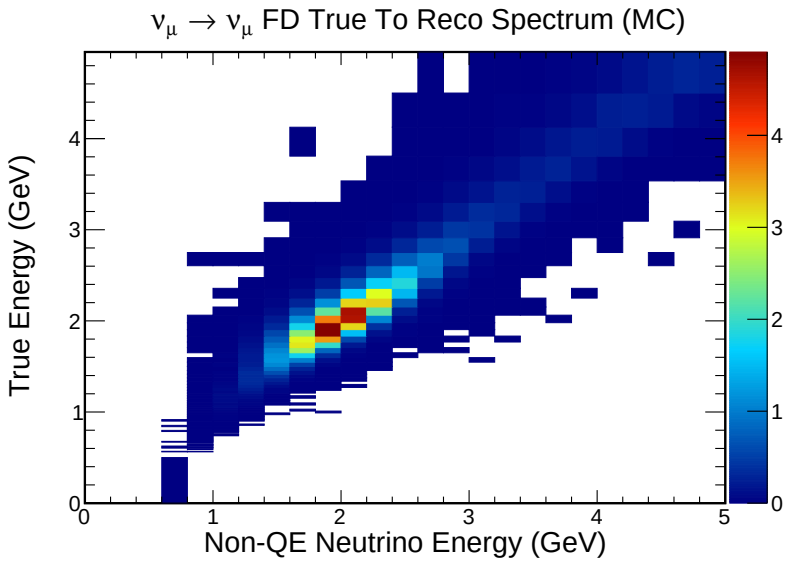


Figure 13.30: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the far detector for ν_μ CC interactions. This plot has been altered to match the extrapolation of near detector data. The color axis is number of interactions. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the nonQE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

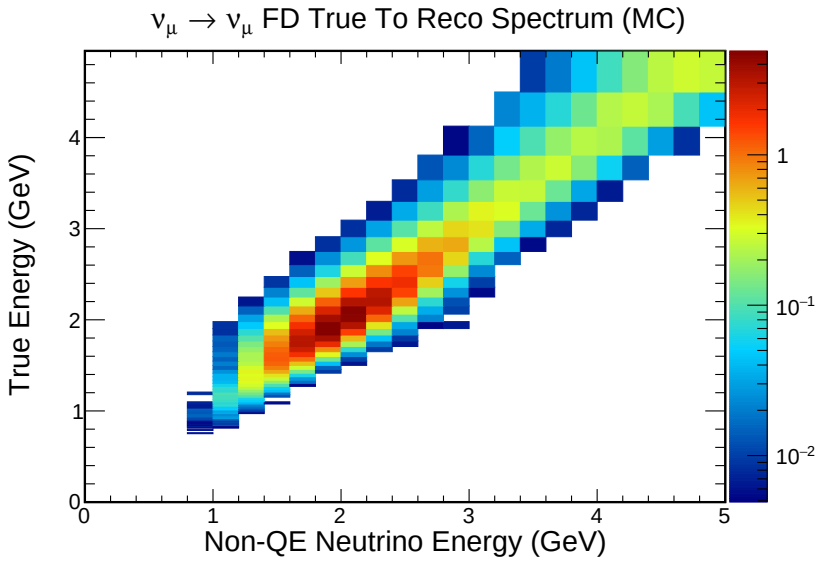


Figure 13.31: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the far detector for ν_μ CC interactions. This plot has been altered to match the extrapolation of near detector data. The color axis is number of interactions and is plotted logarithmically. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the nonQE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

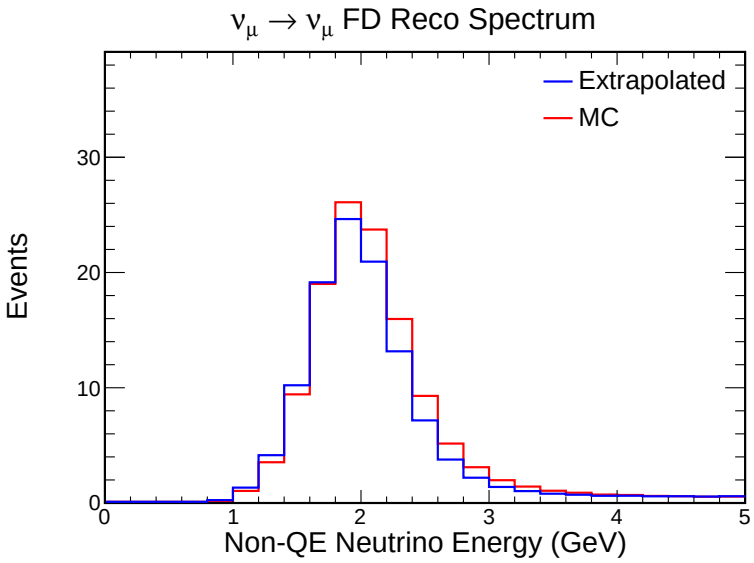


Figure 13.32: Plot comparing unaltered and reweighted-by-data reconstructed neutrino energies in GeV using simulated events in the far detector for ν_μ CC interactions. The solid red histogram is unaltered simulation; the solid blue histogram is the simulation spectrum after reweighting by the extrapolation of near detector data. The extrapolated spectrum is used by the analysis as its unoscillated prediction. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the nonQE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

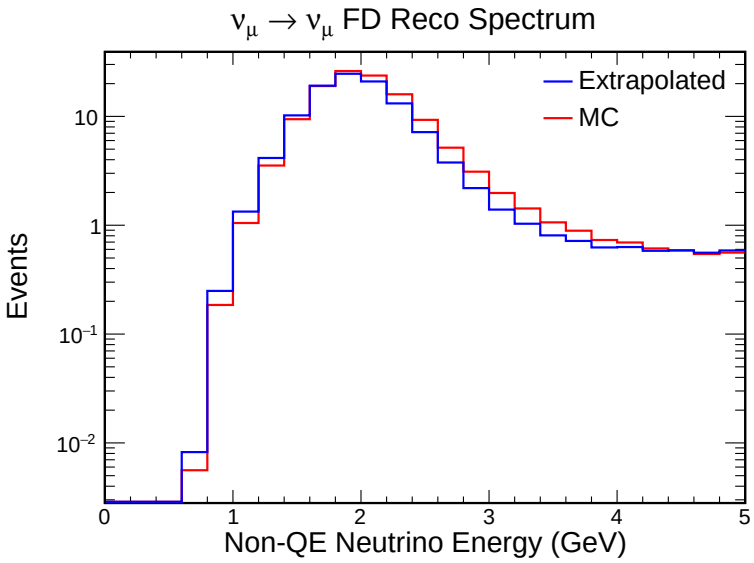


Figure 13.33: Plot comparing unaltered and reweighted-by-data reconstructed neutrino energies in GeV using simulated events in the far detector for ν_μ CC interactions. The vertical axis is plotted logarithmically. The solid red histogram is unaltered simulation; the solid blue histogram is the simulation spectrum after reweighting by the extrapolation of near detector data. The extrapolated spectrum is used by the analysis as its unoscillated prediction. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the nonQE population. The true ν_μ CC population was created using a proportional decomposition, described in Section 13.1.

13.4 $\bar{\nu}_\mu$ CC Interactions

This section contains the plots which illustrate how the far detector simulated prediction for $\bar{\nu}_\mu$ CC interactions is altered due to the near detector data and simulation differences. The $\bar{\nu}_\mu$ CC interactions come from wrong-sign beam contamination; in future analyses, with data from the beam running in the antineutrino mode, we will have many more $\bar{\nu}_\mu$ CC interactions. The true $\bar{\nu}_\mu$ CC population is created from the analysis population by performing a proportional decomposition. This decomposition uses the simulation to determine the proportion of true $\bar{\nu}_\mu$ CC interactions within our analysis sample. See Section 13.1 for more details. This is done independently for the QE and nonQE populations; Section 13.4.1 contains the plots for the QE population and section 13.4.2 contains the plots for the nonQE population.

13.4.1 QE Population

This section presents the plots for the QE population (as defined by the event selection in Chapter 10, specifically Section 10.4). The first step is to compare the reconstructed neutrino energy spectrum for $\bar{\nu}_\mu$ CC interactions in the near detector data and simulation. Figure 13.34 plots this comparison; Figure 13.35 has a logarithmic vertical axis. A 2D histogram of reconstructed neutrino energy vs. true neutrino energy for the near detector is created from the simulation. Figure 13.36 is the unaltered 2D histogram. Figure 13.37 is the same as Figure 13.36, except that it has a logarithmic color axis. The 2D histogram is then altered to make the reconstructed neutrino energy spectrum in the simulation match the near detector data. Figure 13.38 is the adjusted 2D histogram. Figure 13.39 is the same as Figure 13.38, except that it has a logarithmic color axis. The adjusted 2D histogram is projected onto the true neutrino energy axis to create an altered spectrum. Figure 13.40 compares the altered true neutrino energy spectrum and the original true neutrino energy spectrum from simulation. Figure 13.41 is the same as Figure 13.40, except that the vertical axis is displayed logarithmically.

Having an altered true neutrino energy spectrum for the near detector, we now use it to change our far detector prediction. Figure 13.42 compares the original true neutrino energy spectrum for the far detector with one altered to match the extrapolated near detector data. Figure 13.43 is the same, except that it features a logarithmic vertical

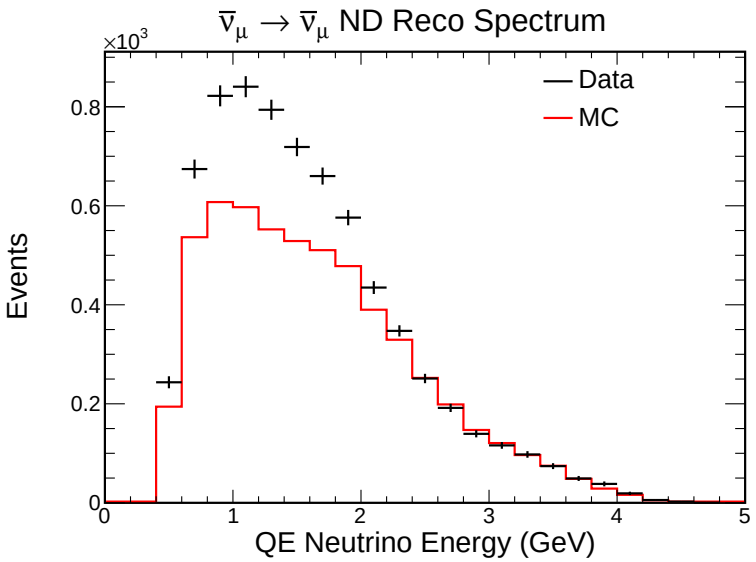


Figure 13.34: Plot comparing data and simulation reconstructed neutrino energies in GeV in the near detector for $\bar{\nu}_\mu$ CC interactions. The solid red histogram is simulation; the black points are data. The error bars displayed on the data are statistical only. The simulation is scaled down to match the POT of the data. This is for the QE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

axis. A 2D histogram of reconstructed neutrino energy vs. true neutrino energy for the far detector is created from the simulation. Figure 13.44 is the unaltered 2D histogram. Figure 13.45 is the same as Figure 13.44, except that it has a logarithmic color axis. The 2D histogram is then altered to make the true neutrino energy spectrum in the simulation match the extrapolated near detector data. Figure 13.46 is the adjusted 2D histogram. Figure 13.47 is the same as Figure 13.46, except that it has a logarithmic color axis. The altered 2D histogram is then projected onto the reconstructed neutrino energy axis. This is the unoscillated prediction used for the analysis. It is compared to the unaltered simulation prediction in Figure 13.48. Figure 13.49 is the same comparison, but featuring a logarithmic vertical axis.

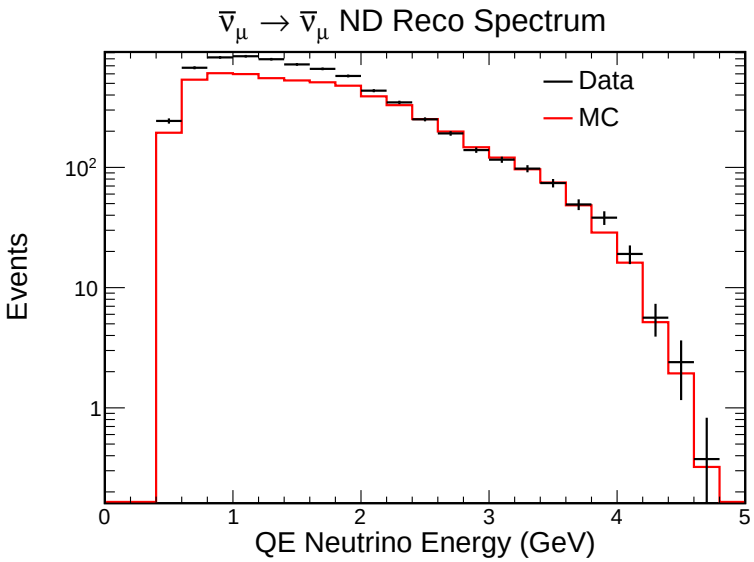


Figure 13.35: Plot comparing data and simulation reconstructed neutrino energies in GeV in the near detector for $\bar{\nu}_\mu$ CC interactions. The vertical axis is plotted logarithmically. The solid red histogram is simulation; the black points are data. The error bars displayed on the data are statistical only. The simulation is scaled down to match the POT of the data. This is for the QE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

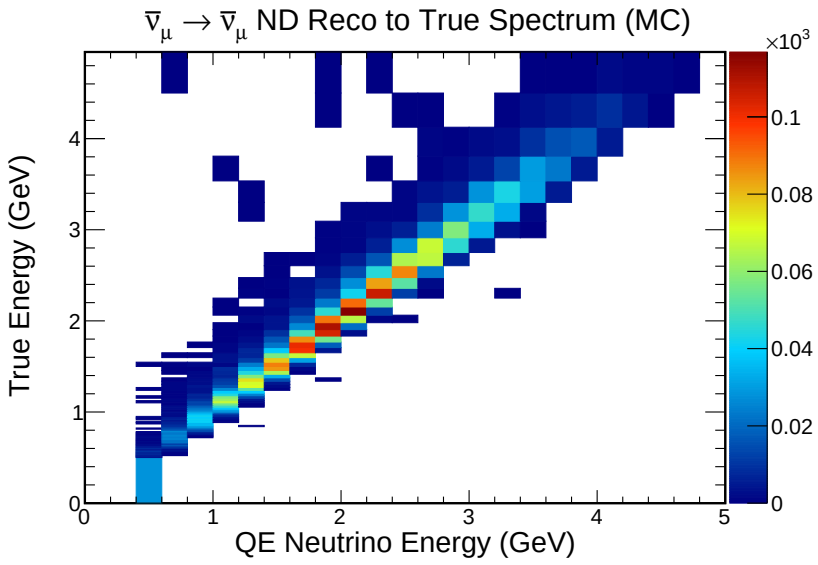


Figure 13.36: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the near detector for $\bar{\nu}_\mu$ CC interactions. This plot has not been altered to match near detector data. The color axis is number of interactions. The simulation is scaled down to match the POT of the data. This is for the QE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

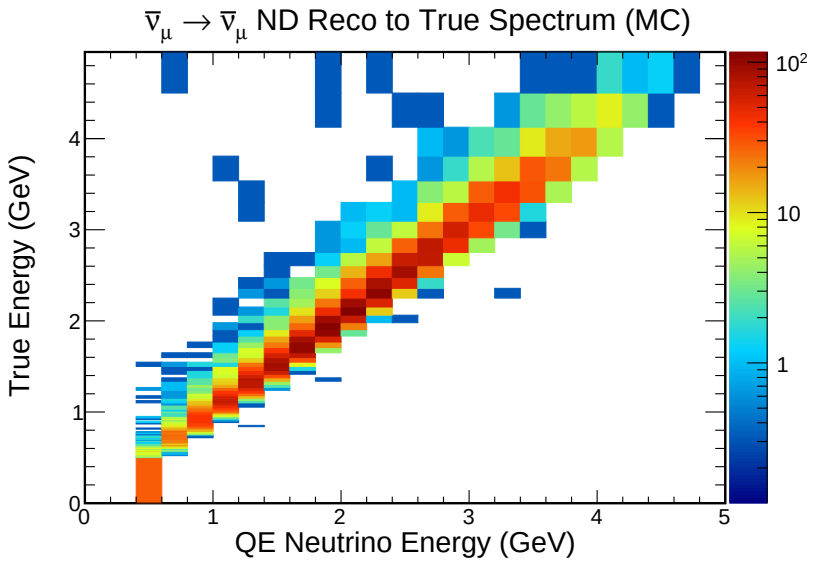


Figure 13.37: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the near detector for $\bar{\nu}_\mu$ CC interactions. This plot has not been altered to match near detector data. The color axis is number of interactions and is plotted logarithmically. The simulation is scaled down to match the POT of the data. This is for the QE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

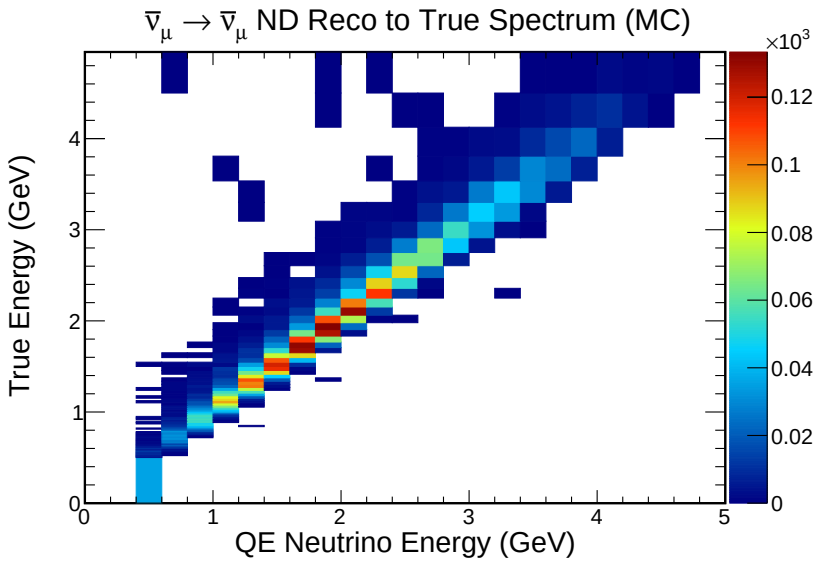


Figure 13.38: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the near detector for $\bar{\nu}_\mu$ CC interactions. This plot has been altered to match near detector data. The color axis is number of interactions. The simulation is scaled down to match the POT of the data. This is for the QE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

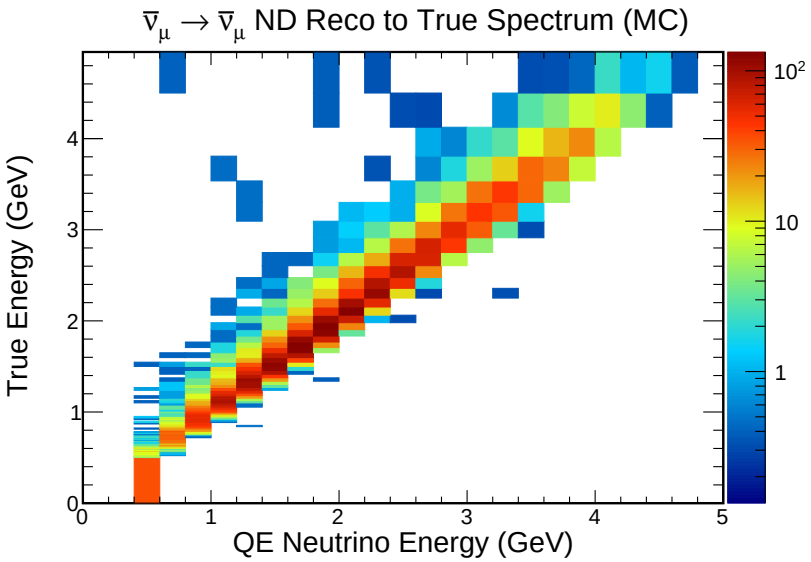


Figure 13.39: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the near detector for $\bar{\nu}_\mu$ CC interactions. This plot has been altered to match near detector data. The color axis is number of interactions and is plotted logarithmically. The simulation is scaled down to match the POT of the data. This is for the QE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

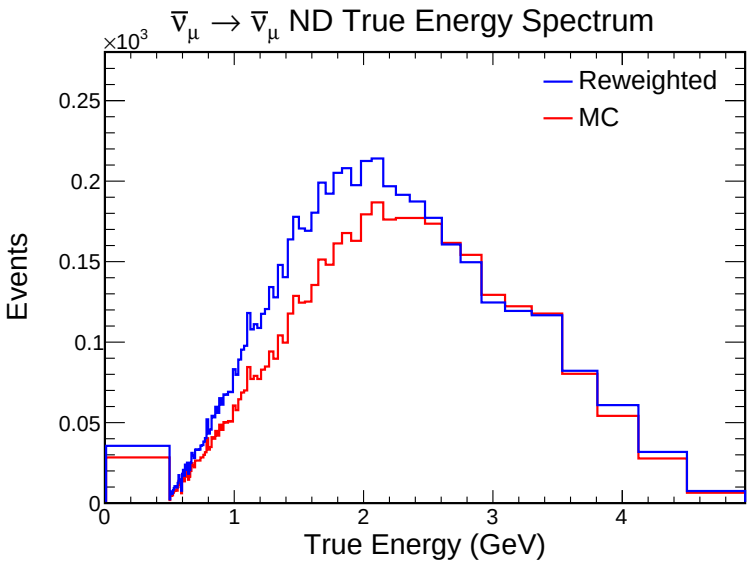


Figure 13.40: Plot comparing unaltered and reweighted-by-data true neutrino energies in GeV using simulated events in the near detector for $\bar{\nu}_\mu$ CC interactions. The solid red histogram is unaltered simulation; the solid blue histogram is the simulation spectrum after reweighting by data. The simulation is scaled down to match the POT of the data. This is for the QE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

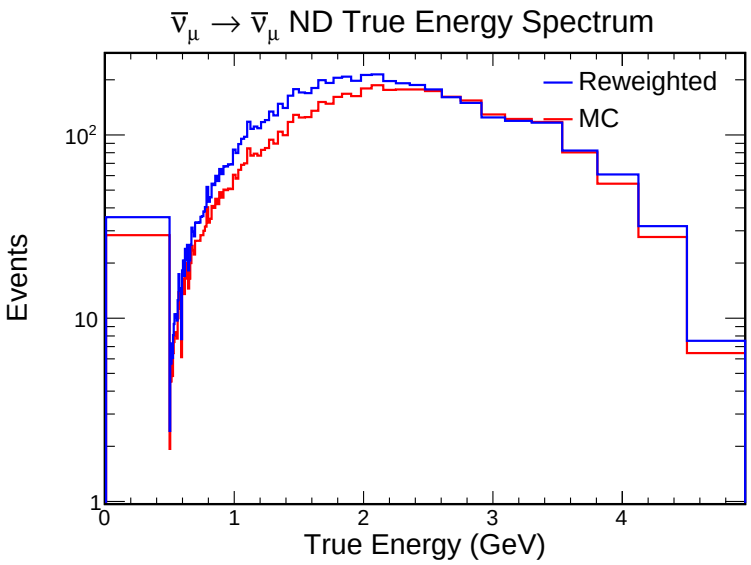


Figure 13.41: Plot comparing unaltered and reweighted-by-data true neutrino energies in GeV using simulated events in the near detector for $\bar{\nu}_\mu$ CC interactions. The vertical axis is plotted logarithmically. The solid red histogram is unaltered simulation; the solid blue histogram is the simulation spectrum after reweighting by data. The simulation is scaled down to match the POT of the data. This is for the QE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

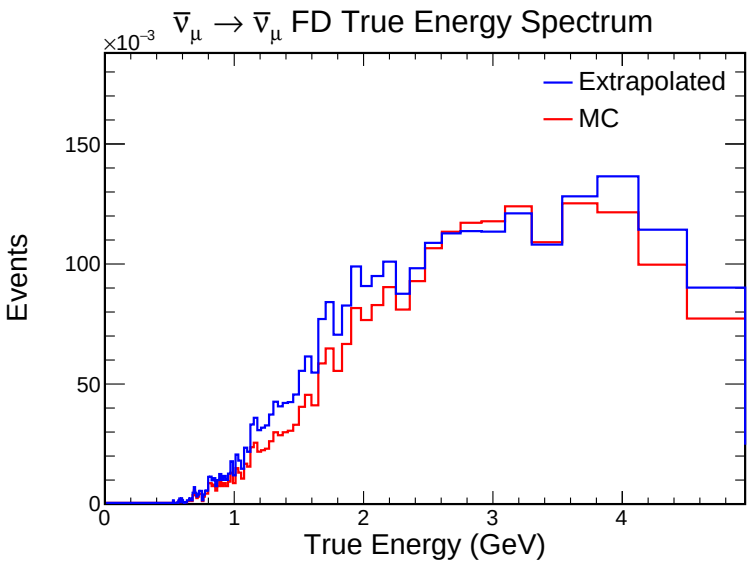


Figure 13.42: Plot comparing unaltered and reweighted-by-data true neutrino energies in GeV using simulated events in the far detector for $\bar{\nu}_\mu$ CC interactions. The solid red histogram is unaltered simulation; the solid blue histogram is the simulation spectrum after reweighting by the extrapolation of near detector data. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the QE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

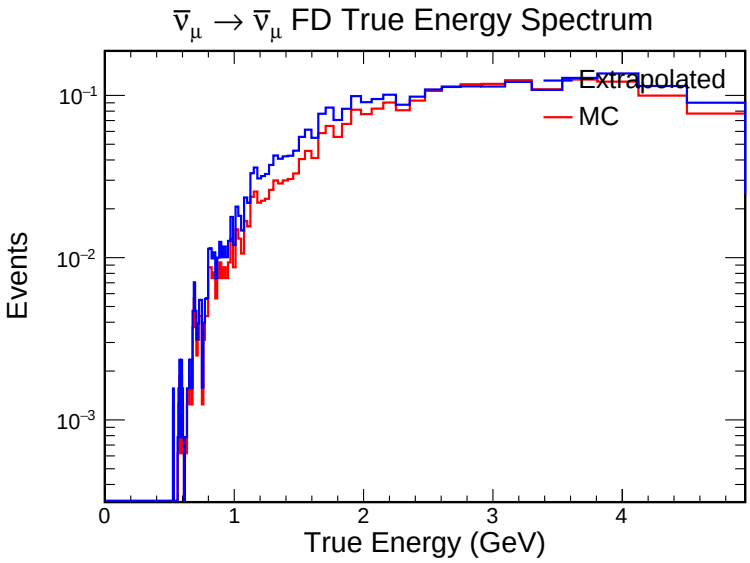


Figure 13.43: Plot comparing unaltered and reweighted-by-data true neutrino energies in GeV using simulated events in the far detector for $\bar{\nu}_\mu$ CC interactions. The vertical axis is plotted logarithmically. The solid red histogram is unaltered simulation; the solid blue histogram is the simulation spectrum after reweighting by the extrapolation of near detector data. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the QE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

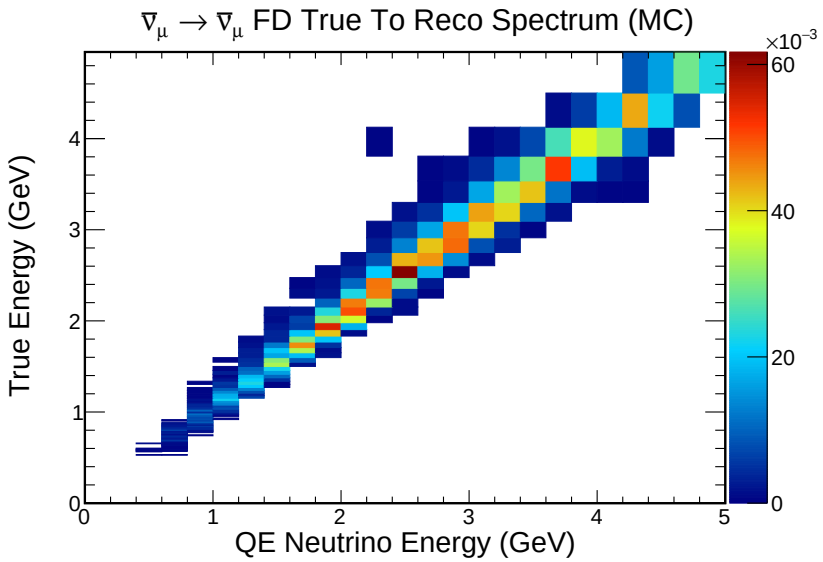


Figure 13.44: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the far detector for $\bar{\nu}_\mu$ CC interactions. This plot has not been altered to match the extrapolation of near detector data. The color axis is number of interactions. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the QE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

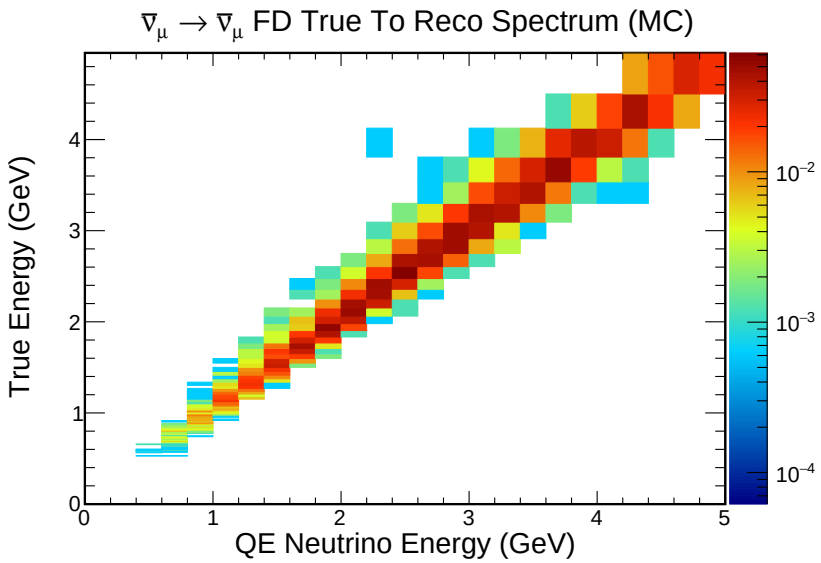


Figure 13.45: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the far detector for $\bar{\nu}_\mu$ CC interactions. This plot has not been altered to match the extrapolation of near detector data. The color axis is number of interactions and is plotted logarithmically. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the QE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

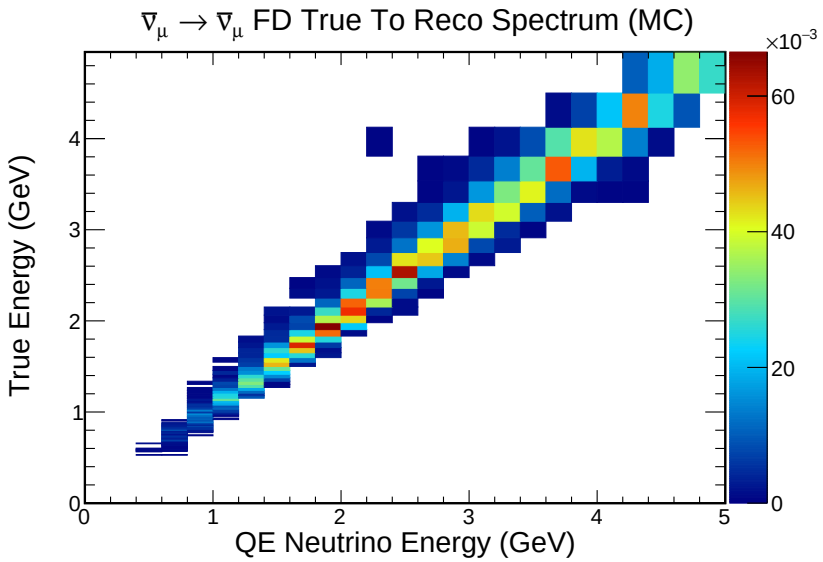


Figure 13.46: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the far detector for $\bar{\nu}_\mu$ CC interactions. This plot has been altered to match the extrapolation of near detector data. The color axis is number of interactions. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the QE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

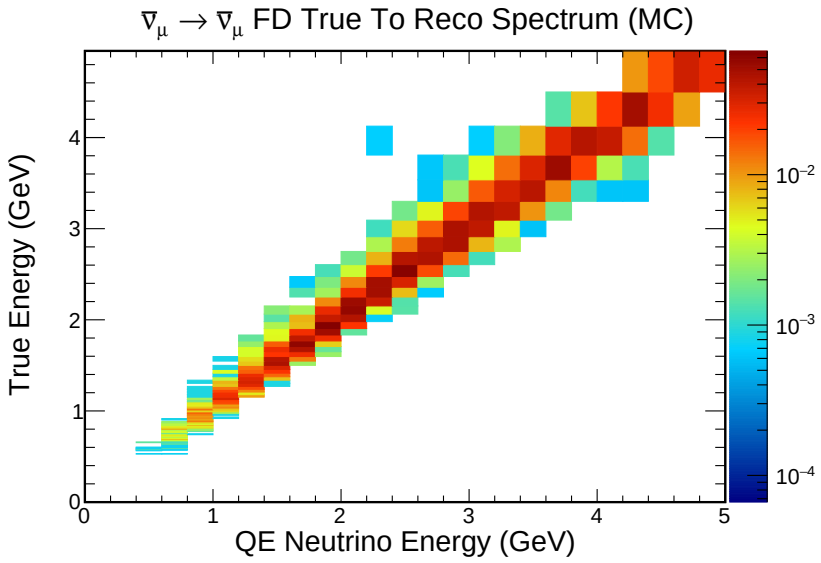


Figure 13.47: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the far detector for $\bar{\nu}_\mu$ CC interactions. This plot has been altered to match the extrapolation of near detector data. The color axis is number of interactions and is plotted logarithmically. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the QE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

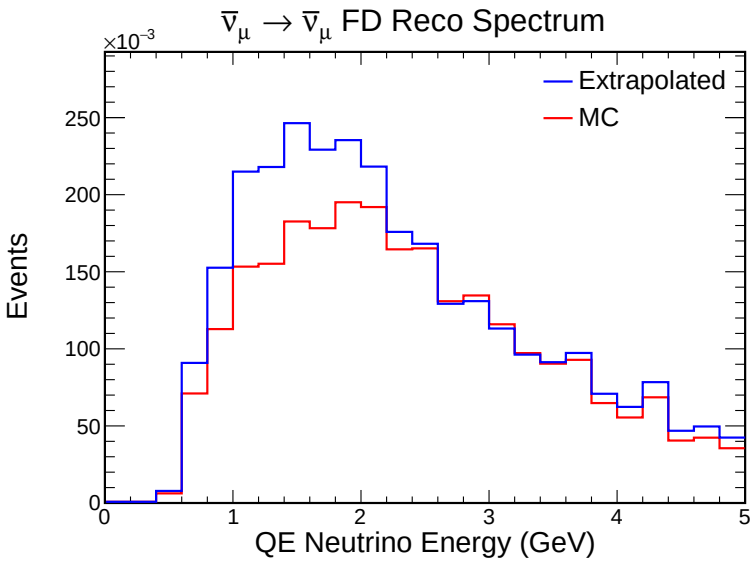


Figure 13.48: Plot comparing unaltered and reweighted-by-data reconstructed neutrino energies in GeV using simulated events in the far detector for $\bar{\nu}_\mu$ CC interactions. The solid red histogram is unaltered simulation; the solid blue histogram is the simulation spectrum after reweighting by the extrapolation of near detector data. The extrapolated spectrum is used by the analysis as its unoscillated prediction. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the QE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

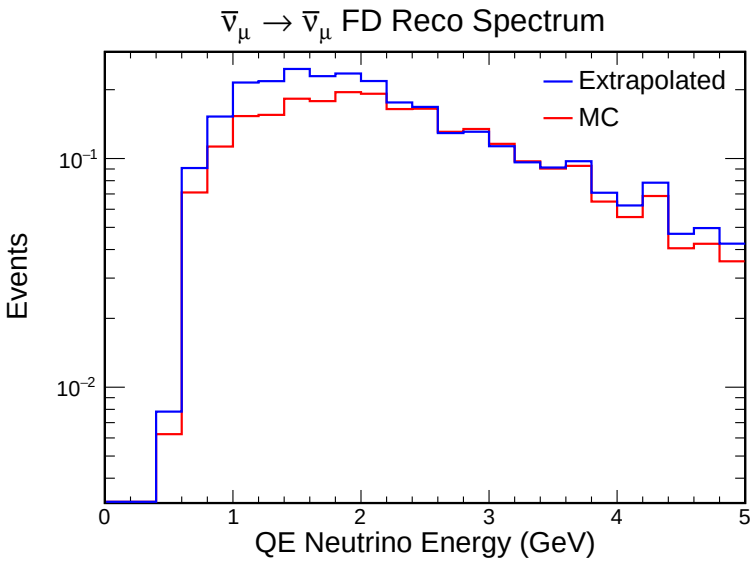


Figure 13.49: Plot comparing unaltered and reweighted-by-data reconstructed neutrino energies in GeV using simulated events in the far detector for $\bar{\nu}_\mu$ CC interactions. The vertical axis is plotted logarithmically. The solid red histogram is unaltered simulation; the solid blue histogram is the simulation spectrum after reweighting by the extrapolation of near detector data. The extrapolated spectrum is used by the analysis as its unoscillated prediction. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the QE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

13.4.2 NonQE Population

This section presents the plots for the nonQE population (as defined by the event selection in Chapter 10, specifically Section 10.4). The first step is to compare the reconstructed neutrino energy spectrum for $\bar{\nu}_\mu$ CC interactions in the near detector data and simulation. Figure 13.50 plots this comparison; Figure 13.51 has a logarithmic vertical axis. A 2D histogram of reconstructed neutrino energy vs. true neutrino energy for the near detector is created from the simulation. Figure 13.52 is the unaltered 2D histogram. Figure 13.53 is the same as Figure 13.52, except that it has a logarithmic color axis. The 2D histogram is then altered to make the reconstructed neutrino energy spectrum in the simulation match the near detector data. Figure 13.54 is the adjusted 2D histogram. Figure 13.55 is the same as Figure 13.54, except that it has a logarithmic color axis. The adjusted 2D histogram is projected onto the true neutrino energy axis to create an altered spectrum. Figure 13.56 compares the altered true neutrino energy spectrum and the original true neutrino energy spectrum from simulation. Figure 13.57 is the same as Figure 13.56, except that the vertical axis is displayed logarithmically.

Having an altered true neutrino energy spectrum for the near detector, we now use it to change our far detector prediction. Figure 13.58 compares the original true neutrino energy spectrum for the far detector with one altered to match the extrapolated near detector data. Figure 13.59 is the same, except that it features a logarithmic vertical axis. A 2D histogram of reconstructed neutrino energy vs. true neutrino energy for the far detector is created from the simulation. Figure 13.60 is the unaltered 2D histogram. Figure 13.61 is the same as Figure 13.60, except that it has a logarithmic color axis. The 2D histogram is then altered to make the true neutrino energy spectrum in the simulation match the extrapolated near detector data. Figure 13.62 is the adjusted 2D histogram. Figure 13.63 is the same as Figure 13.62, except that it has a logarithmic color axis. The altered 2D histogram is then projected onto the reconstructed neutrino energy axis. This is the unoscillated prediction used for the analysis. It is compared to the unaltered simulation prediction in Figure 13.64. Figure 13.65 is the same comparison, but featuring a logarithmic vertical axis.

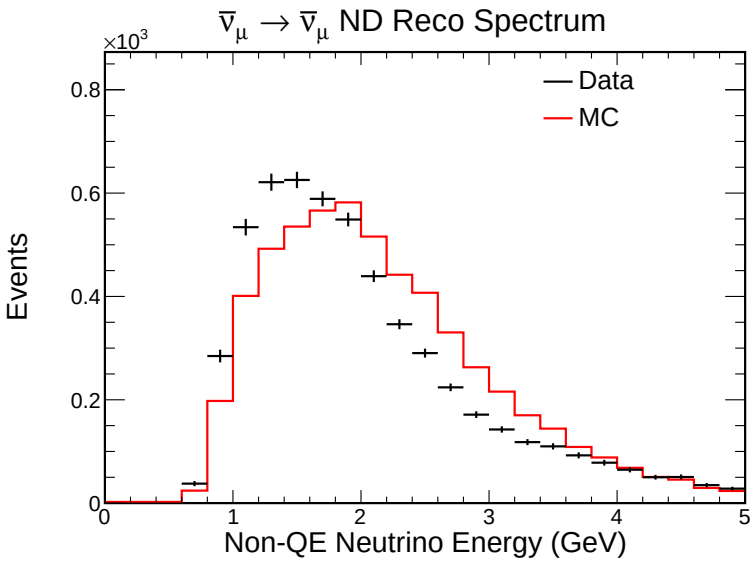


Figure 13.50: Plot comparing data and simulation reconstructed neutrino energies in GeV in the near detector for $\bar{\nu}_\mu$ CC interactions. The solid red histogram is simulation; the black points are data. The error bars displayed on the data are statistical only. The simulation is scaled down to match the POT of the data. This is for the nonQE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

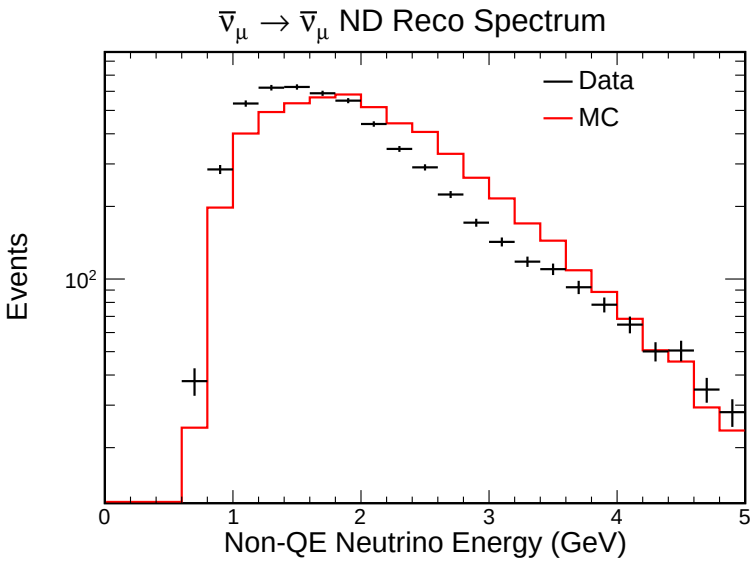


Figure 13.51: Plot comparing data and simulation reconstructed neutrino energies in GeV in the near detector for $\bar{\nu}_\mu$ CC interactions. The vertical axis is plotted logarithmically. The solid red histogram is simulation; the black points are data. The error bars displayed on the data are statistical only. The simulation is scaled down to match the POT of the data. This is for the nonQE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

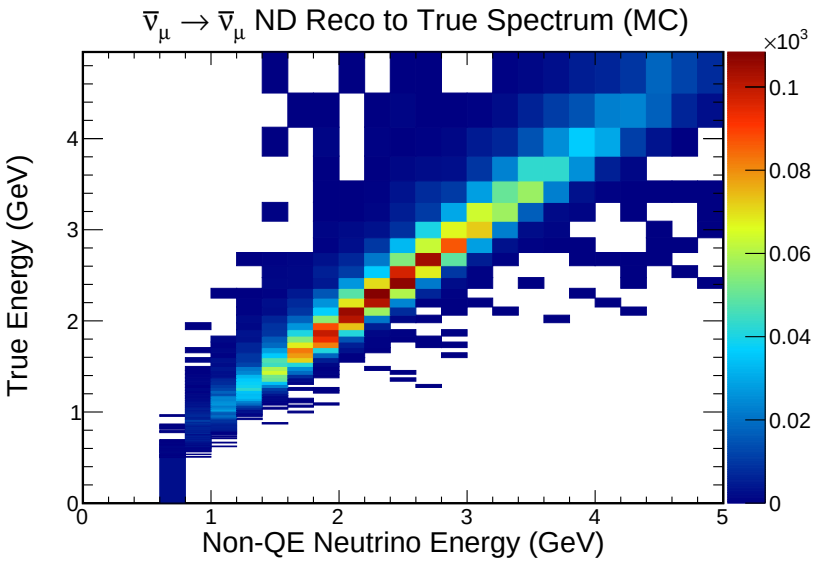


Figure 13.52: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the near detector for $\bar{\nu}_\mu$ CC interactions. This plot has not been altered to match near detector data. The color axis is number of interactions. The simulation is scaled down to match the POT of the data. This is for the nonQE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

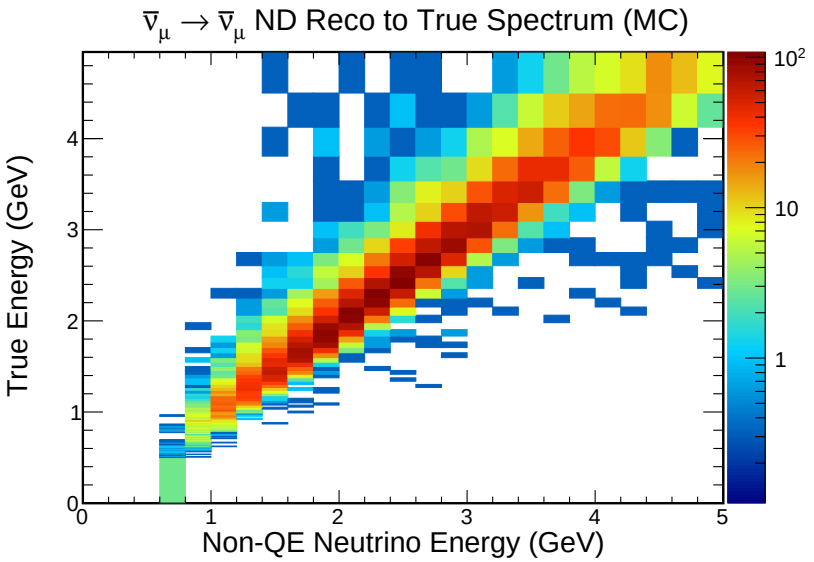


Figure 13.53: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the near detector for $\bar{\nu}_\mu$ CC interactions. This plot has not been altered to match near detector data. The color axis is number of interactions and is plotted logarithmically. The simulation is scaled down to match the POT of the data. This is for the nonQE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

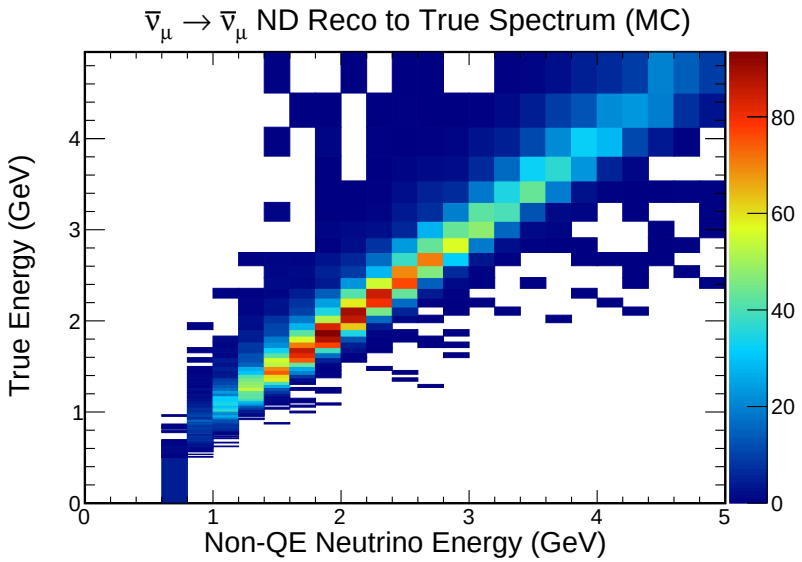


Figure 13.54: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the near detector for $\bar{\nu}_\mu$ CC interactions. This plot has been altered to match near detector data. The color axis is number of interactions. The simulation is scaled down to match the POT of the data. This is for the nonQE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

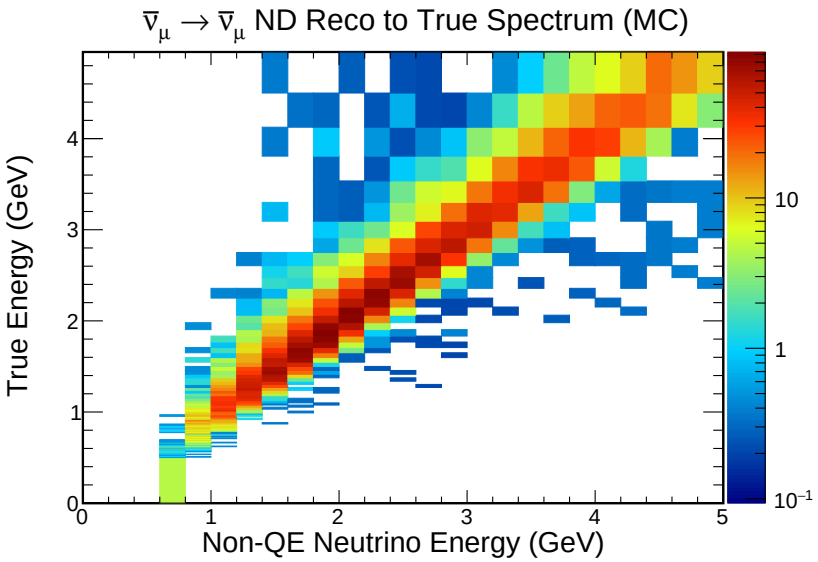


Figure 13.55: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the near detector for $\bar{\nu}_\mu$ CC interactions. This plot has been altered to match near detector data. The color axis is number of interactions and is plotted logarithmically. The simulation is scaled down to match the POT of the data. This is for the nonQE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

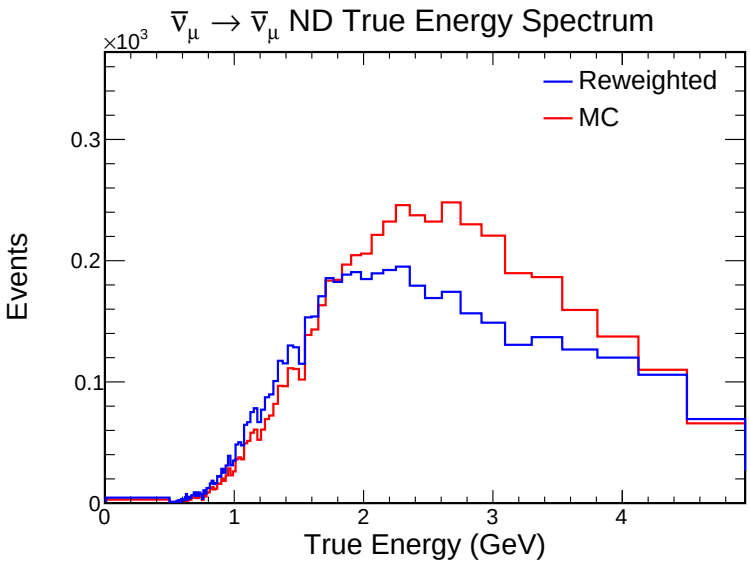


Figure 13.56: Plot comparing unaltered and reweighted-by-data true neutrino energies in GeV using simulated events in the near detector for $\bar{\nu}_\mu$ CC interactions. The solid red histogram is unaltered simulation; the solid blue histogram is the simulation spectrum after reweighting by data. The simulation is scaled down to match the POT of the data. This is for the nonQE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

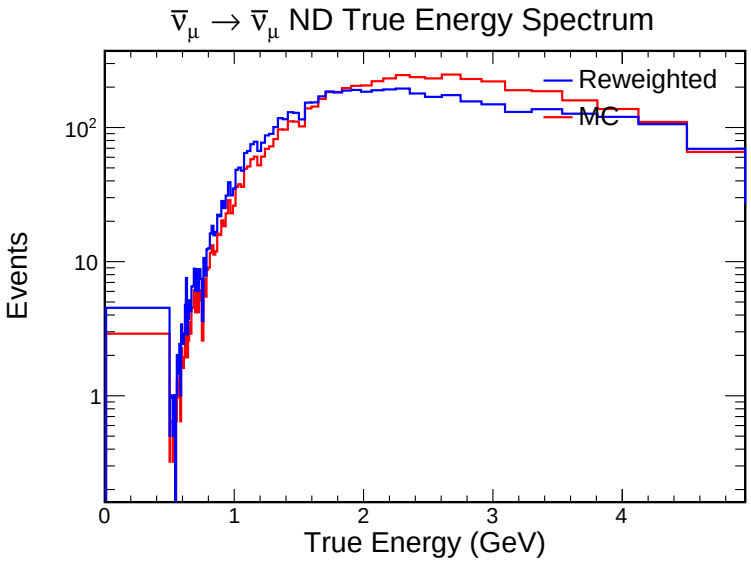


Figure 13.57: Plot comparing unaltered and reweighted-by-data true neutrino energies in GeV using simulated events in the near detector for $\bar{\nu}_\mu$ CC interactions. The vertical axis is plotted logarithmically. The solid red histogram is unaltered simulation; the solid blue histogram is the simulation spectrum after reweighting by data. The simulation is scaled down to match the POT of the data. This is for the nonQE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

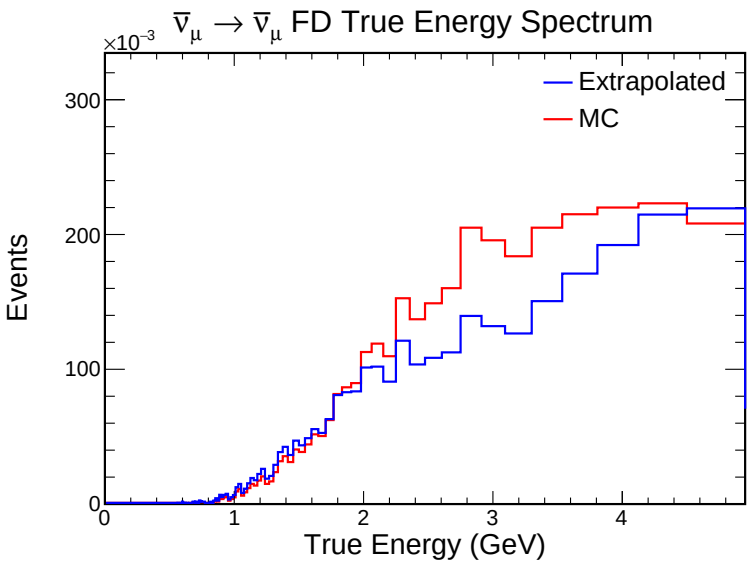


Figure 13.58: Plot comparing unaltered and reweighted-by-data true neutrino energies in GeV using simulated events in the far detector for $\bar{\nu}_\mu$ CC interactions. The solid red histogram is unaltered simulation; the solid blue histogram is the simulation spectrum after reweighting by the extrapolation of near detector data. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the nonQE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

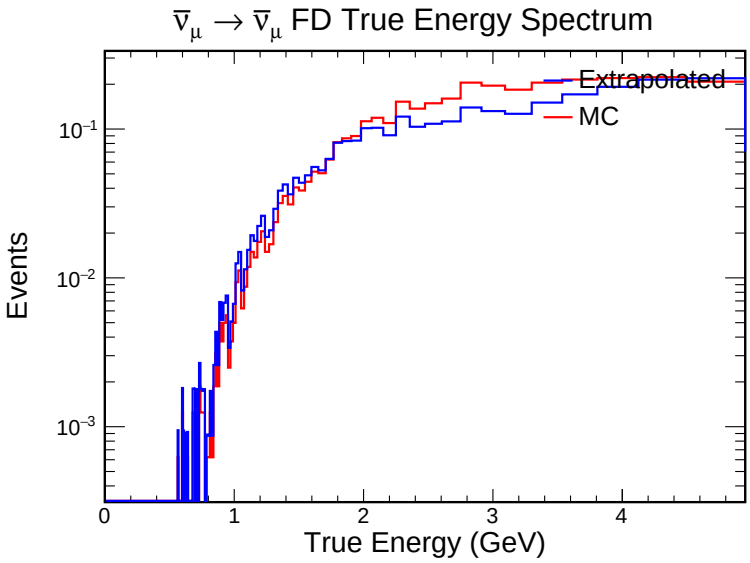


Figure 13.59: Plot comparing unaltered and reweighted-by-data true neutrino energies in GeV using simulated events in the far detector for $\bar{\nu}_\mu$ CC interactions. The vertical axis is plotted logarithmically. The solid red histogram is unaltered simulation; the solid blue histogram is the simulation spectrum after reweighting by the extrapolation of near detector data. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the nonQE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

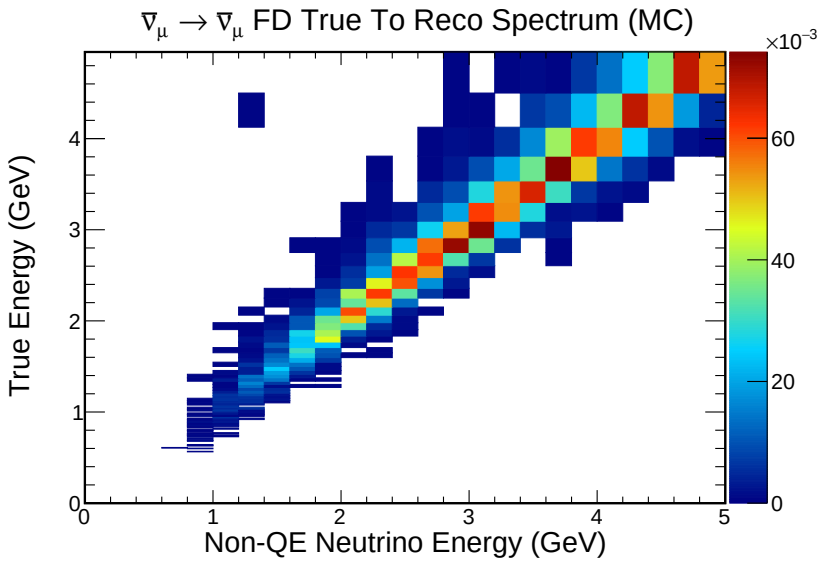


Figure 13.60: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the far detector for $\bar{\nu}_\mu$ CC interactions. This plot has not been altered to match the extrapolation of near detector data. The color axis is number of interactions. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the nonQE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

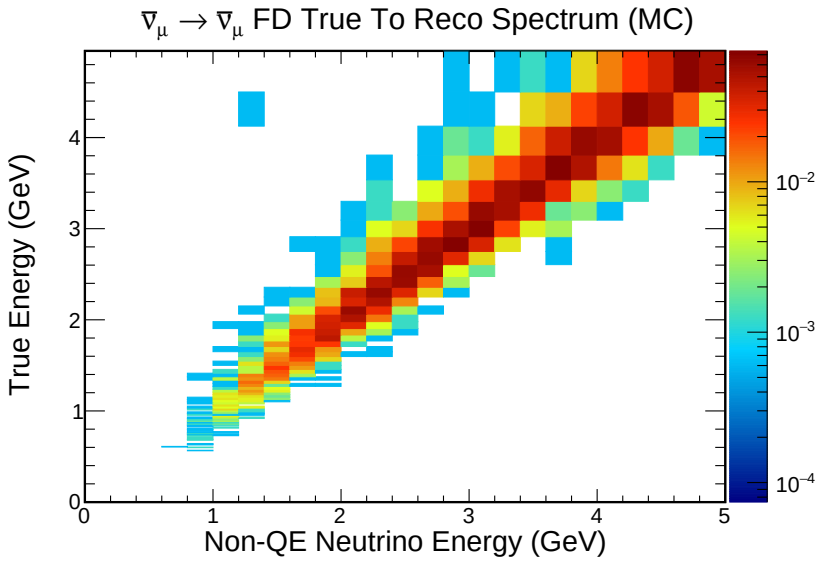


Figure 13.61: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the far detector for $\bar{\nu}_\mu$ CC interactions. This plot has not been altered to match the extrapolation of near detector data. The color axis is number of interactions and is plotted logarithmically. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the nonQE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

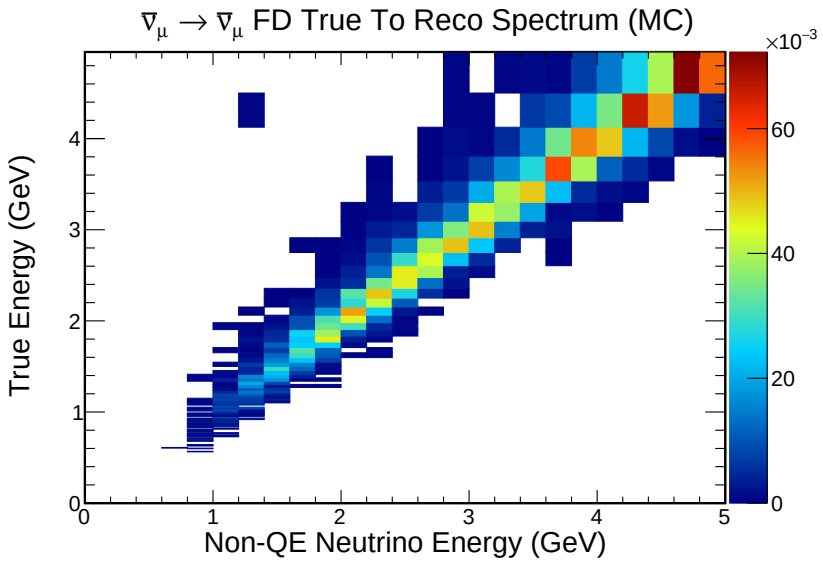


Figure 13.62: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the far detector for $\bar{\nu}_\mu$ CC interactions. This plot has been altered to match the extrapolation of near detector data. The color axis is number of interactions. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the nonQE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

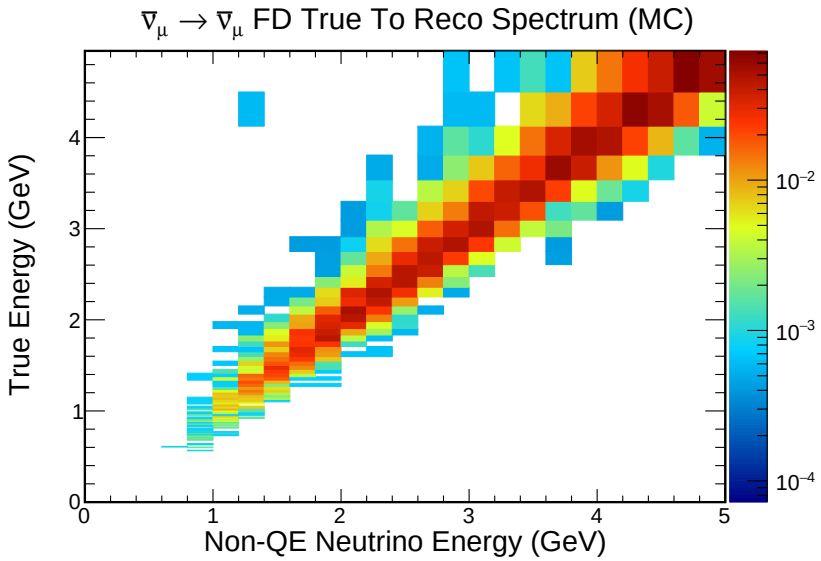


Figure 13.63: Reconstructed neutrino energy in GeV vs. true neutrino energy in GeV using simulated events in the far detector for $\bar{\nu}_\mu$ CC interactions. This plot has been altered to match the extrapolation of near detector data. The color axis is number of interactions and is plotted logarithmically. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the nonQE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

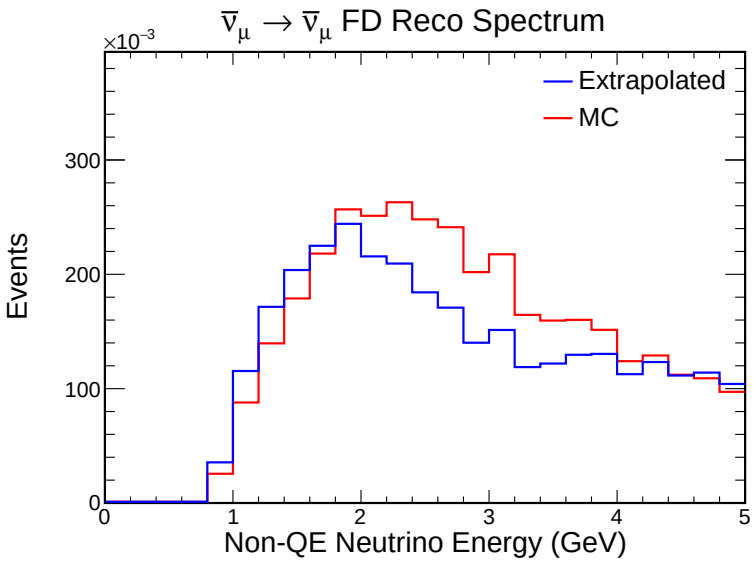


Figure 13.64: Plot comparing unaltered and reweighted-by-data reconstructed neutrino energies in GeV using simulated events in the far detector for $\bar{\nu}_\mu$ CC interactions. The solid red histogram is unaltered simulation; the solid blue histogram is the simulation spectrum after reweighting by the extrapolation of near detector data. The extrapolated spectrum is used by the analysis as its unoscillated prediction. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the nonQE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

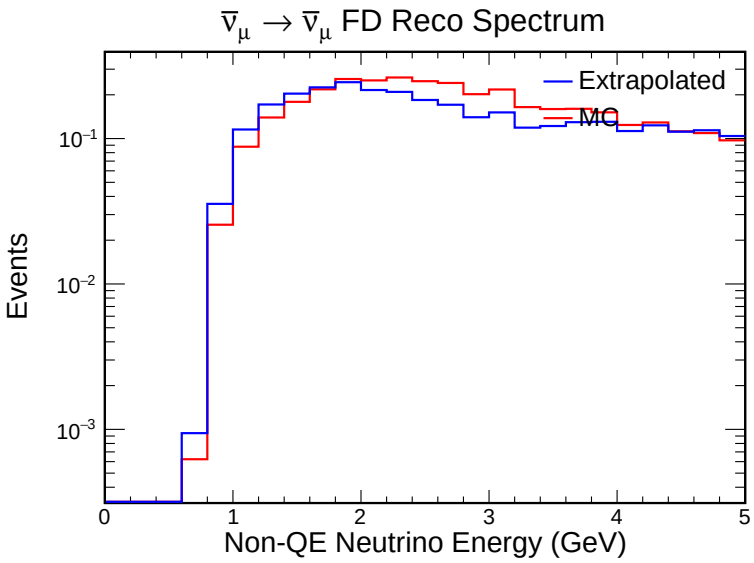


Figure 13.65: Plot comparing unaltered and reweighted-by-data reconstructed neutrino energies in GeV using simulated events in the far detector for $\bar{\nu}_\mu$ CC interactions. The vertical axis is plotted logarithmically. The solid red histogram is unaltered simulation; the solid blue histogram is the simulation spectrum after reweighting by the extrapolation of near detector data. The extrapolated spectrum is used by the analysis as its unoscillated prediction. The simulation is scaled down to match the POT of the far detector data and is unoscillated. This is for the nonQE population. The true $\bar{\nu}_\mu$ CC population was created using a proportional decomposition, described in Section 13.1.

13.5 NC Interactions

This section contains the plots which illustrate how the far detector simulated prediction for NC interactions is altered due to the near detector data and simulation differences. NC interactions are the primary component of the background for this analysis from neutrino interactions that are not ν_μ CC interactions. The true NC population is created from the analysis population by performing a proportional decomposition. This decomposition uses the simulation to determine the proportion of true NC interactions within our analysis sample. See Section 13.1 for more details. This is done independently for the QE and nonQE populations; Section 13.5.1 contains the plots for the QE population and section 13.5.2 contains the plots for the nonQE population.

13.5.1 QE Population

This section presents the plots for the QE population (as defined by the event selection in Chapter 10, specifically Section 10.4). The first step is to compare the reconstructed neutrino energy spectrum for NC interactions in the near detector data and simulation. Figure 13.66 plots this comparison; Figure 13.67 has a logarithmic vertical axis.

Having an altered reconstructed neutrino energy spectrum for the near detector, we now use it to change our far detector prediction. This is the prediction used for the analysis. It is compared to the unaltered simulation prediction in Figure 13.68. Figure 13.69 is the same comparison, but featuring a logarithmic vertical axis.

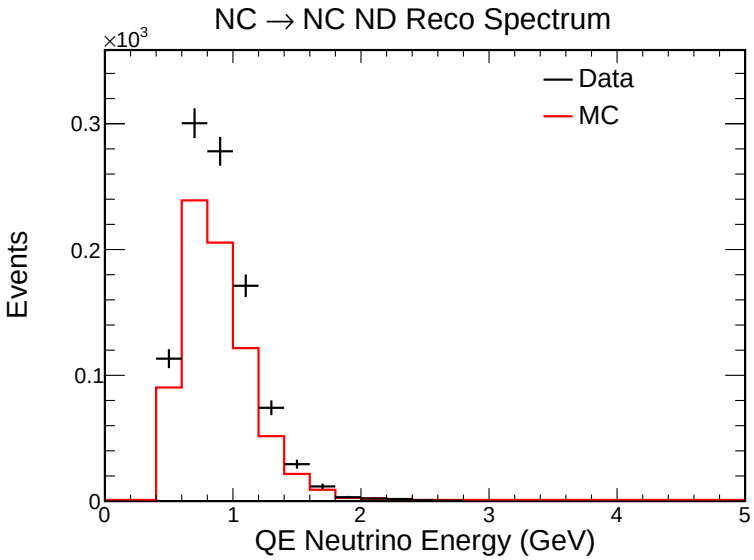


Figure 13.66: Plot comparing data and simulation reconstructed neutrino energies in GeV in the near detector for NC interactions. The solid red histogram is simulation; the black points are data. The error bars displayed on the data are statistical only. The simulation is scaled down to match the POT of the data. This is for the QE population. The true NC population was created using a proportional decomposition, described in Section 13.1.

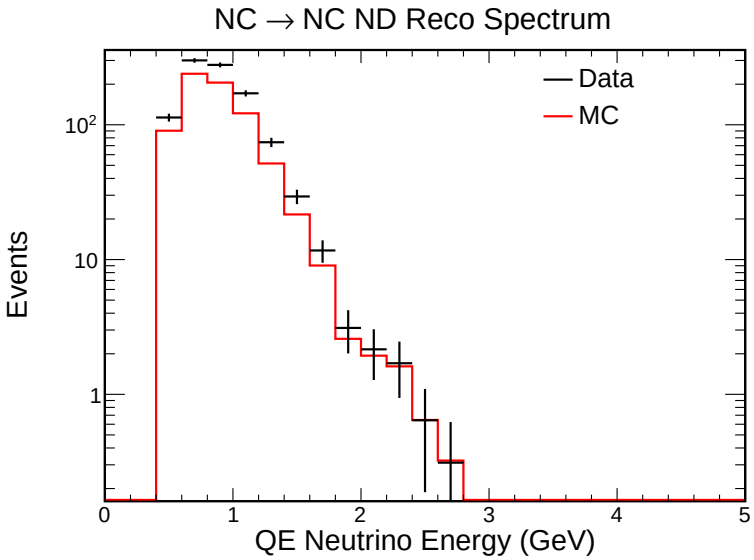


Figure 13.67: Plot comparing data and simulation reconstructed neutrino energies in GeV in the near detector for NC interactions. The vertical axis is plotted logarithmically. The solid red histogram is simulation; the black points are data. The error bars displayed on the data are statistical only. The simulation is scaled down to match the POT of the data. This is for the QE population. The true NC population was created using a proportional decomposition, described in Section 13.1.

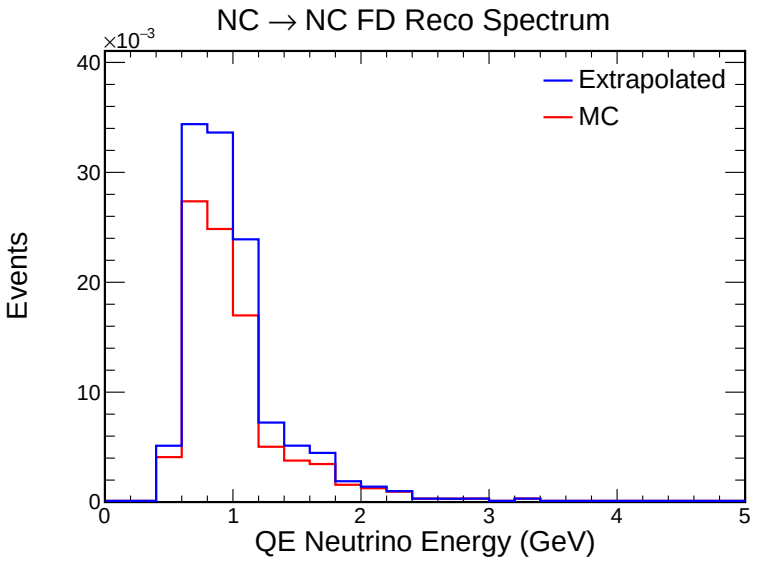


Figure 13.68: Plot comparing unaltered and reweighted-by-data reconstructed neutrino energies in GeV using simulated events in the far detector for NC interactions. The solid red histogram is unaltered simulation; the solid blue histogram is the simulation spectrum after reweighting by the extrapolation of near detector data. The extrapolated spectrum is used by the analysis as its prediction. The simulation is scaled down to match the POT of the far detector data. This is for the QE population. The true NC population was created using a proportional decomposition, described in Section 13.1.

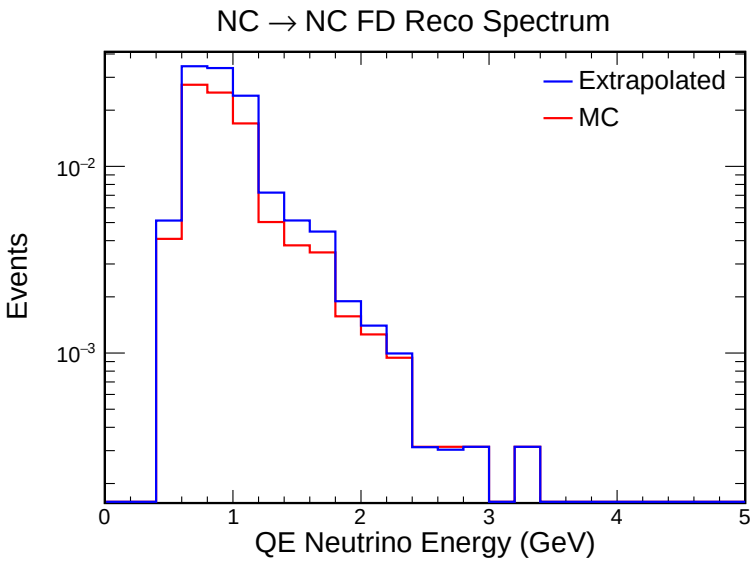


Figure 13.69: Plot comparing unaltered and reweighted-by-data reconstructed neutrino energies in GeV using simulated events in the far detector for NC interactions. The vertical axis is plotted logarithmically. The solid red histogram is unaltered simulation; the solid blue histogram is the simulation spectrum after reweighting by the extrapolation of near detector data. The extrapolated spectrum is used by the analysis as its prediction. The simulation is scaled down to match the POT of the far detector data. This is for the QE population. The true NC population was created using a proportional decomposition, described in Section 13.1.

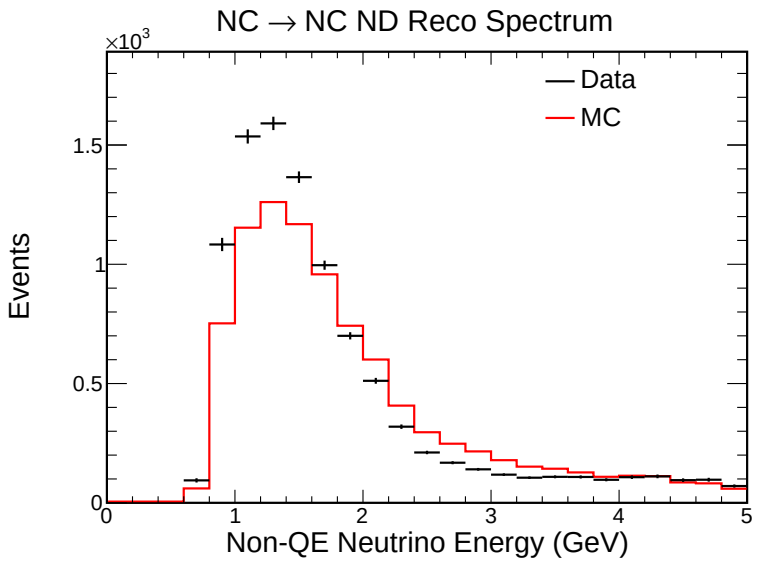


Figure 13.70: Plot comparing data and simulation reconstructed neutrino energies in GeV in the near detector for NC interactions. The solid red histogram is simulation; the black points are data. The error bars displayed on the data are statistical only. The simulation is scaled down to match the POT of the data. This is for the nonQE population. The true NC population was created using a proportional decomposition, described in Section 13.1.

13.5.2 NonQE Population

This section presents the plots for the nonQE population (as defined by the event selection in Chapter 10, specifically Section 10.4). The first step is to compare the reconstructed neutrino energy spectrum for NC interactions in the near detector data and simulation. Figure 13.70 plots this comparison; Figure 13.71 has a logarithmic vertical axis.

Having an altered true neutrino energy spectrum for the near detector, we now use it to change our far detector prediction. This is the prediction used for the analysis. It is compared to the unaltered simulation prediction in Figure 13.72. Figure 13.73 is the same comparison, but featuring a logarithmic vertical axis.

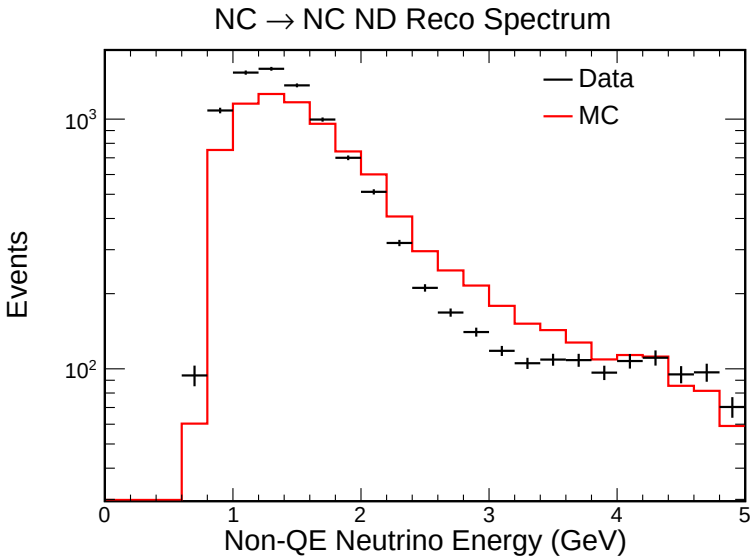


Figure 13.71: Plot comparing data and simulation reconstructed neutrino energies in GeV in the near detector for NC interactions. The vertical axis is plotted logarithmically. The solid red histogram is simulation; the black points are data. The error bars displayed on the data are statistical only. The simulation is scaled down to match the POT of the data. This is for the nonQE population. The true NC population was created using a proportional decomposition, described in Section 13.1.

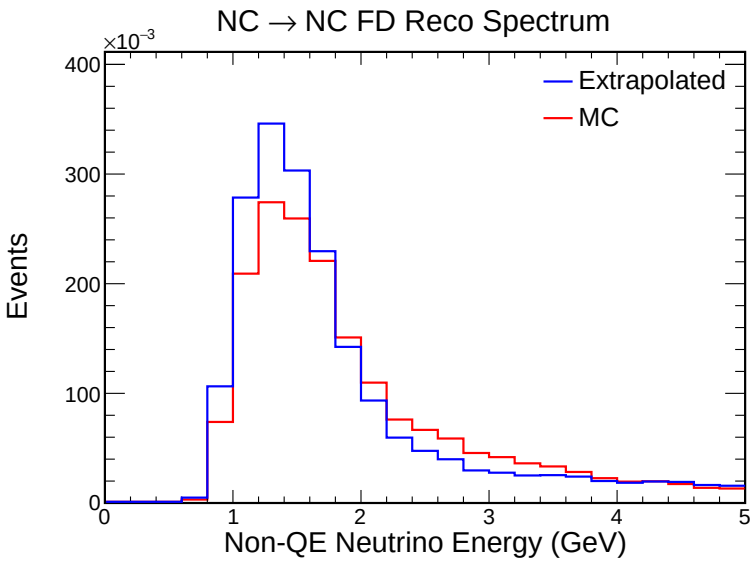


Figure 13.72: Plot comparing unaltered and reweighted-by-data reconstructed neutrino energies in GeV using simulated events in the far detector for NC interactions. The solid red histogram is unaltered simulation; the solid blue histogram is the simulation spectrum after reweighting by the extrapolation of near detector data. The extrapolated spectrum is used by the analysis as its prediction. The simulation is scaled down to match the POT of the far detector data. This is for the nonQE population. The true NC population was created using a proportional decomposition, described in Section 13.1.

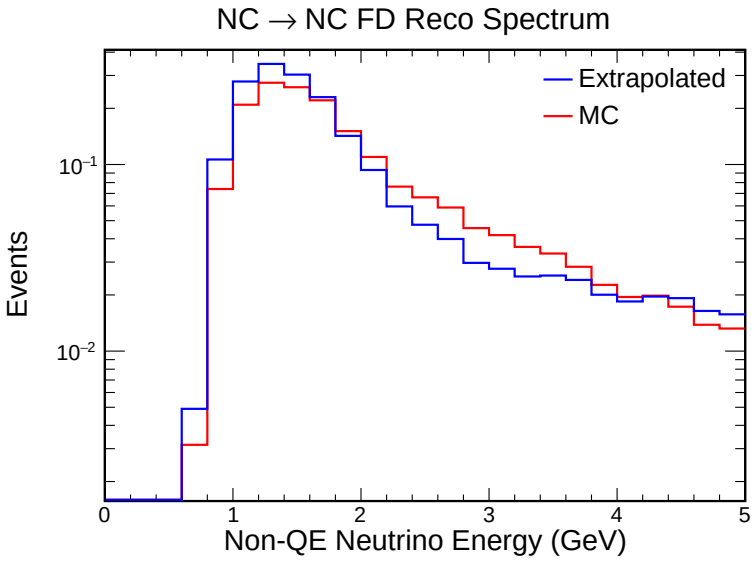


Figure 13.73: Plot comparing unaltered and reweighted-by-data reconstructed neutrino energies in GeV using simulated events in the far detector for NC interactions. The vertical axis is plotted logarithmically. The solid red histogram is unaltered simulation; the solid blue histogram is the simulation spectrum after reweighting by the extrapolation of near detector data. The extrapolated spectrum is used by the analysis as its prediction. The simulation is scaled down to match the POT of the far detector data. This is for the nonQE population. The true NC population was created using a proportional decomposition, described in Section 13.1.

Chapter 14

Systematic Errors

To perform the analysis, we must first understand and quantify the systematic errors. Systematic errors allow us to determine the uncertainty our measurement has due to approximations and unknowns in our analysis. Only the systematic errors with the largest impact to the overall normalization of the analysis sample, the energy resolution of the sample, or the absolute energy scale will need to be included; smaller errors will not have a large impact on the final uncertainty of our measurement.

We can look at the effect of these systematic errors on the reconstructed neutrino energy spectra from simulation. Many systematic uncertainties, such as the overall normalization of charged current quasielastic interactions, stem from imperfect knowledge included in our simulations. Often, these uncertainties will affect the near and far detectors in the same way. If neutrinos have quasielastic interactions more often in the near detector than our simulation predicts, they will also have more quasielastic interactions in the far detector than our simulation predicts. By using the extrapolation procedure described in Chapter 13, we can use our near detector data to "fix" our far detector prediction from simulation. For uncertainties that effect the near and far detectors in the same way, if one of parameters was mis-modeled, the near detector data would show an altered spectrum. Extrapolation would then allow us to account for for this mis-modeling and, to first order, we would have no systematic error. However, this procedure doesn't perfectly remove systematic uncertainty from our simulation when the two detectors do not see exactly identical effects. For instance, since the near detector is smaller, the contained analysis population of the near detector is shifted to have

proportionally more low energy events than the far detector. Differences of this type can result in an imperfect canceling of systematic error.

Some sources of systematic error are not correlated between the two detectors. For instance, we independently calibrate the absolute energy scale in each detector. Therefore, one would not expect that if the near detector has an absolute energy scale 3% too high in the simulation relative to the data, that the far detector would also have an absolute energy scale 3% too high. It could just as well be 5% too high or 4% too low. In general, systematic errors of this type tend to have a larger effect on the analysis than ones that can be partially or wholly mitigated with the extrapolation procedure.

This chapter discusses the systematic errors incorporated in the analysis, along with other systematic errors that were considered but deemed negligible.

14.1 Neutrino Beam Simulation Systematic Error

Systematic errors can arise from mis-modeling the NuMI neutrino beam. The NuMI neutrino beam is created by smashing accelerated protons onto a fixed target. These collisions create many particles - most notably for the creation of a neutrino beam, pions and kaons. These charged hadrons are focused into a beam where they decay into neutrinos. See Chapter 5 for more information about the NuMI neutrino beam. If the beam simulation is mis-modeled, the expected neutrino energy distribution for the analysis will be different. By extrapolating near detector data to the far detector, most of the mis-modeling is fixed. However, the near detector does sample a larger solid angle of the neutrino beam than the far detector and it sees the neutrino beam as a line source, while the far detector sees a point source. These differences mean that the extrapolation does not perfectly cancel out all systematic differences in the beam simulation.

The neutrino beam systematic errors considered by this analysis are grouped broadly into two categories: hadron production systematic errors and beam transport systematic errors. Hadron production systematic errors deal with uncertainties in the initial creation of the pions and kaons that can alter the expected neutrino flux. The beam transport systematic errors cover uncertainties in the focusing of the charged hadrons to create a neutrino beam, which can also alter the expected neutrino flux. Uncertainties

from hadron production in the beam are discussed in Section 14.1.1; Section 14.1.2 covers the uncertainties from beam transportation. Finally, these errors are combined into one overall systematic error for neutrino beam simulation. The effect of this total error is presented in Section 14.1.3. More information about the neutrino beam simulation systematic errors can be found in Mur's presentation[69].

14.1.1 Hadron Production

Hadron production systematic errors deal with uncertainties in the initial creation of the charged pions and kaons which decay to create the NuMI neutrino beam. An error band is created based on the differences between NA49 data results and simulation for hadron production. NA49 was a fixed target experiment at the CERN accelerator facility. It used two large volume, fine-grained time projection chambers (TPC's) to capture the evolution of hadronic shower development and two intermediate sized TPC's to capture vertex behavior. The data used for comparison was published in the European Physics Journal[70]. See the NA49 website for more information about the collaboration, experiment, and findings[71].

The differences seen by hadron type in transverse momentum vs. invariant differential cross-section between the beam simulation and NA49 data is used to motivate alternative parameterizations of the simulation. These alternative parameterizations are used to create 1 σ error bands on the spectrum of true neutrino energy seen at each detector. See Radovic's presentations for more information about using NA49 results to constrain the NOνA flux[72][73]. This systematic error has a much larger effect than those discussed in the beam transport section. However, although it has a large effect individually on the neutrino energy spectrum for each detector, these differences largely cancel out when extrapolating.

14.1.2 Beam Transport

Beam transport systematic errors cover the uncertainties in the focusing of the charged hadrons used to create a neutrino beam. First, accelerated protons are smashed onto a fixed target. The location of the target can alter the final neutrino spectrum, as well as the size of the proton beam. The target interactions create charged hadrons, which

are focused using two horns that produce a magnetic field. The electrical current in the horns is fundamental to the focusing effect. The position of each horn, relative to the target, can also alter the final neutrino spectrum. 8 individual systematic errors were considered. These errors are:

1. Beam position ± 0.5 mm in horizontal direction
2. Beam position ± 0.5 mm in vertical direction
3. Beam spot size ± 0.2 mm in both the horizontal and vertical direction
4. Target position shifted + 2 mm closer to the beam
5. Horn current ± 1 kAmp
6. Model magnetic field as an exponential distribution in the surface of the horns, instead of a linear distribution
7. Position of first horn ± 2 mm in both the horizontal and vertical direction
8. Position of second horn ± 2 mm in both the horizontal and vertical direction

For plots justifying the size of the shifts, see Goodenough's presentation[66]. For more information on the beam transport systematic errors, see Maan's presentation[74]. These systematic errors have a much smaller effect than the error discussed in Section 14.1.1. The effect of these errors also largely cancels out when extrapolating.

14.1.3 Total Systematic Error

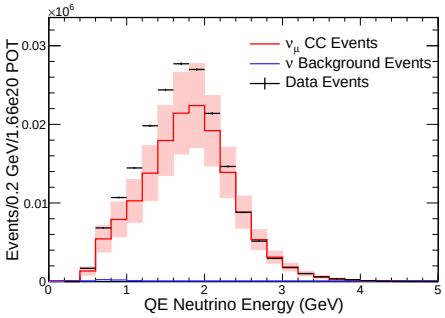
Instead of implementing each systematic uncertainty individually in the analysis, it was easier to implement one total neutrino beam simulation uncertainty. To create a total neutrino beam simulation systematic uncertainty to use for the analysis, envelopes of each error were combined in quadrature. It was assumed that the errors were not correlated. The envelope of each error is defined as, for each bin in true neutrino energy, the maximum of the absolute values of the +1 and -1 σ shifts. The total $\sigma_i^{(tot,\pm)}$ for each bin i in true neutrino energy is given by:

$$\sigma_i^{(tot,\pm)} = \sum_{n=0}^{9} \left[max\left(|\sigma_i^{(n,+)}|, |\sigma_i^{(n,-)}| \right) \right]^2. \tag{14.1}$$

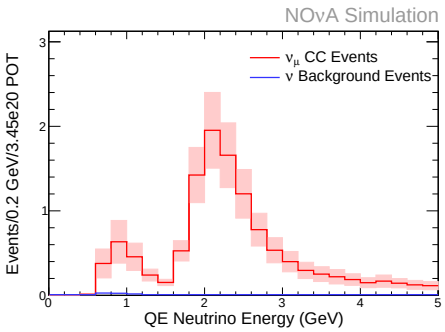
For ease and convenience, it has the same absolute value for +1 and -1 σ shifts. n is one of the nine individual beam simulation systematic errors.

We can look at the effect of the systematic uncertainty on the reconstructed neutrino energy spectra from simulation. For each population, the near detector spectrum is shown with the systematic error band. Then the far detector prediction before extrapolation is shown. Without extrapolation, the effect of the uncertainty is as large as it would be if we only had one detector. Finally, the far detector prediction with extrapolation is shown. For this plot, the effect of the systematic error on the analysis is dramatically reduced. If the beam simulation was mis-modeled, the near detector data would show an altered spectrum. Then, by using the the near detector data to alter our far detector prediction through the extrapolation process, we would account for that mis-modeling. Therefore, the final uncertainty on the analysis is significantly lessened. This procedure doesn't perfectly remove systematic uncertainty from our simulation since the two detectors do not see exactly identical effects. For instance, the near detector subtends a larger solid angle of the beam than the far detector does. Also, the near detector sees a line source of neutrinos, while the more-distance far detector sees a point source. Differences of this type can result in an imperfect canceling of systematic uncertainty.

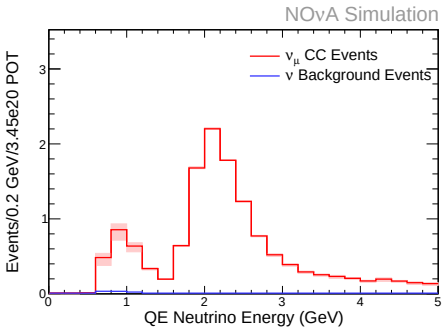
Figure 14.1 displays the systematic error from the beam simulation for the QE population in each detector. Figure 14.2 is for the nonQE population. By comparing the middle and bottom plots for each population, one can see that most of the error cancels when extrapolation is used.



(a) Near Detector

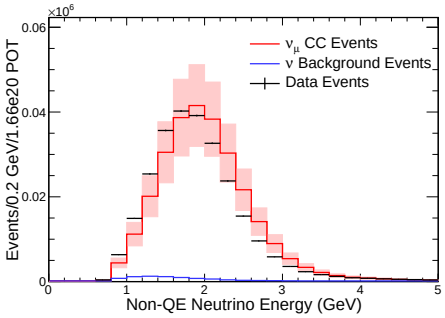


(b) Far Detector without Extrapolation

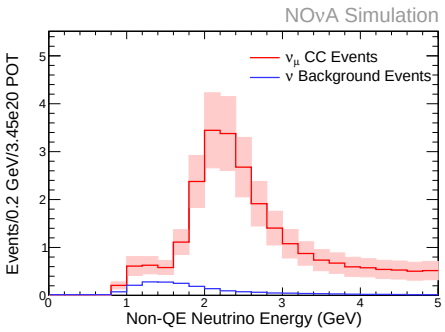


(c) Far Detector with Extrapolation

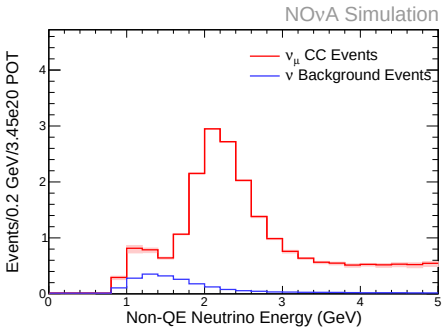
Figure 14.1: Plot of the reconstructed neutrino energy in GeV with the total beam simulation systematic error band for the QE sample. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.



(a) Near Detector



(b) Far Detector without Extrapolation



(c) Far Detector with Extrapolation

Figure 14.2: Plot of the reconstructed neutrino energy in GeV with the total beam simulation systematic error band for the nonQE sample. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.

14.2 Neutrino Interaction Simulation Systematic Error

Having simulated the neutrino beam, we next simulate the neutrinos interacting with the NOνA detectors. This is done using GENIE. For more information about this simulation, see Chapter 6. The model used to simulate neutrino interactions has some uncertainty; for instance, the overall normalization for charged current quasielastic interactions is not perfectly measured by experiments. The GENIE simulation provides handles that allow one to reweight individual simulation interactions by +1 and -1 σ to model many of these uncertainties.

GENIE provides many knobs that can be adjusted, most of which have a very small effect on the analysis. It would be somewhat impractical to use each knob independently. Instead, the six with the largest effects on the signal and background distributions will be used independently. These are discussed in Section 14.2.1. The rest of the systematic errors were added in quadrature and will be applied together as one systematic error. The small systematics are discussed in Section 14.2.2.

14.2.1 Large Neutrino Interaction Simulation Errors

Six adjustments to the neutrino cross section from the GENIE framework had the largest effects on either the signal or background events for this analysis. These adjustments are implemented individually. They are:

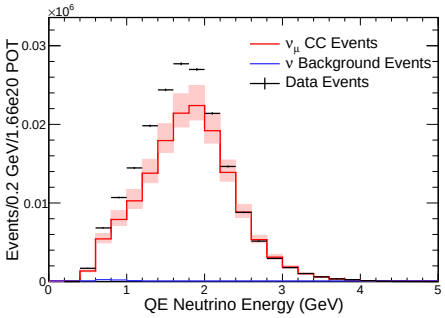
1. Axial mass for charged current quasielastic interactions shifted from -15% to $+25\%$
2. Axial mass for charged current resonance production shifted $\pm 20\%$
3. Vector mass for charged current resonance production shifted $\pm 10\%$
4. Axial mass for neutral current resonance production shifted $\pm 20\%$
5. Vector mass for neutral current resonance production shifted $\pm 10\%$
6. Axial mass for neutral current elastic interactions shifted $\pm 25\%$

The value of the shifts was taken from Chapter 8 of the GENIE Physics and User Manual[75].

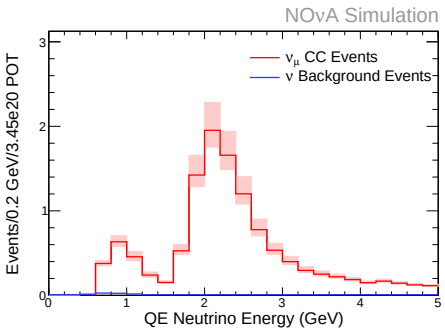
We can look at the effect of these adjustments on the reconstructed neutrino energy spectra from simulation. For each, the near detector spectrum is shown with the associated uncertainty shown as a systematic error band. Then the far detector prediction before extrapolation is shown. Without extrapolation, the effect of the uncertainty is as large as it would be if we only had one detector. Finally, the far detector prediction with extrapolation is shown. From this plot, one can see that the effect of this systematic error on the analysis is dramatically reduced. If one of parameters were mis-modeled, the near detector data would show an altered spectrum. Then, by using the the near detector data to alter our far detector prediction through the extrapolation process, we would account for that mis-modeling. Therefore, the final uncertainty on the analysis is significantly lessened. This procedure doesn't completely remove systematic uncertainty because the two detectors do not see exactly identical effects. For instance, since the near detector is smaller, the contained analysis population of the near detector is shifted to have proportionally more low energy events than the far detector. Differences of this type can result in an imperfect canceling of systematic error.

Figure 14.3 displays the systematic error from uncertainty in the axial mass for CC QE interactions for the QE population in each detector. Figure 14.4 is for the nonQE population. Figure 14.5 displays the systematic error from uncertainty in the axial mass for CC resonance interactions for the QE population in each detector. Figure 14.6 is for the nonQE population. Figure 14.7 displays the systematic error from uncertainty in the vector mass for CC resonance interactions for the QE population in each detector. Figure 14.8 is for the nonQE population. Figure 14.9 displays the systematic error from uncertainty in the axial mass for NC resonance interactions for the QE population in each detector. Figure 14.10 is for the nonQE population. Figure 14.11 displays the systematic error from uncertainty in the vector mass for NC resonance interactions for the QE population in each detector. Figure 14.12 is for the nonQE population. Figure 14.13 displays the systematic error from uncertainty in the axial mass for NC elastic interactions for the QE population in each detector. Figure 14.14 is for the nonQE population.

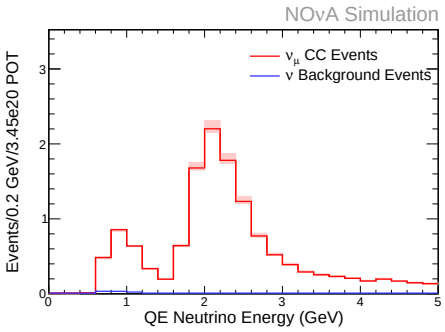
From the figures, we can see that extrapolation removes most of the impact of the systematic uncertainty from these parameters on the analysis. The remaining systematic uncertainty is relatively small.



(a) Near Detector

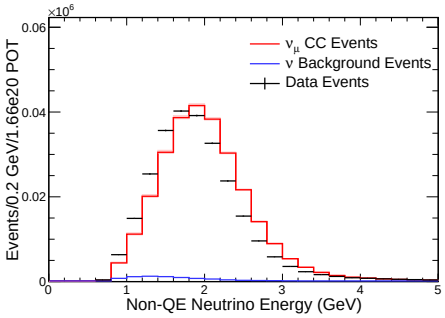


(b) Far Detector without Extrapolation

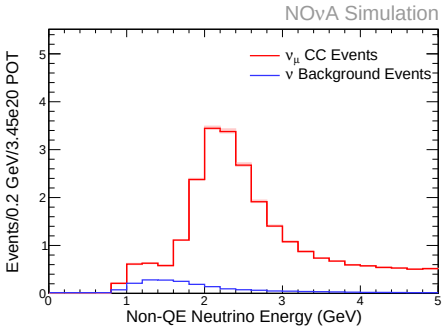


(c) Far Detector with Extrapolation

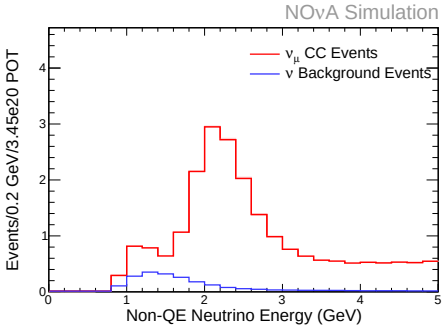
Figure 14.3: Plot of the reconstructed neutrino energy in GeV with the axial mass for CC QE interactions systematic error band for the QE sample. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.



(a) Near Detector

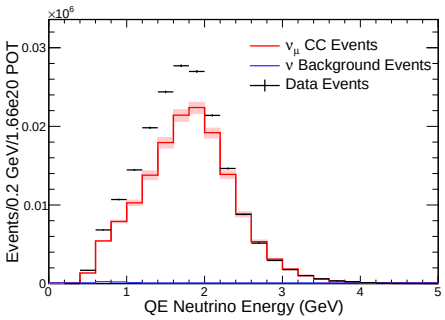


(b) Far Detector without Extrapolation

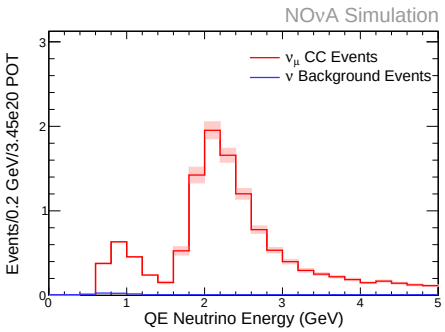


(c) Far Detector with Extrapolation

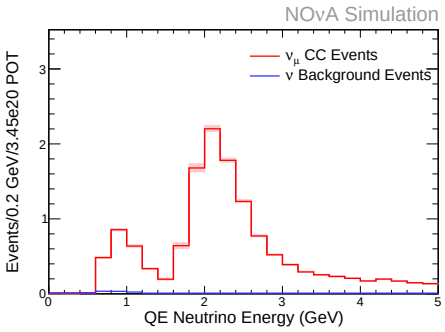
Figure 14.4: Plot of the reconstructed neutrino energy in GeV with the axial mass for CC QE interactions systematic error band for the nonQE sample. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.



(a) Near Detector

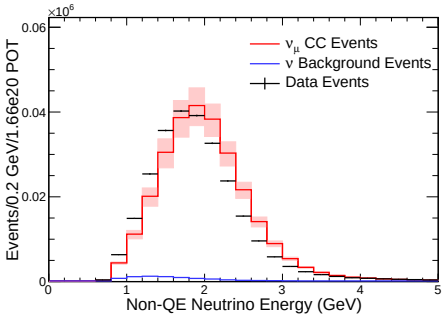


(b) Far Detector without Extrapolation

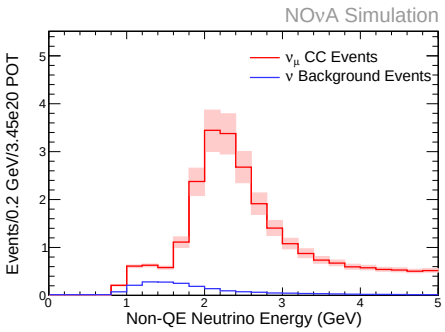


(c) Far Detector with Extrapolation

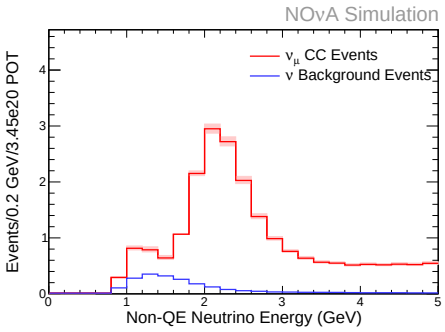
Figure 14.5: Plot of the reconstructed neutrino energy in GeV with the axial mass for CC resonance interactions systematic error band for the QE sample. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.



(a) Near Detector

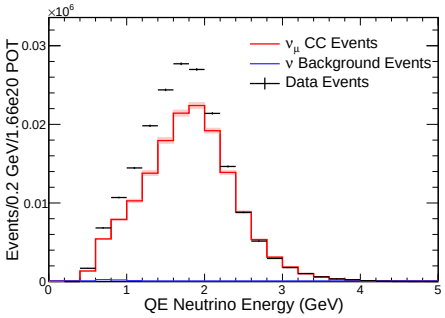


(b) Far Detector without Extrapolation

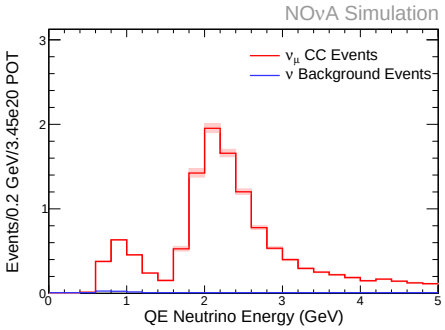


(c) Far Detector with Extrapolation

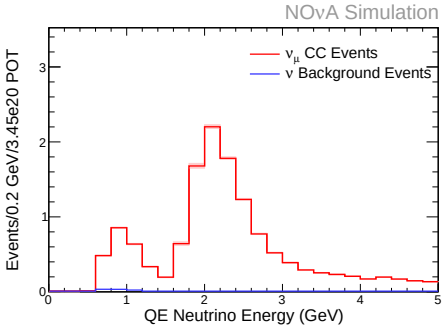
Figure 14.6: Plot of the reconstructed neutrino energy in GeV with the axial mass for CC resonance interactions systematic error band for the nonQE sample. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.



(a) Near Detector

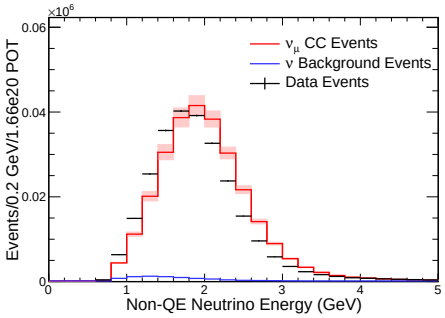


(b) Far Detector without Extrapolation

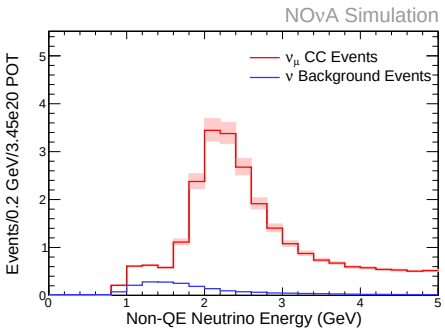


(c) Far Detector with Extrapolation

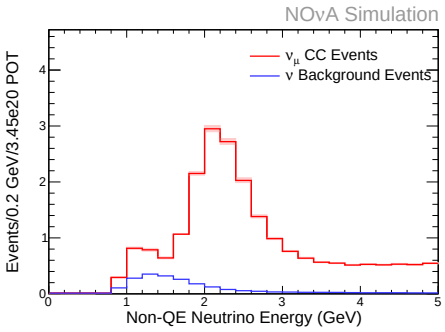
Figure 14.7: Plot of the reconstructed neutrino energy in GeV with the vector mass for CC resonance interactions systematic error band for the QE sample. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.



(a) Near Detector

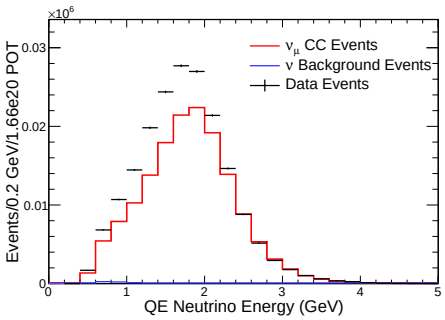


(b) Far Detector without Extrapolation

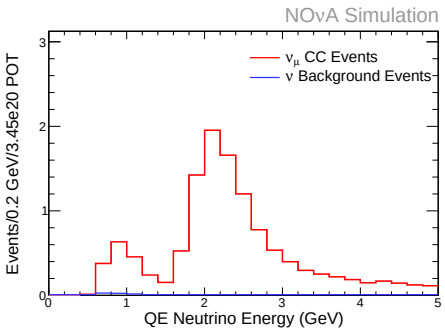


(c) Far Detector with Extrapolation

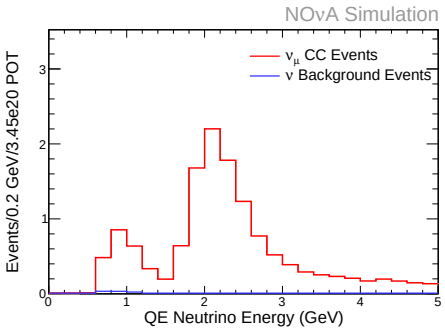
Figure 14.8: Plot of the reconstructed neutrino energy in GeV with the vector mass for CC resonance interactions systematic error band for the nonQE sample. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.



(a) Near Detector

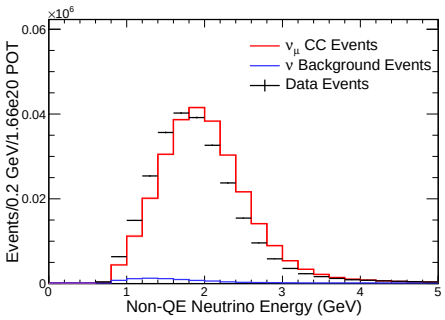


(b) Far Detector without Extrapolation

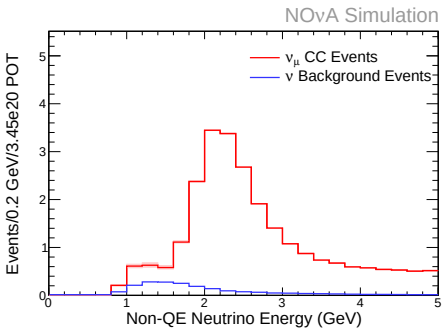


(c) Far Detector with Extrapolation

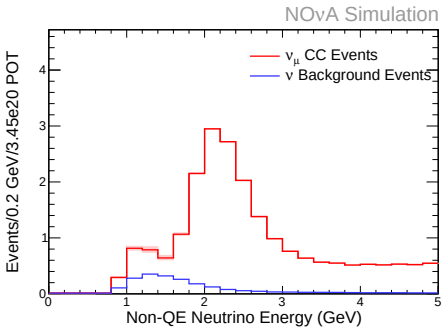
Figure 14.9: Plot of the reconstructed neutrino energy in GeV with the axial mass for NC resonance interactions systematic error band for the QE sample. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.



(a) Near Detector

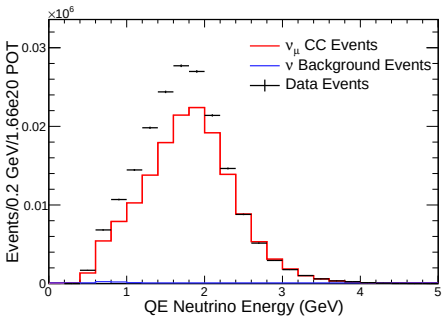


(b) Far Detector without Extrapolation

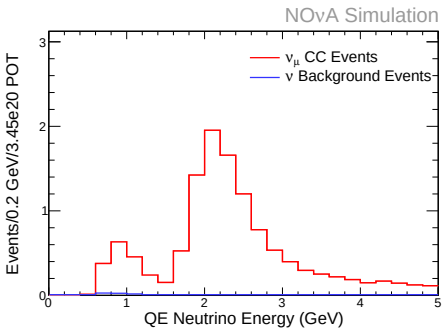


(c) Far Detector with Extrapolation

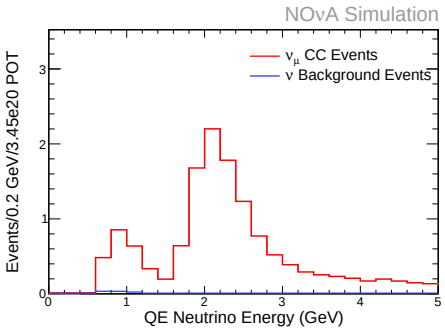
Figure 14.10: Plot of the reconstructed neutrino energy in GeV with the axial mass for NC resonance interactions systematic error band for the nonQE sample. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.



(a) Near Detector

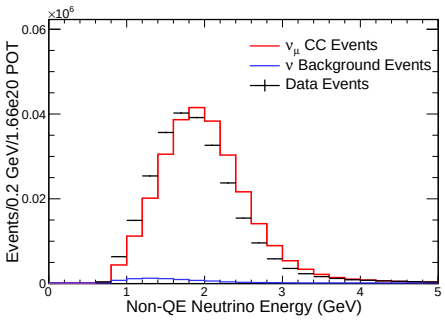


(b) Far Detector without Extrapolation

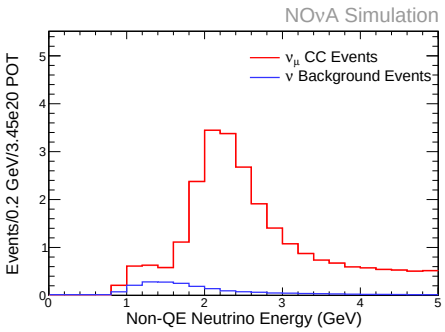


(c) Far Detector with Extrapolation

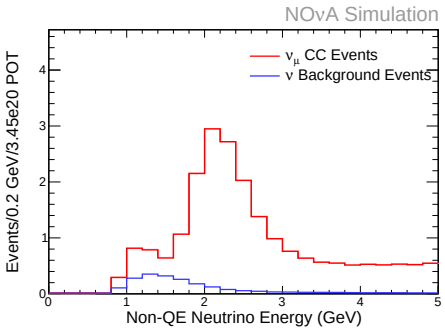
Figure 14.11: Plot of the reconstructed neutrino energy in GeV with the vector mass for NC resonance interactions systematic error band for the QE sample. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.



(a) Near Detector

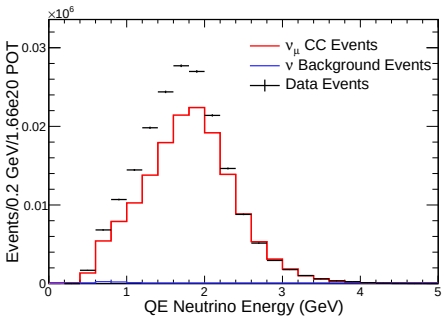


(b) Far Detector without Extrapolation

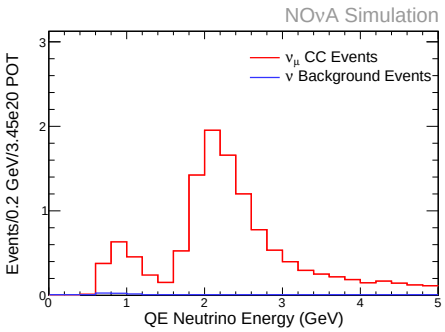


(c) Far Detector with Extrapolation

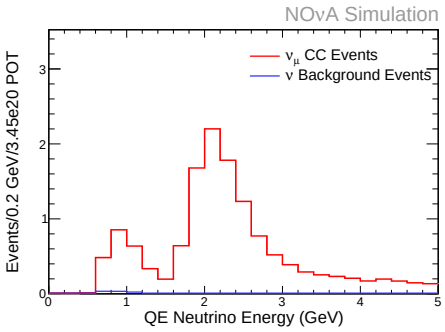
Figure 14.12: Plot of the reconstructed neutrino energy in GeV with the vector mass for NC resonance interactions systematic error band for the nonQE sample. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.



(a) Near Detector

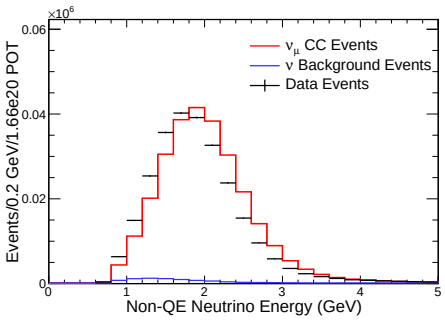


(b) Far Detector without Extrapolation

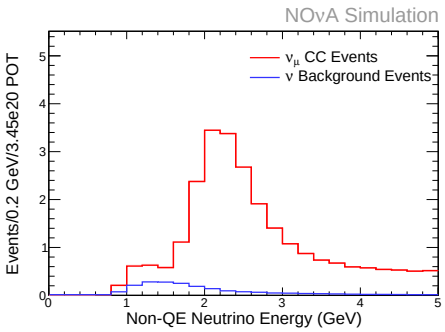


(c) Far Detector with Extrapolation

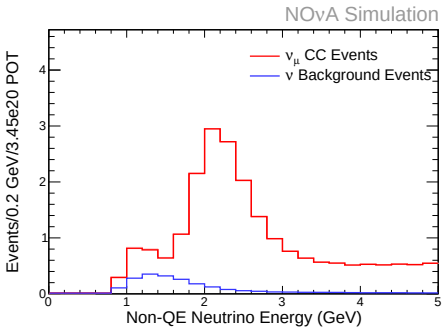
Figure 14.13: Plot of the reconstructed neutrino energy in GeV with the axial mass for NC elastic interactions systematic error band for the QE sample. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.



(a) Near Detector



(b) Far Detector without Extrapolation



(c) Far Detector with Extrapolation

Figure 14.14: Plot of the reconstructed neutrino energy in GeV with the axial mass for NC elastic interactions systematic error band for the nonQE sample. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.

14.2.2 Small Neutrino Interaction Simulation Errors

The six systematic errors from GENIE with the largest effect on the analysis signal and background were treated individually. The rest of the systematic errors from the following adjustments to the parameters determining the neutrino cross section were added in quadrature. The method used is one the discussed in Section 14.1.3. The neutrino interaction cross-section shifts were:

1. Strange axial form factor η for NC elastic interactions shifted $\pm 30\%$
2. CCQE Pauli suppression (via changes in Fermi level k_F) shifted $\pm 35\%$
3. Choice of CCQE vector form factors (BBA05 $\leftrightarrow$ Dipole)
4. Axial mass for CC and NC coherent pion production shifted $\pm 50\%$
5. Nuclear size parameter controlling π absorption in RS model shifted $\pm 10\%$
6. Non-resonance bkg in νp CC 1π reactions shifted $\pm 50\%$
7. Non-resonance bkg in νp CC 2π reactions shifted $\pm 50\%$
8. Non-resonance bkg in νn CC 1π reactions shifted $\pm 50\%$
9. Non-resonance bkg in νn CC 2π reactions shifted $\pm 50\%$
10. Non-resonance bkg in νp NC 1π reactions shifted $\pm 50\%$
11. Non-resonance bkg in νp NC 2π reactions shifted $\pm 50\%$
12. Non-resonance bkg in νn NC 1π reactions shifted $\pm 50\%$
13. Non-resonance bkg in νn NC 2π reactions shifted $\pm 50\%$
14. A_{HT} higher-twist parameter in Bodek-Yang model scaling variable ξ_w shifted $\pm 25\%$
15. B_{HT} higher-twist parameter in Bodek-Yang model scaling variable ξ_w shifted $\pm 25\%$

16. C_{V1u} u valence GRV98 PDF correction parameter in Bodek-Yang model shifted ±30%
17. C_{V2u} u valence GRV98 PDF correction parameter in Bodek-Yang model shifted ±40%
18. Inclusive CC DIS cross-section normalization factor
19. $\bar{\nu}/\nu$ CC ratio
20. DIS nuclear modification (shadowing, anti-shadowing, EMC)

The shifts and their values are reproduced from Table 8.1 of the GENIE Physics and User Manual[75]. The neutrino-induced hadronization and resonance-decay shifts were:

1. Pion transverse momentum (p_T) for $N\pi$ states in AGKY
2. Pion Feynman x (x_F) for $N\pi$ states in AGKY
3. Hadron formation zone shifted ±50%
4. Pion angular distribution in $\Delta \rightarrow \pi N$ (isotropic $\leftrightarrow$ Rein-Sehgal)
5. Branching ratio for radiative resonance decays shifted ±50%
6. Branching ratio for single η resonance decays shifted ±50%

The shifts and their values are reproduced from Table 8.2 of the GENIE Physics and User Manual[75]. The intranuclear hadron transport shifts were:

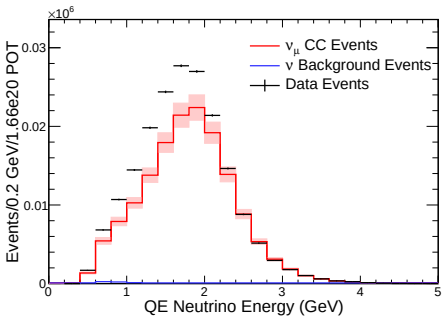
1. Nucleon mean free path (total rescattering probability) shifted ±20%
2. Nucleon charge exchange probability shifted ±50%
3. Nucleon elastic reaction probability shifted ±30%
4. Nucleon inelastic reaction probability shifted ±40%
5. Nucleon absorption probability shifted ±20%
6. Nucleon π-production probability shifted ±20%

7. π mean free path (total rescattering probability) shifted $\pm 20\%$
8. π charge exchange probability shifted $\pm 50\%$
9. π elastic reaction probability shifted $\pm 10\%$
10. π inelastic reaction probability shifted $\pm 40\%$
11. π absorption probability shifted $\pm 20\%$
12. π π-production probability shifted $\pm 20\%$

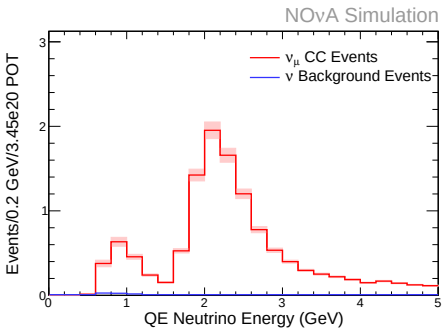
The shifts and their values are reproduced from Table 8.3 of the GENIE Physics and User Manual[75].

We can look at the effect of the systematic error on the reconstructed neutrino energy spectra from simulation. For each population, the near detector spectrum is shown with the resulting uncertainty shown as a band. Then the far detector prediction before extrapolation is shown. Finally, the far detector prediction with extrapolation is shown. For this plot, the effect of the uncertainties in these parameters on the analysis is dramatically reduced. If one of parameters were mis-modeled, the near detector data would show an altered spectrum. Then, by using the the near detector data to alter our far detector prediction through the extrapolation process, we would account for that mis-modeling. Therefore, the final uncertainty on the analysis is significantly lessened. This procedure doesn't completely remove systematic uncertainty from our simulation because the two detectors do not see exactly identical effects. For instance, since the near detector is smaller, the contained analysis population of the near detector is shifted to have proportionally more low energy events than the far detector. Differences of this type can result in an imperfect canceling of systematic error.

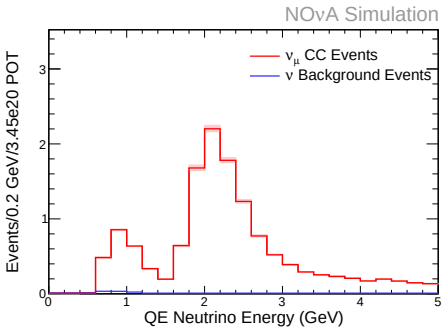
Figure 14.15 displays the systematic error from the summed GENIE effects for the QE population in each detector. Figure 14.16 is for the nonQE population. From the figures, we can see that extrapolation removes most of the impact of the systematic uncertainty from these parameters on the analysis. The remaining systematic uncertainty is relatively small.



(a) Near Detector

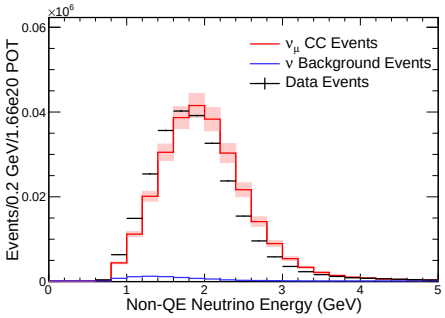


(b) Far Detector without Extrapolation

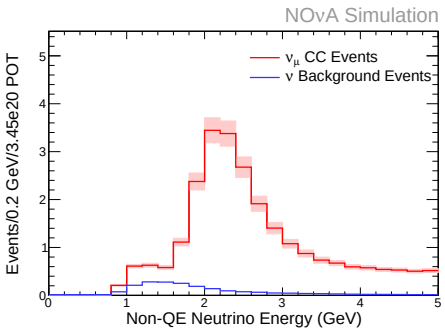


(c) Far Detector with Extrapolation

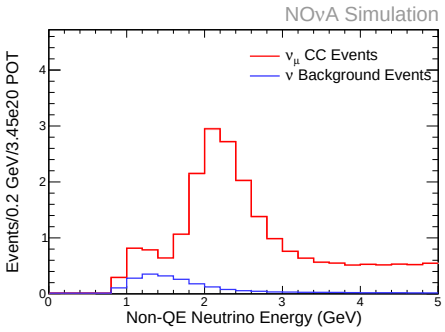
Figure 14.15: Plot of the reconstructed neutrino energy in GeV with the summed GENIE effects systematic error band for the QE sample. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.



(a) Near Detector



(b) Far Detector without Extrapolation



(c) Far Detector with Extrapolation

Figure 14.16: Plot of the reconstructed neutrino energy in GeV with the summed GENIE effects systematic error band for the nonQE sample. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.

14.3 Particle Propagation Simulation Systematic Error

To simulate propagating particles through the NOνA detector after the neutrino interaction, we use Geant4 configured to use the physics list QGSP_BERT_HP. For more information on this simulation, see Section 6.1. The models used in this simulation could be incorrect - this would cause a systematic error in our reconstructed neutrino energy spectrum. To understand the size of this systematic error, a study was conducted by comparing near detector simulations using different Geant4 physics lists.

Three alternative physics lists were used and the results compared to simulation using the nominal physics list, QGSP_BERT_HP. The first alternative list was QGSP_BIC_HP. For this physics list, the Binary cascade model[76] (as opposed to the Bertini cascade model[42][43]) was used to simulate primary protons and neutrons propagating through the detector with energies below 10 GeV. For particles not simulated using the Binary cascade model, the low energy parametrization model[77] was used. Also, the Binary light ion cascade model[78] was used to model inelastic interactions of ions up to a few GeV per nucleon. In other aspects, like the usage of the quark gluon string model[41] and the high precision neutron simulation, this physics list is the same as the nominal physics list.

The second alternative physics list considered was QGSC_BERT. This physics list does not use the high precision neutron simulation. Also, instead of using the G4Precompound model to simulate the de-excitation of nuclei, it uses the Chiral Invariant Phase Space (CHIPS) model[79][80]. In other aspects, like the usage of the quark gluon string model[41] and Bertini cascade model[42][43], this physics list is the same as the nominal physics list.

The third alternative physics list was FTF_BIC. For this list, the FRITIOF description[81] of string excitation and fragmentation was used to model the high energy interactions of protons, neutrons, pions, kaons and nuclei. This model is used for energies greater than 5 GeV. The quark gluon string model[41] is not used. This physics list uses the Binary cascade model[76] to simulate primary protons and neutrons with energies below 10 GeV. The Binary light ion cascade model[76] was used to model inelastic interactions of ions up to a few GeV per nucleon. The high precision neutron simulation was not used. This physics list is the most different from the nominal physics list.

Near detector simulation files with the same underlying GENIE simulation were produced using the nominal physics list and the three variants. A suite of plots was made comparing the simulations after applying the normal event selection cuts (see Chapter 10 for more information about event selection cuts). Looking at plots of reconstructed neutrino energy, the different physics lists changed the mean value ∼0.5% and the overall normalization by ∼1%. The shifts are negligible compared to the discrepancies addressed in the data-driven hadronic number of hits systematic, discussed in Section 14.12. Also, problems with modeling particles propagating through the detector will contribute to the data-driven hadronic number of hits systematic. Therefore, the systematic error seen from using different physics lists will be addressed by the hadronic number of hits systematic. For more information about alternative physics lists used in Geant, see the Geant website[44]. For more information about the study of particle propagation simulation errors, see the presentation by Rocco[82].

14.4 Birks' Law Simulation Systematic Error

As a charged particle travels through the NOνA detector, it interacts with the scintillator to produce light. In general, there is a linear relationship between the energy loss of the particle and the amount of light emitted from the scintillator. However, when the localized amount of energy loss is high, the scintillator can become saturated and the light yield begins to quench. Birks' Law is an empirical description of this non-linear quenching relationship. NOνA performed a measurement of the Birks' parameters by selecting protons in the near detector and measuring the dE/dx of the protons as a function of the distance from the end of the track. These parameter values are used in the simulation. Although the numbers match the NOνA data well, they are four times larger than values typically used by other experiments. This discrepancy lead us to try to evaluate a systematic uncertainty on our parametrization. We created simulation files with two alternative parameterizations of Birks' Law that matched values used by other experiments. The alternative parameterizations had a 10-20% effect on the amount of hadronic energy seen. They also altered some of the other inputs to the analysis, but the effects were small compared to the hadronic energy shift.

This analysis has chosen to not use a systematic error derived from the alternative

simulation files. The hadronic energy effects are already accounted for in the data-driven hadronic number of hits systematic, discussed in Section 14.12. If a systematic uncertainty was taken to account for alternative Birks' parameterizations, it would be double counting.

14.5 Detector Mass Systematic Error

If our detectors are more or less dense than we expect, we will have more or less neutrino interactions as a result. To understand the possible magnitude of this effect, a careful accounting of the uncertainties in our measurement of the densities of the extrusions, scintillator, glue, and fiber was made. The accounting was done independently for each detector. Table 14.1 lists the uncertainty from each detector component. This table is reproduced from the presentation by Raddatz[7]. The extrusions, which account for 36% of the detector mass, and the scintillator, which accounts for 63% of the detector mass, are the largest contributors to the total uncertainty. The uncertainty from each component was added in quadrature. This resulted in a 0.7% uncertainty in detector mass for each detector.

Table 14.1: Table of uncertainties in detector mass from each component. The far detector and near detector uncertainties are listed separately. The total uncertainty is created from adding each component uncertainty in quadrature. This table is reproduced from the presentation by Raddatz[7].

	Far Detector		**Near Detector**	
Component	Mass (kg/m/cell)	Percentage	Mass (kg/m/cell)	Percentage
Extrusions	0.951 ± 0.016	36.4 ± 0.6	0.951 ± 0.016	36.2 ± 0.6
Scintillator	1.641 ± 0.009	62.8 ± 0.3	1.653 ± 0.011	63.0 ± 0.4
Glue	0.019 ± 0.0003	0.07 ± 0.01	0.019 ± 0.0003	0.07 ± 0.01
Fiber	0.001 ± 0.001	0.04 ± 0.04	0.001 ± 0.001	0.04 ± 0.04
Total	2.612 ± 0.018	100 ± 0.7	2.624 ± 0.019	100 ± 0.7

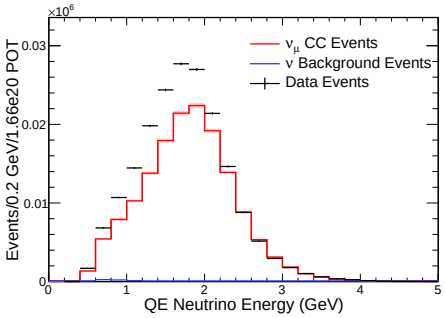
Another source of error stems from the fact that the simulated detector is not identical to the physical detector. Although the correct nominal material densities were used, the dimensions of the extrusions in the simulation are slightly different than those

of the physical extrusions. The PVC simulated volume is too small and the scintillator simulated volume is too large. The amount of simulated glue is half of what the physical detector has. No fiber was simulated, although obviously, the physical detector contains fiber. Table 14.2 lists the differences in total mass by component for the far detector. This table is reproduced from the presentation by Musser[8].

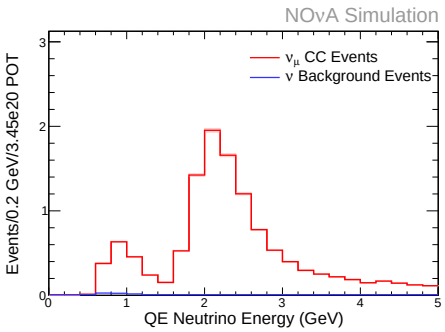
After adding the error from each source in quadrature, we have a -0.82% normalization error. This error could be corrected for in the analysis; however, since the error has a very small effect on the analysis, we instead chose to treat it as a $\pm$ 0.82% systematic uncertainty. This decision was made for convenience. This error, as well as the 0.7% uncertainty in detector mass, affects both detectors in the same way. Therefore, after adding these two errors together, a $\pm$ 1.08% absolute normalization systematic error will be applied. Figure 14.17 displays the systematic error from the absolute normalization for the QE population in each detector. Figure 14.18 is for the nonQE population. The effect of this systematic error on the reconstructed neutrino energy for each detector individually is very small; using extrapolation to correct the far detector prediction based on the near detector data makes the effect almost imperceivable.

Table 14.2: Table of the total mass of each component in the physical and simulated far detector. This table is reproduced from the presentation by Musser[8].

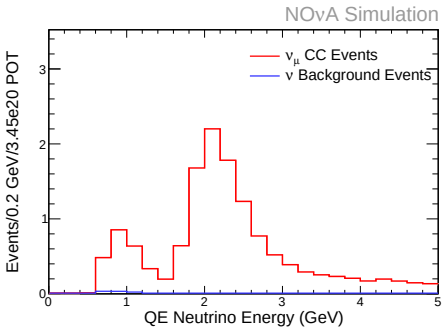
Component	Mass in Physical Det. (T)	Mass in Simulated Det. (T)
Extrusions	3,449.935	3,307.704
Scintillator	10,225.324	10,378,475
Glue	93.3	49.45
Fiber	0.0003	0



(a) Near Detector

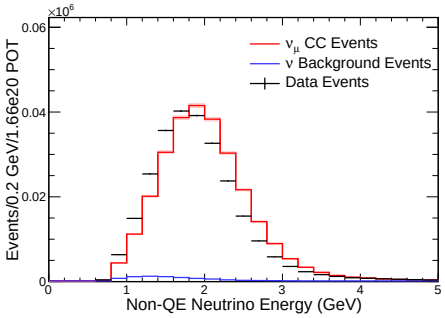


(b) Far Detector without Extrapolation

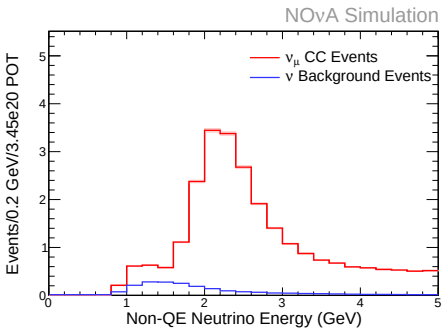


(c) Far Detector with Extrapolation

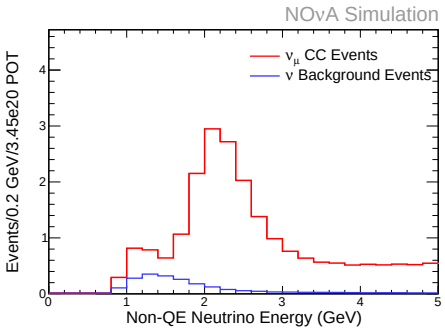
Figure 14.17: Plot of the reconstructed neutrino energy in GeV with the absolute normalization systematic error band for the QE sample. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.



(a) Near Detector



(b) Far Detector without Extrapolation



(c) Far Detector with Extrapolation

Figure 14.18: Plot of the reconstructed neutrino energy in GeV with the absolute normalization systematic error band for the nonQE sample. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.

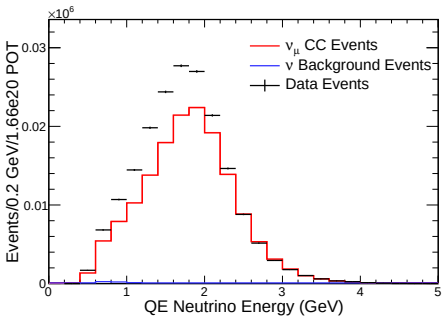
A very small difference between the simulated near and far detectors was found. The extrusion widths coded into our simulation are slightly different but the underlying cell dimensions are identical, so it is currently unclear if this simulation difference actually results in a physics difference. To cover this uncertainty, a ± 0.05% relative normalization systematic error is taken. The magnitude of the simulation difference was presented by Musser[8]. Figure 14.19 displays the systematic error from the relative normalization for the QE population in each detector. Figure 14.20 is for the nonQE population. The effect of this systematic uncertainty on the reconstructed neutrino energy spectra is almost imperceivable.

14.6 Muon Range Systematic Error

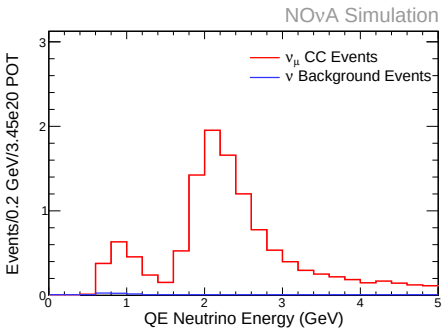
Uncertainties in the mass of the NOνA detector can cause normalization errors; these are discussed in Section 14.5. The uncertainty in the composition of the materials making up the detector can also cause a systematic error when relating the muon range to the muon energy. This would lead to an error in the reconstructed neutrino energy, which directly effects the analysis.

As discussed in Section 14.5, the active region of the detector has a ±0.7% uncertainty in the mass. We will take this as a ±0.7% uncertainty on the length of the muons. For muons in the near detector that traveled through the muon catcher, we need to also account for the uncertainty in the mass of the steel planes. We know the thickness of the steel to ±2%, so we will treat the length of the muons traveling in steel planes as ±2% uncertain. The magnitude of this uncertainty was presented by Raddatz[83].

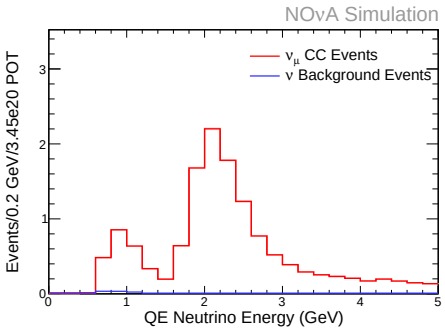
Using these uncertainties in muon length, altered reconstructed neutrino energy spectra were created with $\pm 1\sigma$ shifts. The altered spectra were used to create confidence limits of the oscillation parameters measured by this analysis. There wasn't a visible change in the confidence limits and the error was deemed negligible. This isn't unexpected; the extrapolation procedure should remove most of the effect from the active region uncertainty. The extrapolation procedure doesn't remove uncertainty from the steel thickness, since the far detector doesn't have a muon catcher. However, although the uncertainty in the steel exists, many muons in the near detector never enter the muon catcher or lose only a small portion of their energy in the muon catcher. This



(a) Near Detector

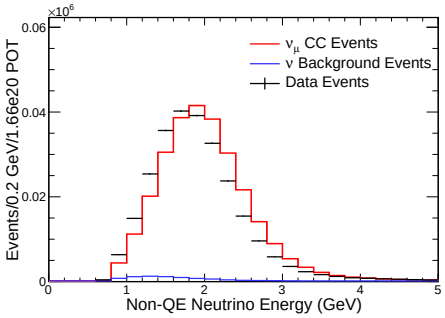


(b) Far Detector without Extrapolation

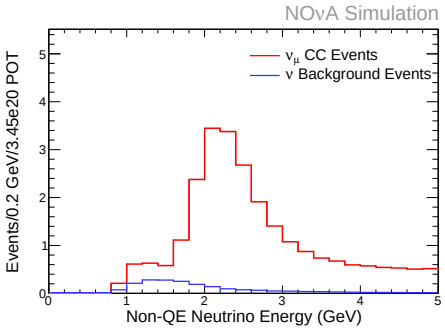


(c) Far Detector with Extrapolation

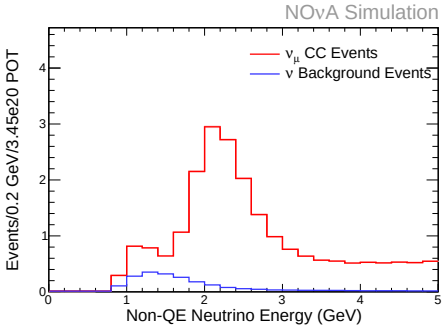
Figure 14.19: Plot of the reconstructed neutrino energy in GeV with the relative normalization systematic error band for the QE sample. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.



(a) Near Detector



(b) Far Detector without Extrapolation



(c) Far Detector with Extrapolation

Figure 14.20: Plot of the reconstructed neutrino energy in GeV with the relative normalization systematic error band for the nonQE sample. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.

systematic error was not used in the final analysis. The analysis of this systematic error was presented by Raddatz[83].

14.7 Detector Alignment Systematic Error

The detector geometry in our simulation does not have the same relative alignment of modules that the physical detector has. The simulation geometry has a regular staggered offset of planes in the same view. This means that if a particle traveling in a straight line happened to travel through a cell wall in one plane of a view, in the next plane of that view, it would be traveling in the middle of a cell. This staggering helps to remove common trajectories of particles that could be hidden from our reconstruction. However, when building the physical detector, this staggering was not as precise as the staggering used by the simulation. Physical modules are large and difficult to position; perfect precision can not be achieved. Geometry differences can cause systematic effects for our analysis. If the cell positions for the physical detector are different than the simulated one, muon tracks might look broken and not be fully reconstructed. This could result in a different muon and hadronic energy estimation, leading to an altered reconstructed neutrino energy spectrum.

In the far detector, the uncertainty in the staggered position of the planes is 0.5 inches for both views. The uncertainty in the tilt of each plane is 1.25 mrad. A set of far detector simulation files was created using an altered geometry for the generation and propagation of neutrino interactions. The altered geometry set the stagger and tilt of each plane relative to nominal by randomly sampling Gaussian distributions with σ values set to the known uncertainty. Then the original geometry, without the jittered shifts, was used to reconstruct the simulated events. The reconstructed neutrino energy spectrum was shifted by less than 0.2% when using the altered geometry. The effect on the ability to measure the oscillation parameters was very small; the uncertainty on $\sin^2\theta_{23}$ was 5×10^{-4} and the uncertainty on $|\Delta m^2_{32}|$ was 5×10^{-8} eV^2. The analysis of the far detector detector alignment effects was presented by Zamorano[84].

For the near detector, a different procedure was used to evaluate the systematic error effects. An improved geometry simulation was created that matched the physical detector better than the simulation used for this analysis. This improved simulated

geometry was not used for the main analysis in the interest of time. The improved geometry set the stagger of each plane to be the value measured by survey data. This was important, because the differences in the near detector stagger are larger and less regular than those in the far detector.

The improved geometry simulation was used to reconstruct the near detector data files. This was compared to near detector data files reconstructed with the original geometry simulation. The differences in the reconstructed neutrino energy spectrum were less than 0.05%. The muon PID distribution, ReMId[1], changed 20% in the selection region. Using the improved geometry, the peak was much sharper. This is because muons had a better chance to be fully reconstructed. This didn't affect the analysis, though, because, within 0.2%, the same number of of events still passed the ReMId cut, which generously includes the entire signal peak region. When extrapolated through the analysis, the improved geometry had a visually negligible effect on the measurement of the oscillation parameters. The analysis of near detector alignment effects was presented by Zamorano[85].

Given that the effects are negligible, no alignment systematic error will be taken for this source of uncertainty.

14.8 POT Accounting Systematic Error

The analysis depends on understanding the total Protons On Target (POT) for the data sets in each detector. The POT allows us to calculate the expected number of neutrino interactions for the dataset. If the true POT was, for instance, higher than we thought, we would see more neutrino interactions than we expected. This change in normalization could alter our measurement of the oscillation parameters. If we only take data when both detectors are active, then any error in the reported POT will cancel with the extrapolation procedure. However, much of the far detector data was taken without corresponding near detector data either existing (9 months of the early running period) or being used (2 months of the later running period). The near detector data from the later running period was not processed in the interest of time. The extra far detector data significantly reduces our statistical error. The measurement sensitivity

[1]See Section 7.6 for more information about ReMId

would be worse if we restricted ourselves to only using running periods with processed data from both detectors. Therefore, we need to consider a systematic error from this unmatched far detector data period.

Monitoring devices exist that can give an estimate on how stable the POT accounting is through time. The POT is reported by the TRTGTD toroid. We are not concerned with mis-calibration of the torrid; only if this mis-calibration changes through time. Therefore, we can check the ratio of the reported POT from the TRTGTD torrid with the backup TR101D toroid and the internal proton beam measurement device DCCT. The ratio with the backup torrid is stable to less than 0.5%. The ratio with the internal proton beam device is stable to about 1%, but the internal proton beam device lost calibration for a period that caused larger than normal disagreement. It was determined that any error from differences in POT accounting would be negligible for the first analysis. The analysis of this systematic error was presented in Chapter 4 of the Beam Technical Note[86].

14.9 Concurrent Neutrino Interactions Systematic Error

The near detector, being much closer to the neutrino beam, experiences a much larger flux of neutrinos than the far detector. This means that, while the far detector sees less than 1 neutrino interaction a day, the near detector sees about 4 neutrino interactions for each beam spill. These concurrent neutrino interactions could cause problems for our analysis. If Slicer (see Section 7.3 for more information about our reconstruction) can not separate individual neutrino interactions with a high purity and efficiency, then the reconstructed neutrino energy spectrum for the near detector will be altered. By grouping two neutrino interactions together, we will reconstruct too high of a neutrino energy. If instead Slicer divides mistakenly divides a single neutrino event, we will reconstruct too low of a neutrino energy.

A study using the simulation was conducted to investigate how well Slicer separates neutrino interactions. Purity was defined as the visible energy in the slice from the primary neutrino interaction, divided by the total visible energy in the slice. Completeness was defined as the visible energy in the slice from the primary neutrino interaction, divided by the total visible energy deposited by the primary neutrino in the detector. The

primary neutrino is defined as the neutrino which contributes the most visible energy to the slice. Truth information was used to associate visible energy and true neutrino interactions. After applying cuts that required 10 hits in each view and that the true neutrino interaction vertex was well-contained in the detector, the near detector mean purity was found to be 98.8%. The mean completeness for the near detector was 96.8%. For the far detector (where concurrent neutrino interactions isn't an effect), the mean purity was 99.3% and the mean completeness was 99.3%. Purity is the metric of most concern to concurrent neutrino interaction effects. The purity for the near detector compared to the far detector is close enough that this effect was deemed negligible for this analysis. For future analyses with smaller errors, a better treatment is warranted. The analysis of concurrent neutrino interactions was presented by Baird[87].

14.10 Bad Channels Systematic Error

Bad channel masks are used to remove channels from the data that might contain suspect information, due to difficulties with the electronics or DAQ software. Channels that never report are removed from the analysis so the reconstruction knows that the channel is malfunctioning and that the particle track didn't have a true gap. Channels that constantly create noise hits are removed so that the reconstruction is not confused by these spurious hits. For more information about bad channel masks, see Section 7.2.

These channel masks have some uncertainty. Channels might be border-line cases; for some runs they are considered good and for some runs they are considered bad, although throughout they display marginal behavior. Some channels might never create hits that correspond to real physics but instead report noise hits at a rate that makes them look like a good channel.

The effects of alternative channel masks on the analysis was investigated. For each detector, a particular mask seen in the dataset was picked that represented a high amount of bad channels. For the near detector, the particular mask had 5% of the channels marked as bad, while the average channel mask marks 2% of the channels as bad. For the far detector, a mask was picked that marked 0.6% of the channels as bad. The average channel masks for the far detector marks 0.3% of the channels as bad.

An alternative set of simulation files was created for each detector which applied

only the particularly bad channel mask. Using the alternative set of simulation files, we can compare the reconstructed neutrino energy spectrum to that seen using the nominal simulation files. For the near detector, the energy shifted by 1%. For the far detector, the energy shifted 0.2%. Since worst-case masks were applied, the real error is considered to be much smaller. For this analysis, this source of systematic error is neglected. For future analyses with smaller statistical errors, a better treatment of this error is warranted. The bad channels systematic error was evaluated by Sepulveda-Quiroz[88].

14.11 Calibration Systematic Error

If the calibration procedure resulted in a systematic bias between the simulation and the data, our analysis would also be biased. See Chapter 8 for more information about calibration. Three types of bias, described below, were considered; only one was determined to have a large enough effect to not neglect.

14.11.1 Overly-uniform Simulation Calibration

The calibration procedure for the simulation treats the detector as if all the X-direction planes behaved the same and all the Y-direction planes behaved the same. In the near detector, the muon catcher is a slight complication. For the active region, all the X-direction planes are treated as the same and all the Y-direction planes are treated as the same. Then, in the muon catcher, all the X-direction planes are treated as the same, but different than the active planes. Also, all the Y-direction planes are treated as the same, but again, different than the active planes. For the data, the behavior of each cell is unique. Different cells will have different PVC reflectivity, scintillator composition, APD gain levels, and different fiber quality. This means that the data has a less uniform calibration than the simulation. To see if this introduces a bias, the simulation calibration was randomly jittered. To determine the size of the the jitter, a comparison between the calibrated energy per cm for simulation and data was made. After shifting the simulation distribution to the same mean location as the data, the RMS of the simulation distribution was smaller than the data distribution. After introducing a random jitter of 8% to the simulation files, the RMS of the two

distributions was the same. One could then consider the reconstructed neutrino energy spectrum using the jittered simulation and nominal simulation. The two spectra were within 1% for energies less than 5 GeV. Next, the impact to evaluating the oscillation parameters was considered. It had almost no impact and will be neglected. The effect of calibration jitter on this analysis was presented by Tamsett[89].

14.11.2 Calibration Effects Due to Distance from Readout

The second calibration systematic considered looked at the differences between the data and simulation detector response as a function of distance from the electronic readout. For the far detector, the calibrated detector response at the far end of the cell is 20% different between the data and the simulation, while it is within a few percent at the near end of the cell. This is less of an effect for the near detector, where distance from the readout is at most 4 m.

We created altered simulation files to evaluate the effects of this discrepancy. The difference between the data and simulation calibrated detector response with respect to distance from the end of the cell was fit with a second-order polynomial. This function was used to create altered simulation files, who have a detector response in distance from the readout that matches that seen in data. Using these altered simulation files, one can determine the differences in the reconstructed neutrino energy. Although a small shape difference can be seen, the energy difference is less than a 5% effect for energy regions of interest. When evaluating the impact to measuring the oscillation parameters, this shift has a larger impact than applying random jitters. However, it is still a relatively small impact, especially compared to other systematic effects taken into account. This systematic effect will be neglected. The analysis evaluating the effect of this systematic on the analysis was presented by Tamsett[89].

14.11.3 Absolute Energy Scale Offsets

The final calibration systematic considered was that of absolute energy scale. If the calibration procedure doesn't work correctly, the same energy deposition in data and simulation could be reconstructed as different amounts of energy. For instance, if the selection of events used for the calibration had more shower events contaminating the

data sample than in the simulation, the calibration for data would be wrong. This would create a bias in the analysis. It wouldn't effect the reconstructed muon energy, which is based on track length, but it would shift the reconstructed hadronic energy and therefore the reconstructed neutrino energy. It can also hamper efforts to identify muons by looking at the energy deposition along the length of a track.

We have evidence that, at least for the near detector, some bias exists. The most convincing study looked at muon tracks in the near detector which passed the official selection cuts in the beam data and simulation files of beam interactions. A few additional cuts were made, restricting the start and stop positions to ensure that no rock muons contaminated the sample and that the muons stopped before entering the muon catcher. Also, the muon tracks were required to have no overlapping hadronic vertex energy. A plot was made of the deposited energy per path length as a function of the distance from the end of the muon track. The curve for data is systematically lower than the curve for simulation. Shifting the simulation by a scale factor of 96.4% minimized the χ^2 between the two curves. Some small shape discrepancies remained, but the overall agreement was much better. For more information on the inconsistent absolute energy scales in data and simulation, see Raddatz's presentation[90].

It is believed that this bias in the calibration process was created due to problems in the underlying cosmic ray events used. The simulation files used to create the calibration constants did not allow cosmic rays to shower and interact in the rock above the near detector. This created cleaner tracks in the simulation files than were seen in the data. The cuts to create the calibration sample were optimized primarily using simulation files and were not severe enough to remove showering events from the data sample. Also, the data files were contaminated with some neutrino beam activity. This activity is different from the clean cosmic rays present in the simulation and could bias the result.

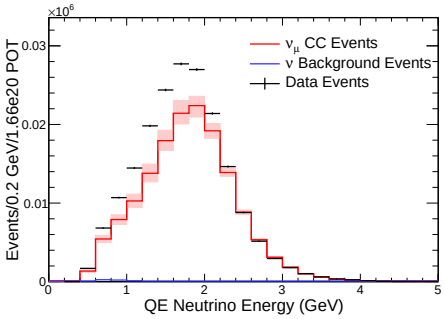
For future analysis, the cosmic ray simulation for the near detector has been improved and should allow for more shower-like events to be simulated. Also, the cuts for future rounds of calibration have been altered to remove showering events. Finally, the filter that removes neutrino beam activity now has tighter timing cuts which remove all beam events. These improvements, although shown to improve the agreement between data and simulation calibration results, were not used for this analysis in the interest of time.

For this analysis, the calibration for the near detector data has been scaled up by 3.7% to match the simulation. This number was determined from the study mentioned above. This study was considered the most in-depth and precise of the studies comparing the data and simulation absolute energy scales. The choice was made to scale up the data, as opposed to scaling down the simulation, because it is believed the data calibration was plagued by showering event contamination. However, a systematic error must still be applied to the absolute energy scale. Other studies found differences in the absolute energy scale, but at different magnitudes. One study, presented by Sachdev[91], was similar to the one used to determine the 3.7% scaling was done on near detector data and simulation, but it used slightly difference choices for selection cuts. This second study favored scaling the data up by 4.7%, which is larger than the scaling used. Also using near detector data and simulation, a study was conducted which selected a sample of π^0 candidates. The invariant mass of these candidates was calculated and plotted for both data and simulation. Again, the data sample was low, this time, by 5.1%. For more information about the π^0 calibration study, see the presentation by Davies[92]. Lastly, a study of Michel electrons was conducted. Comparing data and simulation energies in the near detector, this study found the data was either 8 or 10% low, depending on what metric was used. This study was also conducted for the far detector. For the far detector study, the detector was divided into regions close to the electronics readout and far from the electronics readout. In the near region, the data was 2% lower than the simulation. However, in the far region, the data was actually 6.2% higher than the simulation. For more information about the Michel electron calibration study, see Patterson's notes[93].

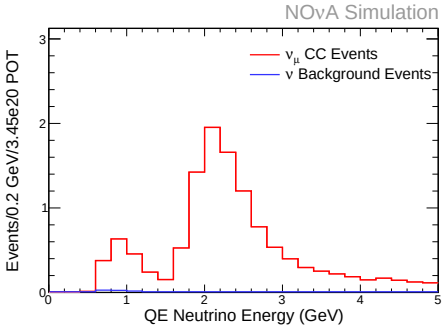
Given the general disagreement as to the level of agreement in calorimetric energy scale for the simulation and data, a systematic must be taken. Shifts in absolute calorimetric energy scale of $\pm 5\%$ will be applied to simulation files for both detectors and reprocessed through the entire chain of reconstruction. Histograms of the differences in the reconstructed neutrino energy spectrum for each population were created and used as templates for $\pm 1\sigma$ systematic errors. First, the peaks of reconstructed neutrino energy for each sample were fit with a Gaussian. A scale factor applied to the reconstructed hadronic energy was found that caused the peaks in reconstructed neutrino

energy to align. However, besides simply shifting the hadronic energy, the altered simulation shows overall normalization shifts (due to failing to identify the muon energy deposition) as well as small shape changes. To account for this, after the hadronic energy shift is made, a correlated shape correction will be made bin-by-bin in reconstructed neutrino energy.

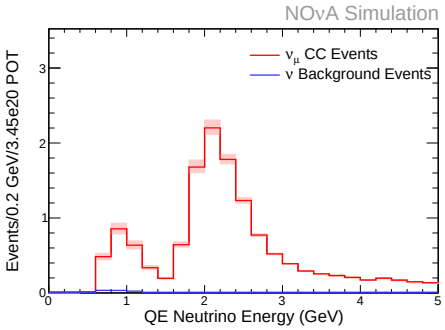
The systematic errors were applied in such a way that the value of the error for each detector is allowed to move independently of the other detector. The extrapolation procedure will not be able to mitigate the effect. Figure 14.21 displays the systematic error from the near detector absolute energy scale offset for the QE population in each detector. The plot for the far detector without extrapolation doesn't have an error band, because without extrapolation, the systematic error from the near detector absolute energy scale offset doesn't impact the far detector spectrum. Figure 14.22 is for the nonQE population. Again, the plot for the far detector without extrapolation doesn't have an error band, because without extrapolation, the systematic error from the near detector absolute energy scale offset doesn't impact the far detector spectrum. Figure 14.23 displays the systematic error from the far detector absolute energy scale offset for the QE population in each detector. Since this is the systematic error from the far detector absolute energy scale offset, the near detector spectrum doesn't have an error band. Figure 14.24 is for the nonQE population. Again, since this is the systematic error from the far detector absolute energy scale offset, the near detector spectrum doesn't have an error band. The effect of this systematic error on the precision of measuring the oscillation parameters is much larger than the other two calibration systematics considered. It is one of the largest systematic errors of this entire analysis - only the errors discussed in Section 14.12 rival the effect of this systematic uncertainty. Future analyses should work to reduce the uncertainty in the calorimetric energy scale.



(a) Near Detector

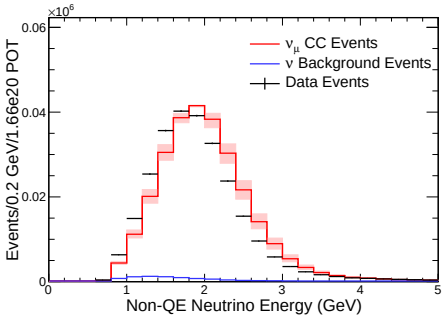


(b) Far Detector without Extrapolation

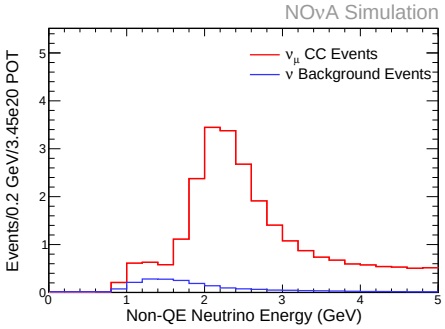


(c) Far Detector with Extrapolation

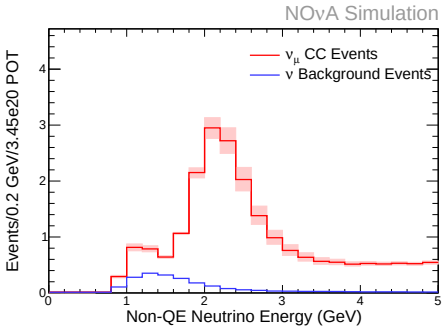
Figure 14.21: Plot of the reconstructed neutrino energy in GeV with the near detector absolute energy scale systematic error band for the QE sample. Figure 14.21b doesn't have an error band because the near detector systematic error doesn't affect the far detector without extrapolation. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.



(a) Near Detector

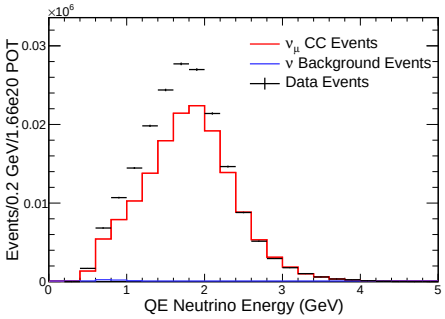


(b) Far Detector without Extrapolation

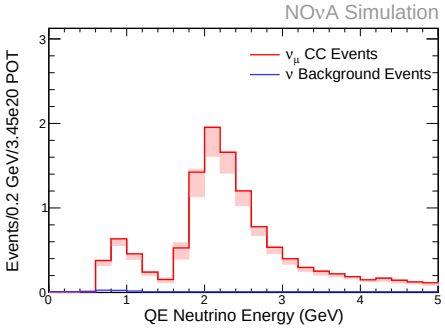


(c) Far Detector with Extrapolation

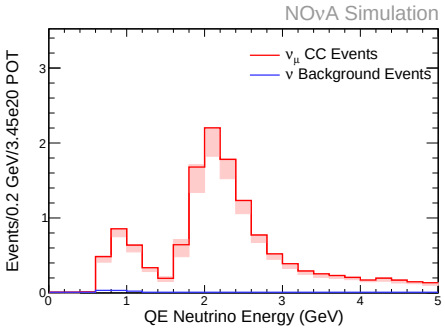
Figure 14.22: Plot of the reconstructed neutrino energy in GeV with the near detector absolute energy scale systematic error band for the nonQE sample. Figure 14.22b doesn't have an error band because the near detector systematic error doesn't affect the far detector without extrapolation. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.



(a) Near Detector

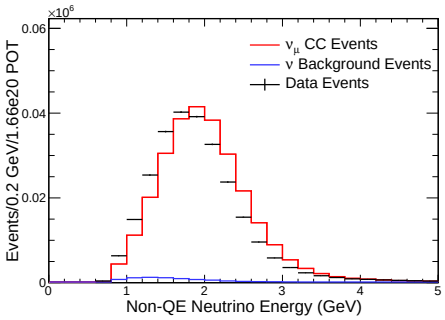


(b) Far Detector without Extrapolation

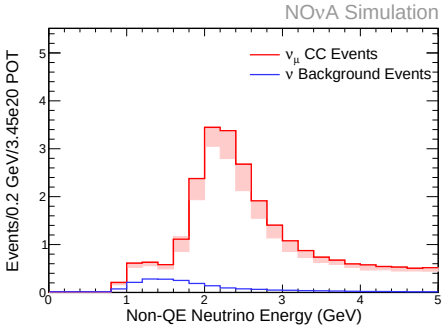


(c) Far Detector with Extrapolation

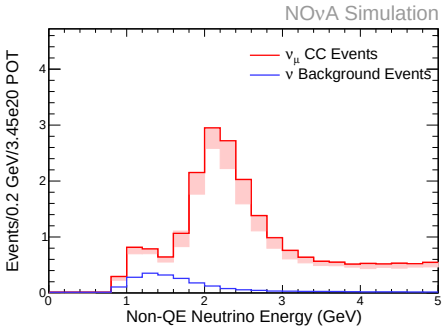
Figure 14.23: Plot of the reconstructed neutrino energy in GeV with the far detector absolute energy scale systematic error band for the QE sample. Figure 14.23a doesn't have an error band because the far detector systematic error doesn't affect the near detector. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.



(a) Near Detector



(b) Far Detector without Extrapolation



(c) Far Detector with Extrapolation

Figure 14.24: Plot of the reconstructed neutrino energy in GeV with the far detector absolute energy scale systematic error band for the nonQE sample. Figure 14.24a doesn't have an error band because the far detector systematic error doesn't affect the near detector. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.

14.12 Hadronic Number of Hits Systematic Error

As previously noted in Chapter 12, the data in the near detector has some significant discrepancies with the simulation. Specifically, the number of hadronic hits is different. Figure 12.27 shows that the data distribution has, on average, fewer hadronic hits than the simulation distribution. This causes the reconstructed hadronic energy in the data to be lower than in simulation, which in turn leads to a lower average reconstructed neutrino energy as well. Much investigation has been done to try and understand this difference; so far, no cause has been fully verified.

It is possible that some of the differences stem from failures in the modeling of neutrino interactions in the simulation. In the energy region of interest, the processes contributing to the neutrino interactions are not well-understood. For example, it has been suggested that an interaction channel with two particle-hole pairs needs to be added to the interaction simulation[94]. Early looks at the effect of this channel in explaining the data and simulation differences in NOνA are promising but not yet conclusive.

Many instrumental avenues to explain this discrepancy have also been pursued. Adjustments to the Birks' parametrization (Section 14.4) certainly change the number of hadronic hits seen; however, the effect was not large enough to explain the majority of the difference. Effects due to mis-matched light level and threshold simulation were investigated; these were not shown to have a large effect. Changes in the noise simulation can also impact the number of hadronic hits; so far, this has not been able to explain the bulk of the differences.

A systematic error based on the observed differences between the data and the simulation has been taken. A simple shift in energy is not sufficient for this analysis. QePId (discussed in Section 7.8) makes QE determinations using information that is linked with the number of hadronic hits. For instance, it only considers slices with one or two tracks as potential QE-candidates; if the data displays too few hadronic hits, it is more likely to have slices with only one or two tracks. Therefore, simulation files were created which randomly removed 30% of reconstructed hits which, by truth, were not associated with a muon. There is no reason to believe this is actually what is physically happening in our detector, but it does show the effect of reducing the number of hits associated with a neutrino event independent of the underlying cause. By looking at

comparison plots of near detector data and simulation, one can see that this method does a good job reproducing the effects seen in our data.

First, we can look at comparisons for lower-level inputs to the analysis. Figure 14.25 plots the number of hits for each slice. This plot shows that reducing the number of non-muon hits gives better agreement in the region occupied by the majority of the data. Figure 14.26 shows the number of reconstructed 3D Kalman tracks for each slice. It shows a better agreement between data and simulation with hits removed for events with one reconstructed track. These are potentially QE events. It shows worse agreement for a larger number of tracks. Figure 14.27 shows the number of hits on the primary track; Figure 14.28 shows the track length distribution for the primary tracks. These show worse agreement for short tracks with a small number of hits and better agreement for longer tracks, more likely to be muons. The analysis level cuts, Section 10.4, have not been applied to the previous plots.

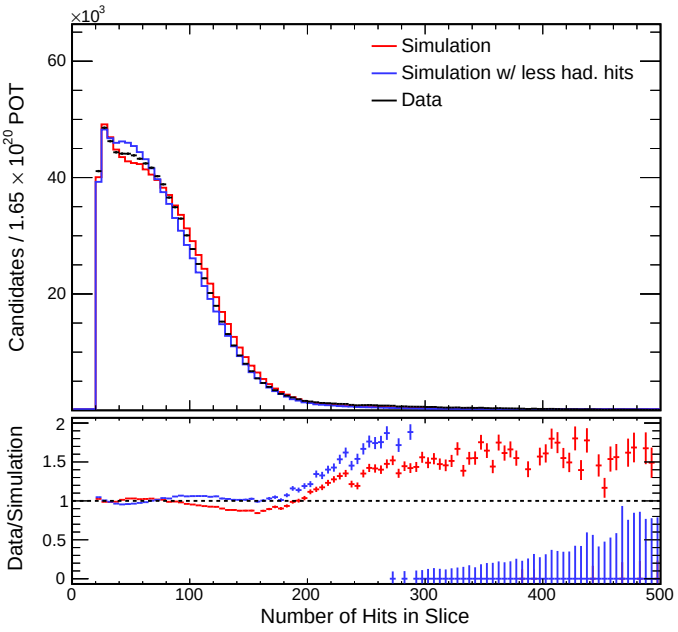


Figure 14.25: Plot of the number of hits in a slice for the near detector. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

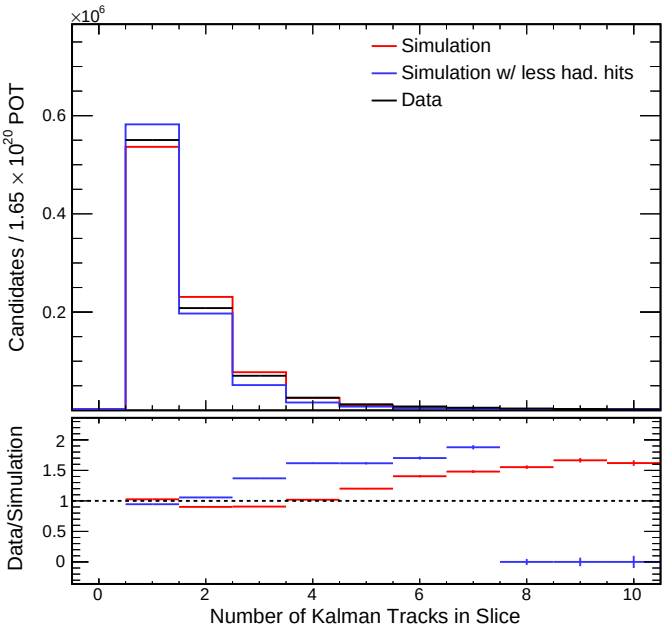


Figure 14.26: Plot of the number of 3D Kalman tracks in a slice for the near detector. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

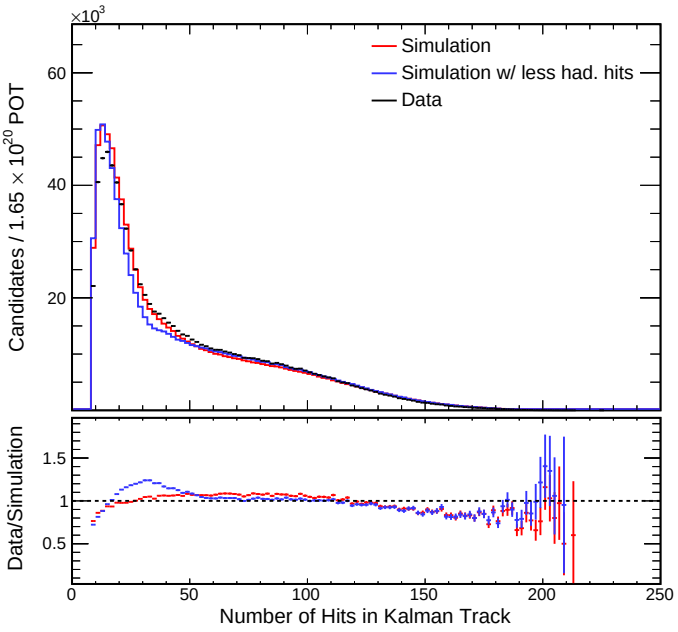


Figure 14.27: Plot of the number of hits on the 3D Kalman track with the highest ReMId value in a slice for the near detector. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

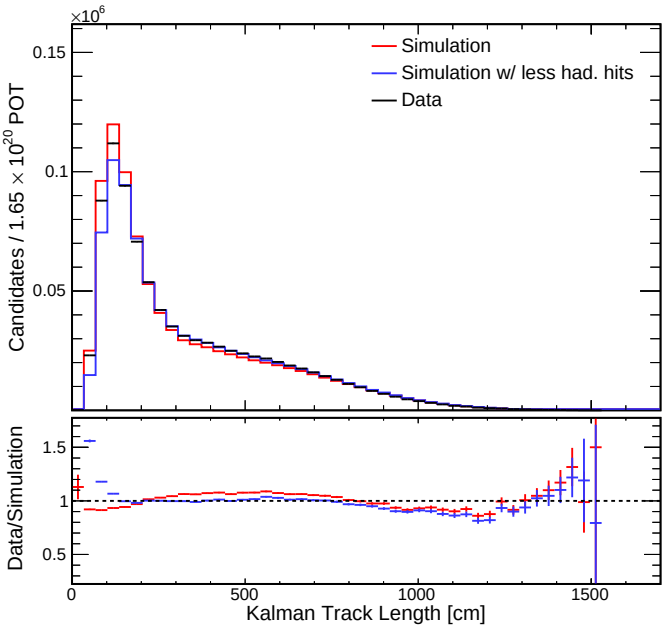


Figure 14.28: Plot of the length, in cm, of the 3D Kalman track with the highest ReMId value in a slice for the near detector. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

Next, in building up to the final selection cuts for the analysis, we can consider the inputs to ReMId. Again, the ReMId cut has not been applied. Each plot has one entry per slice, corresponding to the metric for the 3D Kalman track with the highest ReMId value in the slice. Figure 14.28, the track length of the 3D Kalman track with the highest ReMId value, already shown, is one input. Figures 14.29, 14.30 and 14.31 display the other three inputs to ReMId. Note that ReMId was only trained with the original simulation and not with the simulation with fewer hadronic hits. Finally, Figure 14.32 shows the ReMId distribution. Significant disagreement is seen between the data and the altered simulation for low values of ReMId. These are the least muon-like and are composed mostly of hadrons. This region is most sensitive to the details of the hadronic simulation and disagreement in this region isn't surprising. At cut at 0.7 is made on this distribution to select slices with a reconstructed muon for our analysis. The rest of the plots will have this cut applied.

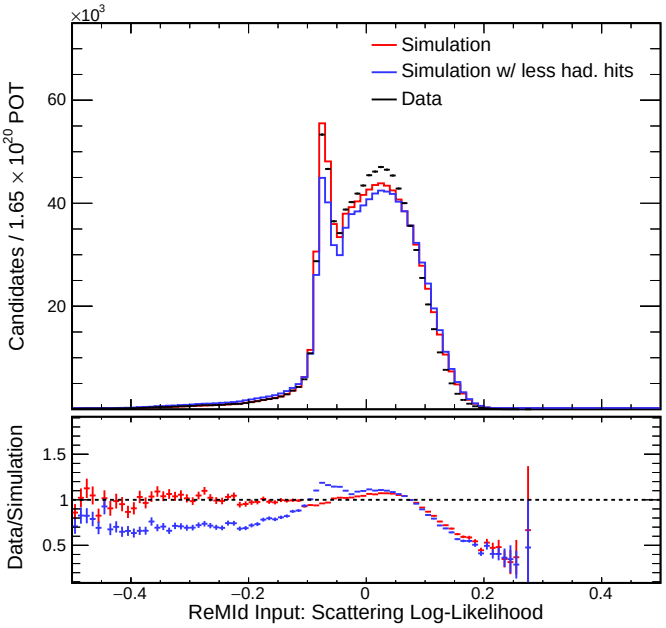


Figure 14.29: Plot of the scattering log-likelihood for the the 3D Kalman track with the highest ReMId value in a slice. This plot is for the near detector populations. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

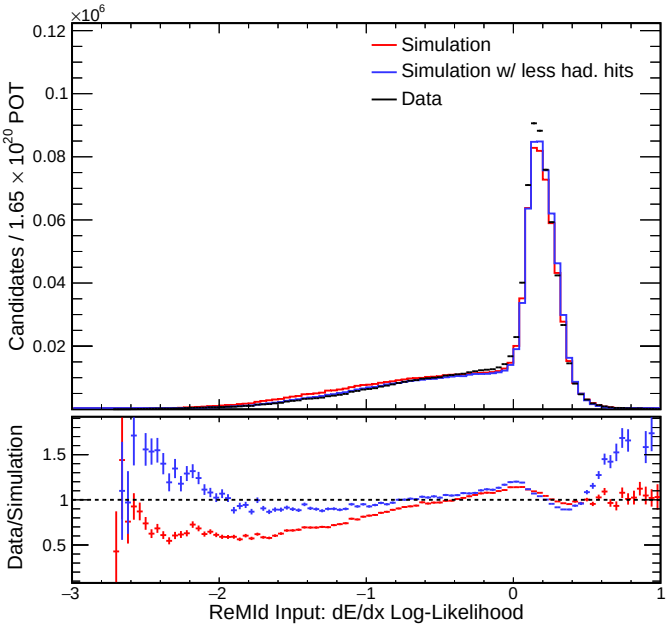


Figure 14.30: Plot of the dE/dx log-likelihood for the the 3D Kalman track with the highest ReMId value in a slice. This plot is for the near detector populations. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

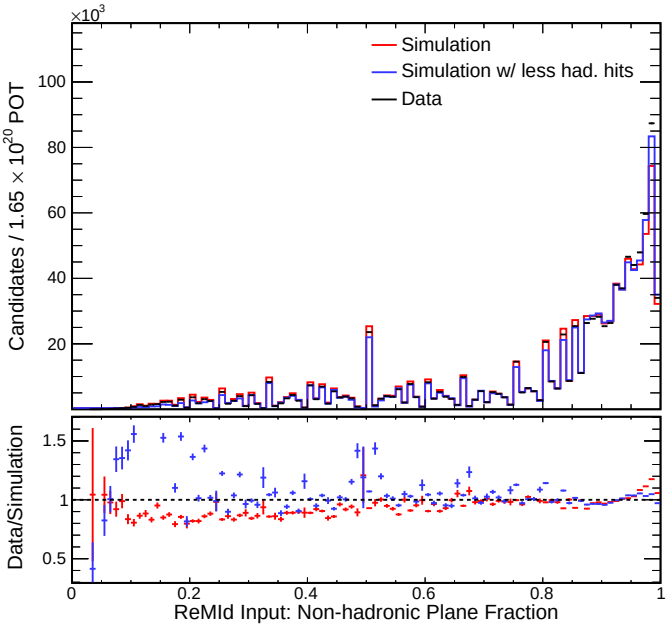


Figure 14.31: Plot of the non-hadronic plane fraction for the the 3D Kalman track with the highest ReMId value in a slice. This plot is for the near detector populations. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

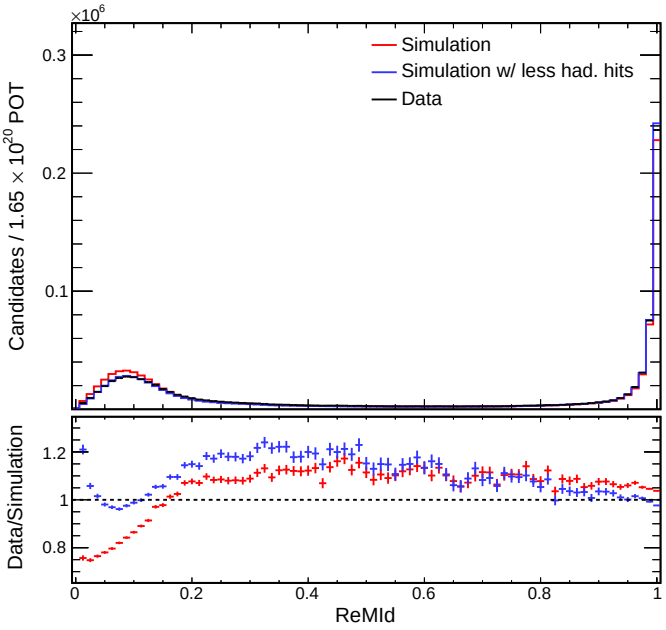


Figure 14.32: Plot of the ReMId value for the the 3D Kalman track with the highest ReMId value in a slice. This plot is for the near detector populations. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. The cuts listed in Section 10.4, notably a cut to select slices with muons, are not applied.

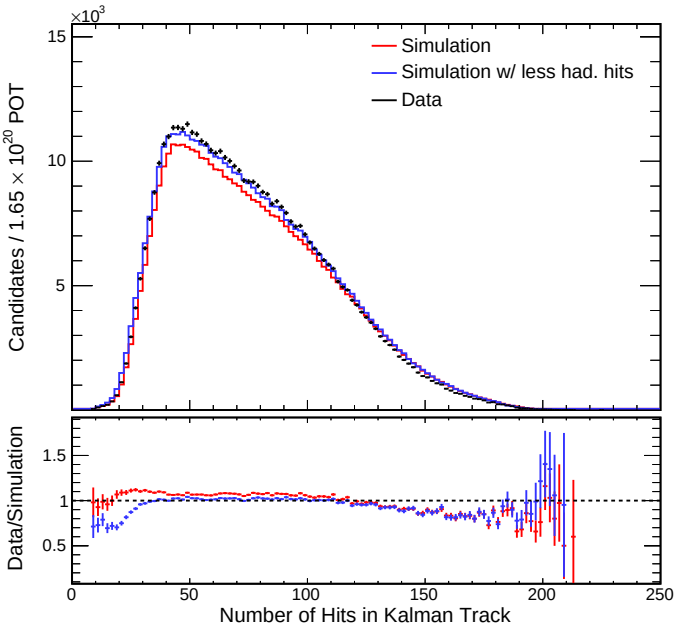


Figure 14.33: Plot of the number of hits on the 3D Kalman track with the highest ReMId value in a slice for the near detector. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

Having made a ReMId cut, we now have a sample which is primarily composed of contained ν_μ CC interactions. We can now look again at the properties of the primary tracks in this sample and inspect the hadronic component of the slice. The samples have not been divided into QE and non-QE populations. Each plot has one entry per slice. When track properties are plotted, only the 3D Kalman track with the highest ReMId value, defined as the reconstructed muon track, is used.

Figure 14.33 shows the number of hits on the primary track; Figure 14.34 shows the track length distribution for the primary tracks. The altered simulation does a surprisingly good job of matching the data. It is possible that with less hadronic activity, a slightly different mix of event topologies passes containment and analysis cuts.

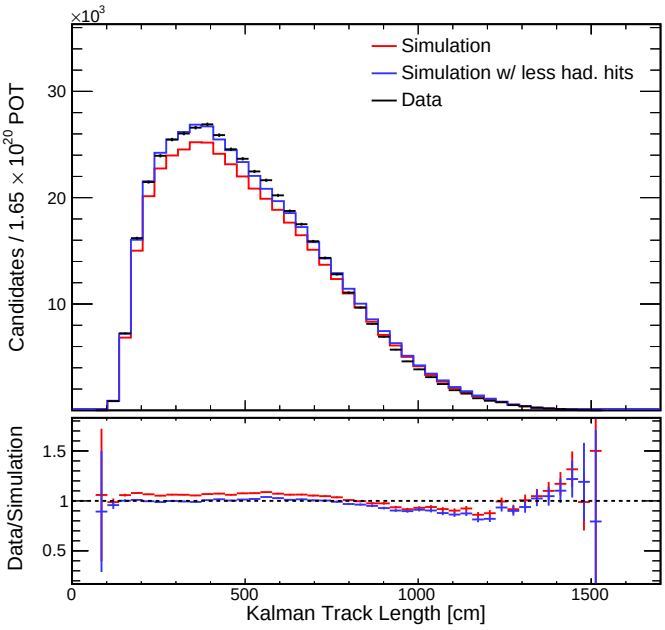


Figure 14.34: Plot of the length, in cm, of the 3D Kalman track with the highest ReMId value in a slice for the near detector. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

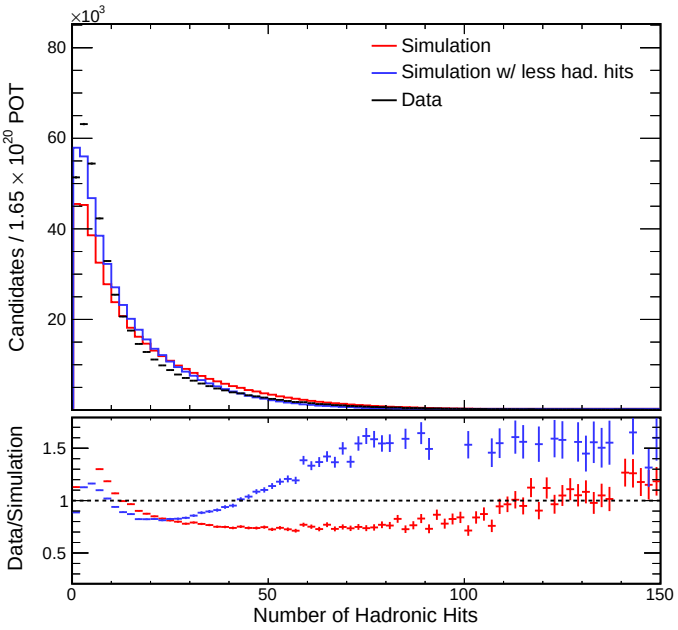


Figure 14.35: Plot of the number of hits in the slice not on the 3D Kalman track with the highest ReMId value for the near detector. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

Next, we can look at the properties of the hadronic part of the slice. Figure 14.35 plots the number of hits in the slice not on the primary track. These hits are defined to be the hadronic hits. This plot demonstrates that by removing hadronic hits in our altered simulation, we can make a distribution that is closer to that of data for the region with the most candidates.

We can look at the visible calorimetric energy of the hadronic sector. Figure 14.36 plots the visible energy, in GeV, for the hadronic hits. Figure 14.37 plots the visible energy, in GeV, associated with hadronic energy in the vertex region of the muon track. These hadronic energies have not had the fits, discussed in Chapter 9, applied. Again, as expected, by removing hadronic hits, the altered simulation is able to mimic the data better than the nominal simulation.

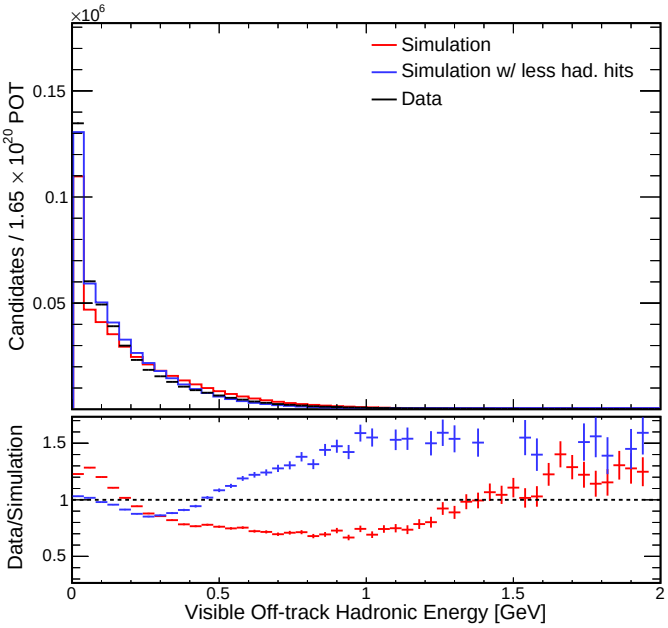


Figure 14.36: Plot of the sum of the visible energy (in GeV) of hits in the slice not on the 3D Kalman track with the highest ReMId value. This plot is for the near detector populations. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

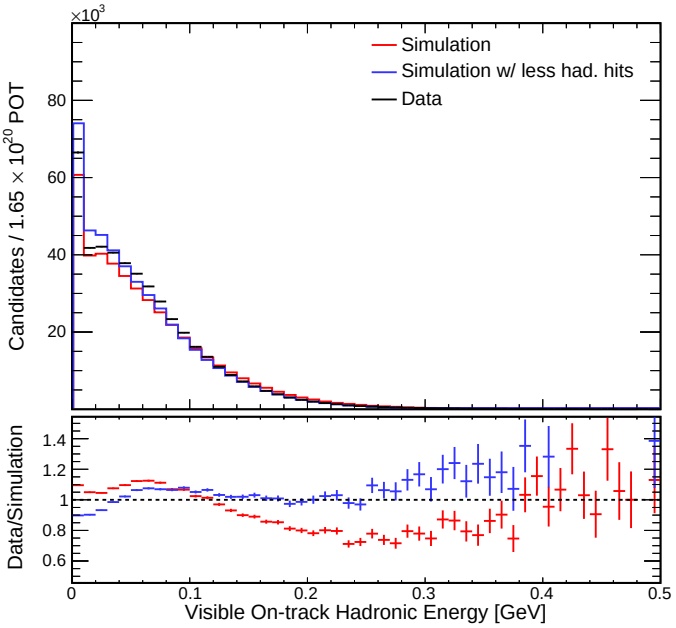


Figure 14.37: Plot of the sum of the visible energy (in GeV) associated with hadronic energy in the vertex region of the 3D Kalman track with the highest ReMId value. This plot is for the near detector populations. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

Average energy per hit is a useful quantity to understand the relationship between total visible energy and number of hits. Figure 14.38 displays, for each slice, the total visible energy in GeV divided by the total number of hits in the slice. Each slice has one entry in the plot. Figure 14.39 plots the total visible energy in GeV for hits on the muon track divided by the total number of hits associated with the muon track. Again, each slice has one entry in the plot. It is possible that the improvement seen from the altered simulation is due to less hadronic hit contamination near the vertex of the muon track. The hadronic energy sector has more structure; Figure 14.40 plots the total visible energy in GeV for hits in the slice not associated with the muon track, divided by the total number of hadronic hits. It is not fully understood why the altered sample shifts to lower energy per hit. It is possible that this is because any muon hits in the hadronic cluster wouldn't have a chance to be randomly removed. This could pull the distribution slightly lower. Alternatively, by removing some hadronic hits, different event topologies could be selected for the analysis sample at higher rates, changing this distribution.

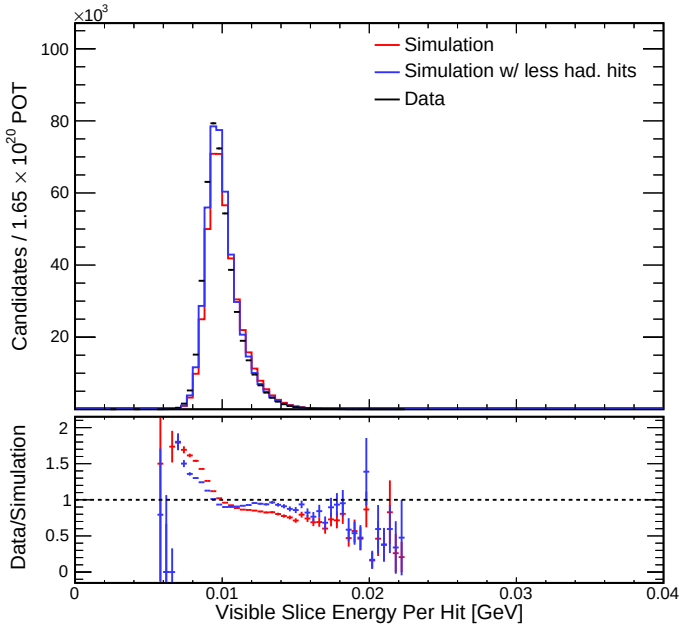


Figure 14.38: Plot of the sum of the visible energy (in GeV) for the slice hits divided by the total number of hits in the slice. Each slice is one entry in the histogram. This plot is for the near detector populations. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

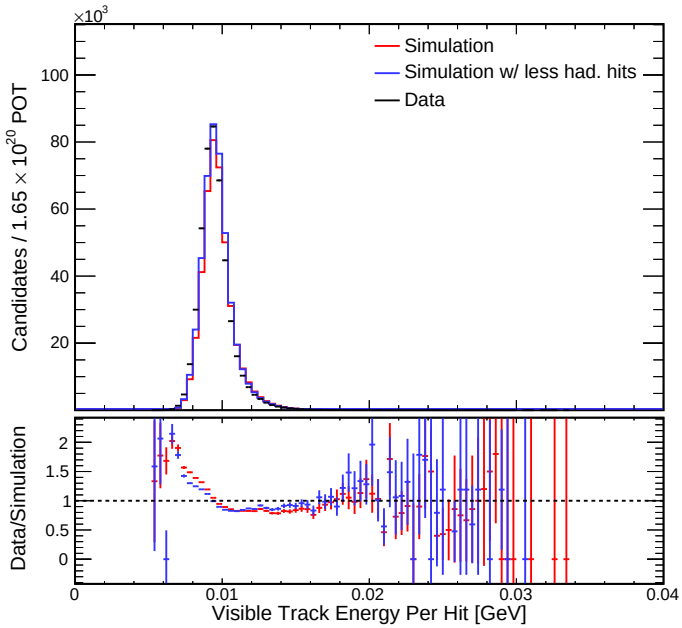


Figure 14.39: Plot of the sum of the visible energy (in GeV) for the hits associated the 3D Kalman track with the highest ReMId value divided by the number of hits associated with the primary track. Each primary track is one entry in the histogram. This plot is for the near detector populations. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

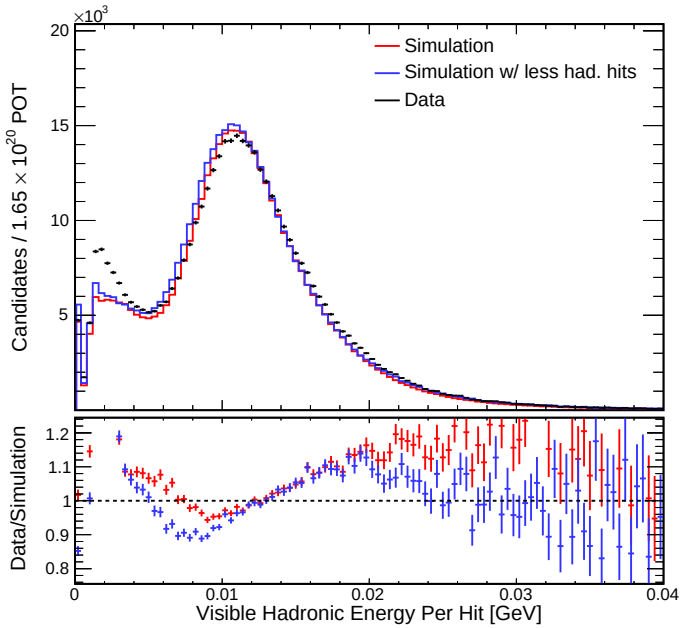


Figure 14.40: Plot of the sum of the visible energy (in GeV) for the slice hits not associated the 3D Kalman track with the highest ReMId value divided by the number of hits in the slice not associated with the primary track. Each slice is one entry in the histogram. This plot is for the near detector populations. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

To make our final analysis sample, we need to apply a QePId cut to create a QE and a nonQE population. First, let us understand the inputs to QePId. Each plot has one entry per slice. Plots are either for the one track or two track sample. For more information on these samples or the inputs to QePId, see Section 7.8.

Figure 14.41, the off-track energy ratio for the one track sample, is one input to QePId. This plot shows remarkable agreement between the altered simulation and data. Figure 14.42 displays the off-track energy ratio, but for the two track sample. In general, the altered simulation does less well capturing all the structure present in the two track sample, compared to the more simple one track sample. Figure 14.43 shows the fractional energy difference for the one track sample and Figure 14.44 is for the two track sample. The fractional energy difference Z-test for the one track sample is plotted in Figure 14.45; the two track sample is shown in Figure 14.46. Figure 14.47 plots the dE/dx ratio, only applicable to the two track sample. For this input to QePId, the altered simulation clearly fails to do better than the original simulation. Randomly removing hits in the hadronic sector is going to have a negative impact on one's ability to make hadronic tracks with expected dE/dx depositions. The approach, in general, works much better for the one track sample. Something more sophisticated would need to be developed if one wanted to properly model effects in samples with reconstructed hadronic tracks.

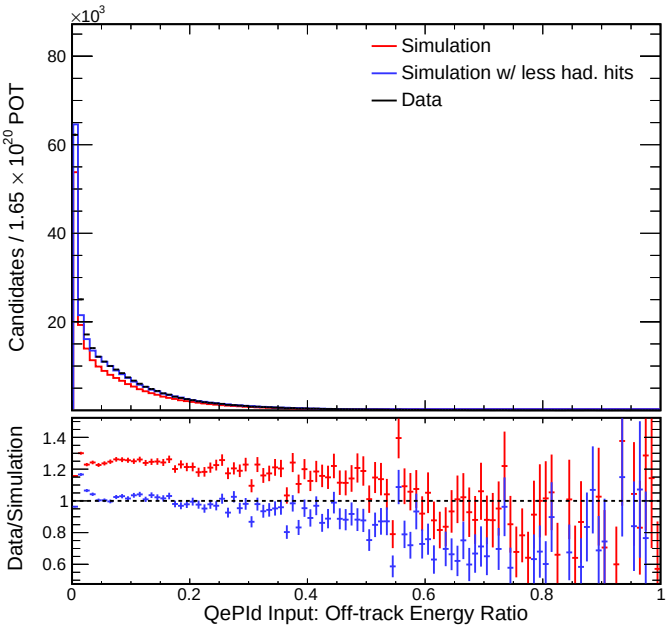


Figure 14.41: Plot of off-track energy ratio for a slice in the near detector. This is for the one track sample. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

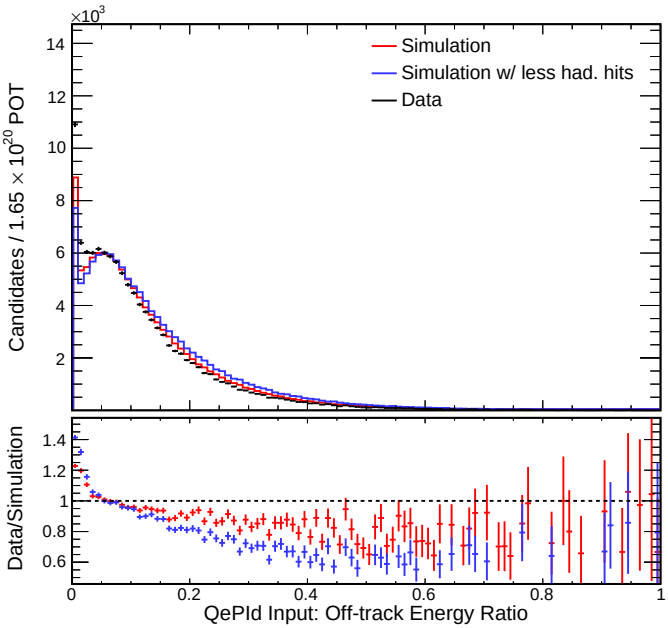


Figure 14.42: Plot of off-track energy ratio for a slice in the near detector. This is for the two track sample. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

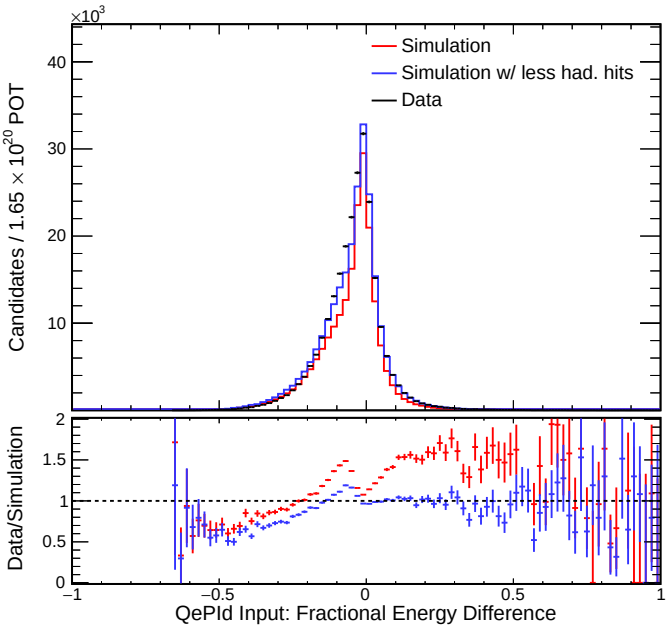


Figure 14.43: Plot of the fractional energy difference for a slice in the near detector. This is for the one track sample. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

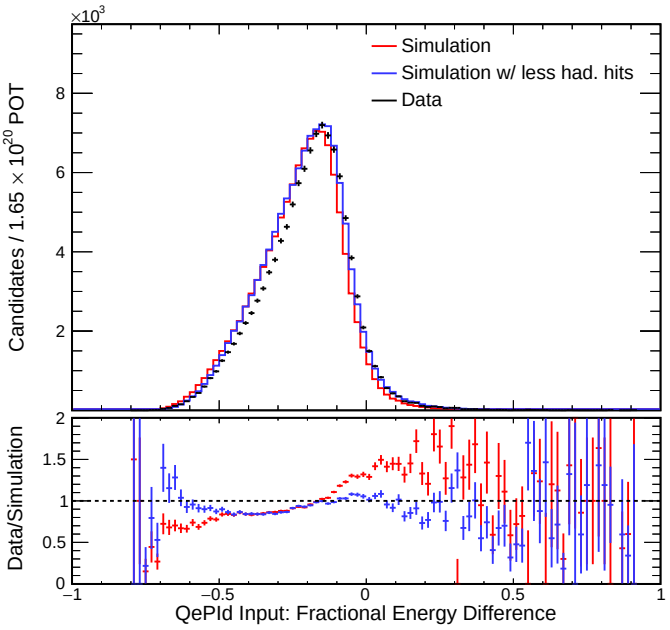


Figure 14.44: Plot of the fractional energy difference for a slice in the near detector. This is for the two track sample. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

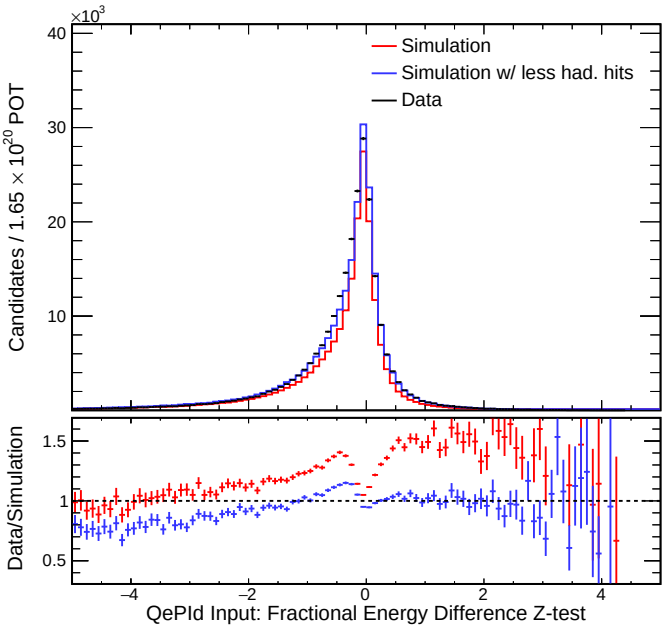


Figure 14.45: Plot of the fractional energy difference Z-test for a slice in the near detector. This is for the one track sample. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

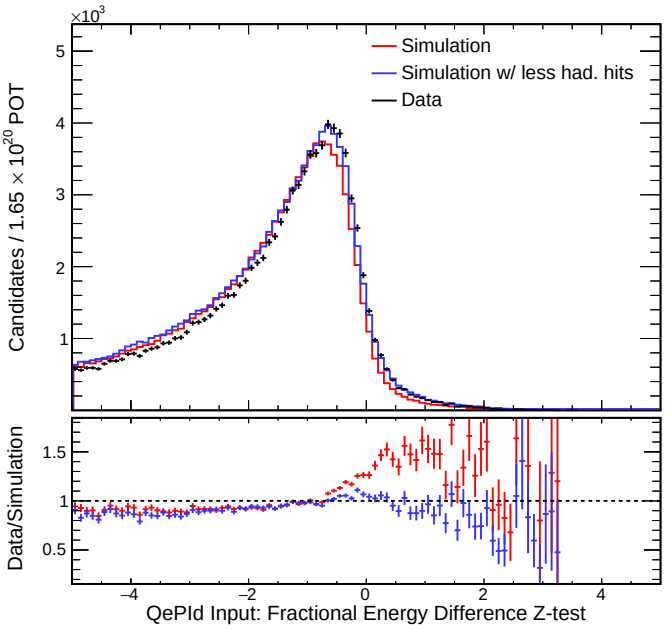


Figure 14.46: Plot of the fractional energy difference Z-test for a slice in the near detector. This is for the two track sample. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

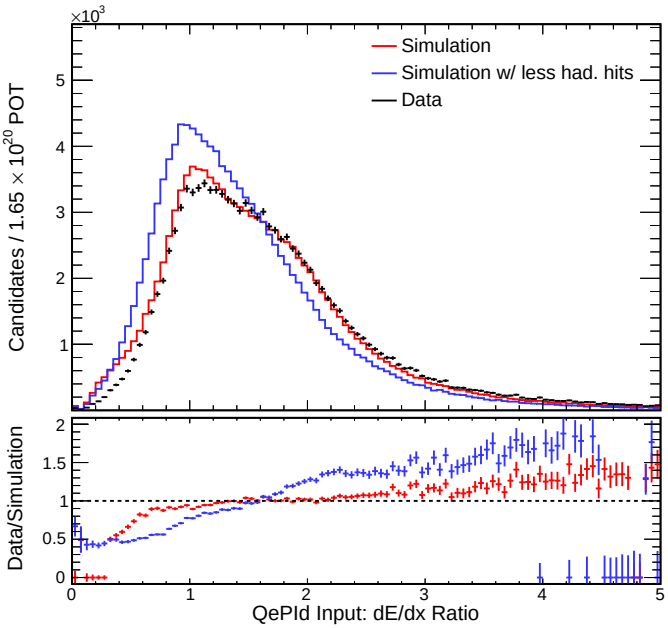


Figure 14.47: Plot of the dE/dx ratio for a slice in the near detector. This is for the two track sample. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

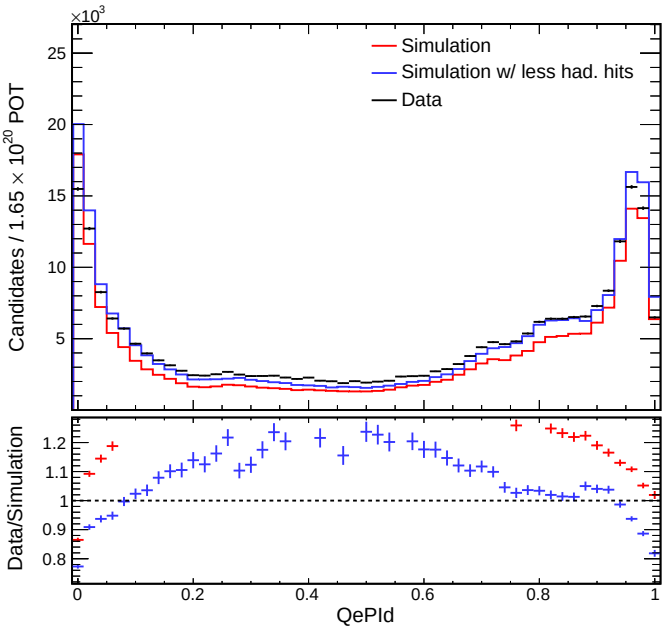


Figure 14.48: Plot of QePId for a slice in the near detector. This is for the one track sample. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. When the ratio is too large for the scale, the point and its error bars are not drawn. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

Finally, Figure 14.48 shows the QePId distribution for the one track sample and Figure 14.49 shows the QePId distribution for the two track sample. For the one track sample, a cut is made at 0.45. Slices with higher values are considered part of the QE sample; slices with lower values are part of the nonQE sample. For the two track sample, the cut is made at 0.4. The altered simulation does a much better job modeling the QePId distribution for the one track sample than the nominal simulation. Both simulations fail to capture the structure seen in the data for the two track sample; however, this sample contributes fewer events to the QE sample than the one track sample and is relatively less important.

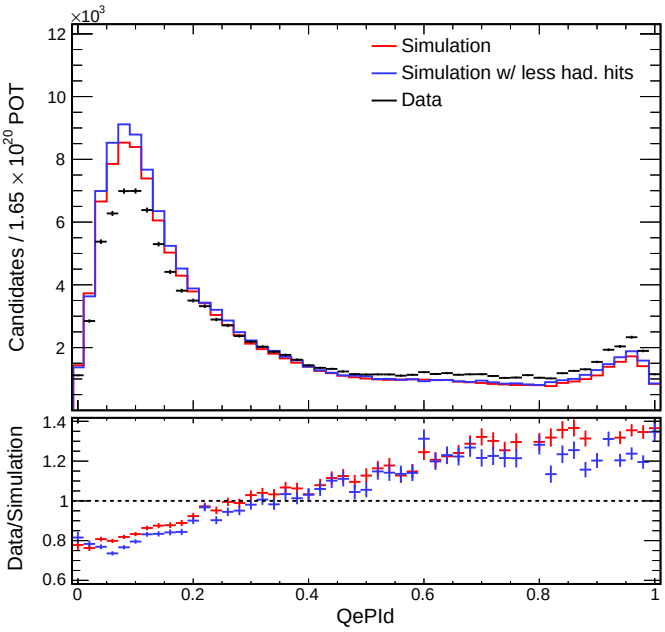


Figure 14.49: Plot of QePId for a slice in the near detector. This is for the two track sample. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2 and 10.3.2 are applied. A cut requiring a ReMId value ≥ 0.7 is also applied.

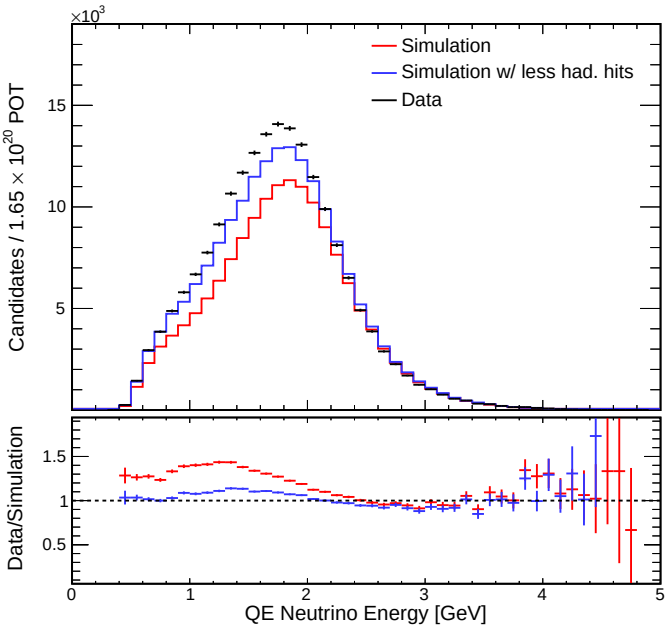


Figure 14.50: Plot of reconstructed neutrino energy in GeV for a slice in the near detector. This is for the QE sample. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2, 10.3.2 and 10.4 are applied.

Having made our final analysis populations, we can now look at the reconstructed neutrino energy. Figure 14.50 displays the reconstructed neutrino energy for the QE sample. Figure 14.51 displays the reconstructed neutrino energy for the nonQE sample. We see that the altered simulation improves the agreements with the data but doesn't perfectly model the relative normalizations or the energy shifts seen in the data. It does capture the behavior well-enough to be used to estimate the systematic error.

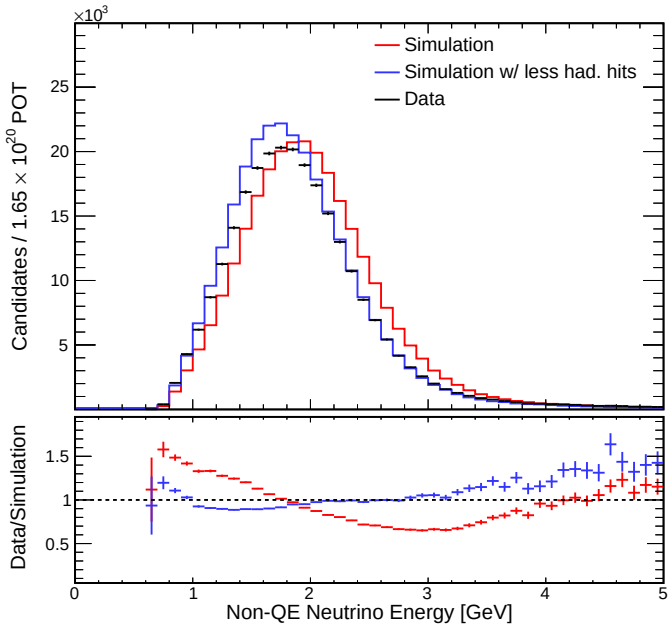


Figure 14.51: Plot of reconstructed neutrino energy in GeV for a slice in the near detector. This is for the nonQE sample. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2, 10.3.2 and 10.4 are applied.

For the curious, we can look at the muon track properties of our final analysis populations. Remember that with a ReMId cut, the muon track length and number of hits matched incredibly well between the altered simulation and the data. However, once we divide our sample into the QE and nonQE populations, we no longer see this striking agreement. Figure 14.52 shows the number of hits on the muon track for the QE population. Figure 14.53 is for the nonQE population. Figure 14.54 displays the muon track length for the QE population. Figure 14.55 shows the nonQE population. The altered simulation still does a better job matching the data than the nominal simulation for the QE population, but the data and altered simulation do not lie on top of each other. This isn't completely unexpected - by looking at our QePId distributions, we already knew that the altered simulation doesn't allow us to make populations that match the data perfectly. These differences carry over into the muon track distributions; for instance, since the data still sees more QE events than the altered simulation, we see more muon tracks in the data sample than in altered simulation.

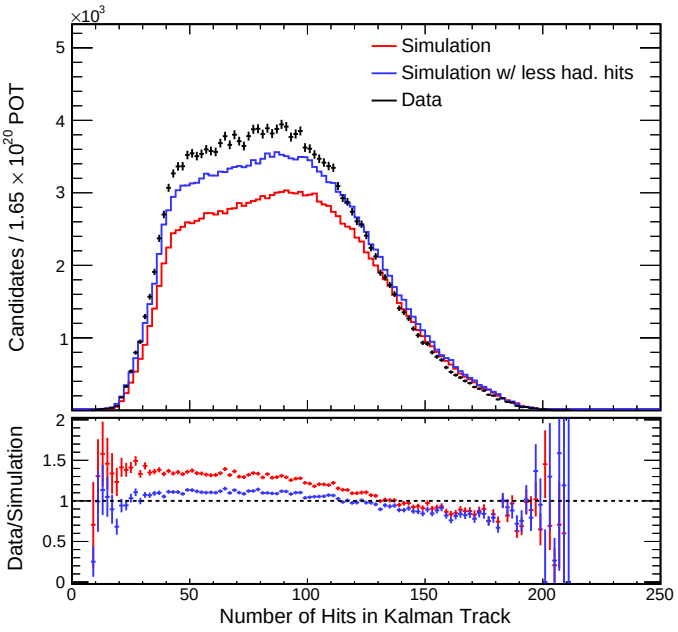


Figure 14.52: Plot of the number of hits on the 3D Kalman track with the highest ReMId value in a slice for the near detector. This is for the QE sample. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2, 10.3.2 and 10.4 are applied.

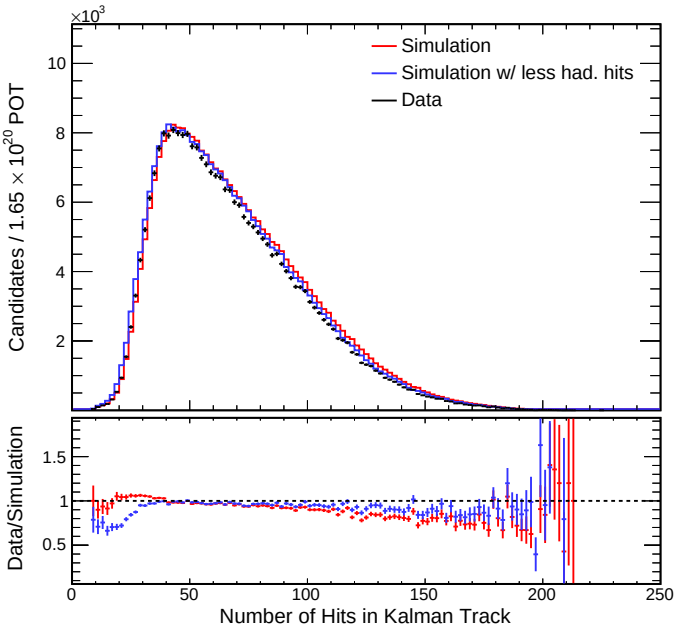


Figure 14.53: Plot of the number of hits on the 3D Kalman track with the highest ReMId value in a slice for the near detector. This is for the nonQE sample. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2, 10.3.2 and 10.4 are applied.

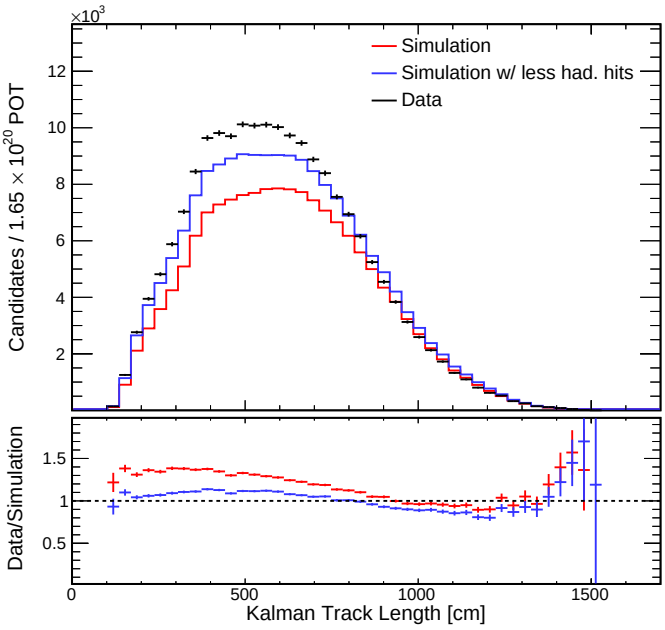


Figure 14.54: Plot of the length, in cm, of the 3D Kalman track with the highest ReMId value in a slice for the near detector. This is for the QE sample. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2, 10.3.2 and 10.4 are applied.

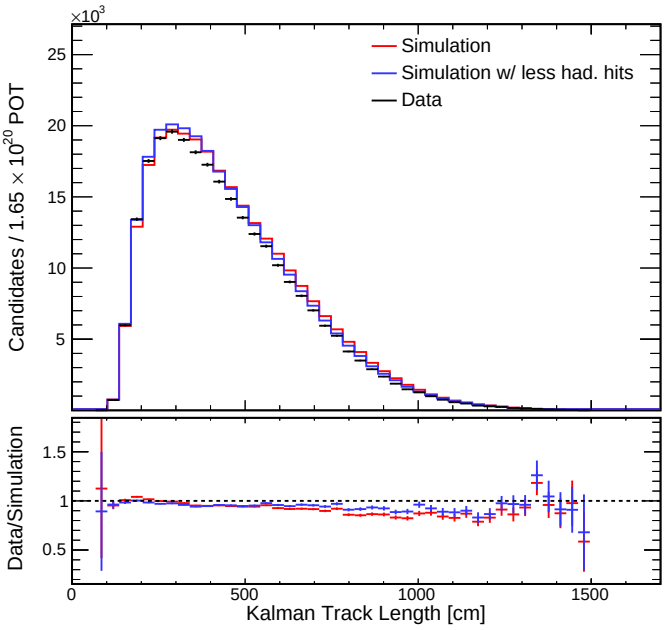


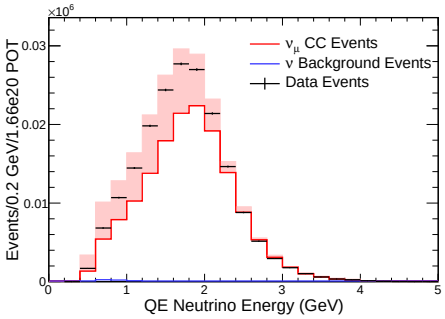
Figure 14.55: Plot of the length, in cm, of the 3D Kalman track with the highest ReMId value in a slice for the near detector. This is for the nonQE sample. The nominal simulation distribution is displayed as a red line. The altered simulation with fewer hadronic hits is drawn as a blue line. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is scaled down by a factor of 6 to 1.65×10^{20} POT, the exposure for the near detector data. All cuts listed in Sections 10.1, 10.2, 10.3.2 and 10.4 are applied.

Having created simulation files which remove hadronic hits for both detectors, we looked at the reconstructed neutrino energy spectra for each population (near and far detectors, QE and nonQE samples) to create $+1\ \sigma$ errors. The peaks of reconstructed neutrino energy for each sample were fit with a Gaussian. A scale factor applied to the reconstructed hadronic energy was found that caused the peaks in reconstructed neutrino energy to align. After this shift is made, a correlated shape correction was made bin-by-bin in reconstructed neutrino energy. The systematic errors are one-sided and $+1\ \sigma$ was defined to be twice the size of the shift seen in the altered simulation files. The shift size was doubled because the shift didn't perfectly encompass all the data and simulation differences. Twice was picked because it conservatively covered the discrepancy.

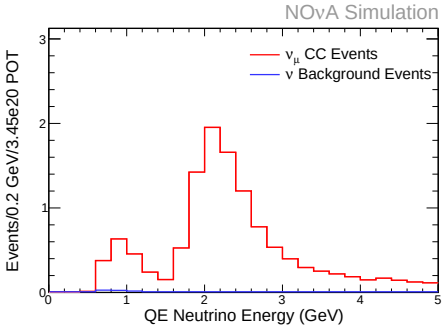
The systematic errors were applied in such a way that the value of the error for each detector is allowed to move independently of the other detector. The extrapolation procedure will not be able to mitigate the effect. Figure 14.56 displays the systematic error from the near detector number of hadronic hits discrepancy for the QE population in each detector. The plot for the far detector without extrapolation doesn't have an error band, because without extrapolation, the systematic error from the near detector number of hadronic hits discrepancy doesn't impact the far detector spectrum. Here, the far detector prediction with extrapolation is actually lower for the QE population due to this systematic error. That is because it is able to undo the effects of the extrapolation, which already increased the expected number of QE events. Figure 14.57 is for the nonQE population. Again, the plot for the far detector without extrapolation doesn't have an error band, because without extrapolation, the systematic error from the near detector number of hadronic hits discrepancy doesn't impact the far detector spectrum. The effect of the systematic error is slightly counter-intuitive; it is again undoing the effects of the extrapolation on this sample. Figure 14.58 displays the systematic error from the far detector number of hadronic hits discrepancy for the QE population in each detector. Since this is the systematic error from the far detector number of hadronic hits discrepancy, the near detector spectrum doesn't have an error band. The shift goes in the expected direction and increases the total predicted number of QE events in the far detector, as well as shifts the reconstructed neutrino energy lower. Figure 14.59 is for the nonQE population. Again, since this is the systematic error from the far detector

number of hadronic hits discrepancy, the near detector spectrum doesn't have an error band. It works as expected in the far detector spectrum, reducing the total number of nonQE events predicted in the far detector and again shifting the reconstructed neutrino energy to lower values.

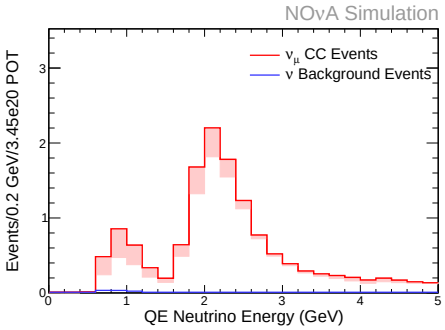
This systematic error, along with the uncertainty in the absolute calorimetric energy scale discussed in Section 14.11.3, has the largest impact on this analysis. It changes the overall normalization of each population and reduces the precision with which we can measure the parameters from the oscillation dip. This systematic error must be reduced for future analyses by understanding the source of the data and simulation differences in the near detector.



(a) Near Detector

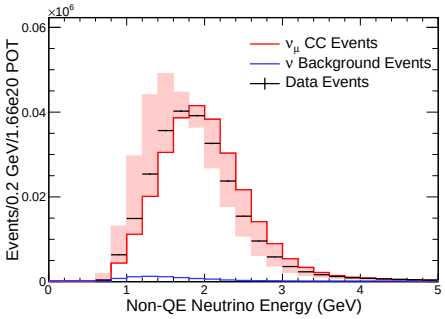


(b) Far Detector without Extrapolation

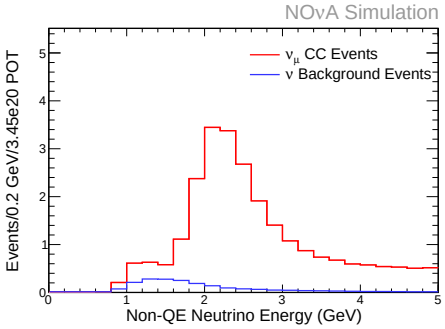


(c) Far Detector with Extrapolation

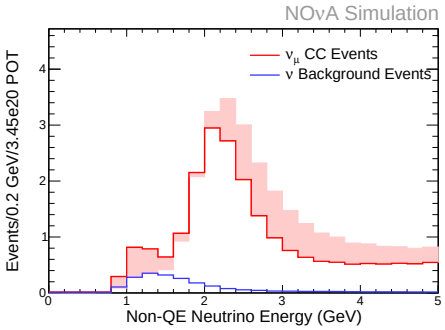
Figure 14.56: Plot of the reconstructed neutrino energy in GeV with the near detector number of hadronic hits systematic error band for the QE sample. Figure 14.56b doesn't have an error band because the near detector systematic error doesn't affect the far detector without extrapolation. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.



(a) Near Detector

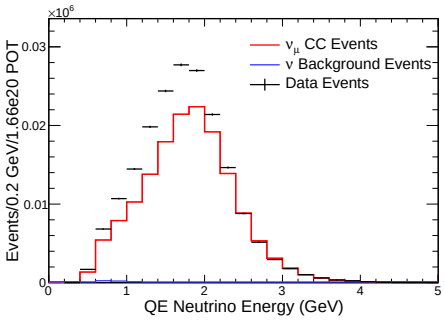


(b) Far Detector without Extrapolation

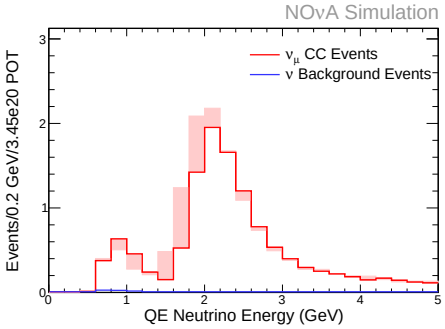


(c) Far Detector with Extrapolation

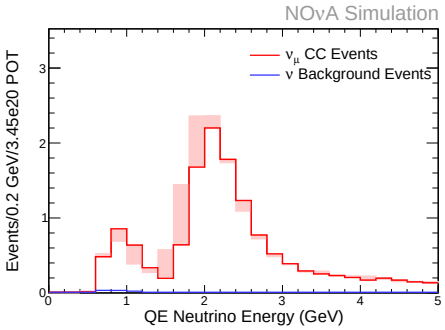
Figure 14.57: Plot of the reconstructed neutrino energy in GeV with the near detector number of hadronic hits systematic error band for the nonQE sample. Figure 14.57b doesn't have an error band because the near detector systematic error doesn't affect the far detector without extrapolation. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.



(a) Near Detector

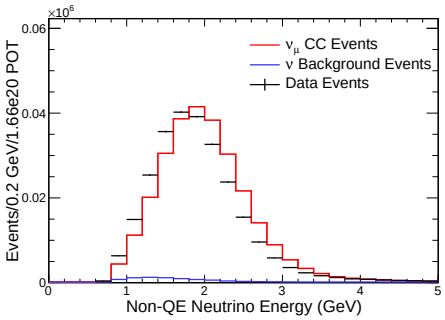


(b) Far Detector without Extrapolation

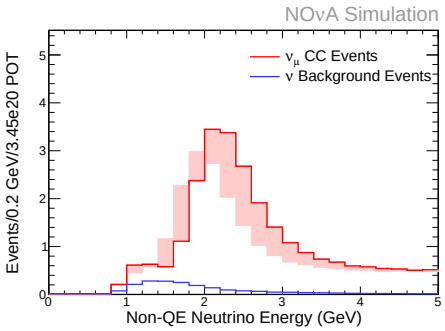


(c) Far Detector with Extrapolation

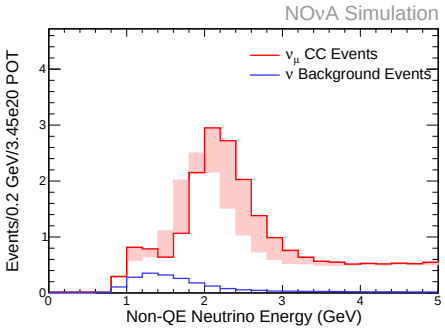
Figure 14.58: Plot of the reconstructed neutrino energy in GeV with the far detector number of hadronic hits systematic error band for the QE sample. Figure 14.58a doesn't have an error band because the far detector systematic error doesn't affect the near detector. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.



(a) Near Detector



(b) Far Detector without Extrapolation



(c) Far Detector with Extrapolation

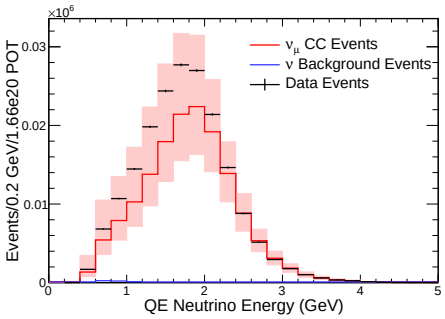
Figure 14.59: Plot of the reconstructed neutrino energy in GeV with the far detector number of hadronic hits systematic error band for the nonQE sample. Figure 14.59a doesn't have an error band because the far detector systematic error doesn't affect the near detector. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.

14.13 Total Systematic Error

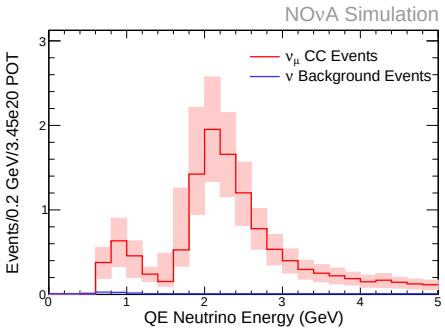
Having considered many systematic errors, some of which this analysis uses and some that are negligible, we can now look at the total systematic error being taken. The two systematic errors that have the largest impact are the absolute calorimetric energy scale error (Section 14.11.3) and the number of hadronic hits error (Section 14.12). Both of these errors are relatively large in each detector. Also, they were treated as uncorrelated between detectors. For the hadronic hits error, this was a conservative choice. Future analyses should be able to reduce this error and its impact. Since the errors were implemented as uncorrelated, the process of using the near detector data to altered the far detector prediction couldn't mitigate the impact of these systematic errors.

Most of the other systematic errors used in this analysis, however, are correlated between detectors. This means that the extrapolation process does mostly remove the effect of the uncertainty from the analysis. The systematic errors from the beam simulation (Section 14.1) are large in each detector, but fairly small when extrapolated. Systematic errors from the neutrino interaction simulation (Section 14.2) are smaller in each detector than the beam simulation errors and even smaller when extrapolated. Normalization systematic errors (Section 14.5) are very tiny and have almost no impact on the total systematic error envelope.

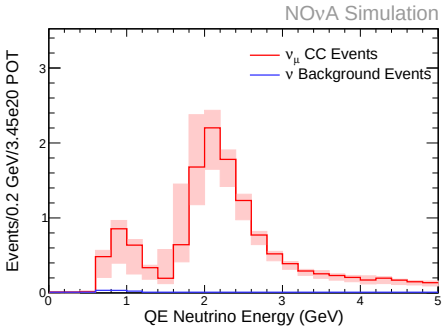
Figure 14.60 displays the total systematic error for the QE population in each detector. Figure 14.61 is for the nonQE population.



(a) Near Detector

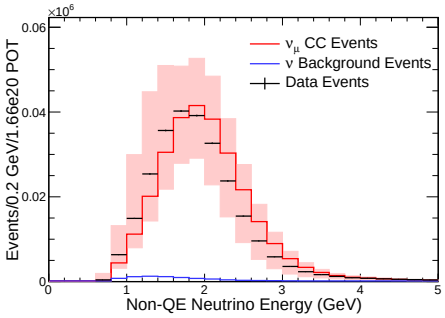


(b) Far Detector without Extrapolation

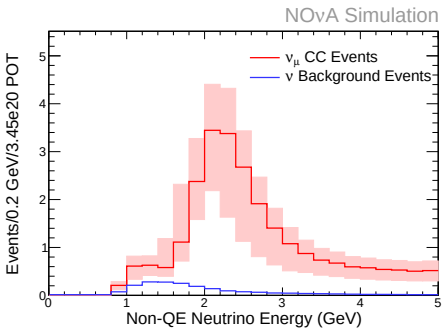


(c) Far Detector with Extrapolation

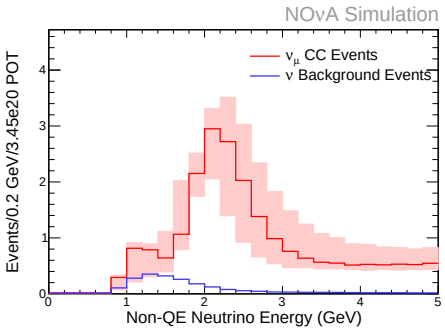
Figure 14.60: Plot of the reconstructed neutrino energy in GeV with the total systematic error band for the QE sample. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.



(a) Near Detector



(b) Far Detector without Extrapolation



(c) Far Detector with Extrapolation

Figure 14.61: Plot of the reconstructed neutrino energy in GeV with the the total systematic error band for the nonQE sample. The simulation distribution is drawn as a red line with red systematic error bands, with neutrino background drawn as blue line. The near detector data is drawn as black points with statistical error bars. The simulation is scaled down to match the data POT. For the near detector, this is 1.66×10^{20} POT. For the far detector, it is 3.45×10^{20} POT. The far detector simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2.

Chapter 15

Results

Having defined a set of cuts to create analysis populations, determined an extrapolation procedure and finalized a set of systematic errors, we can now look at the far detector data. This data corresponds to 3.45×10^{20} POT with variable detector configurations.

If no neutrino oscillations occurred, we would expect roughly 100 events in each analysis population. However, in the case of maximal mixing, we would instead expect about 14 QE events and 24 nonQE events. Table 15.1 lists the event counts for each population in the data as well as predicted event counts from simulation under different hypotheses. To be counted, the event must have a reconstructed neutrino energy between 0 and 5 GeV. This analysis can easily determine if neutrino oscillations occur; since we see 20 QE events and 25 nonQE events in the data, we can rule out the hypothesis of no neutrino oscillations.

The table also lists the event counts predicted for maximal mixing when one varies the systematic errors by 1σ. Systematic errors cause us to expect between 10 and 18 events for the QE population and 13 and 35 events for the nonQE population. Finally, the table also lists the expected background event counts from neutrino interactions that are not ν_μ CC interactions and cosmic ray backgrounds. These are relatively small; for the QE population, we expect about 1 background event and for the nonQE population, we expect about 5 background events.

Table 15.1: Predicted and measured event counts for the far detector. The total POT for the far detector is 3.45×10^{20} POT and this corresponds to variable detector configurations. The counts for each population only include events with reconstructed neutrino energy between 0 and 5 GeV. The total number of data events for each population is listed first. The predicted total event count for each population under the hypothesis of no neutrino oscillations is given next. The rest of the lines correspond to the hypothesis of maximal mixing, setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2 as well as using the oscillation parameters listed in Table 4.2. The event counts predicted by varying the systematic errors by 1σ are listed for the hypothesis of maximal mixing. The expected background event counts are given for each population, with neutrino interactions that are not ν_μ CC interactions and cosmic ray backgrounds separated.

	QE	NonQE
Data	**20**	**25**
Total Events (no osc.)	89.2	122.3
Total Events (max. mix)	14.3	24.2
Total Events w/ syst. 1σ errors (low)	10.3	13.6
Total Events w/ syst. 1σ errors (high)	18.2	34.8
Neutrino Background	0.2	2.1
Cosmic Ray Background	0.7	2.7

Section 15.1 looks at the far detector data events in detail. Then Section 15.2 presents the measurement of the oscillation parameters and compares it to the results of other experiments.

15.1 Far Detector Data Events

We can now look at the far detector data events in some detail. We measured 20 data events for the QE population and 25 for the nonQE population. The event displays for each data event were visually scanned. 2 of the QE events visually looked like possible cosmic ray background events. Both clearly entered or exited from the edge of the detector, but due to a reconstruction error, passed the analysis cuts. This type of failure could be removed from future analyses by adding cuts that looked at slice extents. The QE population was predicted to have 0.7 events from cosmic ray backgrounds. The

probability of measuring 2 or more events when expecting 0.7 was 15%. Table 15.2 lists each selected QE far detector data event.

One QE data event that was visually very signal-like corresponds to run 17953, event 256887. Figure 15.1 is an event display of this event. It shows the entire far detector. Figure 15.2 is a spatially-zoomed version of the event display, showing the slice hits. Figure 15.3 is also spatially-zoomed, but has the Kalman tracks drawn on top of the hits. The long green 3D Kalman track is the muon track; the short blue 2D track is probably a proton. This is a classic example of what a QE event looks like. Note that the muon track does have some gaps. These can be caused by faulty electronics or photo-detector or low hit efficiency when far from the APD readout. Gaps on muon tracks don't hurt our ability to reconstruct the energy of the muon, as long as the reconstruction is still able to reconstruct the entire track.

For an example of a selected event that could be a cosmic-ray background entering the detector, see Figure 15.4 of run 18302, event 413485. It shows the entire far detector. An alternative hypothesis is that this could be a muon from a neutrino interaction which exited the detector. Figure 15.5 is a spatially-zoomed version of the event display, showing the slice hits. Figure 15.6 is also spatially-zoomed, but has the Kalman tracks drawn on top of the hits. Note that, although the slice has hits which are very near to the top of the detector, the curve in the track is not followed by the reconstructed Kalman track. This causes the end of the track to be considered contained and passes the analysis cuts.

Table 15.2: Far detector QE data events. The time is relative to the start of the readout window. The ReMId value listed is for the 3D Kalman track in the slice with the highest ReMId. The energy is reconstructed neutrino energy in GeV. These events were visually scanned and categorized. "S" indicates the event looked signal-like. "I" indicates the event was indeterminate and could be signal or background. "B" indicates the event visually looked like cosmic ray background. "C" indicates that the background was well-contained in the detector. "E" indicates that the event was clearly entering the detector.

Run	Subrun	Event	Slice	Time [μs]	ReMId	Energy [GeV]	Scan
15028	18	124609	37	219.6	1.00	3.91	S
Continued on next page							

Table 15.2 – continued from previous page

Run	Subrun	Event	Slice	Time [μs]	ReMId	Energy [GeV]	Scan
16382	53	302896	25	226.8	1.00	1.80	S
16751	11	45215	41	224.5	1.00	2.21	S
17953	38	256887	29	228.4	0.99	2.60	S
18302	37	413485	32	218.2	0.99	1.20	B, E
18464	40	546039	24	221.3	1.00	1.98	S
18571	03	50129	27	226.2	1.00	1.94	S
18639	10	141206	43	220.6	0.99	3.37	S
18653	16	219065	36	227.0	0.98	2.32	S
18756	37	597960	22	226.1	1.00	2.26	S
18862	51	820267	37	227.7	0.95	0.88	S
18963	15	229864	37	221.4	1.00	1.77	S
19054	26	383867	34	220.8	0.99	3.76	S
19332	16	221786	44	221.7	1.00	3.96	S
19347	03	50163	34	218.9	0.74	0.65	I
19350	46	664010	29	218.4	0.99	1.41	B, E
19420	38	561111	32	225.0	1.00	2.63	S
19422	22	334692	34	226.8	1.00	2.77	S
19425	61	903289	26	220.7	1.00	1.86	S
19468	30	512430	25	227.3	1.00	2.34	S

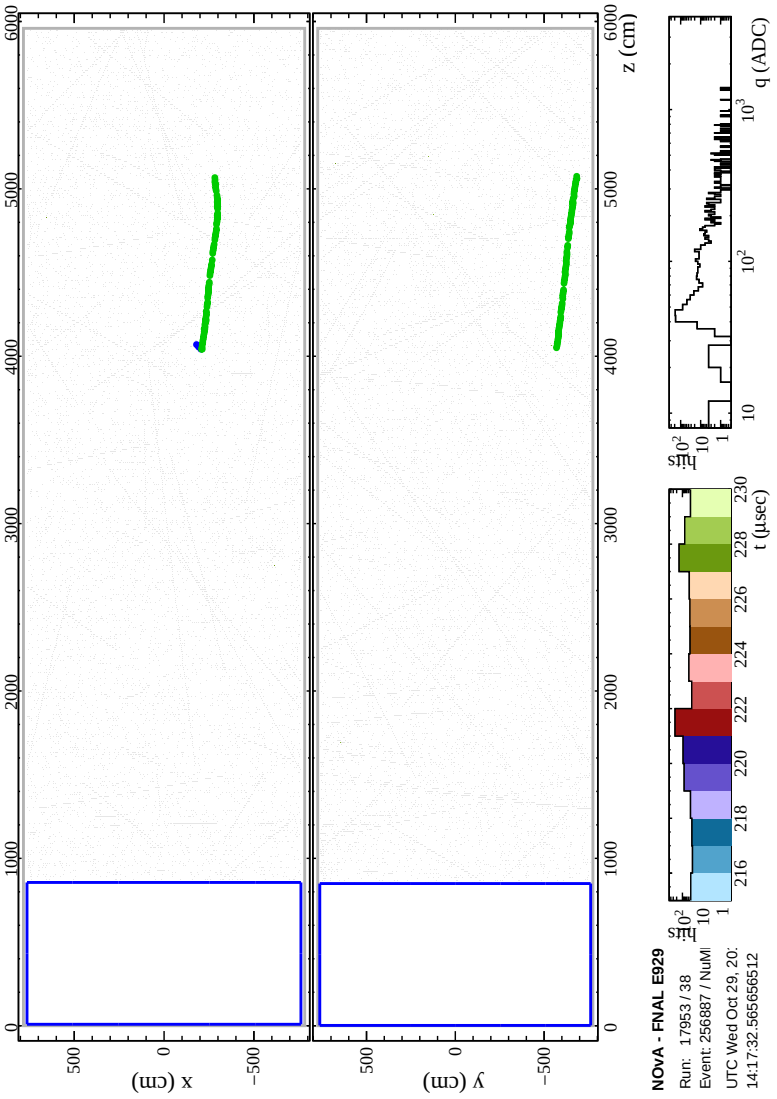


Figure 15.1: Event display of far detector data QE event. This event corresponds to run 17953, event 256887. Note that the figure is rotated. The view shows the entire far detector. Colored dots are drawn over hits in the slice associated with Kalman tracks. The green dots represent a long, 3D Kalman track. This slice also had a short 2D Kalman track, drawn with blue dots. Hits in the readout window not associated with the slice are grayed out. The blue box indicates the region of the detector not considered active.

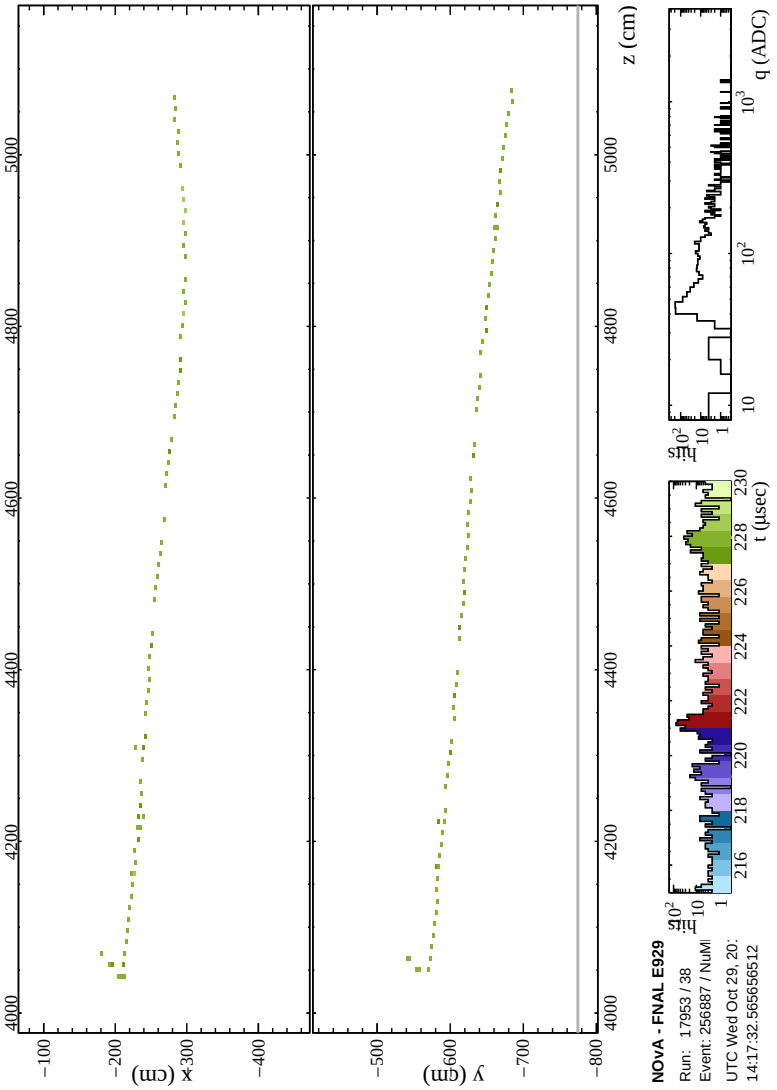


Figure 15.2: Event display of far detector data QE event. This event corresponds to run 17953, event 256887. Note that the figure is rotated. The display is spatially zoomed in to region of interest. Hits in the readout window not associated with the slice are not drawn. Hits are colored by their time relative to the readout window; the left bottom inset relates hit time and colors.

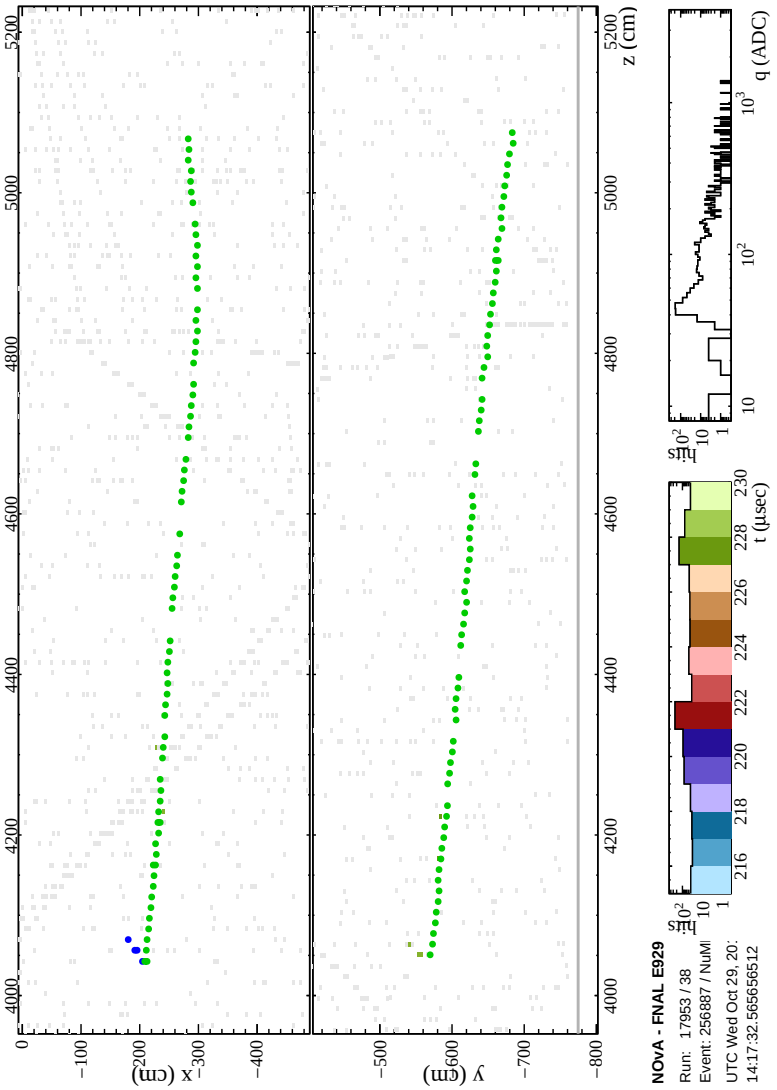


Figure 15.3: Event display of far detector data QE event. This event corresponds to run 17953, event 256887. Note that the figure is rotated. The display is spatially zoomed in to region of interest. Colored dots are drawn over hits in the slice associated with Kalman tracks. The green dots represent a long, 3D Kalman track. This slice also had a short 2D Kalman track, drawn with blue dots. Hits in the readout window not associated with the slice are grayed out.

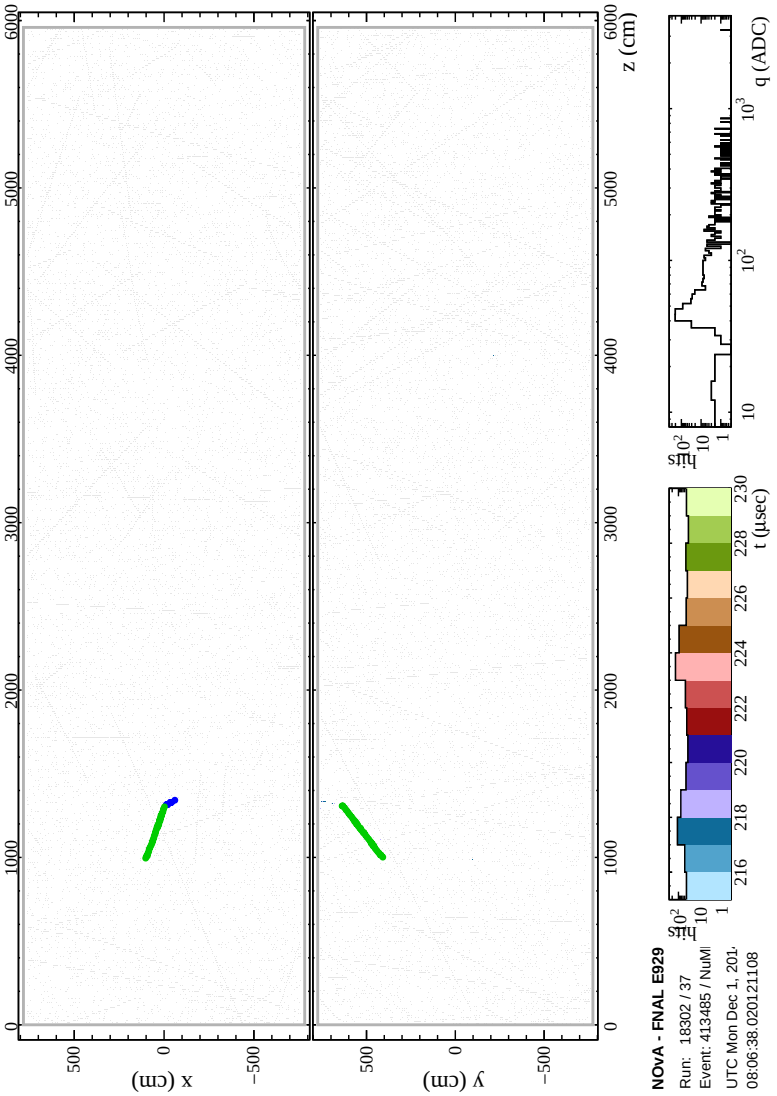


Figure 15.4: Event display of far detector data QE event. This event corresponds to run 18301, event 413485. Note that the figure is rotated. The view shows the entire far detector. Colored dots are drawn over hits in the slice associated with Kalman tracks. The green dots represent a 3D Kalman track. This slice also had a short 2D Kalman track, drawn with blue dots. Hits in the readout window not associated with the slice are grayed out.

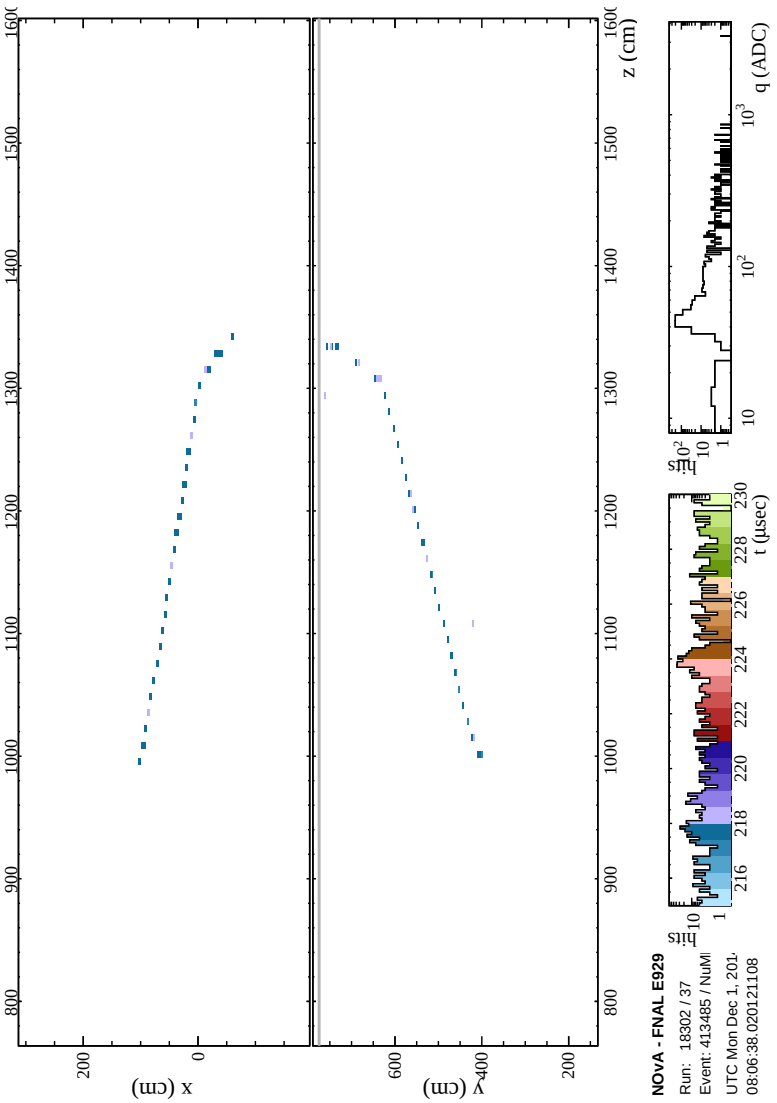


Figure 15.5: Event display of far detector data QE event. This event corresponds to run 18301, event 413485. Note that the figure is rotated. The display is spatially zoomed in to region of interest. Hits in the readout window not associated with the slice are not drawn. Hits are colored by their time relative to the readout window; the left bottom inset relates hit time and colors.

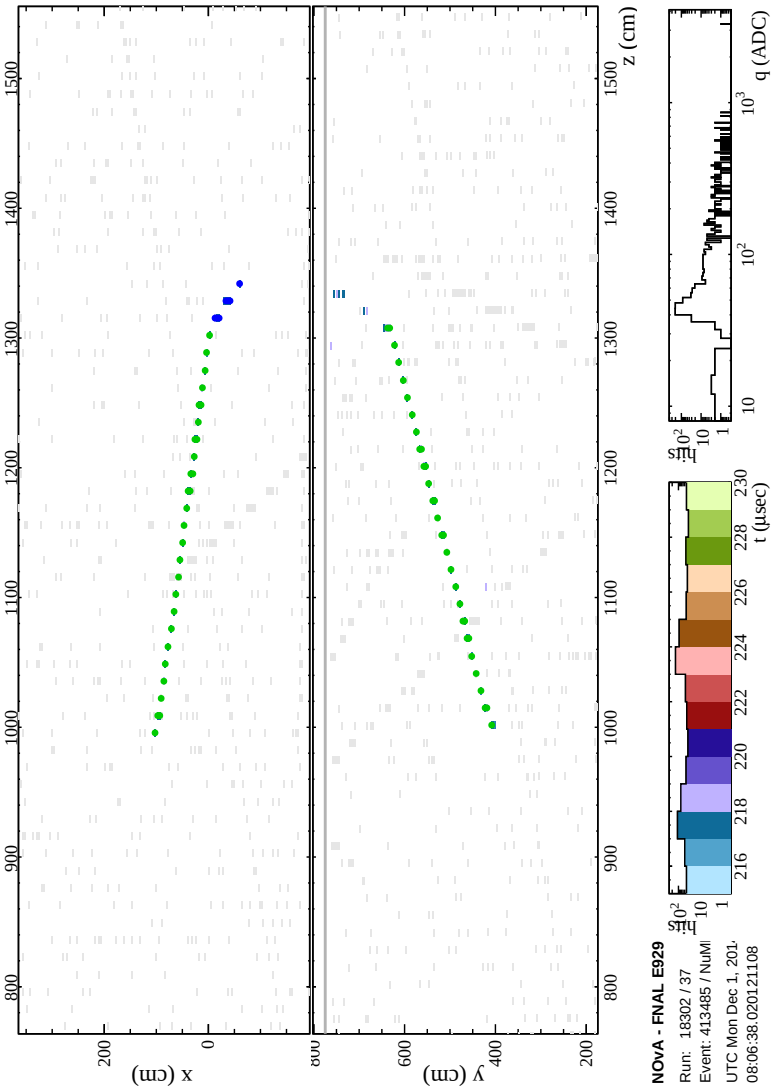


Figure 15.6: Event display of far detector data QE event. This event corresponds to run 18301, event 413485. Note that the figure is rotated. The display is spatially zoomed in to region of interest. Colored dots are drawn over hits in the slice associated with Kalman tracks. The green dots represent a 3D Kalman track. This slice also had a short 2D Kalman track, drawn with blue dots. Hits in the readout window not associated with the slice are grayed out.

Table 15.3 lists each selected nonQE far detector data event. 3 of the nonQE events visually looked like cosmic ray background events. One clearly entered from the edge of the detector, but due to a reconstruction error, passed the analysis cuts. This type of failure could be removed from future analyses by adding cuts that looked at slice extents. This is similar to the QE background event shown in Figure 15.4. The other two events that could be cosmic ray background were well-contained. Cuts that evaluated if the hadronic cluster was closer to the start or end of the muon track might help to remove this class of event; more sophisticated transverse momentum cuts could also help. The nonQE population was predicted to have 2.7 events from cosmic ray backgrounds. The probability of measuring 3 or more events when expecting 2.7 was 50%.

One of the best nonQE data events corresponds to run 18791, event 765587. Figure 15.7 is an event display of this event. It shows the entire far detector. Figure 15.8 is a spatially-zoomed version of the event display, showing the slice hits. Figure 15.9 is also spatially-zoomed, but has the Kalman tracks drawn on top of the hits. The long green 3D track is the muon; the shorter blue 3D track is probably a charged pion. It has a kink, representing a hard scatter, within the track, one of the hallmarks of a charged pion track. Also, the blue track is relatively high energy. This is what makes it classified as a nonQE interaction instead of a QE one.

For an example of a selected event that looked like cosmic-ray background that was well-contained in the detector, see Figure 15.10. It shows the entire far detector. Figure 15.11 is a spatially-zoomed version of the event display, showing the slice hits. Figure 15.12 is also spatially-zoomed, but has the Kalman tracks drawn on top of the hits. Note that what appears to be vertex activity is associated with the end of the muon track. Beam-induced neutrinos would usually have the vertex activity associated with the low-Z side of the track. This event could be caused by a neutrino from a cosmic ray interaction. An alternative hypothesis is that this event could be caused by a neutron, although due to the depth in the detector, this isn't the most likely explanation. Events with similar topology are seen in the data which doesn't correspond to a NuMI beam spill and their rate and source is not fully understood.

Table 15.3: Far detector nonQE data events. The time is relative to the start of the readout window. The ReMId value listed is for the 3D Kalman track in the slice with the highest ReMId. The energy is reconstructed neutrino energy in GeV. These events were visually scanned and categorized. "S" indicates the event looked signal-like. "I" indicates the event was indeterminate and could be signal or background. "B" indicates the event visually looked like cosmic ray background. "E" indicates that the background event was clearly entering the detector.

Run	Subrun	Event	Slice	Time [μs]	ReMId	Energy [GeV]	Scan
14828	38	192569	32	226.1	1.00	2.76	S
15085	00	1746	19	228.4	1.00	4.23	S
15679	16	142934	19	287.6	0.83	1.42	I
15974	14	88744	15	287.6	0.83	2.15	S
16315	04	17937	36	286.7	1.00	2.84	S
16450	17	93029	39	289.7	0.80	2.07	B, C
16453	33	178062	37	292.2	0.94	1.38	B, C
16675	54	249520	26	221.0	0.75	0.81	B, E
16730	25	101478	31	220.5	1.00	3.33	S
18342	47	609061	34	226.4	0.99	2.11	S
18401	02	35501	34	225.7	0.75	1.49	S
18417	43	582977	30	226.1	0.81	0.97	S
18572	19	255330	24	226.8	0.95	1.81	S
18791	48	765587	29	224.9	1.00	2.80	S
19004	22	318354	37	225.8	1.00	2.74	S
19058	39	568646	29	222.9	1.00	1.99	S
19084	62	908450	22	222.0	1.00	3.92	S
19107	52	756898	32	226.0	0.79	1.33	I
19154	45	663979	30	228.1	0.74	0.90	I
19296	35	482441	19	223.1	0.93	1.56	S
19327	18	248695	22	220.0	0.86	1.08	S
19356	11	156460	34	220.5	0.99	1.65	S
19423	01	20137	36	220.4	1.00	2.24	S

Continued on next page

Table 15.3 – continued from previous page

Run	Subrun	Event	Slice	Time [μs]	ReMId	Energy [GeV]	Scan
19476	06	110767	25	222.0	1.00	1.75	S
19485	61	1068905	32	228.6	0.79	2.16	S

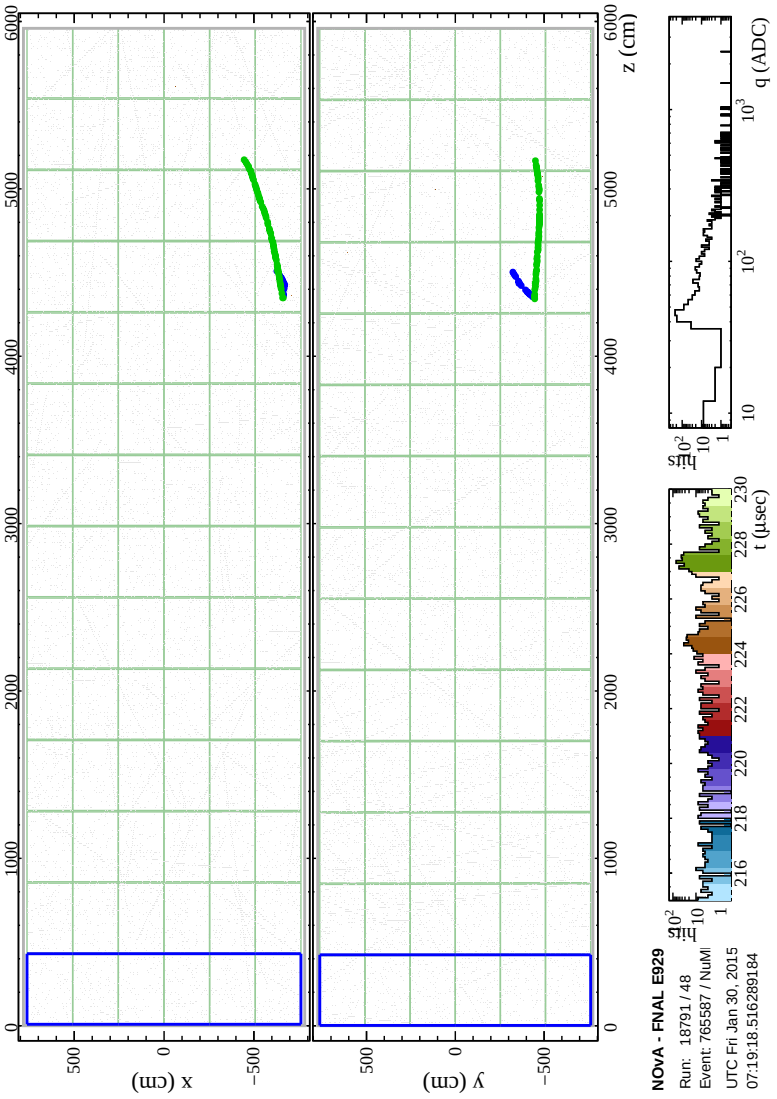


Figure 15.7: Event display of far detector data nonQE event. This event corresponds to run 18791, event 765587. Note that the figure is rotated. The view shows the entire far detector. Colored dots are drawn over hits in the slice associated with Kalman tracks. The green dots represent a long, 3D Kalman track. This slice also had a shorter 3D Kalman track, drawn with blue dots. Hits in the readout window not associated with the slice are grayed out. The blue box indicates the region of the detector not considered active. The green boxes indicate DCM boundaries.

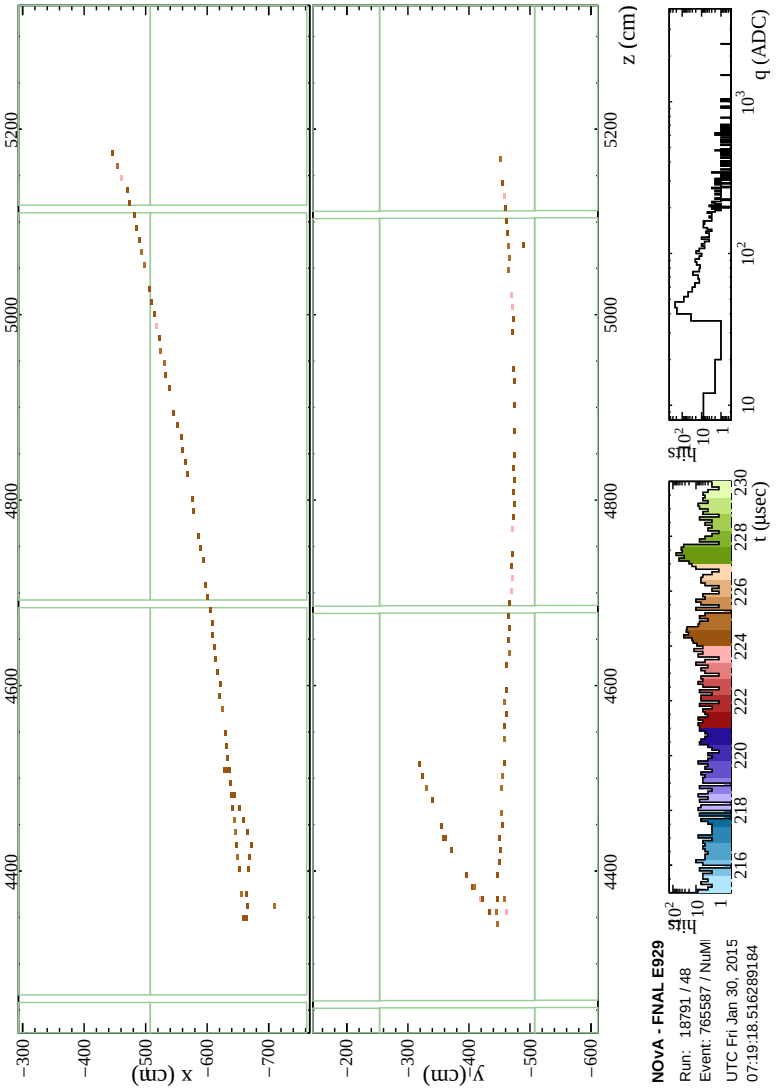


Figure 15.8: Event display of far detector data nonQE event. This event corresponds to run 18791, event 765587. Note that the figure is rotated. The display is spatially zoomed in to region of interest. Hits in the readout window not associated with the slice are not drawn. Hits are colored by their time relative to the readout window; the left bottom inset relates hit time and colors. The green boxes indicate DCM boundaries.

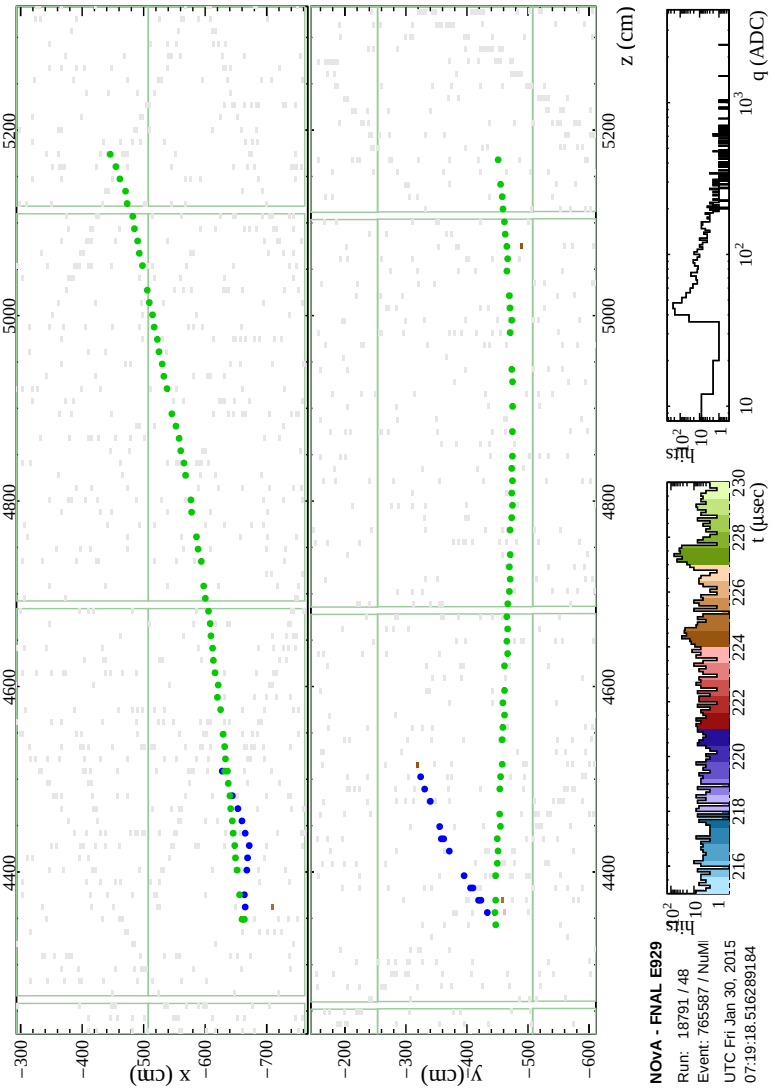


Figure 15.9: Event display of far detector data nonQE event. This event corresponds to run 18791, event 765587. Note that the figure is rotated. The display is spatially zoomed in to region of interest. Colored dots are drawn over hits in the slice associated with Kalman tracks. The green dots represent a long, 3D Kalman track. This slice also had a shorter 3D Kalman track, drawn with blue dots. Hits in the readout window not associated with the slice are grayed out. The green boxes indicate DCM boundaries.

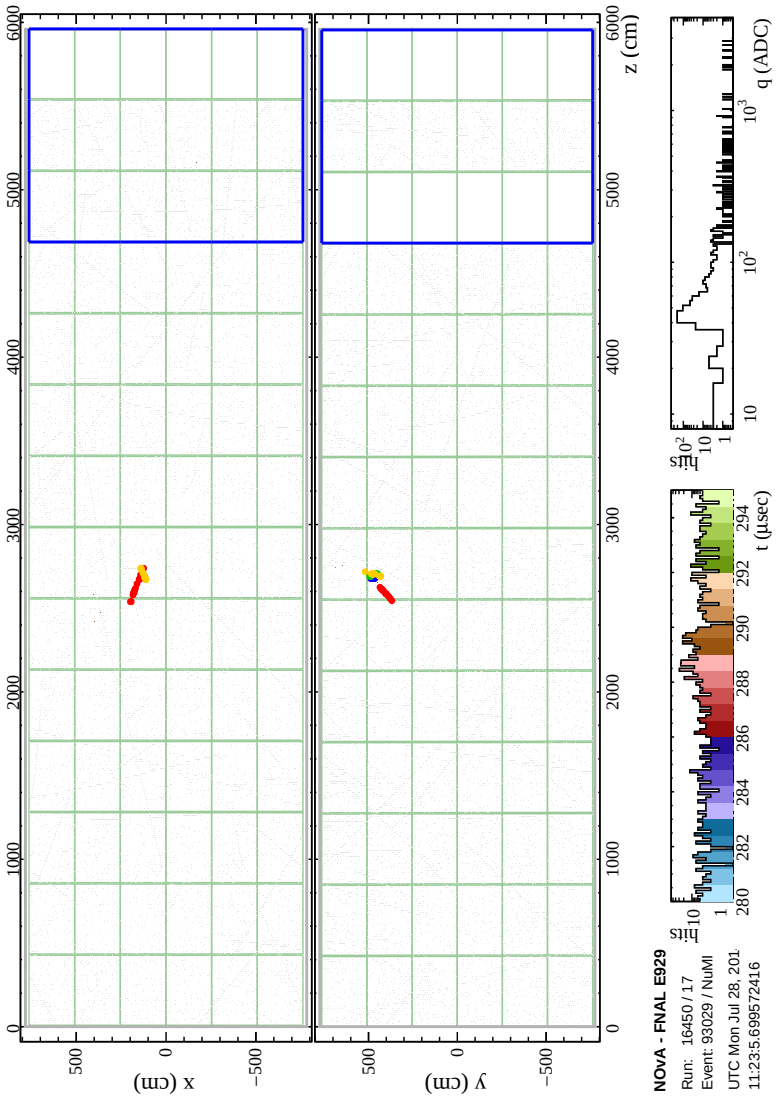


Figure 15.10: Event display of far detector data nonQE event. This event corresponds to run 16450, event 93029. Note that the figure is rotated. The view shows the entire far detector. Colored dots are drawn over hits in the slice associated with Kalman tracks. The red dots represent the longest 3D Kalman track. Yellow dots are associated with a second 3D Kalman track. This slice also had two 2D Kalman track, drawn with blue dots and green dots. Hits in the readout window not associated with the slice are grayed out. The blue box indicates the region of the detector not considered active. The green boxes indicate DCM boundaries.

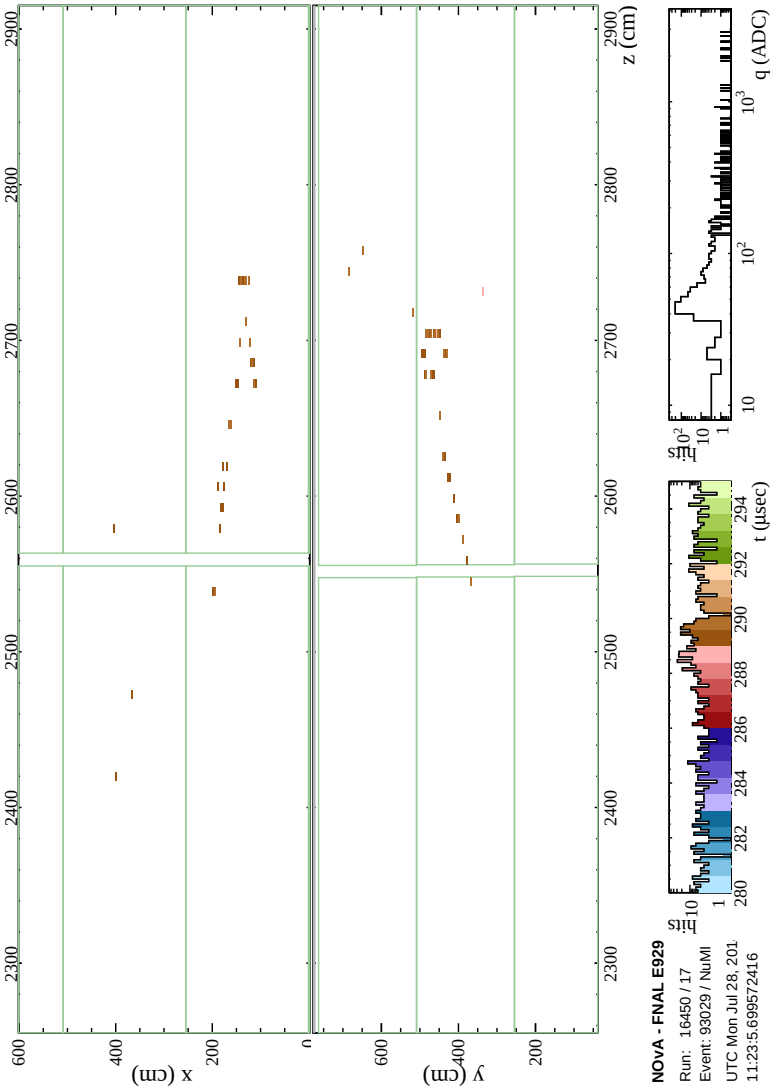


Figure 15.11: Event display of far detector data nonQE event. This event corresponds to run 16450, event 93029. Note that the figure is rotated. The display is spatially zoomed in to region of interest. Hits in the readout window not associated with the slice are not drawn. Hits are colored by their time relative to the readout window; the left bottom inset relates hit time and colors. The green boxes indicate DCM boundaries.

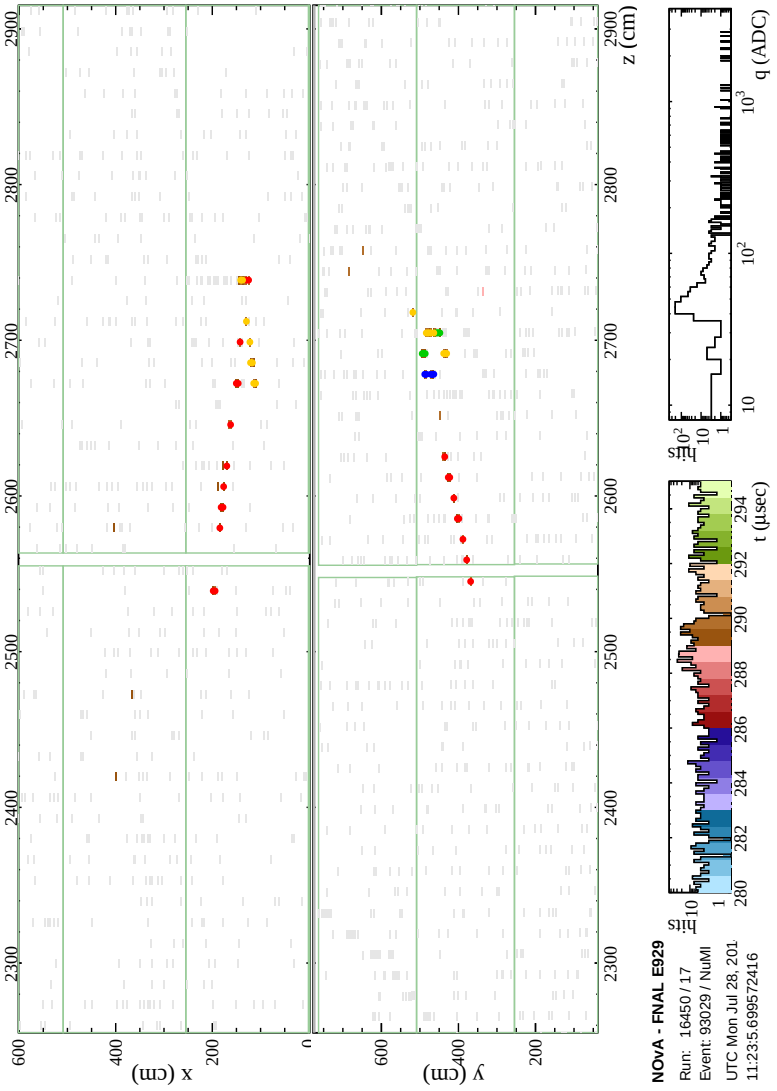


Figure 15.12: Event display of far detector data nonQE event. This event corresponds to run 16450, event 93029. Note that the figure is rotated. The display is spatially zoomed in to region of interest. Colored dots are drawn over hits in the slice associated with Kalman tracks. The red dots represent the longest 3D Kalman track. Yellow dots are associated with a second 3D Kalman track. This slice also had two 2D Kalman track, drawn with blue dots and green dots. Hits in the readout window not associated with the slice are grayed out. The green boxes indicate DCM boundaries.

15.1.1 Far Detector Data Distributions

We can compare the far detector data distributions to our simulation. For these plots, the simulation has been oscillated using the best fit values described in Section 15.2. We oscillated using the values listed in Table 4.2 and setting $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The best fit values for the systematic errors were not used.

For all the distributions examined, the agreement between the far detector data and the simulation is good. With the increased statistics of future analyses, it will be easier to see if any subtle discrepancies exist.

Figure 15.13 displays the number of hits in a slice for the QE population; Figure 15.14 is for the nonQE population. The number of 3D Kalman tracks in a slice is shown in Figure 15.15 for the QE population and in Figure 15.16 for the nonQE population. The number of hits on the primary Kalman track for the QE population is plotted in Figure 15.17; Figure 15.18 is for the nonQE population. The next three plots, Figures 15.19, 15.21 and 15.23, show the X, Y and Z detector coordinate positions for the start of the primary Kalman track for the QE population. Figures 15.20, 15.22 and 15.24 show the track start positions for the nonQE population. Figures 15.25, 15.27 and 15.29, show the X, Y and Z detector coordinate positions for the end of the primary Kalman track for the QE population. Figures 15.26, 15.28 and 15.30 show the track end positions for the nonQE population. The angle between the primary Kalman track and the NuMI beam line is shown in Figure 15.31 for the QE population and Figure 15.32 for the nonQE population. All show excellent agreement between data and simulation.

We can next look at some of the ReMId inputs. Figure 15.33 is the scattering log-likelihood for the primary Kalman track for the QE population; Figure 15.34 is for the nonQE population. Similarly, Figure 15.35 is the dE/dx log-likelihood for the primary Kalman track for the QE population; Figure 15.34 is for the nonQE population. The ReMId values for the QE population are shown in Figure 15.37 and in Figure 15.38 for the nonQE population. Finally, Figure 15.39 plots the visible hadronic energy for the QE population and Figure 15.40 plots the same for the nonQE population. Again, all show excellent agreement between data and simulation.

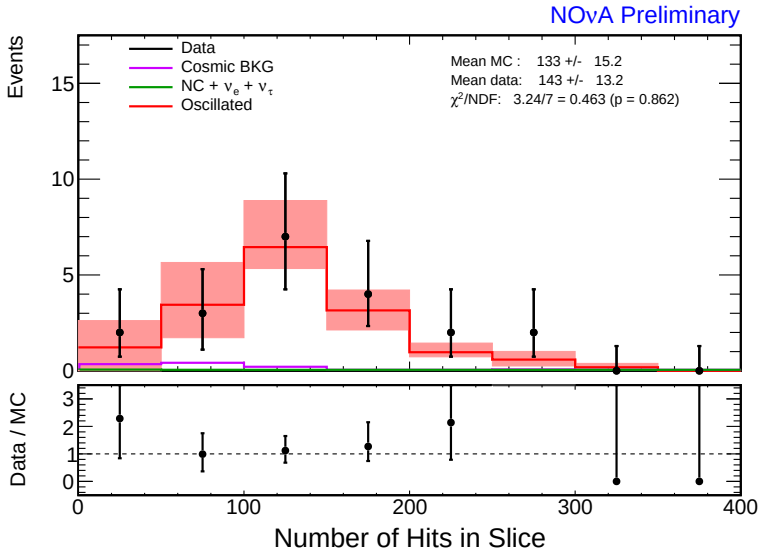


Figure 15.13: Plot of the number of hits in a slice for the far detector for the QE population. The total prediction is drawn as a red line with red total systematic error bands, with neutrino background drawn as a green line and the cosmic ray background drawn as a magenta line. The cosmic ray background distribution was determined from the out-of-time data in the NuMI trigger files. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is oscillated using the values listed in Table 4.2 and setting $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The simulation is scaled down to match the exposure for the far detector data, 3.45×10^{20} POT.

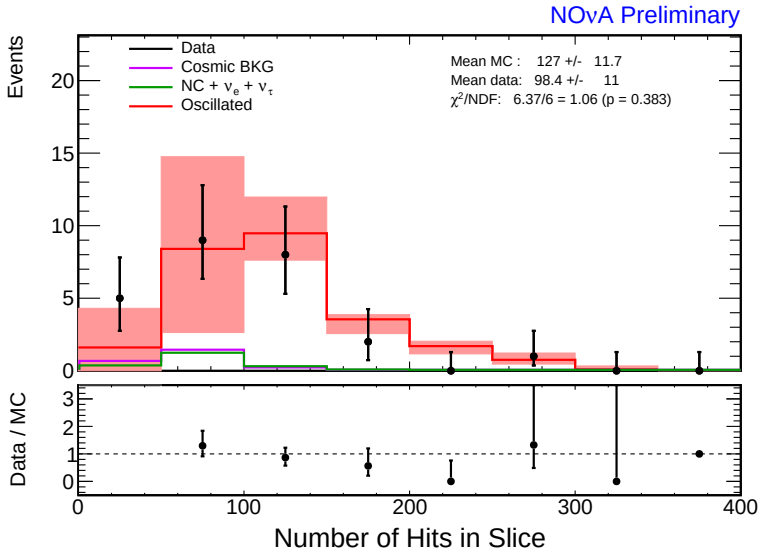


Figure 15.14: Plot of the number of hits in a slice for the far detector for the nonQE population. The total prediction is drawn as a red line with red total systematic error bands, with neutrino background drawn as a green line and the cosmic ray background drawn as a magenta line. The cosmic ray background distribution was determined from the out-of-time data in the NuMI trigger files. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is oscillated using the values listed in Table 4.2 and setting $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The simulation is scaled down to match the exposure for the far detector data, 3.45×10^{20} POT.

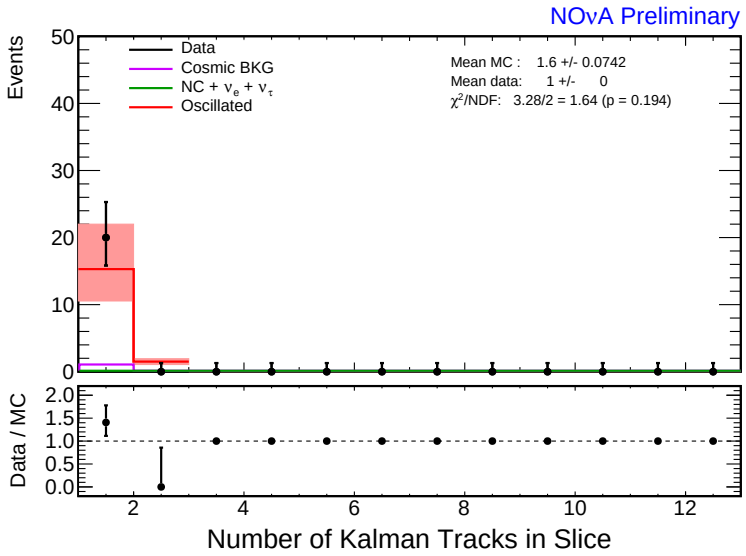


Figure 15.15: Plot of the number of 3D Kalman tracks in a slice for the far detector for the QE population. The total prediction is drawn as a red line with red total systematic error bands, with neutrino background drawn as a green line and the cosmic ray background drawn as a magenta line. The cosmic ray background distribution was determined from the out-of-time data in the NuMI trigger files. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is oscillated using the values listed in Table 4.2 and setting $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The simulation is scaled down to match the exposure for the far detector data, 3.45×10^{20} POT.

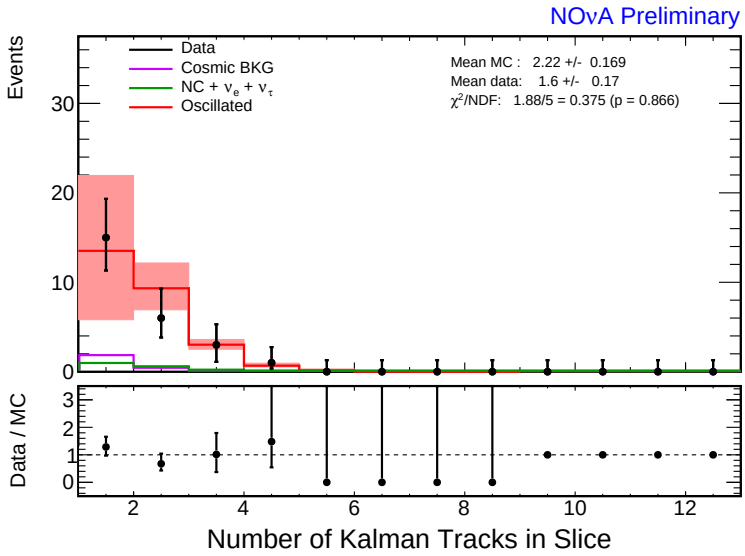


Figure 15.16: Plot of the number of 3D Kalman tracks in a slice for the far detector for the nonQE population. The total prediction is drawn as a red line with red total systematic error bands, with neutrino background drawn as a green line and the cosmic ray background drawn as a magenta line. The cosmic ray background distribution was determined from the out-of-time data in the NuMI trigger files. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is oscillated using the values listed in Table 4.2 and setting $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The simulation is scaled down to match the exposure for the far detector data, 3.45×10^{20} POT.

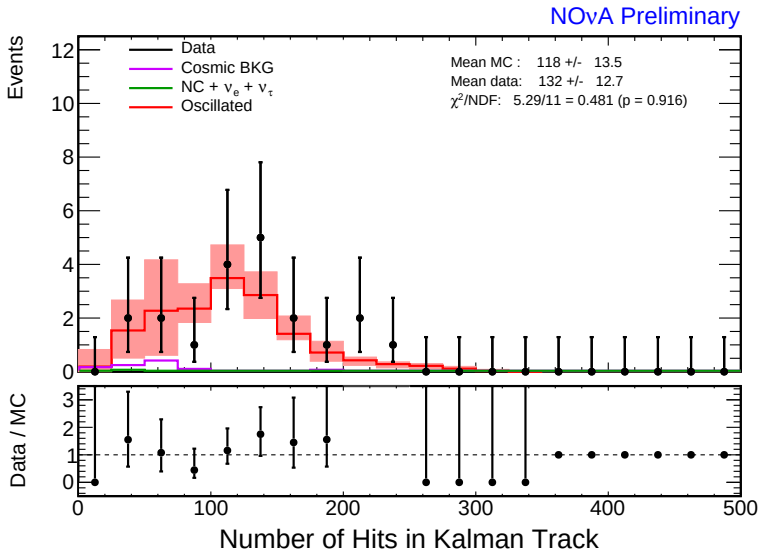


Figure 15.17: Plot of the number of hits on the 3D Kalman track with the highest ReMId value in the slice. This plot is for the far detector QE population. The total prediction is drawn as a red line with red total systematic error bands, with neutrino background drawn as a green line and the cosmic ray background drawn as a magenta line. The cosmic ray background distribution was determined from the out-of-time data in the NuMI trigger files. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is oscillated using the values listed in Table 4.2 and setting $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The simulation is scaled down to match the exposure for the far detector data, 3.45×10^{20} POT.

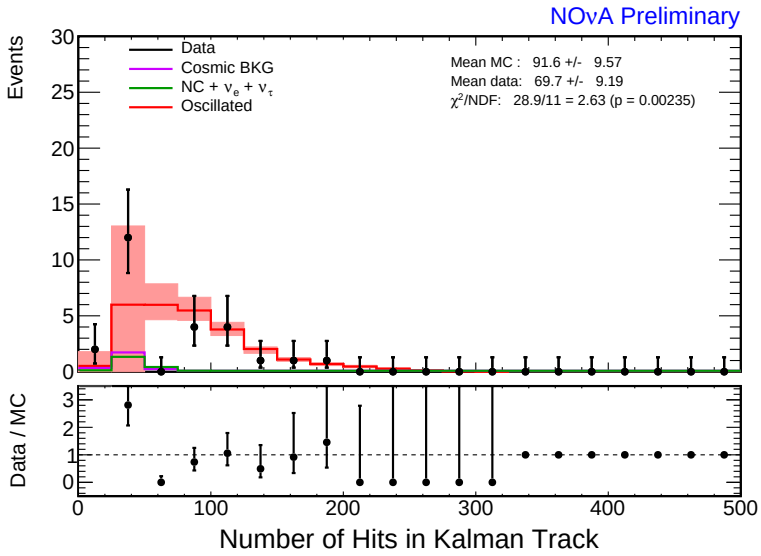


Figure 15.18: Plot of the number of hits on the 3D Kalman track with the highest ReMId value in the slice. This plot is for the far detector nonQE population. The total prediction is drawn as a red line with red total systematic error bands, with neutrino background drawn as a green line and the cosmic ray background drawn as a magenta line. The cosmic ray background distribution was determined from the out-of-time data in the NuMI trigger files. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is oscillated using the values listed in Table 4.2 and setting $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The simulation is scaled down to match the exposure for the far detector data, 3.45×10^{20} POT.

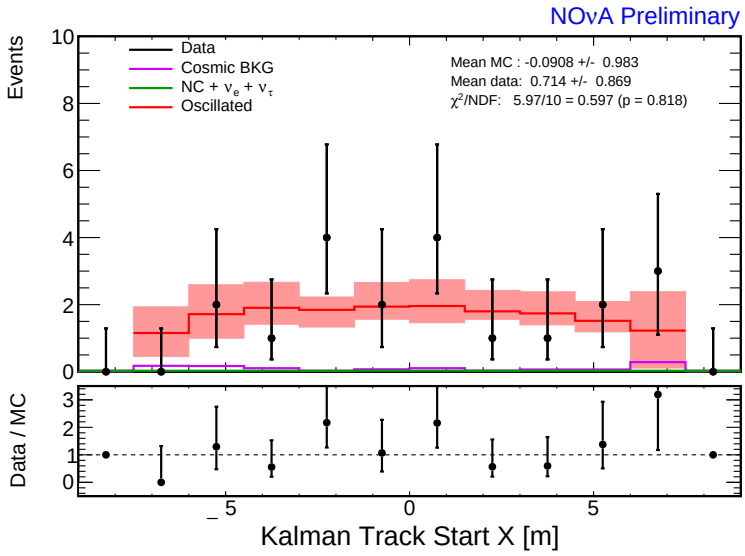


Figure 15.19: Plot of the start position in the detector X coordinate in m for the 3D Kalman track with the highest ReMId value in a slice for the far detector for the QE population. The total prediction is drawn as a red line with red total systematic error bands, with neutrino background drawn as a green line and the cosmic ray background drawn as a magenta line. The cosmic ray background distribution was determined from the out-of-time data in the NuMI trigger files. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is oscillated using the values listed in Table 4.2 and setting $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The simulation is scaled down to match the exposure for the far detector data, 3.45×10^{20} POT.

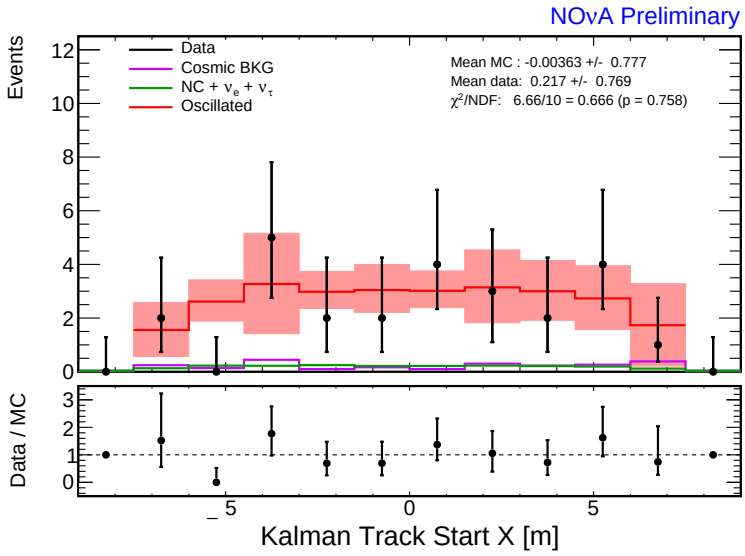


Figure 15.20: Plot of the start position in the detector X coordinate in m for the 3D Kalman track with the highest ReMId value in a slice for the far detector for the nonQE population. The total prediction is drawn as a red line with red total systematic error bands, with neutrino background drawn as a green line and the cosmic ray background drawn as a magenta line. The cosmic ray background distribution was determined from the out-of-time data in the NuMI trigger files. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is oscillated using the values listed in Table 4.2 and setting $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The simulation is scaled down to match the exposure for the far detector data, 3.45×10^{20} POT.

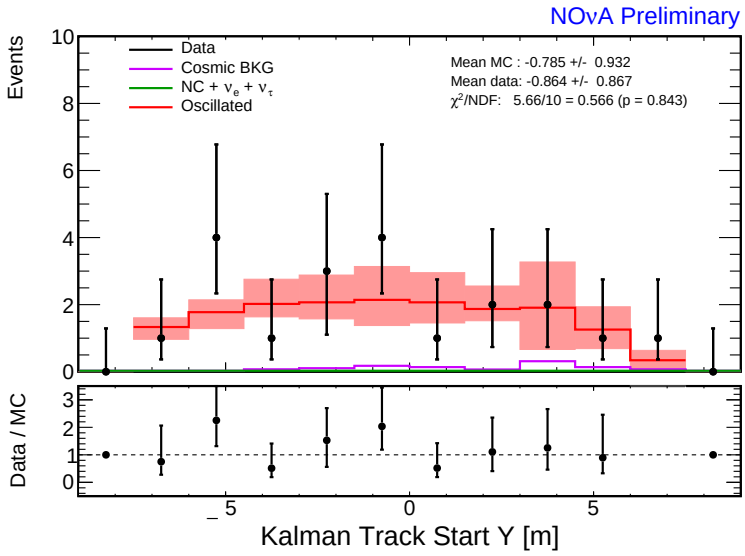


Figure 15.21: Plot of the start position in the detector Y coordinate in m for the 3D Kalman track with the highest ReMId value in a slice for the far detector for the QE population. The total prediction is drawn as a red line with red total systematic error bands, with neutrino background drawn as a green line and the cosmic ray background drawn as a magenta line. The cosmic ray background distribution was determined from the out-of-time data in the NuMI trigger files. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is oscillated using the values listed in Table 4.2 and setting $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The simulation is scaled down to match the exposure for the far detector data, 3.45×10^{20} POT.

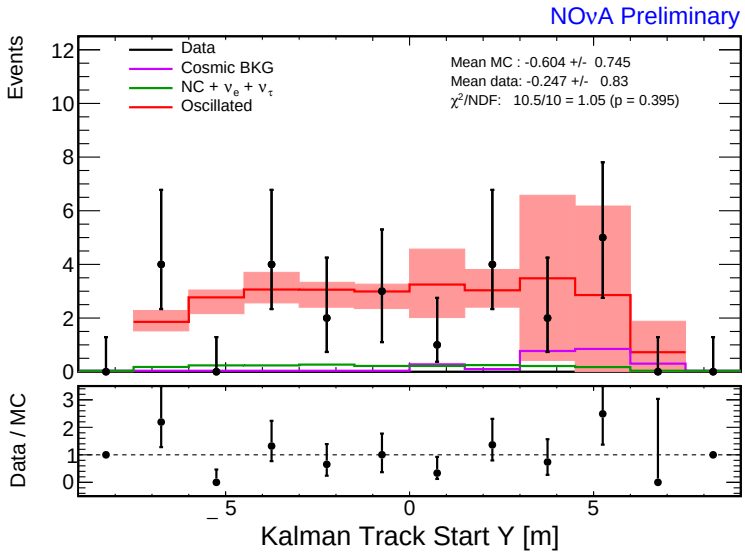


Figure 15.22: Plot of the start position in the detector Y coordinate in m for the 3D Kalman track with the highest ReMId value in a slice for the far detector for the nonQE population. The total prediction is drawn as a red line with red total systematic error bands, with neutrino background drawn as a green line and the cosmic ray background drawn as a magenta line. The cosmic ray background distribution was determined from the out-of-time data in the NuMI trigger files. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is oscillated using the values listed in Table 4.2 and setting $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The simulation is scaled down to match the exposure for the far detector data, 3.45×10^{20} POT.

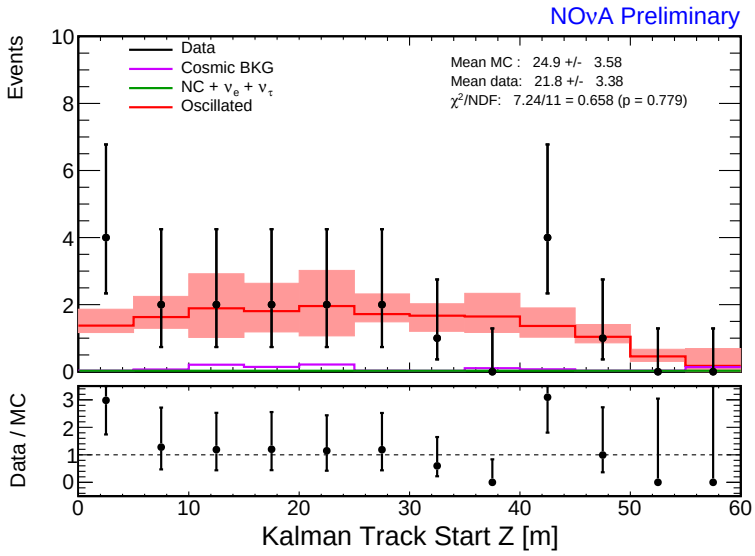


Figure 15.23: Plot of the start position in the detector Z coordinate in m for the 3D Kalman track with the highest ReMId value in a slice for the far detector for the QE population. The total prediction is drawn as a red line with red total systematic error bands, with neutrino background drawn as a green line and the cosmic ray background drawn as a magenta line. The cosmic ray background distribution was determined from the out-of-time data in the NuMI trigger files. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is oscillated using the values listed in Table 4.2 and setting $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The simulation is scaled down to match the exposure for the far detector data, 3.45×10^{20} POT.

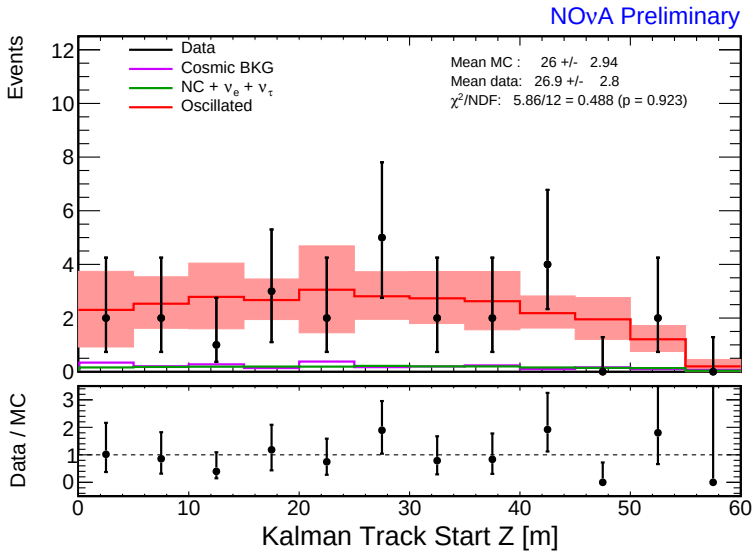


Figure 15.24: Plot of the start position in the detector Z coordinate in m for the 3D Kalman track with the highest ReMId value in a slice for the far detector for the nonQE population. The total prediction is drawn as a red line with red total systematic error bands, with neutrino background drawn as a green line and the cosmic ray background drawn as a magenta line. The cosmic ray background distribution was determined from the out-of-time data in the NuMI trigger files. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is oscillated using the values listed in Table 4.2 and setting $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The simulation is scaled down to match the exposure for the far detector data, 3.45×10^{20} POT.

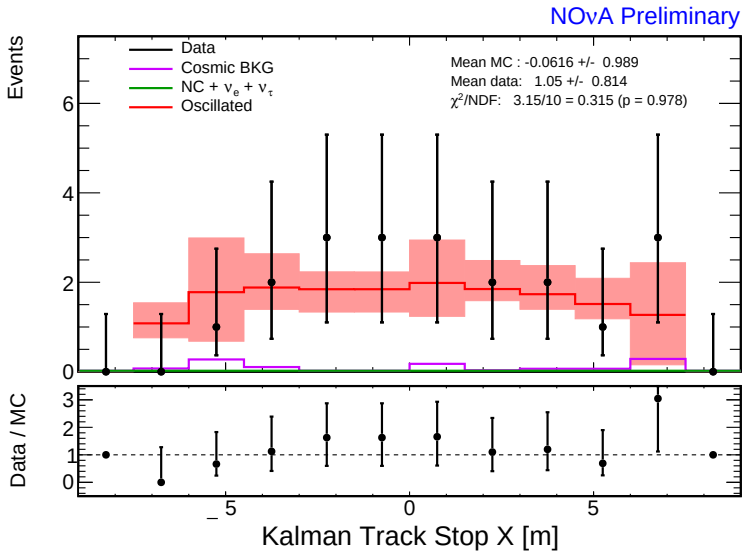


Figure 15.25: Plot of the end position in the detector X coordinate in m for the 3D Kalman track with the highest ReMId value in a slice for the far detector for the QE population. The total prediction is drawn as a red line with red total systematic error bands, with neutrino background drawn as a green line and the cosmic ray background drawn as a magenta line. The cosmic ray background distribution was determined from the out-of-time data in the NuMI trigger files. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is oscillated using the values listed in Table 4.2 and setting $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The simulation is scaled down to match the exposure for the far detector data, 3.45×10^{20} POT.

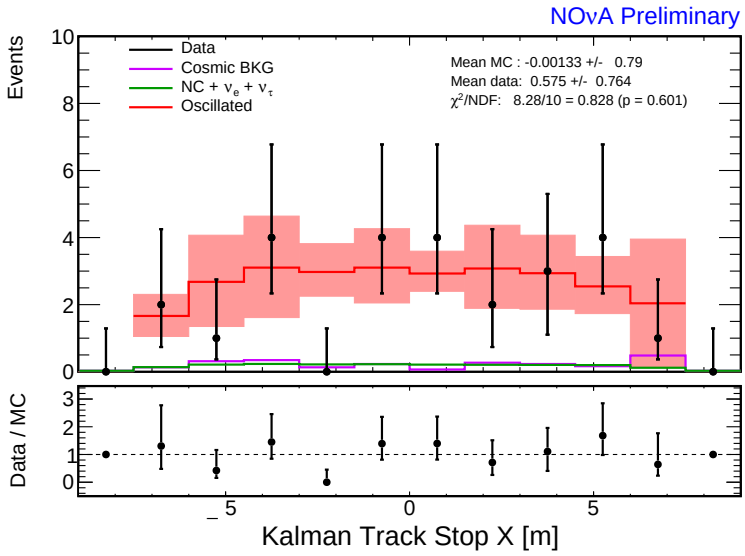


Figure 15.26: Plot of the end position in the detector X coordinate in m for the 3D Kalman track with the highest ReMId value in a slice for the far detector for the nonQE population. The total prediction is drawn as a red line with red total systematic error bands, with neutrino background drawn as a green line and the cosmic ray background drawn as a magenta line. The cosmic ray background distribution was determined from the out-of-time data in the NuMI trigger files. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is oscillated using the values listed in Table 4.2 and setting $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The simulation is scaled down to match the exposure for the far detector data, 3.45×10^{20} POT.

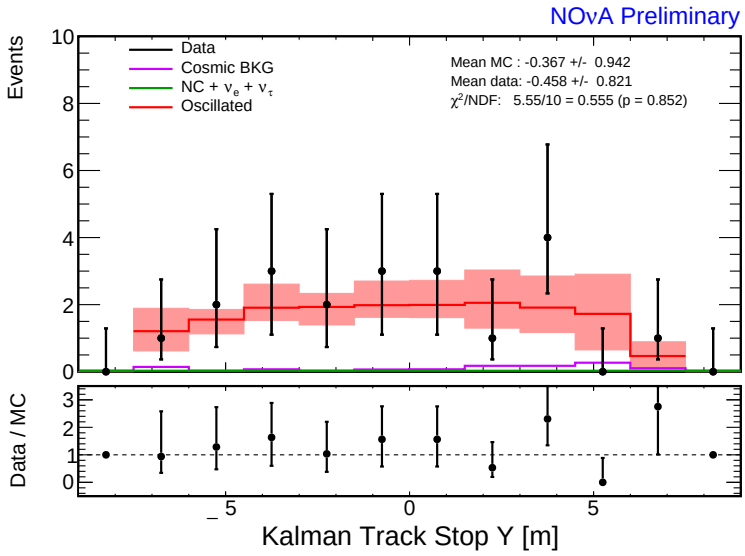


Figure 15.27: Plot of the end position in the detector Y coordinate in m for the 3D Kalman track with the highest ReMId value in a slice for the far detector for the QE population. The total prediction is drawn as a red line with red total systematic error bands, with neutrino background drawn as a green line and the cosmic ray background drawn as a magenta line. The cosmic ray background distribution was determined from the out-of-time data in the NuMI trigger files. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is oscillated using the values listed in Table 4.2 and setting $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The simulation is scaled down to match the exposure for the far detector data, 3.45×10^{20} POT.

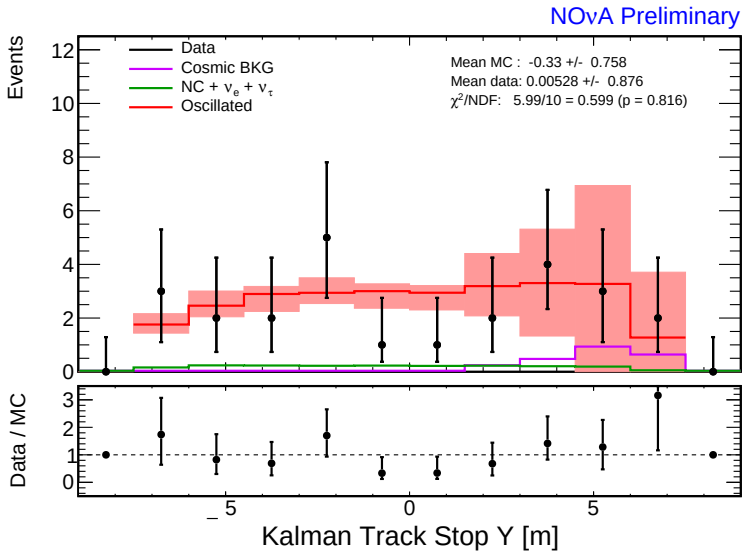


Figure 15.28: Plot of the end position in the detector Y coordinate in m for the 3D Kalman track with the highest ReMId value in a slice for the far detector for the nonQE population. The total prediction is drawn as a red line with red total systematic error bands, with neutrino background drawn as a green line and the cosmic ray background drawn as a magenta line. The cosmic ray background distribution was determined from the out-of-time data in the NuMI trigger files. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is oscillated using the values listed in Table 4.2 and setting $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The simulation is scaled down to match the exposure for the far detector data, 3.45×10^{20} POT.

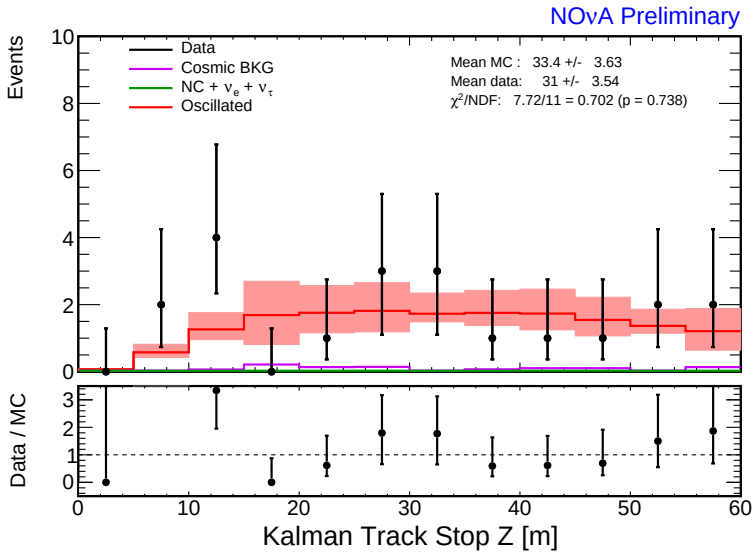


Figure 15.29: Plot of the end position in the detector Z coordinate in m for the 3D Kalman track with the highest ReMId value in a slice for the far detector for the QE population. The total prediction is drawn as a red line with red total systematic error bands, with neutrino background drawn as a green line and the cosmic ray background drawn as a magenta line. The cosmic ray background distribution was determined from the out-of-time data in the NuMI trigger files. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is oscillated using the values listed in Table 4.2 and setting $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The simulation is scaled down to match the exposure for the far detector data, 3.45×10^{20} POT.

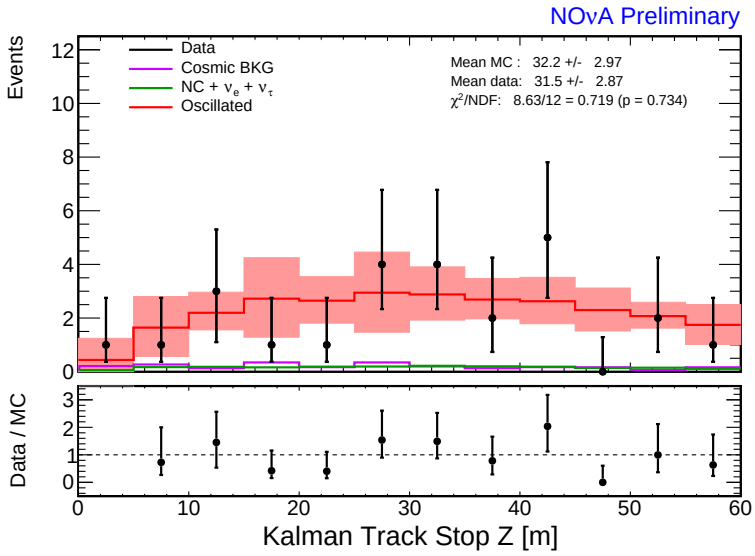


Figure 15.30: Plot of the end position in the detector Z coordinate in m for the 3D Kalman track with the highest ReMId value in a slice for the far detector for the nonQE population. The total prediction is drawn as a red line with red total systematic error bands, with neutrino background drawn as a green line and the cosmic ray background drawn as a magenta line. The cosmic ray background distribution was determined from the out-of-time data in the NuMI trigger files. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is oscillated using the values listed in Table 4.2 and setting $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The simulation is scaled down to match the exposure for the far detector data, 3.45×10^{20} POT.

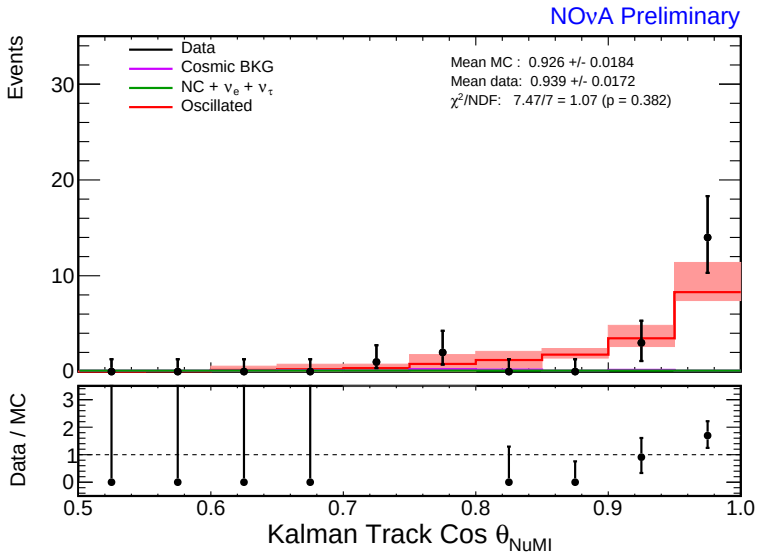


Figure 15.31: Plot of $\cos\theta_{NuMI}$, where θ_{NuMI} is the angle between the 3D Kalman track with the highest ReMId value in the slice and the NuMI beam direction. This plot is for the far detector QE population. The total prediction is drawn as a red line with red total systematic error bands, with neutrino background drawn as a green line and the cosmic ray background drawn as a magenta line. The cosmic ray background distribution was determined from the out-of-time data in the NuMI trigger files. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. When the ratio is too large for the scale, the point and its error bars are not drawn. The simulation is oscillated using the values listed in Table 4.2 and setting $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The simulation is scaled down to match the exposure for the far detector data, 3.45×10^{20} POT.

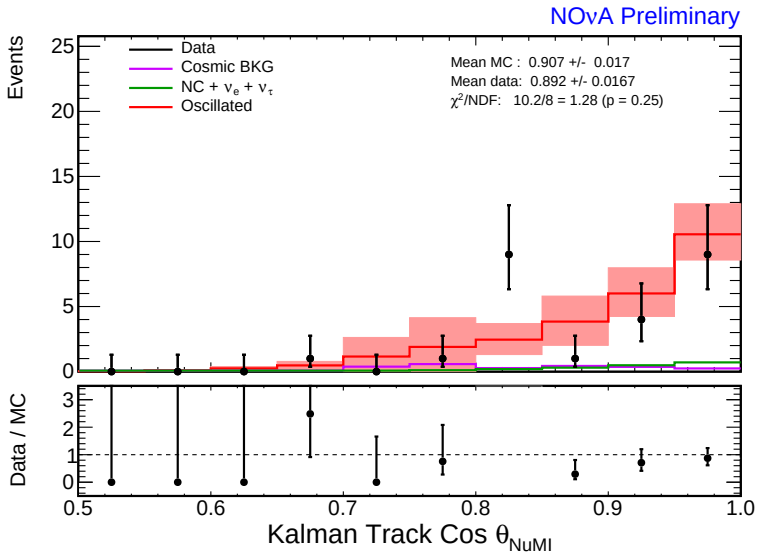


Figure 15.32: Plot of $\cos\theta_{NuMI}$, where θ_{NuMI} is the angle between the 3D Kalman track with the highest ReMId value in the slice and the NuMI beam direction. This plot is for the far detector nonQE population. The total prediction is drawn as a red line with red total systematic error bands, with neutrino background drawn as a green line and the cosmic ray background drawn as a magenta line. The cosmic ray background distribution was determined from the out-of-time data in the NuMI trigger files. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. When the ratio is too large for the scale, the point and its error bars are not drawn. The simulation is oscillated using the values listed in Table 4.2 and setting $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The simulation is scaled down to match the exposure for the far detector data, 3.45×10^{20} POT.

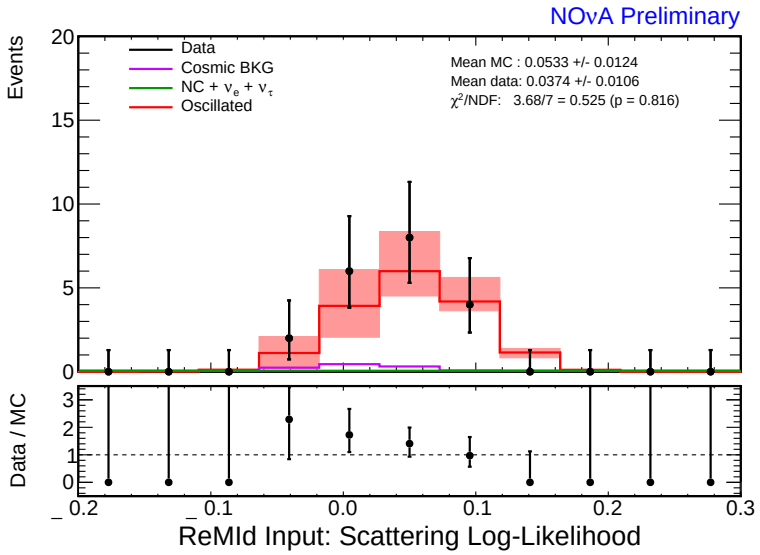


Figure 15.33: Plot of the scattering log-likelihood for the 3D Kalman track with the highest ReMId value in the slice. This plot is for the far detector QE population. The total prediction is drawn as a red line with red total systematic error bands, with neutrino background drawn as a green line and the cosmic ray background drawn as a magenta line. The cosmic ray background distribution was determined from the out-of-time data in the NuMI trigger files. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is oscillated using the values listed in Table 4.2 and setting $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The simulation is scaled down to match the exposure for the far detector data, 3.45×10^{20} POT.

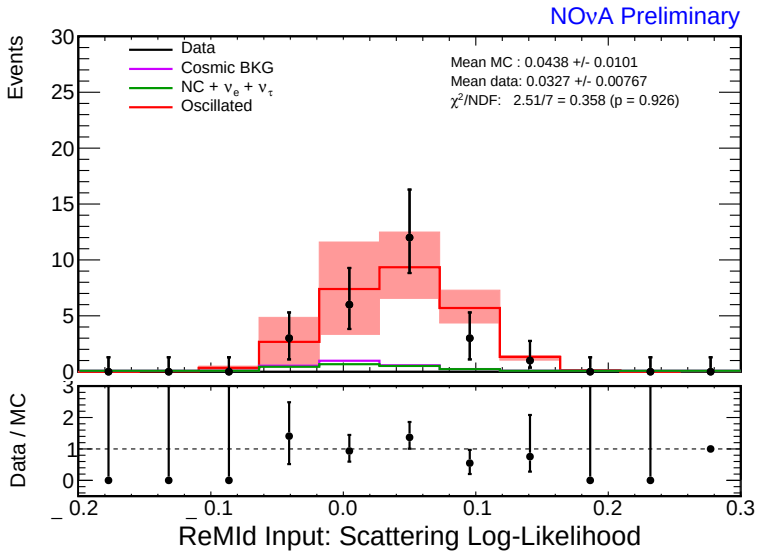


Figure 15.34: Plot of the scattering log-likelihood for the 3D Kalman track with the highest ReMId value in the slice. This plot is for the far detector nonQE population. The total prediction is drawn as a red line with red total systematic error bands, with neutrino background drawn as a green line and the cosmic ray background drawn as a magenta line. The cosmic ray background distribution was determined from the out-of-time data in the NuMI trigger files. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is oscillated using the values listed in Table 4.2 and setting $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The simulation is scaled down to match the exposure for the far detector data, 3.45×10^{20} POT.

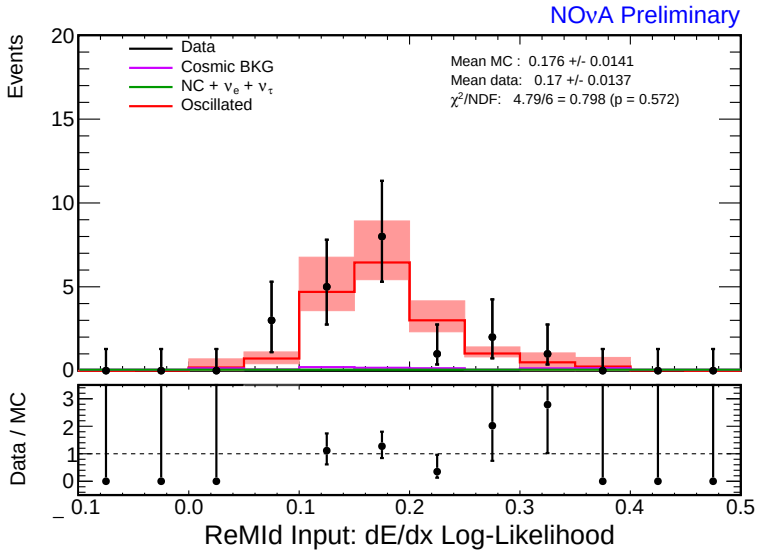


Figure 15.35: Plot of the dE/dx log-likelihood for the 3D Kalman track with the highest ReMId value in the slice. This plot is for the far detector QE population. The total prediction is drawn as a red line with red total systematic error bands, with neutrino background drawn as a green line and the cosmic ray background drawn as a magenta line. The cosmic ray background distribution was determined from the out-of-time data in the NuMI trigger files. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. When the ratio is too large for the scale, the point and its error bars are not drawn. The simulation is oscillated using the values listed in Table 4.2 and setting $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The simulation is scaled down to match the exposure for the far detector data, 3.45×10^{20} POT.

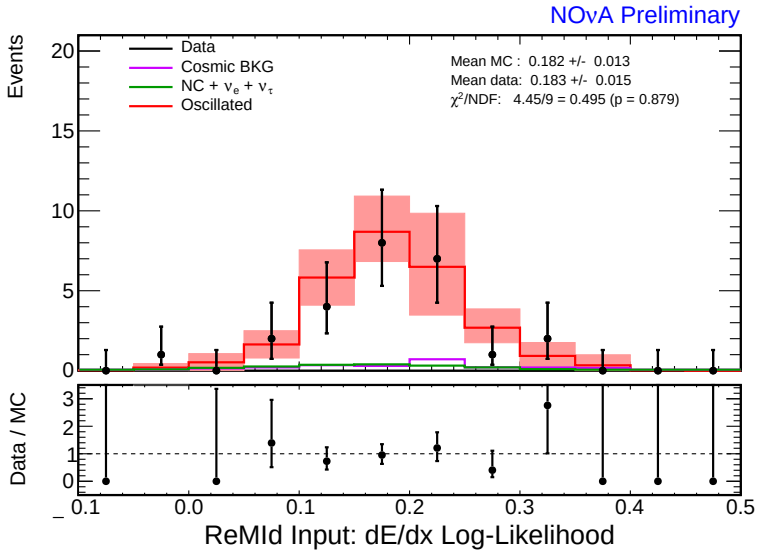


Figure 15.36: Plot of the dE/dx log-likelihood for the 3D Kalman track with the highest ReMId value in the slice. This plot is for the far detector nonQE population. The total prediction is drawn as a red line with red total systematic error bands, with neutrino background drawn as a green line and the cosmic ray background drawn as a magenta line. The cosmic ray background distribution was determined from the out-of-time data in the NuMI trigger files. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. When the ratio is too large for the scale, the point and its error bars are not drawn. The simulation is oscillated using the values listed in Table 4.2 and setting $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The simulation is scaled down to match the exposure for the far detector data, 3.45×10^{20} POT.

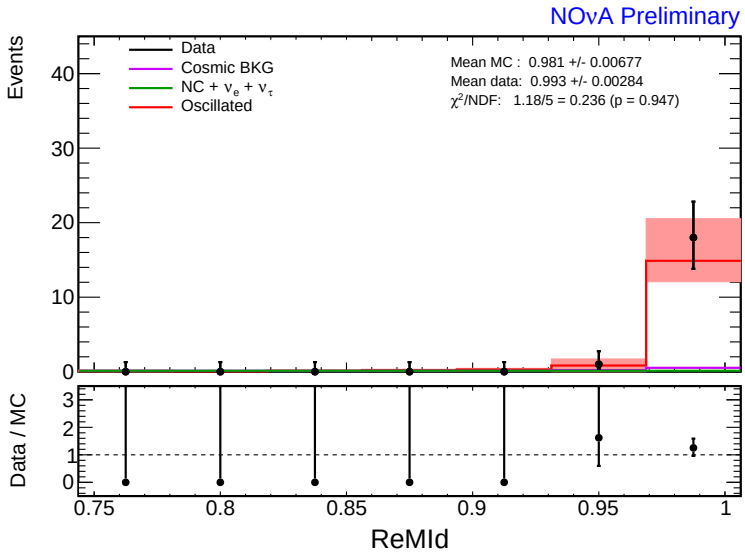


Figure 15.37: Plot of the ReMId value for the 3D Kalman track with the highest ReMId value in the slice. This plot is for the far detector QE population. The total prediction is drawn as a red line with red total systematic error bands, with neutrino background drawn as a green line and the cosmic ray background drawn as a magenta line. The cosmic ray background distribution was determined from the out-of-time data in the NuMI trigger files. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is oscillated using the values listed in Table 4.2 and setting $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The simulation is scaled down to match the exposure for the far detector data, 3.45×10^{20} POT.

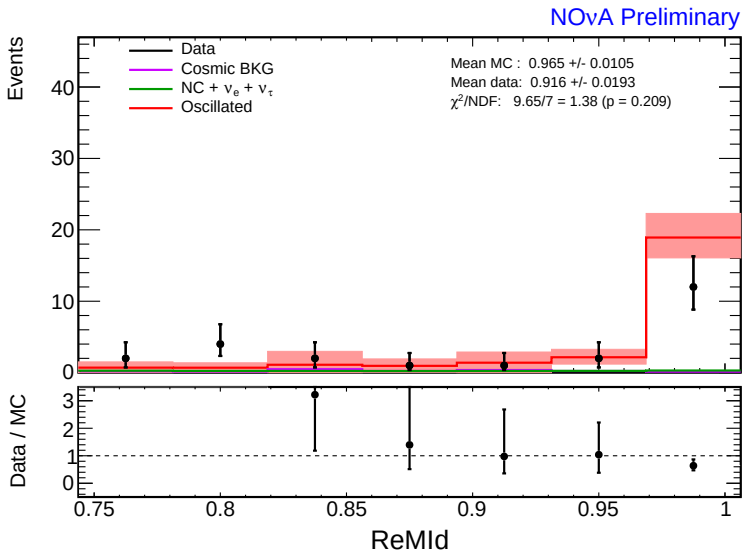


Figure 15.38: Plot of the ReMId value for the 3D Kalman track with the highest ReMId value in the slice. This plot is for the far detector nonQE population. The total prediction is drawn as a red line with red total systematic error bands, with neutrino background drawn as a green line and the cosmic ray background drawn as a magenta line. The cosmic ray background distribution was determined from the out-of-time data in the NuMI trigger files. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. When the ratio is too large for the scale, the point and its error bars are not drawn. The simulation is oscillated using the values listed in Table 4.2 and setting $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The simulation is scaled down to match the exposure for the far detector data, 3.45×10^{20} POT.

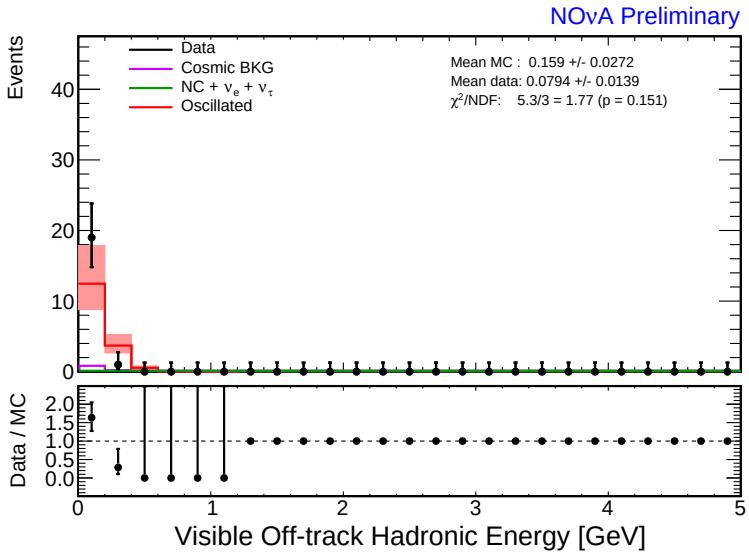


Figure 15.39: Plot of the sum of the visible energy (in GeV) of hits in the slice not on the 3D Kalman track with the highest ReMId value. This plot is for the far detector QE population. The total prediction is drawn as a red line with red total systematic error bands, with neutrino background drawn as a green line and the cosmic ray background drawn as a magenta line. The cosmic ray background distribution was determined from the out-of-time data in the NuMI trigger files. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is oscillated using the values listed in Table 4.2 and setting $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The simulation is scaled down to match the exposure for the far detector data, 3.45×10^{20} POT.

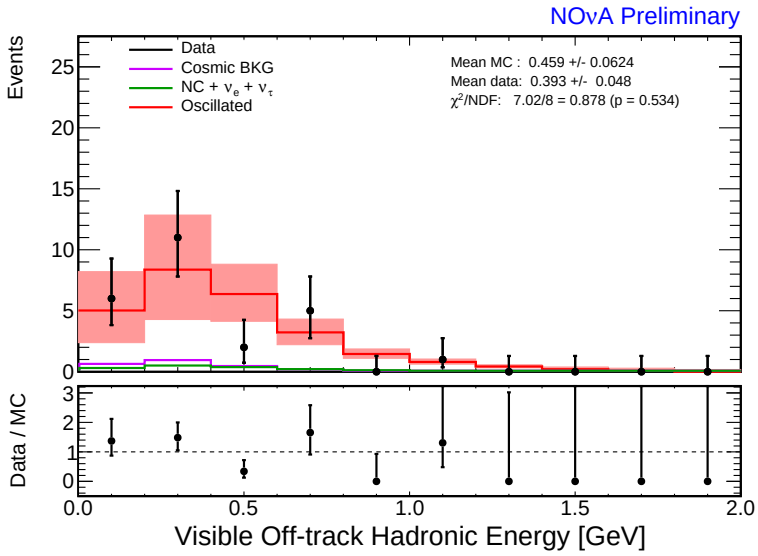


Figure 15.40: Plot of the sum of the visible energy (in GeV) of hits in the slice not on the 3D Kalman track with the highest ReMId value. This plot is for the far detector nonQE population. The total prediction is drawn as a red line with red total systematic error bands, with neutrino background drawn as a green line and the cosmic ray background drawn as a magenta line. The cosmic ray background distribution was determined from the out-of-time data in the NuMI trigger files. The data distribution is drawn as black points with statistical error bars. The bottom plot displays the ratio between the data and simulation distributions. The simulation is oscillated using the values listed in Table 4.2 and setting $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The simulation is scaled down to match the exposure for the far detector data, 3.45×10^{20} POT.

15.2 Results

Figure 15.41 displays the reconstructed neutrino energy spectrum for the data and the simulation with systematic error bands, for the QE population, assuming maximal oscillations. Figure 15.42 is for the nonQE population.

To measure $\sin^2\theta_{23}$ and $|\Delta m^2_{32}|$, we can perform a binned maximum-likelihood fit to the reconstructed energy spectra, marginalizing over our systematic errors. When we marginalize over a systematic error, we scan the allowed values and use the one which minimizes the χ^2 value. The χ^2 is penalized for values of the systematic error large with respect to the σ for each systematic error. Figure 15.43 displays the 90% 2D Gaussian confidence limits for this measurement. The best-fit for the combined sample was found to be $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49\times10^{-3}$ eV^2. The 90% confidence limits for these values are given by Figure 15.43. The space is symmetric about $\sin^2\theta_{23} = 0.5$, so it should be understood that there is another local minimum located near $\sin^2\theta_{23} = 0.4$.

Assuming a Gaussian distribution, however, is not the correct analysis technique. This is a low-statistics analysis that is constrained by the physical boundary of maximal mixing. Instead, the Feldman-Cousins approach[95] should be used. This approach involves creating thousands of pseudo-experiments with true values of $\sin^2\theta_{23}$ and $|\Delta m^2_{32}|$ distributed across the parameter space. Each of these pseudo-experiments uses fake "data" drawn with the proper Poisson statistics from the simulation. Also, each set of fake "data" is assigned random "true" values for each of the systematic errors. The maximum-likelihood fitting procedure is then performed for each pseudo-experiment. This allows us to build a map of $\Delta\chi^2$ distributions for each bin in the space of $\sin^2\theta_{23}$ and $|\Delta m^2_{32}|$. Finally, we can construct confidence limits for our real data that properly account for the low statistics and physical boundaries involved in the analysis. For more information about the Feldman-Cousins approach, see the paper[95] or the excellent discussion by Backhouse in his thesis[96]. Figure 15.44 displays the 90% 2D Feldman-Cousin confidence limits for this measurement, with the Gaussian confidence limits also drawn. Although the two approaches are philosophically different, in this case, they gave very similar results. The best-fit for the combined sample using the Feldman-Cousin approach was found to be $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49\times10^{-3}$ eV^2. These are the same values found using the Gaussian method. Again, the space

is symmetric about $\sin^2\theta_{23} = 0.5$, so there is another local minimum located near $\sin^2\theta_{23} = 0.4$. Table 15.4 lists the best fit values for the systematic errors from the marginalization procedure. None of the values are larger in magnitude than 1 σ. The systematic error for the far detector number of hadronic hits sees the largest shift at $+0.55\sigma$.

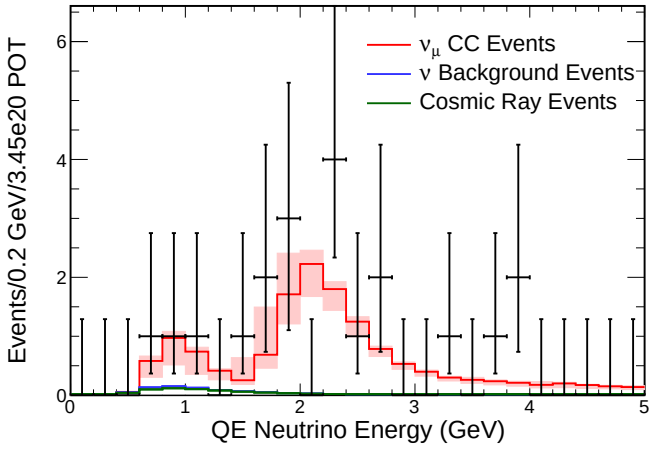


Figure 15.41: Plot of the reconstructed neutrino energy in GeV. This is for the QE population in the far detector. The simulation distribution is drawn as a red line with red total systematic error bands, with neutrino background drawn as a blue line and the cosmic ray background drawn as a green line. These are displayed as stacked distributions. The data is drawn as black points with statistical error bars. The simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2. The simulation is scaled down to match the data POT, 3.45×10^{20} POT.

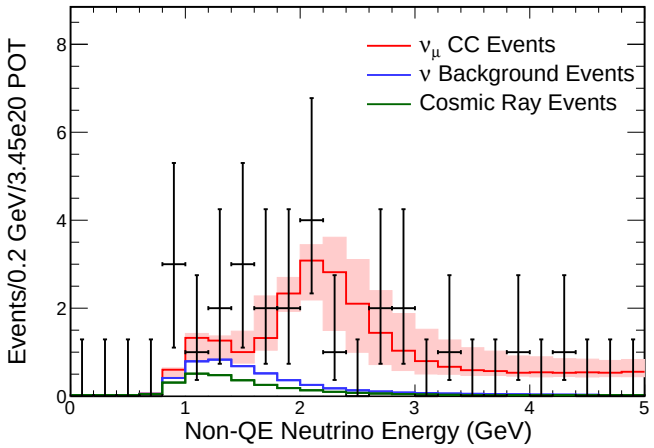


Figure 15.42: Plot of the reconstructed neutrino energy in GeV. This is for the nonQE population in the far detector. The simulation distribution is drawn as a red line with red total systematic error bands, with neutrino background drawn as a blue line and the cosmic ray background drawn as a green line. These are displayed as stacked distributions. The data is drawn as black points with statistical error bars. The simulation is oscillated using the values listed in Table 4.2 and setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2. The simulation is scaled down to match the data POT, 3.45×10^{20} POT.

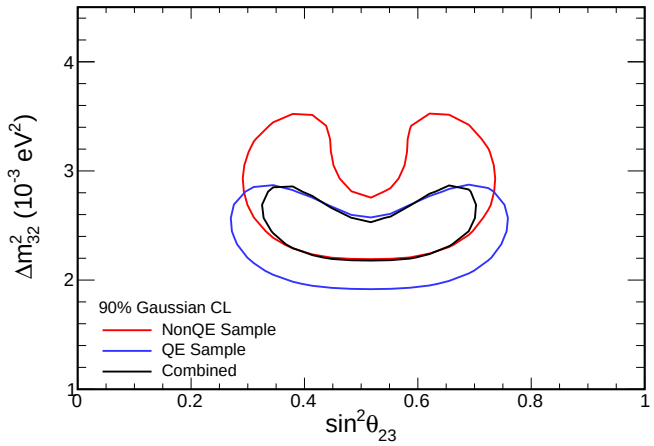


Figure 15.43: Plot of the 90% 2D Gaussian confidence limits for the analysis measurement of $\sin^2\theta_{23}$, on the horizontal axis, and Δm^2_{32}, on the vertical axis in units of 10^{-3} eV^2. This plot assumes normal hierarchy. The red line is for the nonQE population alone; the blue line is for the QE population. The black line shows the result when fitting both samples together.

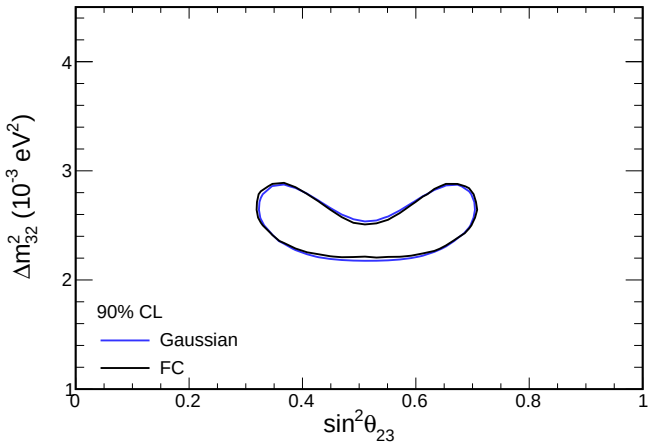


Figure 15.44: Plot of the 90% 2D confidence limits for the analysis measurement of $\sin^2\theta_{23}$, on the horizontal axis, and Δm^2_{32}, on the vertical axis in units of 10^{-3} eV^2. This plot assumes normal hierarchy. The black line shows the results of using the Feldman-Cousin approach. The blue line uses Gaussian assumptions.

Table 15.4: Values of the best fit values for each systematic error from the marginalization procedure using the Feldman-Cousin approach. The best fit value for each systematic error is given in terms of the σ defined for that error. For more information on the systematic errors, see Chapter 14.

Systematic Error	Best Fit [σ]
Beam Sim.	+0.15
Axial Mass - CC QE	+0.06
Axial Mass - CC Res.	−0.04
Vector Mass - CC Res.	−0.02
Axial Mass - NC Res.	+0.06
Vector Mass - NC Res.	+0.01
Axial Mass - NC Elas.	+0.04
Small Neutrino Sim.	−0.03
Abs. Norm.	+0.00
Rel. Norm.	+0.01
ND Abs. E Scale	+0.29
FD Abs. E Scale	−0.13
ND Had. Hits	−0.23
FD Had. Hits	+0.55

We can look at what the predicted event count would be using the best fit values. Table 15.5 lists the expected event counts using the best fit values for the oscillation parameters, as well as the expected event counts when using the best fit values for both the oscillation parameters and the systematic errors. For the QE population, using the best fit values increases the expected event count to be closer to that seen in the data. For the nonQE population, the effect on the expected event count is somewhat smaller but still consistent with the data.

Table 15.5: Predicted and measured event counts for the far detector. The total POT for the far detector is 3.45×10^{20} POT and this corresponds to variable detector configurations. The counts for each population only include events with reconstructed neutrino energy between 0 and 5 GeV. The total number of data events for each population is listed first. The predicted total event count for each population under the hypothesis of no neutrino oscillations is given next. The next line corresponds to the hypothesis of maximal mixing, setting $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2 as well as using the oscillation parameters listed in Table 4.2. When using the oscillation parameters from the best fit to the data, we instead set $\sin^2 \theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. When using the best fit values for the systematic errors, the values listed in Table 15.4 are used.

	QE	NonQE
Data	**20**	**25**
No Oscillations	89.2	122.3
Maximal Mixing	14.3	24.2
Best Fit Oscillation Parm.	15.5	25.5
Best Fit Oscillation Parm. and Systematic Errors	17.2	23.9

We can also look at the energy spectra for these different parameter choices. Figure 15.45 displays the spectra for the QE population; Figure 15.46 is for the nonQE population. We can see that by using the best fit oscillation parameters, which don't correspond to maximal mixing, we expect more events in the dip region. By also using the best fit values for the systematic errors, we expect more QE events in the dip region and that the energy for the nonQE events is shifted to lower values.

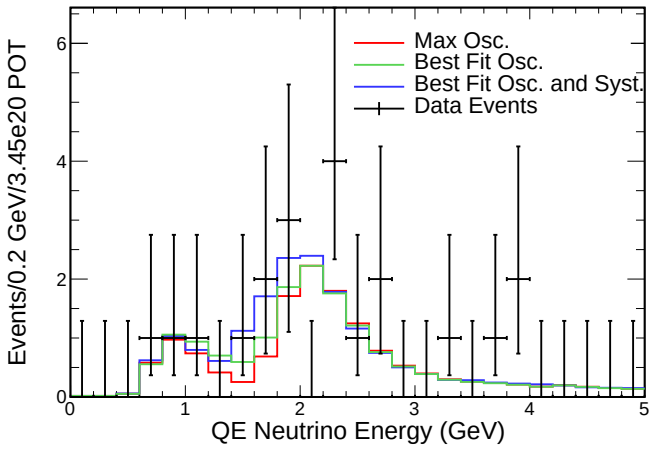


Figure 15.45: Plot of the reconstructed neutrino energy in GeV. This is for the QE population in the far detector. The total predicted simulated spectrum using maximal mixing is drawn as a red line. The total predicted simulated spectrum using the best fit values for the oscillation is drawn as a green line. The total predicted simulated spectrum using the best fit values for the oscillation and systematic error values is drawn as a blue line. The data is drawn as black points with statistical error bars. The simulation is oscillated using the values listed in Table 4.2. For the case of maximal mixing, $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2. When using best fit values for the oscillation, $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. When using the best fit values for the systematic errors, the values listed in Table 15.4 are used. The simulation is scaled down to match the data POT, 3.45×10^{20} POT.

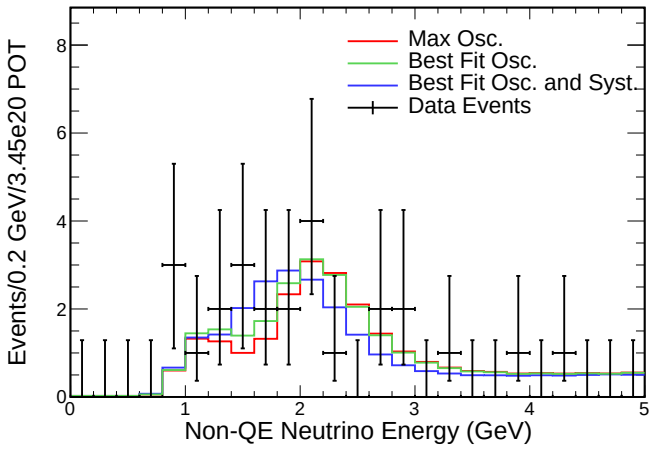


Figure 15.46: Plot of the reconstructed neutrino energy in GeV. This is for the nonQE population in the far detector. The total predicted simulated spectrum using maximal mixing is drawn as a red line. The total predicted simulated spectrum using the best fit values for the oscillation is drawn as a green line. The total predicted simulated spectrum using the best fit values for the oscillation and systematic error values is drawn as a blue line. The data is drawn as black points with statistical error bars. The simulation is oscillated using the values listed in Table 4.2. For the case of maximal mixing, $\theta_{23} = \pi/4$ and $|\Delta m^2_{32}| = 2.4 \times 10^{-3}$ eV^2. When using best fit values for the oscillation, $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. When using the best fit values for the systematic errors, the values listed in Table 15.4 are used. The simulation is scaled down to match the data POT, 3.45×10^{20} POT.

Finally, we can compare this result to results from other experiments. Figure 15.47 displays the 90% confidence limits for this thesis as well as the results from the MINOS experiment[10] and the T2K experiment[11]. Note the restricted axes in relation to previous plots of the confidence limits for this thesis. The result from this thesis is less precise than the other results; however, it is already quite comparable. The MINOS and T2K experiments are both mature experiments with a relatively large amount of statistics and time to fully understand their detectors and simulation. This result uses only 7% of NOνA's expected exposure; with increased statistics, the measurement will become much more precise. Also, understanding of the differences in hadronic response seen between the data and simulation will allow a large reduction in the systematic error of this result. Work to understand the difference is actively ongoing; a resolution is expected to be ready in time for the next round of NOνA results.

NOνA will soon be able to improve the world's knowledge of these two parameters. Figure 15.48 shows the predicted sensitivity to the oscillation parameters with increased amounts of data. Note the different axes in relation to previous plots of confidence limits shown in this thesis. These predicted sensitivities do not include systematic errors or cosmic background events, although the cuts to remove cosmic background have been applied to the signal. The sensitivities are created using only the simulation. Two truth assumptions are shown. These sensitivities show that NOνA will soon have results more precise than those currently published by other experiments. Also, if the value of $\sin^2\theta_{23}$ is approximately 0.4 or 0.6, NOνA will be able to rule out maximal mixing. If the value is closer to 0.5, this will become more difficult.

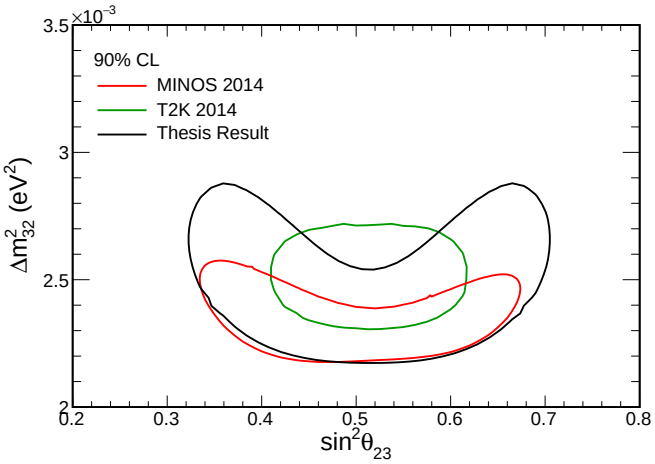


Figure 15.47: Plot of the 90% 2D confidence limits for the analysis measurement of $\sin^2\theta_{23}$, on the horizontal axis, and Δm^2_{32}, on the vertical axis in units of 10^{-3} eV^2. This plot assumes normal hierarchy. The black line shows the results of this thesis. The red line shows the 2014 results from the MINOS experiment[10]. The green line shows the 2014 results from the T2K experiment[11].

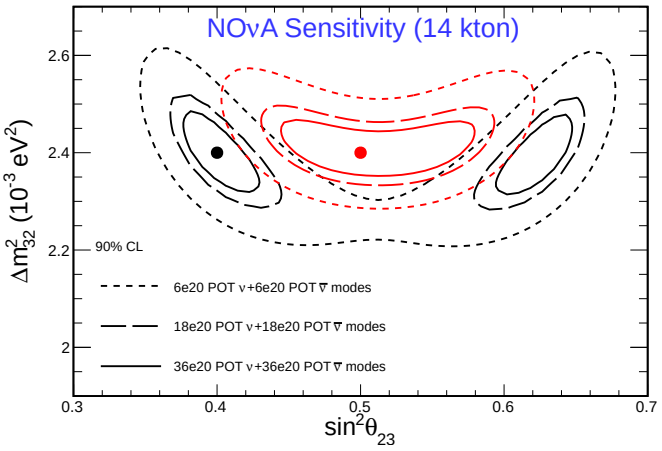


Figure 15.48: Plot of the predicted sensitivity for the NOνA measurement of $\sin^2\theta_{23}$, on the horizontal axis, and Δm^2_{32}, on the vertical axis in units of 10^{-3} eV^2. It shows three different exposures of neutrino and anti-neutrino mode running, including 36 $\times 10^{20}$ POT for each beam configuration. This corresponds to six years of running with the full 14 kton detector and a beam power of 700 kW. Two truth assumptions are included. They are color-coded and the dots indicate the true values. No systematic errors are included. The anti-neutrino mode analysis cuts are assumed to be exactly the same as the neutrino mode analysis cuts. Cosmic background events are not included, although the cuts to reject cosmic background are applied to the signal to account for inefficiencies. The sensitivity is based on simulation only. This plot assumes normal hierarchy.

Chapter 16

Conclusion and Discussion

The NOνA detectors are now constructed and fully operational. The NuMI neutrino beam has successfully undergone its first phase of upgrades. It regularly operated with a beam power of 400 kW and reached a record beam power of 521 kW. Future upgrades are on track to increase the power to the final specification of 700 kW. The NuMI beam had 85% uptime for this period of data-taking and exceeded design goals for POT delivered. The NOνA far detector was completed in July 2014. It is the largest free-standing plastic structure in the world. It has more than 99.5% operational channels from a total of 344,064 and a greater than 95% uptime. The NOνA near detector was completed August 2014. It runs with an average of 98% of channels considered usable for physics analysis. Neutrino interactions were observed within seconds of turning on the near detector. Both detectors are working as expected and performing well.

This analysis uses the first data from the NOνA experiment to make a 2D measurement of $\sin^2\theta_{23}$ and $|\Delta m^2_{32}|$. The hypothesis that neutrinos do not oscillate is easily ruled out. The best fit parameters were found to be $\sin^2\theta_{23} = 0.61$ and $|\Delta m^2_{32}| = 2.49 \times 10^{-3}$ eV^2. The 90% confidence limits using the Feldman-Cousin approach are shown in Figure 15.44. A comparison of these results with those of other, more mature experiments is shown in Figure 15.47. We can see that this analysis agrees well with previous results. This shows that the detectors and analysis techniques are working as expected.

This analysis did not make more precise measurements than those of the other experiments. However, it is competitive for a "first light" measurement and shows the detector is functioning as designed. It used only 7% of the total expected exposure

for NOνA. Also, work to fully understand the differences seen between the data and the simulation will be needed for future analyses to be able to make precision measurements of the oscillation parameters by reducing systematic errors. This work is actively ongoing.

Figure 15.48 shows expected measurement sensitivities for NOνA with more data. It shows that NOνA will soon make measurements of these parameters more precise than the current results of other experiments. If $\sin^2\theta_{23}$ is 0.4 or 0.6, NOνA will be able to rule out maximal mixing with its expected final dataset. If $\sin^2\theta_{23}$ is closer to 0.5, this will be more difficult.

16.1 Future Improvements

There are a number of issues that could be improved for future analyses. For instance, when creating a fit for muon energy from track length, one could use more splines for higher precision or one could switch to using a Bethe-Bloch curve. If one wanted use Bethe-Bloch, one should look at the implementations already existent in Calibration. This implementation could be improved; it currently makes the approximation that the whole detector is made out of density-weighted scintillator.

The hadronic energy estimation could be improved. It is important to first understand the differences between the data and simulation in the hadronic sector before making more complicated energy estimation schemes. After these differences are resolved, one could probably make gains in energy resolution by using more sophisticated reconstruction on the hadronic hits. For instance, one could run a second tracking attempt on only the hadronic cluster. This reconstruction could use the muon start position as extra information. The second tracker could be optimized to find shorter tracks with more sharp kinks than the muon tracks generally have. With hadronic tracks, one could then attempt to identify particle types such as a proton or charged pion. With reliable tracks and particle identification in the hadronic cluster, one could then develop a proton energy estimation, etc.

Cosmic background rejection could be significantly improved with further study. I think it is very likely that a few common failure modes of either reconstruction or bad data quality make up a large portion of the remaining cosmic background. Additional

event display scanning could help to understand and categorize these failure modes and more powerful cuts could probably be crafted to remove the background. Then, one might be able to loosen cuts currently being made, like the transverse momentum cut, to regain low energy signal. A cut on the maximum y position of the slice might be particularly useful. Also, utilizing more information about if the non-muon energy is better aligned with the start or end of the track could help to reject background neutrino interactions from cosmic rays.

Some improvements could be made to the software that detects low probability electronics failures. For example, the DCM edge metric is not sensitive enough in its present form to find cases where only a few DCMs are out of sync. This is an important failure mode that future analyses should reject.

Work is actively ongoing for the most important improvements to the NOνA analysis. A second round of analysis results from NOνA will be published next year. The future of NOνA promises to hold important scientific results soon!

References

[1] Particle Data Group, K. Olive et al., Chin.Phys. C38 (2014) 090001.

[2] C. Zhang et al., Recent Results from Daya Bay, 2014, Slides located at https://indico.fnal.gov/getFile.py/access?contribId=256&sessionId=15&resId=0&materialId=slides&confId=8022.

[3] J.W. Cooper, CD-4 Closeout Project Overview Presentation, NOvA Collaboration Internal Document 12152-v1, 2014, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=12152.

[4] T. Miao, As-Build Near Detector Locations, NOvA Collaboration Internal Document 11902-v3, 2014, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=11902.

[5] J.W. Cooper, NOνA (E929) NuMI Off-Axis ν_e Apperance Experiment Technical Design Report, 2007, This document can be located at http://www-nova.fnal.gov/nova_cd2_review/tdr_oct_23/tdr.htm.

[6] K. Kephart, Near Detector Under Ground (NDUG) Parameters, NOvA Collaboration Internal Document 9860-v2, 2014, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=9860.

[7] N. Raddatz, Fiducial Mass, NOvA Collaboration Internal Document 13237-v1, 2015, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=13237.

[8] J. Musser, Detector Mass Modeling and Systematic Errors, NOvA Collaboration Internal Document 13655-v3, 2015, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=13655.

[9] R.M. Zwaska, Accelerator systems and instrumentation for the NuMI neutrino beam, PhD thesis, University of Texas, Austin, 2005.

[10] P. Adamson et al., Phys. Rev. Lett. 112 (2014) 191801.

[11] K. Abe et al., Phys. Rev. Lett. 112 (2014) 181801.

[12] W. Pauli, Offener Brief an die Gruppe der Radioaktiven bei der Gauvereins-Tagung zu Tübingen, [Open letter to the group of radioactive people at the Gauverein meeting in Tübingen], 1930.

[13] C.L. Cowan et al., Science 124 (1956) 103.

[14] G. Danby et al., Phys. Rev. Lett. 9 (1962) 36.

[15] ALEPH, D. Decamp et al., Phys. Lett. B231 (1989) 519.

[16] DONUT, K. Kodama et al., Phys. Lett. B504 (2001) 218, hep-ex/0012035.

[17] ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration, SLD Collaboration, LEP Electroweak Working Group, SLD Electroweak Group, SLD Heavy Flavour Group, S. Schael et al., Phys.Rept. 427 (2006) 257, hep-ex/0509008.

[18] B. Pontecorvo, Sov. Phys. JETP 7 (1958) 172.

[19] Z. Maki, M. Nakagawa and S. Sakata, Prog. Theor. Phys. 28 (1962) 870.

[20] R. Davis, Jr., D.S. Harmer and K.C. Hoffman, Phys. Rev. Lett. 20 (1968) 1205.

[21] Super-Kamiokande, Y. Fukuda et al., Phys. Rev. Lett. 82 (1999) 2644, hep-ex/9812014.

[22] SNO, J. Boger et al., Nucl. Instrum. Meth. A449 (2000) 172, nucl-ex/9910016.

[23] KamLAND, K. Eguchi et al., Phys. Rev. Lett. 90 (2003) 021802, hep-ex/0212021.

[24] Particle Data Group, J. Beringer et al., Phys. Rev. D 86 (2012) 010001.

[25] B.P. Speakman, Atmospheric Electron Neutrinos in the MINOS Far Detector, PhD thesis, University of Minnesota, 2007.

[26] P. Adamson et al., Phys. Rev. Lett. 112 (2014) 191801.

[27] H. Nunokawa, S. Parke and R.Z. Funchal, Phys. Rev. D 72 (2005) 013009.

[28] D.I. Scully, Neutrino Induced Coherent Pion Production, PhD thesis, University of Warwick, 2013.

[29] S.L. Mufson, Scintillator Update, NOvA Collaboration Internal Document 8541-v2, 2013, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=8541.

[30] V. Guarino, Near Detector Muon Steel Construction and Assembly, NOvA Collaboration Internal Document 2436-v1, 2007, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=2436.

[31] D.F. Friend, NDUG Block, MUON Catcher and Mini-Block Transport Layouts, NOvA Collaboration Internal Document 9238-v5, 2013, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=9238.

[32] E. Arrieta-Diaz, Plans for Fiducial Mass Calculation, NOvA Collaboration Internal Document 12362-v2, 2014, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=12362.

[33] E. Arrieta-Diaz, Fiducial Mass Calculation, NOvA Collaboration Internal Document 12468-v1, 2014, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=12468.

[34] L.M. Mualem, APDs, Vertical Slice, Electronics and Noise, NOvA Collaboration Internal Document 6950-v1, 2012, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=6950.

[35] T.T. Bohlen et al., Nuclear Data Sheets 120 (2014) 211.

[36] A. Ferrari et al., CERN-2005-10 INFN/TC_05/11, SLAC-R-773 (2005).

[37] C. Andreopoulos et al., Nucl. Instrum. Meth. A614 (2010) 87.

[38] C. Hagmann et al., 2 (2007) 1143.

[39] S. Agostinelli et al., Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 506 (2003) 250 .

[40] J. Allison et al., Nuclear Science, IEEE Transactions on 53 (2006) 270.

[41] A. Kaidalov and K. Ter-Martirosyan, Physics Letters B 117 (1982) 247, cited By 212.

[42] M. Guthrie et al., Nucl. Instr. Meth. 66 (1968).

[43] H. Bertini et al., Nucl. Phys. A 169 (1971).

[44] Geant 4 Collaboration, Reference Physics Lists, 2013, This website can be located at http://geant4.cern.ch/support/proc_mod_catalog/physics_lists/referencePL.shtml.

[45] BackTracker Class Reference, 2015, This document can be located at http://nusoft.fnal.gov/nova/novasoft/doxygen/html/classcheat_1_1BackTracker.html.

[46] S. Lein, CloseInTime Backtracker Function, NOvA Collaboration Internal Document 9092-v1, 2013, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=9092.

[47] J.M. Paley, The NOvA BadChannels Package, NOvA Collaboration Internal Document 12771-v1, 2015, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=12771.

[48] M.D. Baird, A Side By Side Comparison of Slicer, Cosmic Slicer, and Slicer4D, NOvA Collaboration Internal Document 9195-v1, 2013, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=9195.

[49] R. Fruhwirth, Nucl.Instrum.Meth. A262 (1987) 444.

[50] M. Regler and R. Frhwirth, Reconstruction of charged tracks, Techniques and Concepts of High-Energy Physics V, edited by T. Ferbel, , NATO ASI Series Vol. 20, pp. 407–499, Springer US, 1990.

[51] N. Raddatz, Track Reconstruction with KalmanTrack, NOvA Collaboration Internal Document 6828-v1, 2011, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=6828.

[52] N.J. Raddatz, Measurement of Muon Neutrino Disapperance with Non-fiducial Interactions in the NOvA Experiment, PhD thesis, University of Minnesota, 2015.

[53] N.S. Altman, The American Statistician 46 (1992) 175.

[54] N.J. Raddatz, Reconstructed Muon Indentification, NOvA Collaboration Internal Document 11206-v1, 2014, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=11206.

[55] N.J. Raddatz, Quasi-elastic Particle Indentification, NOvA Collaboration Internal Document 11207-v1, 2014, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=11207.

[56] R. Brun et al., Nucl. Inst. and Meth. in Phys. Res. A 389 (1997) 81.

[57] D.R. Rocco, CAF ROOT Macros and PyROOT, 2014, This wiki can be located at https://cdcvs.fnal.gov/redmine/projects/novaart/wiki/CAF_ROOT_Macros_and_PyROOT.

[58] D.R. Rocco, CAF Tree Structure and Variable Listing, 2014, This wiki can be located at https://cdcvs.fnal.gov/redmine/projects/novaart/wiki/CAF_Tree_Structure_and_Variable_Listing.

[59] C. Backhouse and A. Radovic, The Attenuation and Threshold Correction of the NOνA detectors, NOvA Collaboration Internal Document 13579-v12, 2015, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=13579.

[60] L. Vinton, Calorimetric Energy Scale Calibration of the NOνA detectors, NOvA Collaboration Internal Document 13579-v12, 2015, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=13579.

[61] K. Sachdev, MRCC Tech Note, NOvA Collaboration Internal Document 9729-v3, 2013, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=9729.

[62] M. Betancourt, Study of the Quasi-Elastic Scattering in the NOνA Detector Prototype, PhD thesis, University of Minnesota, 2013.

[63] J. Coelho, FarDet Good Data Selection, NOvA Collaboration Internal Document 12733-v1, 2015, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=12733.

[64] B. Chowdhury, ND Good Run Update, NOvA Collaboration Internal Document 12941-v1, 2015, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=12941.

[65] X. Bu et al., Spill Level Data Quality, NOvA Collaboration Internal Document 12437-v2, 2015, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=12437.

[66] L. Goodenough, Discussion of Beam Quality Cuts for Use in POT Counting, NOvA Collaboration Internal Document 12474-v1, 2014, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=12474.

[67] M. Messier, Timing Shift Summary, NOvA Collaboration Internal Document 13364-v1, 2015, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=13364.

[68] J. Lozier, ModularExtrap Technical Note, NOvA Collaboration Internal Document 12563-v1, 2014, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=12563.

[69] E.C. Mur, Flux Systematics and Extrapolation, NOvA Collaboration Internal Document 13271-v1, 2015, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=13271.

[70] NA49, C. Alt et al., Eur.Phys.J. C49 (2007) 897, hep-ex/0606028.

[71] . CERN, NA49 - Large Acceptance Hadron Detector for an Investigation of Pb-induced Reactions at the CERN SPS, 1997, This document can be located at http://na49info.web.cern.ch/na49info/.

[72] A. Radovic, Fixed Target Motivated Hadron Production Uncertainties at NOvA, NOvA Collaboration Internal Document 12181-v2, 2014, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=12181.

[73] A. Radovic, NA49 vs. FLUKA Motivated Hadron Production Uncertainties For the First Analysis, NOvA Collaboration Internal Document 12691-v2, 2015, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=12691.

[74] K.K. Maan, Flux Systematics for NOvA, NOvA Collaboration Internal Document 12961-v1, 2015, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=12961.

[75] L. Alverez-Ruso et al., The GENIE Neutrino Monte Carlo Generator Physics and User Manual, 2015, This document can be located at http://genie.hepforge.org/manuals/GENIE_PhysicsAndUserManual_v2.10.00a.pdf.

[76] M. Pia, Proceeding of the CHEP 2000 Conference, 2000.

[77] H. Fesefeld, Simulation of hadronic showers, physics and applications, 1985, Technical Reprot PITHA 85-02.

[78] D.H. Wright et al., AIP Conference Proceedings 896 (2007) 11.

[79] P. Degtyarenko, M. Kossov and H.P. Wellisch, European Physical Journal A 8 (2000) 217, cited By 40.

[80] P. Degtyarenko, M. Kossov and H.P. Wellisch, European Physical Journal A 9 (2000) 411, cited By 28.

[81] V. Uzhinsky et al., Proceeding of the MC2010 Monte Carlo Conference, 2010.

[82] D. Rocco, Alternative GEANT Physics Lists, NOvA Collaboration Internal Document 12989-v2, 2015, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=12989.

[83] N. Raddatz, Muon Energy Scale Systematic, NOvA Collaboration Internal Document 13637-v1, 2015, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=13637.

[84] B. Zamorano, Misalignment Systematic, NOvA Collaboration Internal Document 12671-v1, 2015, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=12671.

[85] B. Zamorano, Near Detector Misalignment Systematic, NOvA Collaboration Internal Document 13295-v2, 2015, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=13295.

[86] A. Radovic et al., A Technote Describing the Derivation and Size of NuMI Flux Uncertainties Used in the First NOvA Analyses, NOvA Collaboration Internal Document 13584-v6, 2015, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=13584.

[87] M. Baird, Summary of Pileup Problems for the First Analysis, NOvA Collaboration Internal Document 13293-v1, 2015, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=13293.

[88] J.A. Sepulveda-Quiroz, On Bad Channel Systematics, NOvA Collaboration Internal Document 13663-v3, 2015, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=13663.

[89] M. Tamsett, Calibration systematic uncertainties for the disappearance analysis, NOvA Collaboration Internal Document 13410-v1, 2015, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=13410.

[90] N. Raddatz, Data/MC dE/dx Differences, NOvA Collaboration Internal Document 13355-v1, 2015, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=13355.

[91] K. Sachdev, dE/dx Comparison in ND Data and MC, NOvA Collaboration Internal Document 13322-v2, 2015, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=13322.

[92] G.S. Davies, Pi-zero invariant mass cross-check, NOvA Collaboration Internal Document 13323-v2, 2015, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=13323.

[93] R. Patterson, Quick look at Michel electrons in the ND and FD (calibration), NOvA Collaboration Internal Document 13340-v1, 2015, This document can be located at http://nova-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=13340.

[94] J. Nieves, I. Ruiz Simo and M.J. Vicente Vacas, Phys. Lett. B721 (2013) 90, 1302.0703.

[95] G. Feldman and R. Cousins, Phys. Rev. D 57 (1998) 3873, physics/9711021.

[96] C. Backhouse, Measuring neutrino oscillation parameters using ν_μ disappearance in MINOS, PhD thesis, Oriel College, Oxford, 2011.

[97] C. Bassin et al., The Current Limits of Resolution for Surface Wave Tomography in North America, EOS Trans AGU, 81, F897, 2000, The main webpage is located at http://igppweb.ucsd.edu/ gabi/rem.html and the specific webpage for CRUST 2.0 is located at http://igppweb.ucsd.edu/ gabi/crust2.html.

[98] M. Ester et al., pp. 226–231, AAAI Press, 1996.

Appendix A

Glossary and Acronyms

Care has been taken in this thesis to minimize the use of jargon and acronyms, but this cannot always be achieved. This appendix defines jargon terms in a glossary, and contains a table of acronyms and their meaning.

A.1 Glossary

AngleQEE – Angle Quasielastic Energy is a reconstructed neutrino energy based on the muon track length and the angle of the muon track relative to the beam direction under the assumption of a quasielastic ν_μ CC interaction.

Module – A module is a C++ class which performs a specific task on the input file.

Run – A period of detector running where the manually-set detector configuration does not change. It is composed of subruns and ends when either there are 64 subruns, the total duration of the run is 24 hours, or the detector stops taking data.

PID – A particle or event identification metric.

Subrun – A subdivision of a run. Subruns end when the duration of the subrun is 1 hour, the file size is 1 GB, or the detector stops taking data.

TrkQEE – Track Quasielastic Energy is a reconstructed neutrino energy based on the muon track length and the visible hadronic energy. The hadronic energy fit uses the assumption of a quasielastic ν_μ CC interaction.

TrkNonQEE – Track Non-Quasielastic Energy is a reconstructed neutrino energy based on the muon track length and the visible hadronic energy. The hadronic energy fit uses the assumption of a non-quasielastic ν_μ CC interaction.

A.2 Acronyms

Table A.1: Acronyms

Acronym	Meaning
ADC	Analog-to-digital converter
AngleQEE	Angle Quasielastic Energy
CC	Charged Current
NC	Neutral Current
NDF	Number of Degrees of Freedom
PID	Particle Identification
reco	reconstructed
TDC	Time-to-digital converter
TrkNonQEE	Track Non-Quasielastic Energy
TrkQEE	Track Quasielastic Energy
ν_e	Electron Neutrino
ν_μ	Muon Neutrino
ν_τ	Tau Neutrino

Appendix B

Density of the Earth

Matter effects, discussed in Section 4.3, rely on ρ, the average density of the earth through which the neutrinos travel. NOνA uses the value 2.84 g/cm^3. This value assumes a spherical earth and is based on the CRUST 2.0 model[97]. This model provides measured average earth densities as a function of distance below the surface of the earth. Alexander Radovic created a simple polynomial fit to the measured values. The average neutrino depth underground going from the near detector to the far detector is 9.38 km. When this value is used in the simple polynomial fit, 2.84 g/cm^3 is the result. This is also the answer Alexander obtains when he integrates over the path of the neutrinos instead of simply using an average depth.

Appendix C

Slicer Algorithm Details

To group cell hits into clusters, Slicer uses a score function to determine the distance between cell hits in space and time. Hits deemed to be close are grouped together and those deemed far away are not. The algorithm was inspired by this paper[98]. Each hit is determined to be either a core point or a border point. Core points are those that have more than the minimum number of neighbors; border points do not. Border points can only be included in a cluster if they are a neighbor of a core point. Neighbors are defined as having $\epsilon < 2.0$ for far detector and $\epsilon < 5.0$ for near detector. ϵ is defined as:

$$\epsilon = \left(\frac{\Delta T - \Delta \overrightarrow{r}/c}{T_{res}}\right)^2 + \left(\frac{\Delta Z}{D_{pen}}\right)^2 + \left(\frac{\Delta XY}{D_{pen}}\right)^2 + \left(\frac{PE_{pen}}{PE}\right)^5$$

where T_{res} is the timing resolution of two hits, D_{pen} is a distance penalty, PE_{pen} is an energy penalty, and PE is the number of photoelectrons for both hits added in quadrature. The ΔT and $\Delta \overrightarrow{r}$ values are the distance between the two hits in time and space. ΔZ is the distance in the z direction between the two points; ΔXY is the distance in the other direction for hits in the same view. For hits in different views, ΔXY is zero.

The algorithm begins by looping over all points. When it finds a core point, it begins a cluster. The core point and all of its neighbors are added to the cluster. The algorithm then asks if any of the neighbors added to the cluster are core points. If they are, then the neighbors of the new core points are added to the cluster. This process continues until the neighbors of all core points are contained within the cluster. The

algorithm then determines if any remaining hits not within a cluster are core points, creates a new cluster, and continues. Any point that does not get clustered is defined as a noise hit. For more information, see the technical note[48].

Appendix D

Fitting Peaks, not Means

There are various ways to take a 2D histogram and condense it into a 1D graph that can be fit. Originally, this was done by using ROOT's Profile function, which takes each bin in x and finds the mean value in y. The fits that result from this method describe the bulk distribution best but can fail to describe the peak of the distribution well. Means are, by construction, sensitive to the tails of the distribution. Another approach is to use the peak of the y distribution for each bin in x. This is conceptually better if one distrusts the tails and does not want the fit to be overly influenced by them. This is the approach that has taken.

The procedure used to make the 1D graph is to first loop over each bin in x. For each bin slice in x, quality cuts are made requiring at least 30 entries in the slice to make a point. Then the y bin with the most contents is found and the width at half maximum on each side of the highest bin is defined. These values are then used to initialize a Gaussian fit. The mean of the resulting Gaussian fit is set as the graph point; the error on the mean, reported by the fit, is the size of the error bars used.

The fit range is chosen to be a multiplier times the width at half height. With high statistics samples (approx. 10 million neutrino interactions), a multiplier of 1.5 often works well. For the hadronic fits, variable binning is used to reduce the effect of lower statistics in the high visible energy regions. Still, there are a few regions where a multiplier of 1.5 is not sufficient. This is primarily seen in the high visible energy region of QE hadronic energy. In these bins, the distribution does not peak strongly enough so that width at half-height is a reliable metric. For now, this has been dealt with by

looking bin-by-bin in regions of concern and testing the effects of different multipliers. Then regional multipliers are assigned as high as 25. In the future, it would be nice if this could be a more automated and robust system. One way to do this would simply have even more statistics but one would probably need at least 20 million neutrino interactions to do a significantly better job in the troublesome regions. Otherwise, one could increase the x bin size even more or make the variable binning choices based directly on number of entries.

Appendix E

Using Spline Fits

Arbitrary functional forms can be described by piecing together straight-line segments required to connect at join locations. The spline fits I used have an offset, the slopes of the lines, and the join locations as free parameters of the fit.

For each fit, I looked individually at using a single line up to 4 spline fits. I found that using 4 spline fits described the far detector muon track length distribution and both far detector hadronic energy distributions significantly better than fits with less free parameters. The near detector muon catcher fit required only a simple linear fit to capture the structure. The near detector hadronic fits worked best with 2 spline fits.

This is the functional form of an arbitrary linear fit:

```
 y = slope1 * x + offset ;
```

This is the functional form of an arbitrary two spline fit:

```
if (x < stitch1 ){
 y = slope1 * x + offset ;
}
else {
 y = slope2 * x +
   ((slope1-slope2) * stitch1 + offset) ;
```

```
}
```

This is the functional form of an arbitrary four spline fit:

```
if (x < stitch1 ){
 y = slope1 * x + offset ;
}
elseif (x < stitch2 ){
 y = slope2 * x +
   ((slope1-slope2) * stitch1 + offset) ;
}
elseif (x < stitch3 ){
 y = slope3 * x +
   ((slope2-slope3) * stitch2 +
    (slope1-slope2) * stitch1 + offset) ;
}
else {
 y = slope4 * x +
   ((slope3-slope4) * stitch3 +
    (slope2-slope3) * stitch2 +
    (slope1-slope2) * stitch1 + offset) ;
}
```

When doing the fits, one must consider carefully if the fit worked well or if better initial estimates of parameter values would result in a better fit. This is one reason to make the progression from linear up to 4 spline fits - it allows the user to track the subtle changes in slope and join locations. Sometimes the fits can fail to fit well if not provided with very enlightened initial parameters.

Appendix F

Fit Populations

When one performs a fit, they need to choose what population to fit over. One could fit a super-pure population that was chosen by utilizing truth information or one could fit a population using our actual reconstruction-based cuts. Optimizing fits over the actual reconstruction-based population that one will use in the analysis will result in the best energy resolution. However, this can also cause problems. The reconstruction selection used depends in some part on the energy fits (for instance, QePId compares TrkQEE and AngleQEE energies to determine if a slice is QE). This means that to do the fits and selections properly, one would need to find energy fits, determine selection criteria, and then repeat until convergence was obtained. This method would take more time.

I have used a primarily reconstruction-based population with a few truth cuts to fit over. The reconstruction cuts were picked to match those of the official ν_μ CC analysis group and are not identical to the reconstruction used in the analysis presented in this thesis. However, the effect on final energy resolutions will be small.

There are many criteria that are used to define the fitting populations. Truth information is used to require that the neutrino interaction is a ν_μ charged current interaction. Furthermore, the truth is used to distinguish between if the interaction was QE or non-QE when the hadronic energies are fit. It is also required that the primary muon from the neutrino interaction has at least 3 hits in each view.

For the far detector, the slice must pass quality cuts, be contained, pass the cut to identify ν_μ charged current interactions, and pass cosmic rejection cuts. For the near detector, the definition of contained is different and no cosmic rejection cuts are applied.

The quality cuts used by the official ν_μ CC analysis group were applied. It is required that at least one 3D Kalman track with a ReMId value greater than zero is associated with the slice. Also, the NumuEnergy TrkNonQEE must be greater than zero. The number of hits in the slice must be larger than 20. The number of continuous planes in the slice must be larger than four. The number of Cosmic tracks must be larger than zero.

The far detector containment used by the official ν_μ CC analysis group has a number of criteria. The slice must be more than one cell away from the edge of the detector. The slice must be more than one plane away from the front of the detector and more than one plane away from the back of the detector. The 3D Kalman track with the highest ReMId value must be more than 10 cells away from the edge of the detector, based on the projected track direction at the start and end of the track. Similarly, the Cosmic track must be at least one cell away from the edge of the detector, based on the projected track direction at the start and end of the track.

The near detector containment is somewhat different than that used for the far detector. Again, the slice must be more than one cell away from the edge of the detector. The first plane of the slice must be greater than 1 and the last plane of the slice must be less than 212. The start position of the 3D Kalman track with the highest ReMId value must have a z position of less than 1150 cm. To avoid traveling in uninstrumented area when transitioning from the fully active region to the muon catcher, the z position of the 3D Kalman track with the highest ReMId value must be less than 1275 cm or have a y position in the transition plane of less than 55 cm. The visible off-track hadronic energy in the slice contained in the transition plane and the muon catcher must be less than 0.03 GeV. Finally, the 3D Kalman track with the highest ReMId value must be more than 4 cells away from the edge of the detector, based on the projected track direction at end of the track. The distance from the start of the track to the edge of the detector along the projected track direction must be more than 8 cells.

To be considered a ν_μ charged current interaction, the highest ReMId value of a 3D Kalman track must be greater than 0.75.

Lastly, for the far detector, the cosmic rejection cuts for the official ν_μ CC analysis group were applied. This required that the cosine of the angle between the 3D Kalman track with the highest ReMId value and the NuMI beam direction was greater than 0.5.

The number of slice hits was required to be less than 400. The boosted decision tree trained to separate contained ν_μ CC interactions at the far detector from cosmic ray background events must return a PID value greater than 0.535.

I used unoscillated simulation files to fit with. For the far detector and the near detector, I used 10,000 events each. The files were matched by run and subrun number to the detector configurations and bad channels used for this analysis.