Machine Learning-Based Energy Estimation for Sterile Neutrino Searches in the NOνA Experiment

A DISSERTATION

SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL

OF THE UNIVERSITY OF MINNESOTA

BY

Shaowei Wu

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Advisor: Professor Gregory Pawloski

August, 2025

Acknowledgements

I would like to express my deepest gratitude to my advisor, Professor Pawloski, for his invaluable academic guidance and support throughout my Ph.D. journey. His patient mentorship has been essential to my development as a researcher.

I would like to thank the members of my thesis committee — Professor Coughlin, Professor Heller, and Professor Kapusta — for generously dedicating their time to evaluate my work.

I am deeply thankful to my family, especially my wife, Yuqin Xiao, for her unwavering support and companionship throughout my years as a student.

I am also grateful to my fellow graduate students, especially Ting Gao, for insightful discussions and theoretical guidance during my doctoral studies.

Many thanks to my collaborators in the NOvA collaboration for their support on various aspects of my research. In particular, I would like to acknowledge Dmitrii Torbunov, Joshua Barrow, Shivam, V Hewes, Haejun Oh, Adam Lister, Jeremy Wolcott, Alejandro Yankelevich, Alexander Booth, Adam Aurisano, and Brian Rebel for their contributions.

Special thanks to my friend Liyi Chen for providing technical support in machine learning, which greatly benefited my research.

Abstract

This dissertation presents a search for sterile neutrinos using Monte Carlo datasets and experimental data from the NOνA experiment. This work introduces novel machine learning techniques for energy reconstruction in neutral current events. A deep learning-based energy estimator was developed and integrated into the analysis framework. Using this new energy estimator, the sensitivity to sterile neutrino-induced oscillations is evaluated and presented. Furthermore, the dissertation explores potential methods for further improvement of the analysis.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Over the past few years, my research has focused on applying machine learning algorithms to the problem of neutrino energy reconstruction. This thesis presents my work on reconstructing the energy of the incoming neutrinos in neutral current events using machine learning techniques in the NOνA experiment. This document is organized as follows:

- Chapter 2 briefly presents the history of, and science behind, the subjects presented in this thesis.
- Chapter 3 describes the experimental setup and the detector technology.
- Chapter 4 describes the simulation process used in the analysis.
- Chapter 5 follows the chain of reconstruction software used to obtain meaningful results from data.
- Chapter 6 hashes out the strategy for selecting and analyzing neutral current

neutrino interactions.

- Chapter 7 demonstrates the method used to estimate the incoming energy of neutrinos that are used in this data set. This is the main novel technique developed for this thesis.
- Chapter 8 describes the analysis setup and result.
- Chapter 9 presents a final discussion of the analyses presented in the thesis.

Chapter 2

Physics of Neutrinos

2.1 The Discovery of Neutrinos

In 1911, Lise Meitner and Otto Hahn made a groundbreaking observation concerning the energy spectrum of electrons emitted during beta decay [1]. They discovered that this spectrum was continuous rather than discrete, a result that challenged the prevailing understanding of the conservation of energy. This discovery sparked significant debate and prompted further investigation into the nature of beta decay.

In December 1930, Wolfgang Pauli, in an open letter addressed to a group of physicists gathered for a conference, proposed a bold hypothesis to resolve this conundrum [2]. Pauli suggested the existence of an additional, electrically neutral particle that was emitted alongside the electron during beta decay. He referred to this hypothetical particle as the "neutron," believing it was responsible for preserving the conservation of energy within the process.

Building on this idea, Enrico Fermi, in 1933, developed a comprehensive theory of beta decay. In this framework, Fermi refined Pauli's concept and renamed the particle "neutrino," meaning "little neutral one" in Italian [3]. He proposed that beta decay was mediated by a new fundamental interaction, the weak interaction. This interaction is much weaker than the electromagnetic interaction. Despite its theoretical foundation, detecting neutrinos posed a significant challenge due to their weak interactions with matter.

It was not until 1956 that experimental confirmation of neutrinos was achieved. Clyde Cowan and Frederick Reines successfully detected anti-electron neutrinos ($\overline{\nu}_e$) emitted from a nuclear reactor. This landmark discovery was made through the observation of the reaction $\overline{\nu}_e + p \rightarrow n + e^+$ [4].

Further advancements followed in 1962, when scientists at Brookhaven National Laboratory discovered a second type of neutrino, the muon neutrino (ν_μ). The researchers fired a beam of high-energy protons at a beryllium target. This collision produced pions, which are unstable particles that quickly decay into other particles, including neutrinos. The resulting beam of neutrinos was then directed towards a specialized detector. This detector was heavily shielded to ensure only neutrinos could reach it, preventing interference from other particles. Production of muons was observed in the detector which provided proof that there was a unique type of neutrino specifically linked to muons. This finding, which expanded the family of known neutrinos, was pivotal in demonstrating the existence of distinct neutrino flavors [5].

The discovery of the third and final neutrino, the tau neutrino (ν_τ), occurred much later, in 2000. The DONUT experiment team at Fermilab achieved this

milestone, confirming the existence of all three neutrino generations. In this experiment, a beam of protons was used to generate a stream of particles including tau neutrinos, which were then directed toward a detector designed to capture their rare interactions. The appearance of short-lived tau leptons in the detector provided the first direct observation of ν_τ. This discovery also completed the identification of the last known lepton in the Standard Model of particle physics [6].

Together, these milestones chart the evolution of our understanding of neutrinos from theoretical postulates to experimental verification.

Today, it is well-established that neutrinos are generated from a wide variety of sources, including but not limited to the following:

- Electron neutrinos are continuously produced through nuclear fusion processes within the Sun. These processes occur simultaneously, each contributing to the overall solar neutrino flux. The energy spectrum of solar neutrinos is a superposition of partial spectra, encompassing both discrete mono-energetic components (e.g. ^{7}Be neutrinos) and continuous distributions (e.g. ^{8}B neutrinos).

- When cosmic rays (protons, nuclei, etc) collide with the atoms in the Earth's atmosphere, they initialize "showers" (cascades of particles). Most of the pions and kaons generated in the showers rapidly decay into $\mu^- + \overline{\nu}_\mu$ pairs or $\mu^+ + \nu_\mu$ pairs. The muons and anti-muons are not stable and further decay, producing electron neutrinos and anti-electron neutrinos.

- Nuclear reactors are intense source of anti-electron neutrinos, primarily produced through the beta decay of neutron-rich fission fragments generated during the nuclear fission of isotopes such as ^{235}U, ^{239}Pu, ^{238}U, and ^{241}Pu. Each fission event produces a range of radioactive isotopes, which subsequently undergo a series of beta decays, emitting electron antineutrinos as they stabilize.

- The collision of high-energy proton beams on a target serves as an intense source of neutrinos. When protons strike the target, they produce a cascade of secondary particles, including pions and kaons, which subsequently decay to generate neutrinos. This process is foundational in neutrino physics, enabling the creation of controlled and high-intensity neutrino beams for experiments.

2.2 Neutrinos in the standard model

As of the writing of this thesis, there are three known flavor eigenstates of neutrinos(ν_e, ν_μ and ν_τ), corresponding to three flavors of charged leptons (electron/e^-, muon/μ^- and tauon/τ^-). An electron neutrino carries the same lepton number as an electron. There also exist antineutrinos, which carry the opposite lepton numbers of their counterparts.

In the Standard Model, neutrinos are spin-1/2 massless fermions and they do not participate in the strong interaction and electromagnetic interaction. Besides the gravitational interaction which is too weak to observe and is not a part of the

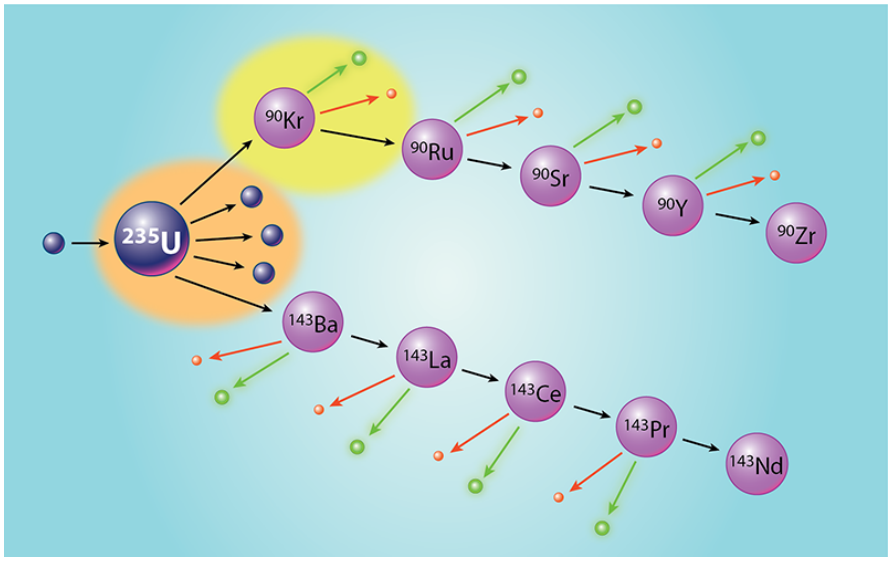


Figure 2.1: In the nuclear fission, anti-electron neutrinos (green) and electrons (red) are produced in a series of radioactive decays. This diagram illustrates an example decay chain resulting in the emission of eight anti-electron neutrinos been emitted [7].

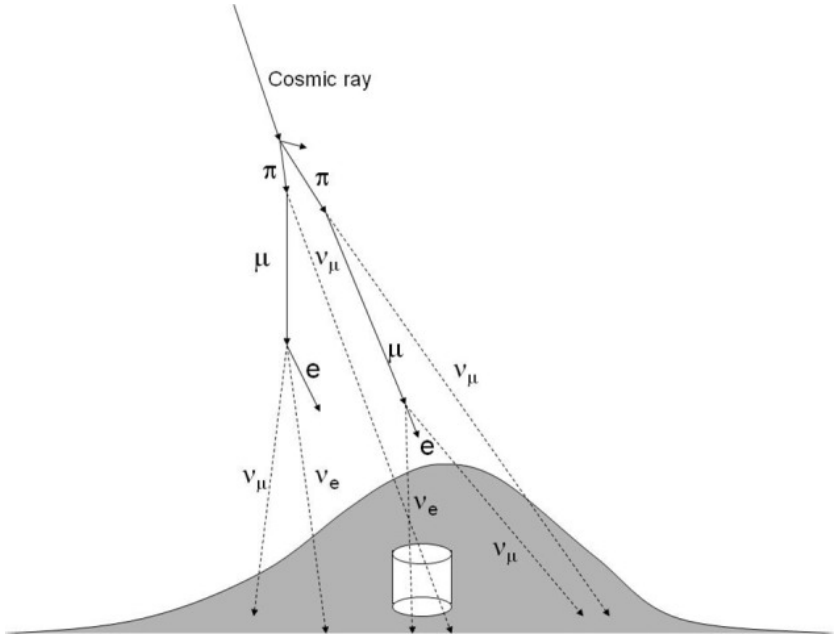
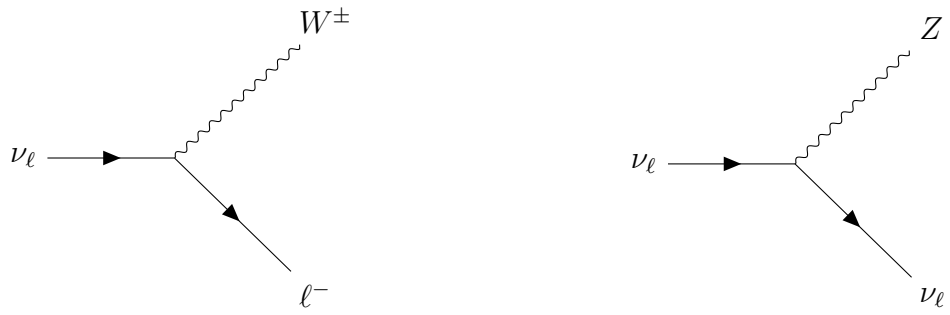


Figure 2.2: A cosmic ray interacts with an air nucleus, producing showers in the atmosphere [8].

Standard Model, neutrinos participate in the weak interaction through charged current (CC) interactions and neutral current (NC) interactions.

Generation	**Neutrino (ν_ℓ)**	**Lepton (ℓ^-)**
First	ν_e	e^-
Second	ν_μ	μ^-
Third	ν_τ	τ^-

Table 2.1: Generations of Neutrinos and Corresponding Leptons in the Standard Model



(a) $W^\pm$ boson interaction vertex. (b) Z boson interaction vertex.

Figure 2.3: Interaction vertices of neutrinos with $W^\pm$ and Z bosons.

The charged current interaction (see Figure 2.3a) is mediated by W^+ and W^- bosons. For instance, when a neutrino interacts with a target particle through a charged current process, it can produce a charged lepton. This reaction can be represented as

$$\nu_\ell + n \rightarrow \ell^- + p$$

where ν_ℓ denotes a neutrino of any flavor, ℓ^- is the corresponding charged lepton, and n and p represent a neutron and a proton, respectively. The flavor of incoming neutrinos participating in the charged current interaction can be identified by the

outgoing charged leptons according to the conservation of lepton numbers.

The neutral current interaction (see Figure 2.3b) is mediated by Z^0 bosons. For instance, in the scattering between a neutrino (of any flavor) and a nucleus ($\mathbb{Z}$) through neutral current interaction:

$$\nu_\ell + \mathbb{Z} \to \nu_\ell + \mathbb{Z}$$

where the flavor of the neutrino stay the same. Unlike charged current interactions, the outgoing leptons are neutrinos and it is impossible to identify the flavors of incoming neutrinos.

In neutrino experiments, both types of interactions occur. However, the energy of incoming neutrinos involved in neutral current (NC) interactions is more challenging to estimate, as a portion of the system's energy is carried away by the invisible outgoing neutrino. We will revisit this issue in subsequent chapters and provide a detailed analysis and discussion.

2.3 Neutrino Oscillation

Neutrino oscillation is one of the most intriguing and exciting discoveries of the past few decades. In this section, we will review the theoretical and experimental aspects of neutrino oscillations and derive some of the most essential formulas.

2.3.1 The History of Neutrino Oscillation Theory

Inspired by the idea of $K^0 \leftrightarrow \overline{K}^0$ oscillations proposed by Gell-Mann and Pais in 1955 [9], Bruno Pontecorvo suggested in 1957 [10] that, under the assumption of one flavor of neutrinos, there would be neutrino $\rightarrow$ antineutrino transitions. In 1967, after the muon neutrinos were discovered, he discussed [11] the oscillation between ν_e and ν_μ and possible violations of leptonic charge conservation such as neutrino-less double β decay. He further predicted that if the oscillation length is much smaller than the radius of the solar region where neutrinos are effectively produced, the flux of observable solar neutrinos would be reduced to half of the total solar neutrino flux. This prediction was particularly relevant to the ongoing Davis solar neutrino experiment, the first experiment which aimed to detect and measure the flux of solar neutrinos. In 1968, Gribov and Pontecorvo [12] presented the model of neutrino masses and mixing, considering four types of neutrino (ν_e, $\overline{\nu}_e$, ν_μ and $\overline{\nu}_\mu$). A few years earlier, in 1962, Ziro Maki, Masami Nakagawa and Shoichi Sakata published a paper [13] that proposed the mixing of neutrinos of different types and introduced formalism for it. This paper was not widely recognized until the discovery of solar neutrino deficit. The neutrino mixing theory was later extended to three flavors and three mass eigen-states and formalized as the Pontecorvo–Maki–Nakagawa–Sakata matrix, which elegantly describes how the neutrino flavor states are quantum mechanical superpositions of the neutrino mass states.

2.3.2 Neutrino Experiments in History

The Standard Solar Model (SSM), originally referred to as calibrated solar models [14], was initially developed in the mid-20th century. It integrates fundamental laws of physics and is constrained by observational data from the Sun. It was used to make predictions of the reactions inside the core of the Sun, including the production of solar neutrinos. In 1964, John Bahcall provided the first theoretical prediction [15] of solar neutrino fluxes based on the Standard Solar Model, laying the groundwork for subsequent experiments.

The pioneering Davis solar neutrino experiment, also known as the "Homestake experiment", used 390,000 liters of liquid tetrachloroethylene (C_2Cl_4) contained in a horizontal cylindrical tank located 4,890 ft underground. The neutrinos can be captured by the abundant ^{37}Cl nuclei in the detector through the following reaction:

$$^{37}\mathrm{Cl} + \nu_e \rightarrow e^- + ^{37}\mathrm{Ar}^+.$$

The deep underground detector ensures that the the production of ^{37}Ar by the cosmic-ray is reduced. The number of neutrinos captured by ^{37}Cl nuclei can be estimated by counting the number of ^{37}Ar produced in the detector. In 1968, Raymond Davis and collaborators published the results of their experiment [16], demonstrating that the observed flux of solar electron neutrinos was approximately one-third of the theoretical prediction by John Bahcall, revealing the solar neutrino problem, a significant discrepancy that challenged the standard solar model and neutrino physics.

The Kamiokande experiment is a neutrino observatory located 1 kilometer

underground in the Mozumi Mine near Kamioka, Japan [17]. The Kamiokande detector is an imaging water-Cherenkov detector with 3,000 tons of pure water contained in a cylindrical tank measuring 16 meters in height and 15.6 meters in diameter. In the first phase of the experiment (Kamiokande-I, 1983 - 1985), it aimed to detect the decay of nucleons. With an accumulated exposure of 3.76 $kt \cdot yr$, although no evidence for nucleon decay has been observed, constraints on the nucleon partial lifetime for various decay modes have been established. The second phase of Kamiokande experiment (Kamiokande-II, 1985 - 1990) started after an upgrade with an electronics system capable of recording the arrival time of photomultiplier tubes (PMT) signal in addition to its charge. Compared with Davis experiment which was conducted two decades ago, Kamiokande-II allowed for real-time observation of solar ^{8}B neutrinos. The ^{8}B neutrino flux is produced by the beta decay of ^{8}B in the sun:

$$^8\mathrm{B} \rightarrow ^8\mathrm{Be}^* + e^+ + \nu_e.$$

Kamiokande-II detected solar electron neutrinos via the elastic scattering:

$$\nu_e + e^- \rightarrow \nu_e + e^-.$$

In 1988, Kamiokande-II reported the first result for the solar neutrino flux, marking the first real-time measurement of solar neutrinos. A 6.7 MeV threshold was set to select high-energy neutrinos with 50% efficiency [17]. With an angular resolution of 28°, Kamiokande-II was able to measure the direction of the incoming

neutrinos, confirming that the detected neutrinos originated from the Sun. The measured flux of solar neutrinos was significantly lower than the predictions of the Standard Solar Model, indicating the existence of the solar neutrino deficit problem.

The Super-Kamiokande experiment (Super-K) [18] is the successor of Kamiokande experiment. It has larger detectors containing 50,000 tons of pure water, with better energy resolution. In 1998, Super-K confirmed the deficit of solar neutrinos, which was consistent with the result from Kamiokande-II [19]. Besides detecting the solar neutrinos, Super-K performs a variety of crucial tasks including the study of atmospheric neutrino, supernova neutrino decay of the nucleon. The atmospheric neutrino study by Super-K showed strong evidence of neutrino oscillations [19].

The Sudbury Neutrino Observatory (SNO) is a solar neutrino observatory located in Creighton Mine in Sudbury, Ontario, Canada. It utilizes 1,000 tons of heavy water (D_2O) as a target. There are three interactions that SNO detects neutrinos through.

1. Charged current (CC) reactions $\nu_e + d \rightarrow e^- + p + p$. Only electron neutrinos take part in this interaction.

2. Neutral current (NC) reactions $\nu_x + d \rightarrow p + n + \nu_x$. The total flux of all neutrino flavors above a threshold above 2.2 MeV is measured.

3. Elastic scattering $\nu_x + e^- \rightarrow \nu_x + e^-$.

Here, x denotes any neutrino flavor: e, μ, or τ. Unlike the CC reaction, the NC reaction involves neutrinos of any flavors, meaning that by measuring the

ratio of the ^8B neutrino CC reaction rate to that of the NC reaction rate, we can directly measure the fraction of neutrino flavor change [20].

The first operation of the SNO detector utilized pure heavy water. Neutrons were captured by deuterium, causing the production of a gamma ray, which then Compton-scattered atomic electrons. These scattered electrons produced Cherenkov light. The second SNO operation dissolved 2 tons of NaCl in heavy water. The chlorine ions increased the neutron capture efficiency.

In 2002, the SNO collaboration published result of fitted ^8B flux (unit = 10^6 cm^{-2} s^{-1}) in the second operation [21]:

$$\phi_{\mathrm{CC}}^{\mathrm{SNO}} = 1.76^{+0.05}_{-0.05}(\mathrm{stat})^{+0.09}_{-0.09}(\mathrm{syst})$$
$$\phi_{\mathrm{NC}}^{\mathrm{SNO}} = 5.09^{+0.44}_{-0.43}(\mathrm{stat})^{+0.46}_{-0.43}(\mathrm{syst}).$$

The result is a strong evidence that electron neutrinos oscillated to other flavors. In addition, the observed NC rate is consistent with the prediction of the SSM.

2.3.3 Neutrino Oscillation in Vacuum

The Standard Model assumes that neutrinos are massless. However, to explain neutrino oscillation, this assumption must be revised. Oscillations between neutrino flavors require that neutrinos have mass and that the flavor eigenstates (denoted by Greek subscripts, e.g., $|\nu_\alpha\rangle$) are superpositions (linear combinations)

of mass eigenstates (denoted by Latin subscripts, e.g., $|\nu_l\rangle$).

$$|\nu_\alpha\rangle = \sum_l U_{\alpha l} |\nu_l\rangle \tag{2.1}$$

where U is a unitary complex matrix ($U^\dagger U = \mathbb{I}$). Inversely,

$$|\nu_l\rangle = \sum_\alpha U^*_{\alpha l} |\nu_\alpha\rangle . \tag{2.2}$$

For three flavor oscillation, U is the 3×3 PMNS matrix.

In the following calculations, we treat neutrinos as a plane wave and assume that different mass eigenstates share the same energy when a neutrino travels. Assume that a neutrino with eigenstate $|\nu_\alpha\rangle$ and energy E is created at ($\mathbf{x} = 0,\ t = 0$). The evolution of neutrino wave function with respect to time and space can be written as:

$$|\nu_\alpha(t)\rangle = \sum_l U_{\alpha l} e^{-i(Et - \mathbf{p_l}\cdot\mathbf{x})} |\nu_l\rangle . \tag{2.3}$$

According to the special relativity, $E^2 = m_l^2 + p_l^2$. For $E \gg m_l$, we have

$$p_l = \sqrt{E^2 - m_l^2} \approx E - \frac{m_l^2}{2E}. \tag{2.4}$$

Let L be the distance that the neutrino travels, we get $|\mathbf{x}| = t = L$. Now equation (Equation (2.3)) can be simplified as:

$$|\nu_\alpha(L)\rangle = \sum_l U_{\alpha l} e^{-i\frac{m_l^2 L}{2E}} |\nu_l\rangle . \tag{2.5}$$

Insert equation Equation (2.2) to equation Equation (2.5) and get the expression of oscillated neutrino state as superpositions of flavor eigenstates:

$$|\nu_\alpha(L)\rangle = \sum_l U_{\alpha l} e^{-i\frac{m_l^2 L}{2E}} (\sum_\beta U^*_{\beta l} |\nu_\beta\rangle) \tag{2.6}$$

$$= \sum_\beta \sum_l (U_{\alpha l} e^{-i\frac{m_l^2 L}{2E}} U^*_{\beta l}) |\nu_\beta\rangle . \tag{2.7}$$

To calculate the flavor transition probability, we need to calculate the amplitude:

$$\langle \nu_\beta | \nu_\alpha(L)\rangle = \sum_l U_{\alpha l} U^*_{\beta l} e^{-i\frac{m_l^2 L}{2E}} \tag{2.8}$$

The probability is the square of the amplitude modulus:

$$\begin{aligned} P(\nu_\alpha \to \nu_\beta) &= |\langle \nu_\beta | \nu_\alpha(L)\rangle|^2 \\ &= \sum_{l,j} U_{\alpha l} U^*_{\beta l} e^{-i\frac{m_l^2 L}{2E}} U^*_{\alpha j} U_{\beta j} e^{i\frac{m_j^2 L}{2E}} \\ &= \sum_{l,j} U_{\alpha l} U^*_{\beta l} U^*_{\alpha j} U_{\beta j} e^{-i\frac{(m_l^2 - m_j^2)L}{2E}} \\ &= \sum_{l,j} U_{\alpha l} U^*_{\beta l} U^*_{\alpha j} U_{\beta j} + \sum_{l,j} U_{\alpha l} U^*_{\beta l} U^*_{\alpha j} U_{\beta j} (e^{-i\frac{(m_l^2 - m_j^2)L}{2E}} - 1). \end{aligned} \tag{2.9}$$

Based on the property of unitary matrix $\sum_i U_{\alpha i} U^*_{\beta i} = \delta_{\alpha\beta}$ and let $\Delta m^2_{lj} = m_l^2 - m_j^2$, the equation Equation (2.9) can be simplified as

$$P(\nu_\alpha \to \nu_\beta) = \delta_{\alpha\beta} + \sum_{l,j} U_{\alpha l} U^*_{\beta l} U^*_{\alpha j} U_{\beta j} (e^{-i\frac{\Delta m_{lj}^2 L}{2E}} - 1). \tag{2.10}$$

We can further simplify the above expression by exploiting the symmetry with respect to the indices. First, when $l = j$, the exponential part becomes 1. Second, for any $x \in \mathbb{R}$, we have

$$e^{-ix} + e^{ix} = 2\mathfrak{R}(e^{-ix}) = 2\mathfrak{R}(e^{ix}) \tag{2.11}$$

where $\mathfrak{R}$ denotes the real part of a complex number and $\mathfrak{I}$ denotes the imaginary part of a complex number(see below). Equation (2.10) now becomes

$$P(\nu_\alpha \to \nu_\beta) = \delta_{\alpha\beta} + 2\mathfrak{R}(\sum_{l>j} U_{\alpha l} U^*_{\beta l} U^*_{\alpha j} U_{\beta j} (e^{-i\frac{\Delta m^2_{lj} L}{2E}} - 1)). \tag{2.12}$$

In addition, we define the following variables:

$$\phi \equiv \frac{\Delta m^2_{lj} L}{2E} \tag{2.13}$$

$$\Omega \equiv U_{\alpha l} U^*_{\beta l} U^*_{\alpha j} U_{\beta j} (e^{-i\phi} - 1). \tag{2.14}$$

Consequently,

$$\begin{aligned}
\mathfrak{R}(\Omega) &= \mathfrak{R}(U_{\alpha l} U^*_{\beta l} U^*_{\alpha j} U_{\beta j} (e^{-i\phi} - 1)) \\
&= \mathfrak{R}(U_{\alpha l} U^*_{\beta l} U^*_{\alpha j} U_{\beta j} (\cos(\phi) - i\sin(\phi) - 1)) \\
&= \mathfrak{R}(U_{\alpha l} U^*_{\beta l} U^*_{\alpha j} U_{\beta j} ((\cos(\phi) - 1) - i\sin(\phi))) \\
&= \mathfrak{R}(U_{\alpha l} U^*_{\beta l} U^*_{\alpha j} U_{\beta j} (-2\sin^2(\frac{\phi}{2}) - i\sin(\phi))) \\
&= -2\mathfrak{R}(U_{\alpha l} U^*_{\beta l} U^*_{\alpha j} U_{\beta j}) \sin^2(\frac{\phi}{2}) + \mathfrak{I}(U_{\alpha l} U^*_{\beta l} U^*_{\alpha j} U_{\beta j}) \sin(\phi).
\end{aligned} \tag{2.15}$$

By putting the above expression Equation (2.15) in Equation (2.12) and define

a new variable [22]

$$\Delta_{lj} \equiv \frac{\Delta m_{lj}^2 L}{4E} = \frac{\phi}{2}, \tag{2.16}$$

we get

$$P(\nu_\alpha \to \nu_\beta) = \delta_{\alpha\beta} - 4\sum_{l>j} \Re(U_{\alpha l} U_{\beta l}^* U_{\alpha j}^* U_{\beta j}) \sin^2 \Delta_{lj} + 2\sum_{l>j} \Im(U_{\alpha l} U_{\beta l}^* U_{\alpha j}^* U_{\beta j}) \sin 2\Delta_{lj} \tag{2.17}$$

From Equation (2.16) and Equation (2.17) , we can observe that besides the unitary matrix U, the flavor transition probability in vacuum depends on the following factors:

1. The squared difference of the neutrino mass Δm_{lj}^2. It is a crucial parameter in neutrino physics and it directly governs the oscillation frequency. If $\Delta m_{lj}^2 = 0$, there will be no oscillation.

2. The distance L that neutrinos travel from their production point to the detection point, often referred to as the baseline. For the NOνA experiment, $L = 810$ km.

3. The energy E of neutrinos. Another parameter that impacts the oscillation frequency. Precise measurements of neutrino energy are essential for determining oscillation probabilities and extracting fundamental parameters. This represents the central idea of this thesis.

It is important to note that the calculations above are performed using the natural unit system, where $c = \hbar = 1$. If we adopt kilometers (km) as the unit of

length, electron volt (eV) as the unit of mass and giga-electronvolts (GeV) as the unit of energy, the expression will take on a different form:

$$\Delta_{lj} = 1.267\frac{\Delta m^2_{lj} L}{E}. \tag{2.18}$$

2.3.4 Two Flavor Neutrino Oscillation in Vacuum

Consider a neutrino oscillation involving two flavors. For simplicity the mixing matrix is written as

$$U = \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \tag{2.19}$$

The matrix U can be a complex one, but we treat it as a real unitary matrix here. $c = \cos\theta$ and $s = \sin\theta$. θ is the mixing angle. Inserting the expression of U into Equation (2.17), one obtains the flavor transition probability as:

$$P(\nu_\alpha \to \nu_\beta) = \begin{cases} 1 - \sin^2(2\theta)\sin^2(\frac{\Delta m^2 L}{4E}), & \text{if } \alpha = \beta \\ \sin^2(2\theta)\sin^2(\frac{\Delta m^2 L}{4E}), & \text{if } \alpha \neq \beta. \end{cases} \tag{2.20}$$

When $\alpha = \beta$, the neutrino retains its original flavor. This probability is referred to as the ν_α survival probability. Conversely, when $\alpha \neq \beta$, the neutrino undergoes a flavor transition, oscillating into a different flavor state. This is referred to as the ν_β appearance probability.

It is obvious that the magnitude of neutrino oscillation probability is controlled by the mixing angle. With a mixing angle $\theta = \frac{\pi}{4} + \frac{n\pi}{2}$ (so that $\sin^2(2\theta) = 1$), we

would expect a maximal flavor mixing and there will be a set of periodic points on the baseline where the neutrinos will fully oscillate to other flavors. The locations of these points are determined by $\sin^2(\frac{\Delta m^2 L}{4E}) = 1$.

2.3.5 Three Flavor Neutrino Oscillation in Vacuum

As mentioned in Section 2.3.3, three flavor oscillation scenario can be described by PMNS matrix:

$$\begin{aligned} U &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \cdot \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{CP}} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta_{CP}} & 0 & c_{13} \end{pmatrix} \cdot \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{CP}} \\ -s_{12}c_{23} - c_{12}s_{13}s_{23}e^{i\delta_{CP}} & c_{12}c_{23} - s_{12}s_{13}s_{23}e^{i\delta_{CP}} & c_{13}s_{23} \\ s_{12}s_{23} - c_{12}s_{13}c_{23}e^{i\delta_{CP}} & -c_{12}s_{23} - s_{12}s_{13}c_{23}e^{i\delta_{CP}} & c_{13}c_{23} \end{pmatrix}, \end{aligned} \tag{2.21}$$

where $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$. It serves the same purpose as the CKM matrix in quark theory. There are 4 free parameters in the matrix. Three of them are mixing angles (θ_{12}, θ_{23} and θ_{13}). The last parameter δ_{CP} is the charge-parity violation phase angle which describes the magnitude of CP asymmetry and introduces the complex nature of the matrix.

We provide some recent measurements of these parameters in Table 2.2. The last mass splitting can be calculated using the relation

$$\Delta m^2_{21} + \Delta m^2_{32} + \Delta m^2_{13} = 0. \tag{2.22}$$

From the data we can see that $\Delta m^2_{32} \approx -\Delta m^2_{13} \gg \Delta m^2_{21}$.

Parameter	Value	Experiment
$\sin^2(\theta_{12})$	$0.307^{+0.013}_{-0.012}$	KamLAND + global solar
$\sin^2(\theta_{23})$	$0.57^{+0.03}_{-0.04}$	NOνA
$\sin^2(\theta_{13})$	$2.80^{+0.28}_{-0.65} \times 10^{-2}$	T2K
Δm^2_{21}	$7.53^{+0.22}_{-0.20}(\times 10^{-5}\ eV^2)$	KamLAND + global solar + SBL + accelerator
Δm^2_{32}	$2.41 \pm 0.07(\times 10^{-3}\ eV^2)$	IceCube
δ_{CP}	$1.37^{+0.31}_{-0.20}(\times \pi\ rad)$	T2K

Table 2.2: Recently measured 3-flavor mixing parameters. Normal mass ordering ($m_1 < m_2 < m_3$) assumed. Data provided by PDG 2024 [23].

If the CP violation phase is zero, the matrix U would be a real matrix. The imaginary part of equation Equation (2.17) vanishes:

$$P(\nu_\alpha \to \nu_\beta) = \delta_{\alpha\beta} - 4 \sum_{l>j} U_{\alpha l} U_{\beta l} U_{\alpha j} U_{\beta j} \sin^2 \Delta_{lj} \tag{2.23}$$

where $\Delta_{lj} = \frac{\Delta m^2_{lj} L}{E}$. For 3-flavor appearance experiment, we remove the first term on the right side and expand the summation:

$$P(\nu_\alpha \to \nu_\beta) = -4(U_{\alpha 1} U_{\beta 1} U_{\alpha 2} U_{\beta 2} \sin^2 \Delta_{21} \tag{2.24}$$

$$+ U_{\alpha 1} U_{\beta 1} U_{\alpha 3} U_{\beta 3} \sin^2 \Delta_{31} \tag{2.25}$$

$$+ U_{\alpha 2} U_{\beta 2} U_{\alpha 3} U_{\beta 3} \sin^2 \Delta_{32}), \tag{2.26}$$

where $\alpha \neq \beta$. To further simplify the expression, let's consider some edge cases.

1. When $\frac{L}{E}$ is small, we can ignore the term Equation (2.24), since $\sin^2 \Delta_{21} \to 0$. Using the approximation $\Delta_{31} \approx \Delta_{32}$, we combine the other two terms and obtain the following:

$$P(\nu_\alpha \to \nu_\beta) = -4 \left(U_{\alpha 1} U_{\beta 1} U_{\alpha 3} U_{\beta 3} + U_{\alpha 2} U_{\beta 2} U_{\alpha 3} U_{\beta 3}\right) \sin^2 \Delta_{32}. \tag{2.27}$$

Insert the PMNS matrix element expressions to the equation above we have the appearance probabilities [24]:

$$P(\nu_\mu \to \nu_\tau) \approx \cos^4(\theta_{13}) \sin^2(2\theta_{23}) \sin^2 \Delta_{23} \tag{2.28}$$

$$P(\nu_e \to \nu_\mu) \approx \sin^2(2\theta_{13}) \sin^2(\theta_{23}) \sin^2 \Delta_{23} \tag{2.29}$$

$$P(\nu_e \to \nu_\tau) \approx \sin^2(2\theta_{13}) \cos^2(\theta_{23}) \sin^2 \Delta_{23}. \tag{2.30}$$

Considering the fact that θ_{13} is very small, we can proceed with further approximations below:

$$P(\nu_\mu \to \nu_\tau) \approx \sin^2(2\theta_{23}) \sin^2 \Delta_{23} \tag{2.31}$$

$$P(\nu_e \to \nu_\mu) \approx 0 \tag{2.32}$$

$$P(\nu_e \to \nu_\tau) \approx 0. \tag{2.33}$$

2. When $\frac{L}{E}$ is large, the oscillation frequency is high and thus $\sin^2 \Delta_{32}$ and $\sin^2 \Delta_{31}$ would average to $\frac{1}{2\pi} \int_0^{2\pi} \sin^2 x dx = \frac{1}{2}$. The appearance probability

now becomes:

$$P(\nu_\alpha \to \nu_\beta) = -4(U_{\alpha1}U_{\beta1}U_{\alpha2}U_{\beta2}\sin^2\Delta_{21} + \frac{1}{2}(U_{\alpha1}U_{\beta1}U_{\alpha3}U_{\beta3} + U_{\alpha2}U_{\beta2}U_{\alpha3}U_{\beta3})). \tag{2.34}$$

Use the expression of PMNS matrix, we can get the following [24] electron neutrino disappearance probability:

$$P(\nu_e \to \nu_{\mu,\tau}) \approx \cos^2(\theta_{13})\sin^2(2\theta_{12})\sin^2\Delta_{21} + \frac{1}{2}\sin^2(2\theta_{13}). \tag{2.35}$$

When $\theta_{13} \to 0$,

$$P(\nu_e \to \nu_{\mu,\tau}) \approx \sin^2(2\theta_{12})\sin^2\Delta_{21}. \tag{2.36}$$

The above expression is only dependent on the mixing angle θ_{12} and mass splitting Δm^2_{21}.

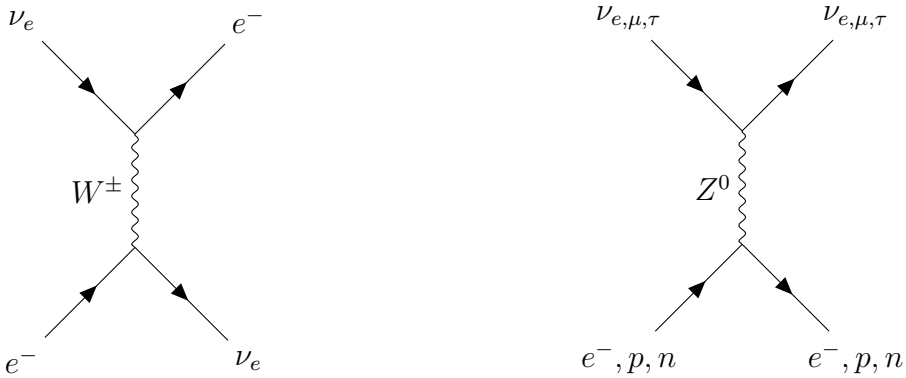
With an extremely long baseline and relatively low energy, the solar neutrinos oscillation is dominated by the mixing angle θ_{12} and mass splitting Δm^2_{21}. On the other hands, the atmospheric neutrino oscillation $\nu_\mu \to \nu_\tau$ is dominated by the mixing angle θ_{23} and mass splitting Δm^2_{32}.

2.3.6 Matter Potential

So far, our calculations only consider neutrino oscillations in a vacuum. In reality, neutrinos interact with matter as they travel, which significantly alters their oscillation behavior. This phenomenon, known as the Mikheyev-Smirnov-Wolfenstein

(MSW) effect, occurs when neutrinos pass through dense matter, such as the Sun's interior, or when they traverse long distances through the Earth's crust in underground experiments.

Neutrinos of all flavors can interact with electrons, protons and neutrons in matter through neutral current (NC) interaction mediated by Z bosons. In addition, the electron neutrino will also interact with the electrons in matter through charged current (CC) [25].



(a) Charged current interaction. (b) Neutral current interaction.

Figure 2.4: Interaction of neutrinos with matter.

The effective Hamiltonian for neutrino CC interaction is

$$H_{CC} = \frac{G_F}{\sqrt{2}}[\overline{e}\gamma_\mu(1-\gamma_5)\nu_e][\overline{\nu}_e\gamma^\mu(1-\gamma_5)e], \tag{2.37}$$

where G_F is Fermi's constant. Do a Fierz transformation on H_{CC}:

$$H_{CC} = \frac{G_F}{\sqrt{2}}[\overline{e}\gamma_\mu(1-\gamma_5)e][\overline{\nu}_e\gamma^\mu(1-\gamma_5)\nu_e]. \tag{2.38}$$

We integrate over electrons and define the matter potential [25]:

$$H_{eff}(\nu_e) = \langle H_{CC} \rangle_{electron} \equiv \overline{\nu}_e V_e \nu_e \tag{2.39}$$

$$\langle \overline{e}\gamma_0 e \rangle = \left\langle e^\dagger e \right\rangle = N_e,\ \langle \overline{e}\boldsymbol{\gamma} e \rangle = \langle \mathbf{v}_e \rangle \tag{2.40}$$

$$\langle \overline{e}\gamma_0\gamma_5 e \rangle = \left\langle \frac{\boldsymbol{\sigma}_e \mathbf{p}_e}{E_e} \right\rangle,\ \langle \overline{e}\boldsymbol{\gamma}\gamma_5 e \rangle = \langle \boldsymbol{\sigma}_e \rangle \tag{2.41}$$

where $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \gamma_3)$, N_e is the electron number density.

Assuming that the matter is unpolarized, one can get the following expression of matter potential for CC:

$$V_e = \sqrt{2} G_F N_e. \tag{2.42}$$

Using a similar approach, we can derive the expression for the matter potential associated with neutral current (NC) interactions. In this case, the contributions from electrons and protons cancel each other out, leaving neutrons as the sole effective "scatterers" in matter.

$$V_n = -G_F N_n/\sqrt{2} \quad , \tag{2.43}$$

where N_n is the neutron number density.

Thus, taking both NC and CC into account, the matter potential for neutrinos are:

$$V_{\nu_e} = \sqrt{2} G_F (N_e - \frac{N_n}{2}) \tag{2.44}$$

$$V_{\nu_\mu} = V_{\nu_\tau} = \sqrt{2} G_F (-\frac{N_n}{2}). \tag{2.45}$$

For antineutrinos, we simply multiply the matter potential with -1.

2.3.7 Neutrino Oscillation in Matter

Let's consider a system of two flavors of neutrinos in vacuum and write the evolution equation in the mass eigenstates $|\nu_m\rangle$:

$$i\frac{d}{dt}|\nu_m\rangle = H_m|\nu_m\rangle\,, \tag{2.46}$$

where H_m is a diagonalized matrix

$$H_m = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix}. \tag{2.47}$$

The equation can be written in the flavor basis using the relation Equation (2.2):

$$i\frac{d}{dt}|\nu_\ell\rangle = UH_mU^\dagger|\nu_\ell\rangle\,. \tag{2.48}$$

Using the relativistic approximation Equation (2.4) and replacing U with the expression Equation (2.19), we obtain the following equation:

$$i\frac{d}{dt}\begin{pmatrix} \nu_e \\ \nu_\mu \end{pmatrix} = \begin{pmatrix} -\frac{\Delta m^2}{4E}\cos 2\theta & \frac{\Delta m^2}{4E}\sin 2\theta \\ \\ \frac{\Delta m^2}{4E}\sin 2\theta & \frac{\Delta m^2}{4E}\cos 2\theta \end{pmatrix}\begin{pmatrix} \nu_e \\ \nu_\mu \end{pmatrix} \tag{2.49}$$

where we have dropped the constant $p + \frac{m_1^2+m_2^2}{4E}$ from the diagonal of the matrix because it doesn't affect the oscillation.

The matter potential contributes to the diagonal element of matrix. The common part (matter potential of neutron, V_n) of V_{ν_e} and V_{ν_μ} does not affect the neutrino mixing angle and is therefore omitted. The new matrix is

$$i\frac{d}{dt}\begin{pmatrix} \nu_e \\ \nu_\mu \end{pmatrix} = \begin{pmatrix} -\frac{\Delta m^2}{4E}\cos 2\theta + \sqrt{2}G_F N_e & \frac{\Delta m^2}{4E}\sin 2\theta \\ \frac{\Delta m^2}{4E}\sin 2\theta & \frac{\Delta m^2}{4E}\cos 2\theta \end{pmatrix} \begin{pmatrix} \nu_e \\ \nu_\mu \end{pmatrix} \tag{2.50}$$

where $\Delta m^2 = m_2^2 - m_1^2$.

The eigenstates can be obtained by diagonalizing the matrix:

$$\nu_A = \nu_e \cos\theta_{eff} + \nu_\mu \sin\theta_{eff} \tag{2.51}$$

$$\nu_B = -\nu_e \sin\theta_{eff} + \nu_\mu \cos\theta_{eff} \tag{2.52}$$

where we defined an effective mixing angle θ_{eff}

$$\tan 2\theta_{eff} = \frac{\frac{\Delta m^2}{2m}\sin 2\theta}{\frac{\Delta m^2}{2m}\cos 2\theta - \sqrt{2}G_F N_e}. \tag{2.53}$$

The probability of neutrino oscillation has similar form as that in vacuum if the density is constant:

$$P(\nu_e \to \nu_\mu; L) = \sin^2(2\theta_{eff})\sin^2(\pi\frac{L}{l_m}) \tag{2.54}$$

where l_m is the effective oscillation length and $\sin^2(2\theta_{eff})$ is the oscillation amplitude [25].

$$l_m = \frac{2\pi}{\sqrt{(\frac{\Delta m^2}{2E}\cos 2\theta - \sqrt{2}G_F N_e)^2 + (\frac{\Delta m^2}{2E})^2 \sin^2 2\theta}} \tag{2.55}$$

$$\sin^2 2\theta_{eff} = \frac{(\frac{\Delta m^2}{2E})^2 \sin^2 2\theta}{(\frac{\Delta m^2}{2E}\cos 2\theta - \sqrt{2}G_F N_e)^2 + (\frac{\Delta m^2}{2E})^2 \sin^2 2\theta} \tag{2.56}$$

When $N_e = 0$, the scenario reduces to the vacuum case. Furthermore, when $\frac{\Delta m^2}{2E}\cos 2\theta > \sqrt{2}G_F N_e$ and N_e gradually increases from 0, the oscillation amplitude increases correspondingly. In other words, the oscillation is enhanced.

The oscillation amplitude reaches the maximum value 1 when

$$\frac{\Delta m^2}{2E}\cos 2\theta - \sqrt{2}G_F N_e = 0. \tag{2.57}$$

This is the resonant neutrino oscillation condition. The mixing angle is 33.44°, so $\cos 2\theta > 0$. Thus, we require that $\Delta m^2 = m_2^2 - m_1^2 > 0$, since we have a positive N_e.

For antineutrinos, we need $\Delta m^2 = m_2^2 - m_1^2 < 0$.

The 3-flavor case of the evolution equation reads [25]

$$i\frac{d}{dt}\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \left[\frac{1}{2E}U\begin{pmatrix} m_1^2 & 0 & 0 \\ 0 & m_2^2 & 0 \\ 0 & 0 & m_3^2 \end{pmatrix}U^\dagger + \begin{pmatrix} V_{CC} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\right]\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix}. \tag{2.58}$$

The calculation of oscillation for 3-flavor neutrino system in matter is much more

complicated. For a more detailed analytical treatment, we recommend referring to specialized works [26], such as the perturbative expansion approach by Denton et al. [27], which calculates the oscillation probability up to the second order.

When consider the fact that $\sin^2_{13} \leq 0.04$ and $\frac{\Delta m^2_{21}}{\Delta m^2_{31}} \approx 0.03$, we get the following leading order expression [25]:

$$\begin{aligned} P(\nu_\mu \to \nu_e) = & \sin^2\theta_{23} \sin^2 2\theta_{13} \frac{\sin^2(\Delta_{31} - aL)}{(\Delta_{31} - aL)^2} \Delta^2_{31} \\ & + \sin 2\theta_{23} \sin^2 2\theta_{13} \sin^2 2\theta_{12} \frac{\sin(\Delta_{31} - aL)}{\Delta_{31} - aL} \Delta_{31} \frac{\sin(aL)}{aL} \Delta_{21} \cos(\Delta_{31} + \delta) \\ & + \cos^2\theta_{23} \sin^2 2\theta_{12} \frac{\sin^2(aL)}{(aL)^2} \Delta^2_{21} \end{aligned} \tag{2.59}$$

where $\Delta_{lj} = \frac{\Delta m^2_{lj} L}{E}$, $a = \frac{G_F N_e}{\sqrt{2}}$.

2.4 Sterile Neutrinos

The hypothesis that more than three light neutrino species contribute to neutrino oscillations has been a subject of investigation for several decades. Early deficits of atmospheric muon neutrinos observed by the Kamiokande and IMB [28] experiments prompted significant interest in the possibility of oscillations from muon neutrinos (ν_μ) to sterile neutrinos (ν_s), as discussed in various studies. Similarly, transitions of active neutrinos into sterile neutrinos were explored as a potential resolution to the solar neutrino problem.

Subsequent research has firmly established that atmospheric and solar neutrino oscillations arise from transitions among active neutrino flavors. However,

several observed anomalies remain unaccounted for within the standard three-flavor oscillation framework. These anomalies have been investigated in a variety of experimental settings, including accelerator-based neutrino beam experiments, nuclear reactors and radioactive isotope sources. We will briefly list them in the following.

1. The Liquid Scintillator Neutrino Detector (LSND) at the Los Alamos Neutron Science Center reported [29] evidence of $\overline{\nu}_\mu \to \overline{\nu}_e$ transition suggesting a mass splitting Δm^2 in the range from 0.2 eV^2 to 2 eV^2, which indicated that there is at least one neutrino mass greater than 0.4 eV^2.

2. The MiniBooNE [30] experiment later reported an event excess in the range from 0.2 GeV to 1.25 GeV, which suggested a mass splitting in the range from 0.2 eV^2 to 2 eV^2.

3. The deficit observed in the Gallium radioactive source experiments [31] has been interpreted as potential evidence for electron neutrino disappearance. Within the effective two-neutrino mixing framework, the parameters are estimated to satisfy $\sin^2 2\theta \gtrsim 0.03$ and $\Delta m^2 \gtrsim 0.1$ eV^2.

Since the estimated mass splitting significantly exceeds the measured values of Δm^2_{32} and Δm^2_{21}, it becomes necessary to consider an extension of the current three-flavor neutrino model. Introducing an additional neutrino mass eigenstate appears to be the simplest and plausible approach.

In 3+1 flavor neutrino system, the mixing matrix is the generalization of 3-flavor PMNS matrix [32]:

$$U = \begin{bmatrix} U_{e1} & U_{e2} & e^{-i\delta_{13}} s_{13} c_{14} & e^{-\delta_{14}} s_{14} \\ U_{\mu 1} & U_{\mu 2} & -e^{-i(\delta_{13}-\delta_{14}+\delta_{24})} s_{13} s_{14} s_{24} + c_{13} s_{23} c_{24} & e^{-i\delta_{24}} c_{14} s_{24} \\ U_{\tau 1} & U_{\tau 2} & -e^{i\delta_{24}} c_{13} s_{23} s_{24} s_{34} + c_{13} c_{23} c_{34} - e^{-i(\delta_{13}-\delta_{14})} s_{13} s_{14} c_{24} s_{34} & c_{14} c_{24} s_{34} \\ U_{s1} & U_{s2} & -e^{i\delta_{24}} c_{13} s_{23} s_{24} c_{34} - c_{13} c_{23} s_{34} - e^{-i(\delta_{13}-\delta_{14})} s_{13} s_{14} c_{24} c_{34} & c_{14} c_{24} c_{34} \end{bmatrix} \tag{2.60}$$

which includes new parameters: $\Delta m^2_{41},\ \theta_{14},\ \theta_{24},\ \theta_{34},\ \delta_{14},\ \delta_{24}$.

By definition, sterile neutrinos do not participate in weak interactions. As previously mentioned, in neutral current (NC) interactions, all three active neutrino flavors interact at the same rate. However, if sterile neutrinos exist, the oscillation of active neutrinos into sterile neutrinos would result in a deficit of detected NC events.

The NC survival probability in the new 4 flavor model can be approximated as (ν_μ are the primary type of neutrinos produced in NOνA) [32]

$$\begin{aligned} 1 - P(\nu_\mu \to \nu_s) \approx & 1 - \cos^4\theta_{14} \cos^2\theta_{34} \sin^2 2\theta_{24} \sin^2\Delta_{41} \\ & - \sin^2\theta_{34} \sin^2 2\theta_{23} \sin^2\Delta_{31} \\ & + \frac{1}{2} \sin\delta_{24} \sin\theta_{24} \sin 2\theta_{23} \sin\Delta_{31}. \end{aligned} \tag{2.61}$$

In the near detector, where baseline L is small, the last two terms approach zero. Consequently, the equation can be further simplified as:

$$1 - P(\nu_\mu \to \nu_s) \approx 1 - \cos^4\theta_{14} \cos^2\theta_{34} \sin^2 2\theta_{24} \sin^2\Delta_{41}. \tag{2.62}$$

On the other hand, if the mass splitting Δm_{41} is large enough, then the oscillation is rapid and $\sin^2 \Delta_{41}$ averages to $\frac{1}{2}$. We have

$$1 - P(\nu_\mu \to \nu_s) \approx 1 - \frac{1}{2} \cos^4 \theta_{14} \cos^2 \theta_{34} \sin^2 2\theta_{24}. \tag{2.63}$$

Another probability we are interested in is the ν_μ survival probability [32]

$$\begin{aligned} P(\nu_\mu \to \nu_\mu) \approx 1 &- \sin^2 2\theta_{23} \sin^2 \Delta_{31} \\ &+ 2 \sin^2 2\theta_{23} \sin^2 \theta_{24} \sin^2 \Delta_{31} \\ &- \sin^2 2\theta_{24} \sin^2 \Delta_{41}, \end{aligned} \tag{2.64}$$

which is independent of θ_{34}. Using the same approximation, the probability for short baseline is

$$P(\nu_\mu \to \nu_\mu) \approx 1 - \sin^2 2\theta_{24} \sin^2 \Delta_{41} \tag{2.65}$$

and it can be further approximated as

$$P(\nu_\mu \to \nu_\mu) \approx 1 - \frac{1}{2} \sin^2 2\theta_{24} \tag{2.66}$$

for large Δm^2_{41}.

Chapter 3

The NOνA Experiment

The NOvA (NuMI Off-Axis ν_e Appearance) experiment is a long-baseline neutrino oscillation experiment based in the United States that leverages the NUMI (Neutrinos at the Main Injector) neutrino source, from Fermilab in Illinois to the NOvA Far Detector in Ash River, Minnesota. NOνA measures the appearance probability of $\nu_e/\overline{\nu}_e$ and disappearance probability of $\nu_\mu/\overline{\nu}_\mu$. NO$\nu$A investigates the following physical parameters by measuring ν_e appearance:

1. The octant of mixing angle θ_{23} (whether $\theta_{23} > \frac{\pi}{4}$ or $< \frac{\pi}{4}$).

2. CP violation phase.

3. Neutrino mass hierarchy.

4. PMNS mixing angle θ_{13}.

With the ν_μ disappearance measurement, NOνA measures the oscillation parameters $|\Delta m^2_{23}|$ and θ_{23}.

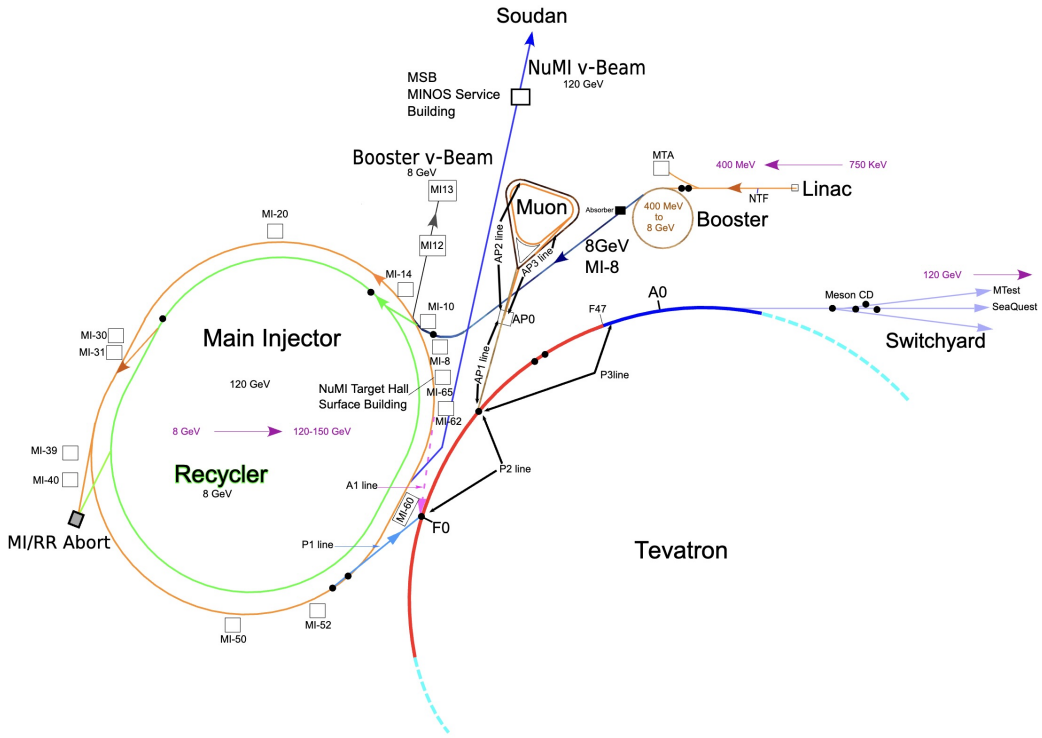


Figure 3.1: Fermilab Accelerator Complex [38]. The AP1, AP2, AP3, A1, P1 lines are no longer in use. Squared boxes labeled MI are service buildings.

In addition, NOνA is also able to measure the neutrino cross sections, supernova neutrinos, sterile neutrinos, etc.

3.1 The NuMI Beam

The NuMI beam was first built for the MINOS [33] experiment and the COSMOS [34] experiment (canceled) at Fermilab. Later, more experiments (MINERνA [35], ArgoNeuT [36], and NOνA [37]) utilized the high energy neutrino beam.

Currently, the NuMI beam is the world's most powerful neutrino beam. The beam power was initially designed to be 400 kW and has been improved over time.

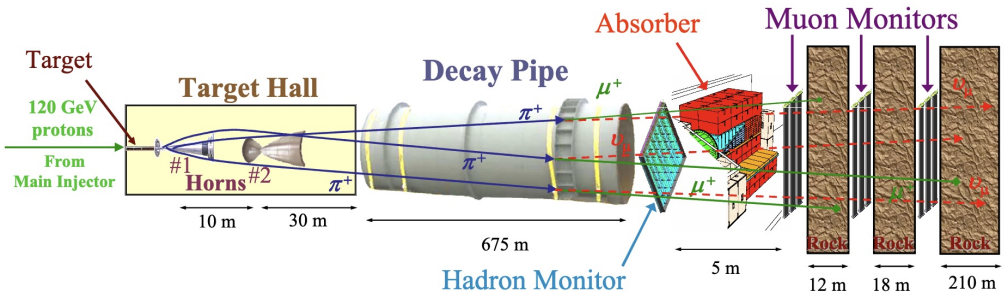


Figure 3.2: Schematic of NuMI [39]

On June 26th, 2024, the power reached 1.018 MW for a short period. A 120 GeV proton beam accelerated by Fermi-lab's Main Injector (MI) collides with a narrow 0.95 meter-long graphite target Figure 3.3 consisting of 47 fins inside the NuMI facility. The neutrino beam is created with a series of particle accelerators. At the first stage, a linear accelerator (Linac) accelerates negatively charged hydrogen ions (H^-) to an energy of 400 MeV (see Figure 3.1). The Booster subsequently strips the electrons, converting the ions into protons, and accelerates them in batches to 8 GeV. These proton batches are then transferred to the Main Injector, which further accelerates them to 120 GeV. The Main Injector's circumference is seven times that of the Booster, enabling it to store and accelerate up to six Booster batches simultaneously. The beam is so powerful that a mis-steering could cause serious damage to the facility along the beam line. A "baffle" is designed to protect the target and horns by degrading mis-steered beams and withstanding short-term exposure to full-intensity beams. It is installed in upstream of the Target. The hadronic interactions between the beam and the Target create mesons (pions, kaons, etc) which are then focused by two pulsed toroidal magnetic

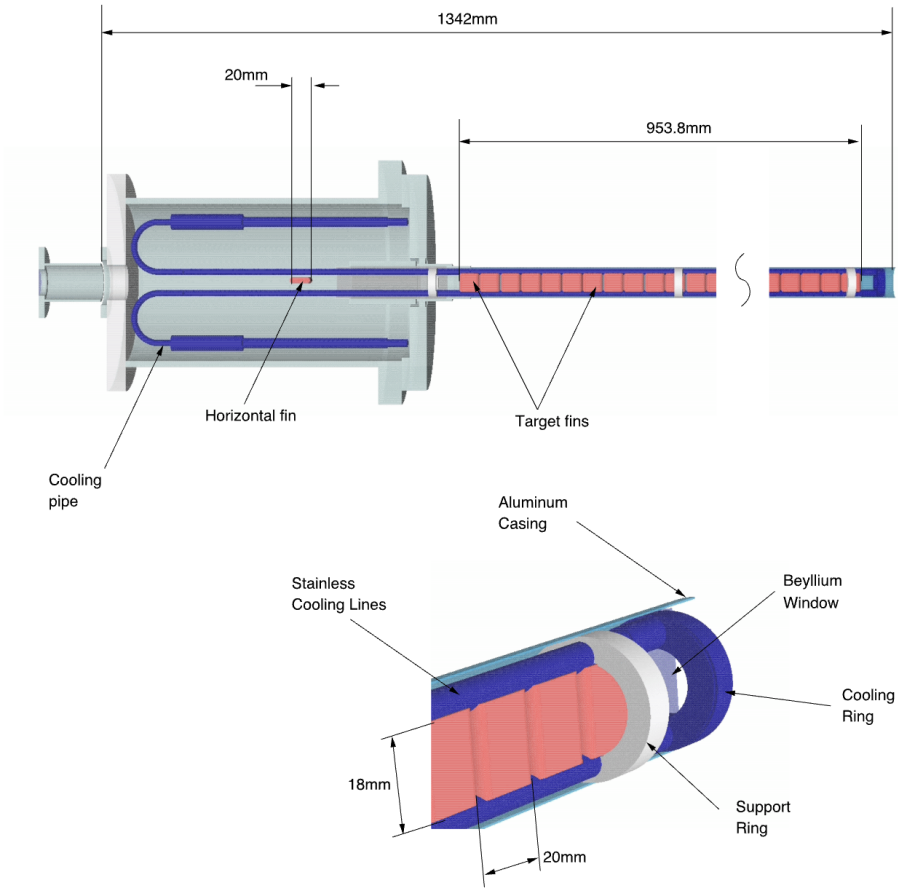


Figure 3.3: Schematic of NuMI Target and the Target Canister [38]. In addition to the 47 fins, a horizontal fin is installed for beam alignment. The target is cooled using a water-cooling system.

horns (Horn1 and Horn2). The magnetic horns play a crucial role in significantly enhancing the hadron flux within the desired energy range while providing flexibility in selecting that energy. Both the distance between the target and the first horn, as well as the separation between the two horns, are adjustable. Most of the charged pions π^+/π^- decay into leptons through $\pi^+ \to \mu^+ + \nu_\mu$ and $\pi^- \to \mu^- + \overline{\nu}_\mu$ in the Decay Pipe with a diameter of 2 m and a length of 675 m [38]. The direction of the current in the magnetic horn is used to select π^+ and π^- mesons. In the Forward Horn Current (FHC) mode, π^+ mesons are predominantly selected, resulting in a beam primarily composed of muon neutrinos. Conversely, in the Reverse Horn Current (RHC) mode, the majority of the outgoing neutrinos are anti-muon neutrinos. The approximated relationship between pion and neutrino energy is [40]

$$E_\nu = \frac{0.427 E_\pi}{1 + \gamma^2 \theta^2} \tag{3.1}$$

where $\gamma = \frac{E_\pi}{m_\pi}$ is the Lorentz boost of the pion and θ is the angle between directions of pion and neutrino. A Hadronic Monitor consisting of a grid of ionization chambers is positioned at the end of the decay pipe to measure the flux of residual protons and hadrons that did not decay or interact, serving as a diagnostic tool for monitoring the integrity of the Target-Baffle system. The primary flux components at the Hadronic Monitor are protons and hadrons originating from the target. Any loss of material from the target would result in an increased flux intensity at the Hadronic Monitor, providing a clear indicator of issues in the Target.

Directly downstream of the target is the Hadron Absorber, which, as its name

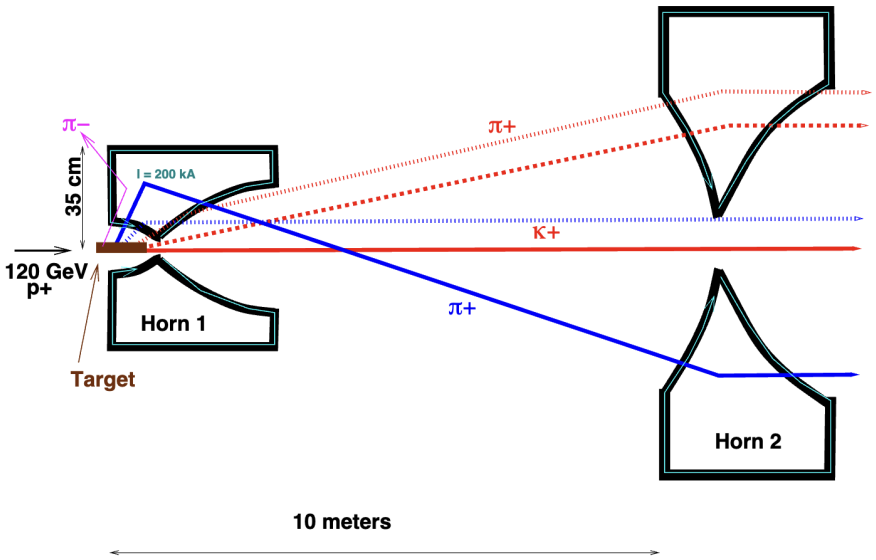


Figure 3.4: Schematic of Marnetic Horns [38]. Possible trajectories of charged mesons are shown.

suggests, is designed to absorb the majority of the protons and hadrons from the beam. It has an aluminum core with heat transferred to a water cooling system. The core is surrounded by steel blocks, which capture hadronic shower tails created inside the core, and an outer layer of concrete, providing radiation shielding by absorbing thermal and low-energy neutrons which cannot be absorbed by iron.

Downstream of the Hadron Absorber, three Muon Monitors are positioned, separated by layers of dolomite rock, and are based on ionization chamber technology [38]. Among all charged particles, only muons are capable of traversing the absorber and the intervening rock to reach the monitors. The placement of each monitor corresponds to a specific muon energy threshold, with monitors located behind more rock layers detecting only higher-momentum muons, as lower-momentum muons are stopped before reaching the detectors.

3.2 Off-Axis Design

Both the Near Detector (ND) and the Far Detector (FD) of NOνA reside 14.6 mrad off the NuMI beam axis.

In addition to the kinetic relation (Equation (3.1)) provided, we have the following flux equation for muon neutrinos produced by pion decay:

$$F_\nu = (\frac{2\gamma}{1+\gamma^2\theta^2})^2 \frac{A}{4\pi L^2} \tag{3.2}$$

where A is the area of detector and L is the distance between detector and neutrino source. Figure Figure 3.5 shows how the muon neutrino flux and energy change as the angle θ increases. Moving the detector off-axis introduces a trade-off: the neutrino energy range becomes more constrained, but this comes at the cost of reduced flux. This effect can be visualized in figure Figure 3.6 and Figure 3.7 made from the simulation that placed detectors in different angles with respect to the beam line. The 14.6 mrad angle is chosen to ensure that the peak of the energy spectrum aligns with the energy ($\approx$1.6 GeV) that maximizes the ν_e appearance probability for a baseline of $L = 810$ km.

Figure 3.8 shows more completed spectra in FD. First, the following spectra are plotted without considering oscillation: 1. Green dashed line: ν_μ CC events at the FD; 2. Magenta line: ν_e CC events at the FD. Then, the following spectra are made considering the oscillation: 1. Green line: ν_μ CC events at the FD; 2. Violet line: ν_e CC events at the FD. All four spectra above use true neutrino energy as the x-axis. For the neutral current spectrum (blue line), the x-axis represents the

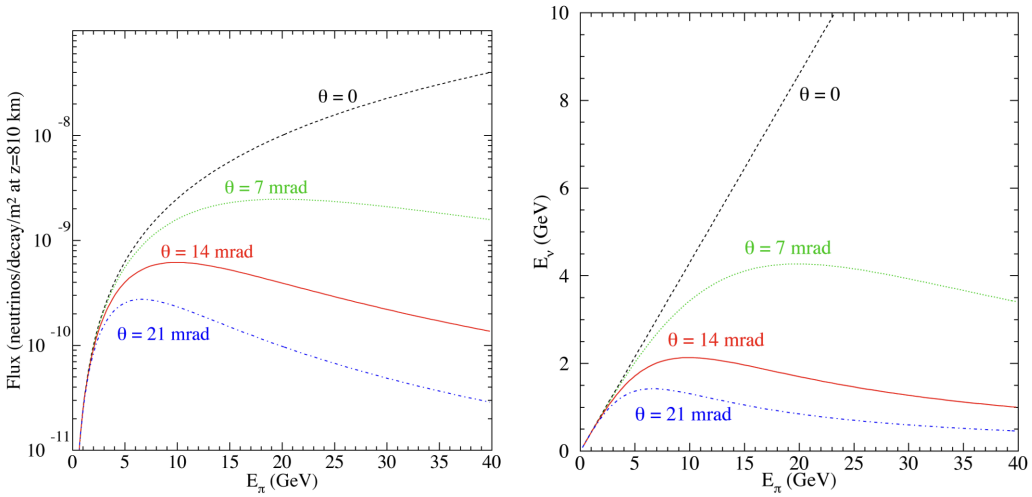


Figure 3.5: Flux and energy of muon neutrinos as a function of pion energy for different off-axis angles [22, 37].

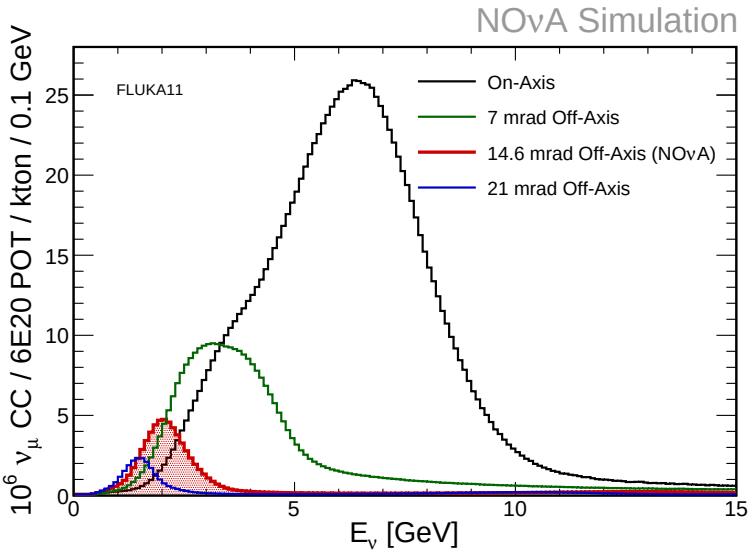


Figure 3.6: ν_μ charged current (numu CC) event energy spectrum for **Near Detector** simulated by Fluka (a particle physics MonteCarlo simulation package) [41].

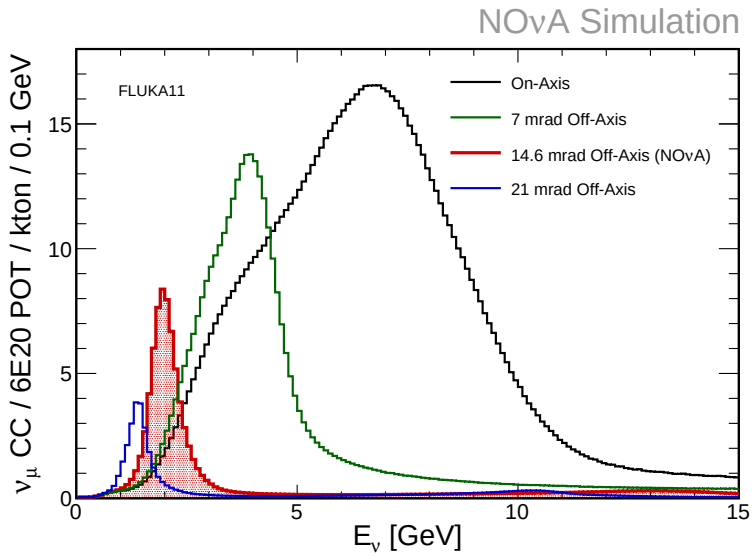


Figure 3.7: ν_μ charged current event energy spectrum for **Far Detector** simulated by Fluka assuming no oscillations [41].

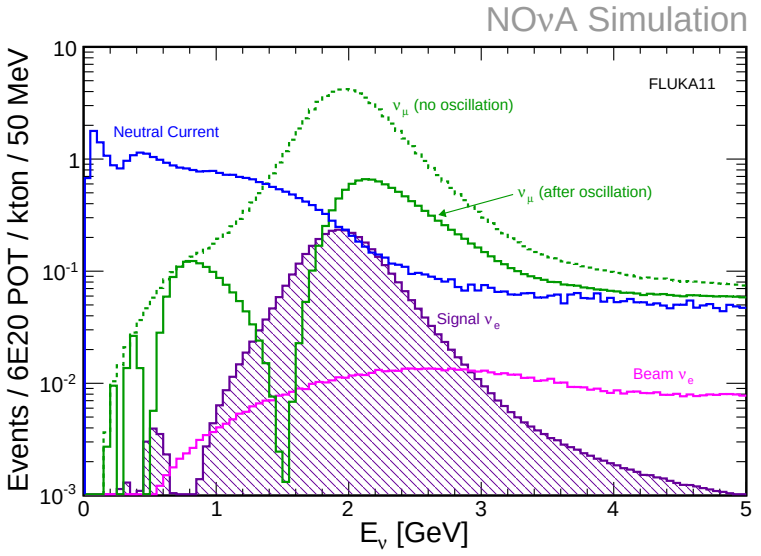


Figure 3.8: A complete set of spectra at the FD [41].

true "visible energy", which corresponds to the true energy not carried away by the outgoing neutrino. We refer to it as "true" because these results are based on Monte Carlo simulations where the energy of each particle is explicitly known.

3.3 The Liquid Scintillator

During the design of the NOνA detectors, several critical considerations are prioritized: optimizing the detector's efficiency for detecting ν_e interactions at approximately 2 GeV, while simultaneously minimizing costs to enable the construction of a large-scale Far Detector. A few candidate designs were considered.

- Cherenkov detectors achieve optimal performance for low-energy neutrinos (below 1 GeV). However, at higher energies, hadronic energy reconstruction becomes less accurate. Moreover, the overlapping Cherenkov rings further complicate particle identification [42].

- The Liquid Argon Time Projection Chamber (TPC) was widely acknowledged for its significant potential in neutrino physics. However, at the time of the NOνA technology decision, the largest existing Liquid Argon TPC was the ICARUS detector [43], with an imaging mass of 500 tons—substantially smaller than the proposed 14.4 kt NOνA Far Detector. Consequently, the technology was deemed immature and was not selected.

- Resistive Plate Chamber (RPC) sampling calorimeters exhibit several limitations. These include relatively lower particle identification efficiency for electron neutrinos, higher construction costs compared to liquid scintillator

detectors, and vulnerabilities for cosmic-ray penetration. Additionally, the performance of RPC systems tends to degrade over time due to the aging of their gas-handling infrastructure, further impacting their reliability and long-term utility [42].

- A segmented liquid scintillator detector is designed to be sensitive over a substantial fraction of its active volume [42], thereby maximizing detection efficiency. A liquid scintillator typically consists of a primary scintillant, waveshifters, wavelength shifting (WLS) fibers and a solvent.

The final design was selected following an evaluation of cost-effectiveness and performance. More specifically, NOνA chose mineral oil as the solvent and pseudocumene (1,2,4-trimethylbenzene) as the scintillant. The scintillator is primarily composed of mineral oil, which constitutes approximately 95% of its total mass, with pseudocumene making up about 5%. Other components are present in smaller proportions. Upon excitation by an ionizing particle, the primary scintillator, pseudocumene, undergoes decay, emitting photons within the 270–320 nm range. These ultraviolet photons excite the first wavelength shifter, PPO (2,5-diphenyloxazole), which subsequently emits photons predominantly in the 340–380 nm range. In the next stage, the down-converted scintillation photons excite the second wavelength shifter, bisMSB (1,4-bis-(o-methyl-styryl)-benzene), leading to emission predominantly in the 390–440 nm range. Photons in the 390–460 nm range are then absorbed by the wavelength shifter in the WLS fiber of 0.7 mm diameter, which further shifts them to green 490-550 nm. These photons are transmitted to the Avalanche Photodiode (APD, a highly sensitive type

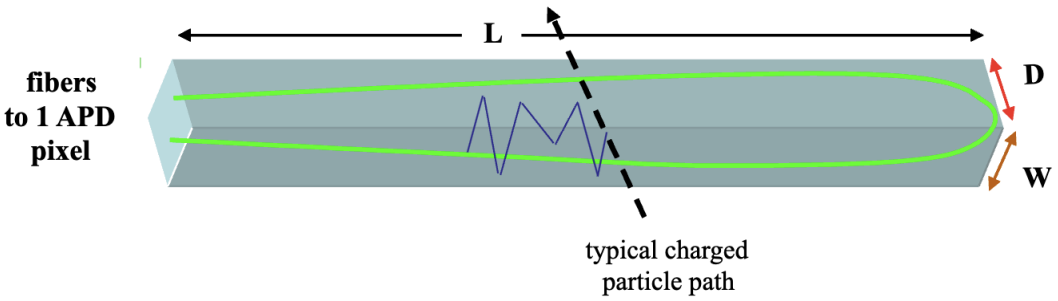


Figure 3.9: Schematic of a PVC cell of dimensions $W \times D \times L$: Photons (blue lines) are generated when a charged particle interacts with the cell. These photons propagate through the cell, reflecting off the walls until either absorbed by the looped wavelength-shifting fiber or lost. Each cell is coupled to a single APD for photon detection.

of photodetector) at the end of fibers. When incident photons hit the APD, they generate electron-hole pairs through the photoelectric effect. The applied high electric field in the depletion region accelerates these carriers, causing secondary ionizations and resulting in an avalanche multiplication process. This mechanism significantly amplifies the initial signal, enabling the detection of low-intensity light. APDs in NOνA are cooled to $-15°$C to reduce the thermal noise.

The basic unit of the detectors is a rigid rectangular PVC plastic cell (see fig Figure 3.9) filled with liquid scintillator. The PVC is loaded with titanium dioxide (TiO_2) to achieve a higher refractive index. The interior width of each cell is 3.8 cm, transverse to the beam direction, and the depth is 5.9 cm, along the beam direction. The cell length is 15.5 m for the Far Detector and 4.1 m for the Near Detector.

The building blocks of the two detectors are PVC modules (see fig Figure 3.10).

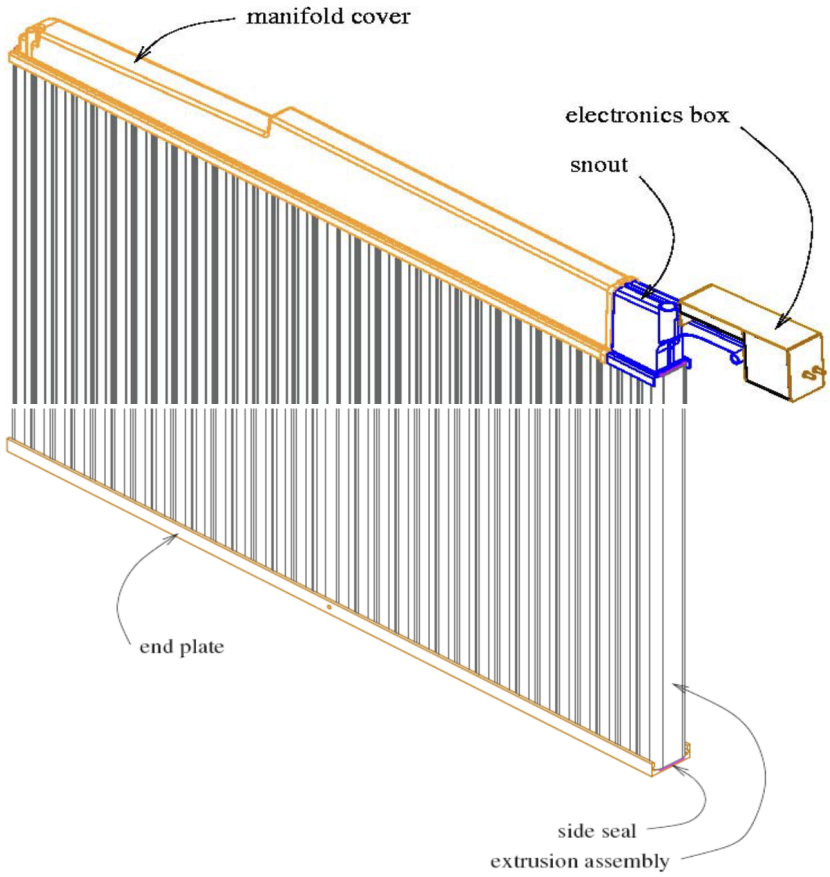


Figure 3.10: Schematic of PVC module assembly [37].

Each module is constructed by gluing together two 16-cell PVC extrusions, forming a 32-cell assembly. It is capped at one end with a PVC end plate and at the other end with a fiber manifold, which channels the fibers to the APDs. The end plate and manifold connect the 32 cells and make them a common liquid volume. The snout on the manifold assembly is a compact component that supports the APD, associated electronics, and cooling system, while efficiently routing fibers. It prevents pseudocumene vapor condensation on the fibers and includes fill and vent ports for introducing liquid scintillator into the cells.

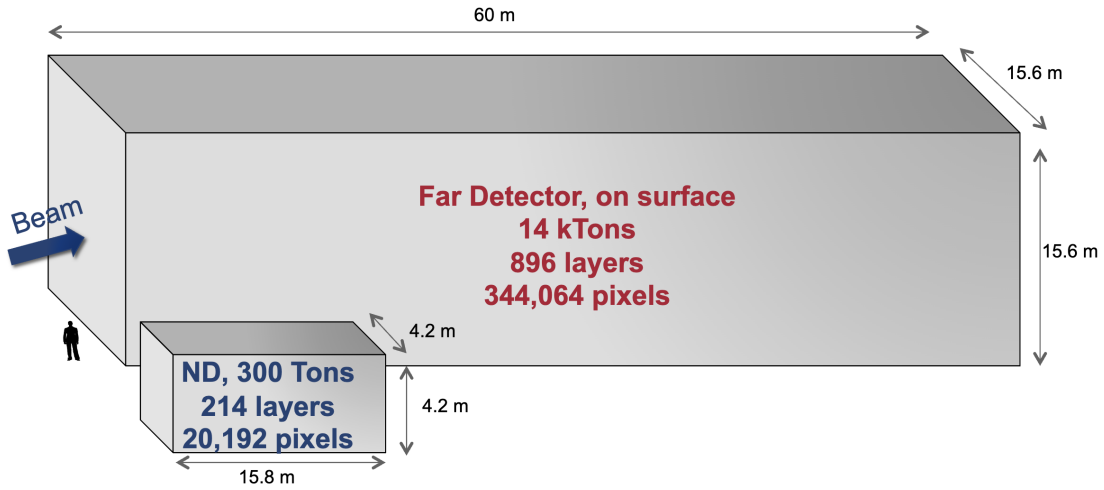


Figure 3.11: Schematic of FD and ND.

3.4 The NOνA Detector

The NOνA experiment employs two detectors (see Figure 3.11) to investigate neutrino oscillations. The Near Detector is located at Fermilab, while the Far Detector is situated at Ash River, Minnesota. Both detector are functionally identical, while the Near Detector is much smaller than the Far Detector.

3.4.1 The Far Detector

The Far Detector is housed within a building at the Earth's surface, with a 16-meter-deep excavation into solid granite made to accommodate the detector. The excavated granite is repurposed as a cosmic ray shield for the above-ground walls of the structure. The roof above the detector comprises 1.22 meters of concrete, overlaid with 0.15 meters of barium sulfate ($BaSO_4$) to enhance cosmic ray shielding [44]. The detector itself contains 344,064 cells, collectively filled with 2,674,041

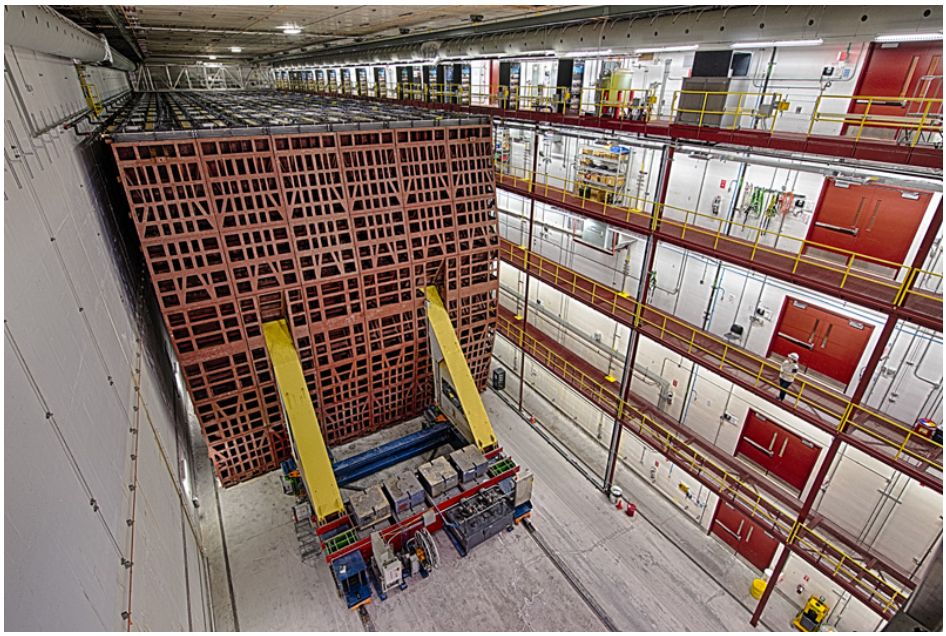


Figure 3.12: NOνA Neutrino Experiment Far Detector at Ash River, MN. Image from Fermilab image database.

gallons of liquid scintillator [45]. The Far Detector weighs 14 kilotons and features 896 alternating vertical and horizontal planes, each perpendicular to the beam line. Each plane has dimensions of 15.6 m $\times$ 15.6 m $\times$ 5.9 cm [46]. The overall size of FD is 15.6 m $\times$ 15.6 m $\times$ 60 m.

The average NuMI beam pulse rate is 0.75 Hz, with each pulse producing a burst of neutrinos (beam spill) lasting approximately 10 μs. Data collected within a 420 μs window surrounding the beam spill are used to measure the cosmic background.

Figure 3.13: NOνA Neutrino Experiment Near Detector. Image from Fermilab image database.

3.4.2 The Near Detector

The Near Detector (ND) is positioned 105 meters below the surface, minimizing cosmic ray interference, and is located approximately 1 km from the neutrino source. The solid angle subtended by the ND at the neutrino source is substantially greater than that of the FD. Consequently, the ND receives a substantially higher neutrino flux than the FD, necessitating the implementation of faster electronics for data acquisition. The ND consists of 206 planes, each with dimensions of 15.6 m $\times$ 15.6 m $\times$ 5.9 cm [37]. The detector has an overall size of 4.2 m $\times$ 4.2 m $\times$ 15.8 m, containing 20,192 cells and holding 40,141 gallons of liquid scintillator. Its total weight is approximately 290 tons. Given its relatively compact design, the ND is significantly shorter than the FD, allowing muons to

traverse its entire length. Downstream of the ND, a "muon catcher" is installed, consisting of 11 pairs of scintillator planes (one vertical and one horizontal) separated by 10 steel planes, aiming to improve muon containment. The steel planes block most of the muons while the scintillator planes can be used to measure the muon energies.

3.5 Data Acquisition System

The huge amount of APD channels in detector requires efficient data acquisition and processing. The Data Acquisition System (DAQ) is designed to concentrate the data into a stream for offline analysis. It provides data buffer for delayed recording or rejection. The signals from the APD arrays are processed by the Front-End Boards (FEBs) with low noise. Each FEB is equipped with a custom Application-Specific Integrated Circuit (ASIC) that amplifies and shapes the signals, which are then sampled at a 2 MHz digitization clock. Data streams from up to 64 FEBs are transmitted to a Data Concentrator Module (DCM). Within the DCM, an FPGA organizes the data into 50 μs "microslices", while an embedded PowerPC system retrieves the data from the FPGA. Subsequently, an event-building algorithm aggregates these microslices into larger 5 ms "millislices". Each DCM is connected to and transmits data to its assigned "buffer node" within a computing cluster via Gigabit Ethernet. The buffer node maintains a circular buffer that stores a fixed number of events. This enables high-latency trigger decisions, including those for supernova neutrino events.

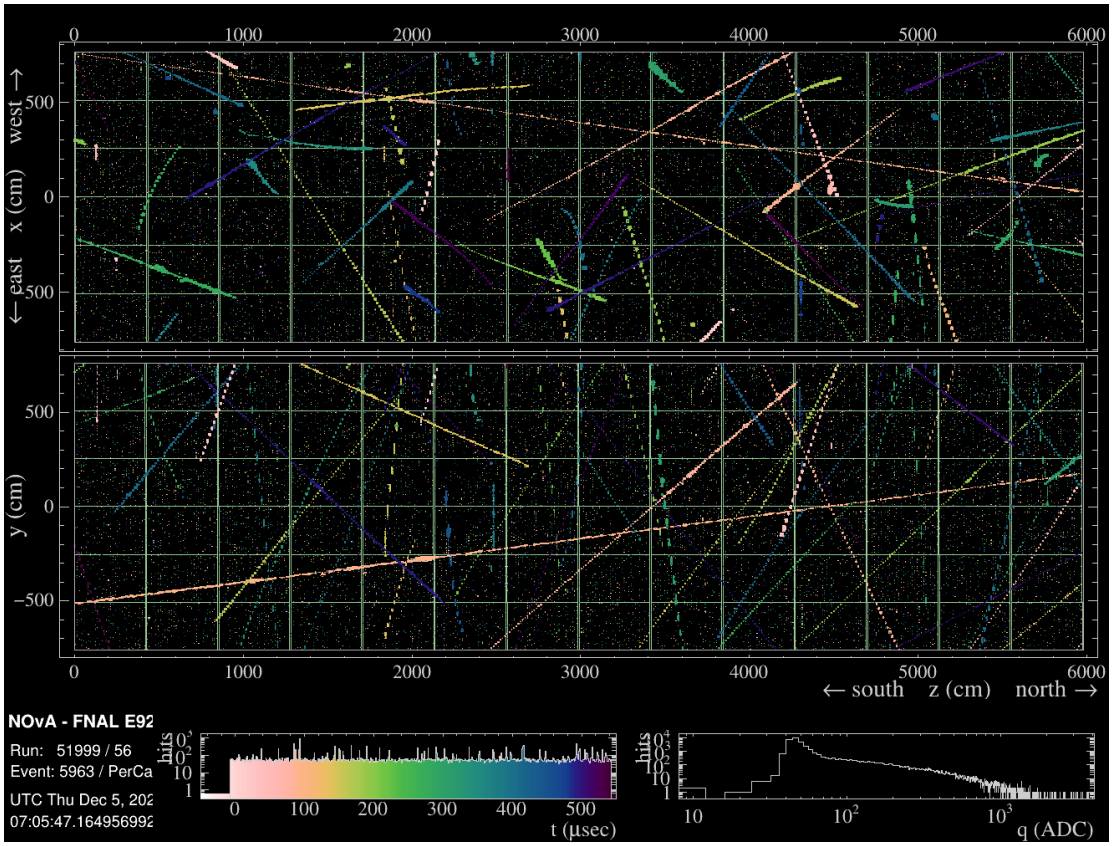


Figure 3.14: The event display of 500 μs window time from the NOvA Far Detector. The XZ view corresponds to vertical aligned planes while the YZ view corresponds to the horizontal aligned planes. The majority of tracks are created by cosmic ray muons. The beam is off.

Chapter 4

Simulation and Calibration

In this chapter, I will describe how the Monte Carlo (MC) samples are produced and how the calibration works in NOνA. The simulation chain comprises multiple stages, beginning with beam simulation and culminating in the generation of MC samples, which serve as an analogue to the data collected from the two detectors[1]. There are two types of calibration in NOνA: energy calibration and time calibration. The energy calibration is a multi-step process that determines the amount of energy deposited in each detector cell. The calibrated energy is then used in subsequent analyses. The time calibration aims to correct time discrepancies between different front-end boards (FEBs), ensuring accurate time information is provided by the detectors.

[1]In this thesis, we refer to the software-simulated samples as MC and the experimentally collected samples as data.

4.1 Beam simulation

The NuMI beam in NOνA is simulated by a package called "FLUGG" which consists of FLUKA [47] and Geant4 [48, 49, 50] geometry. FLUKA and Geant4 are two Monte Carlo simulation frameworks for particle transport and interactions with matter. Geant4 has known limitations in accurately simulating hadron production at medium energies [51]. FLUGG uses Geant4 geometry files that describe the information of targets, horns, baffle, etc. as input and simulate the hadron production inside the target, focusing in the horns and downstream tertiary production including the decay of kaons, muons and pions with FLUKA. Initial particles are 120 GeV protons. The outputs of FLUGG, the flux files, record the resulting neutrinos' flavor, momentum and energy along with information about their parentage.

4.2 Neutrino Event Generation

Once the flux files are generated, they can be used as inputs of GENIE [52] to generate neutrino events (interactions between neutrinos and detectors). Meanwhile, cosmic rays introduce a significant background in the Far Detector. NOνA uses CRY [53] to simulate the cosmic ray events.

4.2.1 Neutrino Event Generation by GENIE

Generates Events for Neutrino Interaction Experiments (GENIE) is the most widely used neutrino event generator. It is adopted by nearly all modern neutrino experiments including NOνA, MicroBooNE, DUNE, SBND, etc. GENIE's physics models cover the dominant scattering mechanisms from a few MeV to several hundred GeV, applicable to all neutrino flavors and target types. These models encompass various essential physical processes and can be categorized into three main groups: nuclear physics models, cross-section models, and hadronization models. GENIE simulates neutrino interactions with detectors by modeling the entire process from the initial neutrino interaction with the nuclei in the detector to the final state particles that escape the nucleus. The initial interaction incorporates various cross-section models such as quasi-elastic (QE) scattering, resonance (RES) pion production, deep inelastic scattering (DIS), meson exchange current(MEC) and coherent (COH) scattering. See Table 4.1 for more information. After the primary interaction, outgoing hadrons undergo secondary interactions within the nuclear environment. This is called final state interactions (FSI). Generally, there are three topological channels: Absorption (ABS, no pions are observed in the final state), Charge Exchange (CX, a single π^0 is observed in the final state) and Quasi-Elastic (QE, single pion with same sign as initial pion is observed). The output of GENIE includes a detailed list of particles generated in the simulation, along with event-level information such as interaction type, differential cross section, and event weights. This data provides a comprehensive description of each simulated neutrino interaction, supporting analysis and

reconstruction in neutrino experiments.

Interaction Type	**Description**
Quasi-Elastic Scattering (QE)	Neutrino interacts with a single nucleon, typically changing its type (CC) or ejecting it (NC). The nucleus may break up.
Resonance Production (RES)	Neutrino excites the target nucleon to a baryon resonance (e.g., $\Delta(1232)$), which subsequently decays, often producing pions.
Deep Inelastic Scattering (DIS)	Neutrino scatters incoherently off quarks within the nucleon at high energy/momentum transfer. Produces a complex hadronic state.
Meson Exchange Current (MEC / 2p2h)	Neutrino interacts with a correlated pair of nucleons (2p2h: two-particle, two-hole). Leads to ejection of multiple nucleons.
Coherent Scattering (COH)	A neutrino scatters off an entire nucleus as a whole.

Table 4.1: Summary of Neutrino Interaction Cross-Section Types.

4.2.2 Neutrino Event Generation by CRY

Cosmic-ray Shower Library (CRY) generates cosmic-ray particle showers at one of three elevations (sea level, 2100m, and 11300m). NOνA experiment assumes sea level instead of the actual altitude when generating cosmic Monte Carlo samples for the Far Detector.

The outputs[2] of either GENIE and CRY are passed to Geant4 for simulations of particle propagation and energy deposits within the detectors.

[2]Note, that the exact modeling of cosmic rays is not needed in sterile neutrino analysis, since we use actual cosmic rays collected in data.

4.3 Detector Simulation

4.3.1 Detector Geometry

The geometry used in the Geant4 simulation includes material compositions, dimensions, and total masses. The Far Detector utilizes two similar scintillator formulations: low-pseudocumene and high-pseudocumene. In contrast, the Near Detector exclusively employs the high-pseudocumene formulation. NOνA uses the average scintillator composition for each detector, except for calibration purposes. The effects of absorbed water, as well as dissolved nitrogen and oxygen, are also taken into account. While the fibers immersed in the scintillator are not explicitly modeled, they are treated as part of the overall scintillator-fiber mixture [54]. The geometry also includes the muon catcher of the Near Detector. The materials and weights of both detector are summarized in table Table 4.2 and Table 4.3.

Detector	Material	Density (g cm^{-3})
Far, Near	Scintillator	0.859
Far, Near	PVC	1.49
Far, Near	Glue	1.34
Far, Near	WLS Fiber	1.0457
Far (Pivoter)	Steel	7.87
Near (Muon Catcher)	Steel	7.87

Table 4.2: Summary of materials in both detectors (adapted from [55]).

Detector Component	weight (metric tons)
Far Detector (total)	13843.59
Far Detector Pivoter	94.2314
Near Detector (total)	284.42358
Near Detector Muon Catcher	93.48968

Table 4.3: Weights of detector components [55].

4.3.2 Birks-Chou model and energy deposite

The deposited energy of each particle needs to be converted to scintillator light. Ideally, we expect a linear relationship between the energy deposited along the particle track and the light yield:

$$\frac{dL}{dx} = L_0 \frac{dE}{dx}, \tag{4.1}$$

where L_0 is a coefficient that connects the deposited energy per unit length ($\frac{dE}{dx}$) and light yield per unit length ($\frac{dL}{dx}$).

However, due to the recombination and quenching effects, the final light yield is reduced in organic scintillators. Therefore, a more complicated non-linear model is needed. NOνA adopted Birks-Chou model which provides up to second order correction term to describe the actual light yield:

$$\frac{dL}{dx} = \frac{L_0 \frac{dE}{dx}}{1 + k_B \frac{dE}{dx} + k_C (\frac{dE}{dx})^2}, \tag{4.2}$$

where k_B is the Birks' constants and k_C is the Chou constants. NOνA determines these two empirical constant by matching data and MC in near detector events.

The best fit values are k_B = 0.040 g MeV^{-1}cm^{-2} and k_C = -0.0005 cm^2 MeV^{-2} [55].

4.3.3 Photon transport

Photon transport includes the collection of scintillation photons by wavelength shifting fibers, transport of photons inside fibers, and the APD response. These processes are parameterized in NOνA by a template to reduce the simulation cost, given the identical function and structure of readout cells.

Photon capture by fibers

A ray-tracing simulation was developed to model the capture of scintillation photons by the fiber, tracking each photon from its point of emission to its collection. To analyze the photon collection efficiency, a two-dimensional histogram (see Figure 4.1) is employed, illustrating the collection rates of scintillation photons as a function of ΔT and ΔZ, which represents the time difference and projected position difference (along the fiber) between the energy deposition and photon collection. In addition, NOνA incorporates a model of the cell wall reflectivity as a function of wavelength (see Figure 4.2), further refining the simulation's accuracy in describing photon behavior.

Fiber Transport

When a photon is captured by the fiber, it can propagate in either of two possible directions along the fiber. We assume that half of the photons take the short

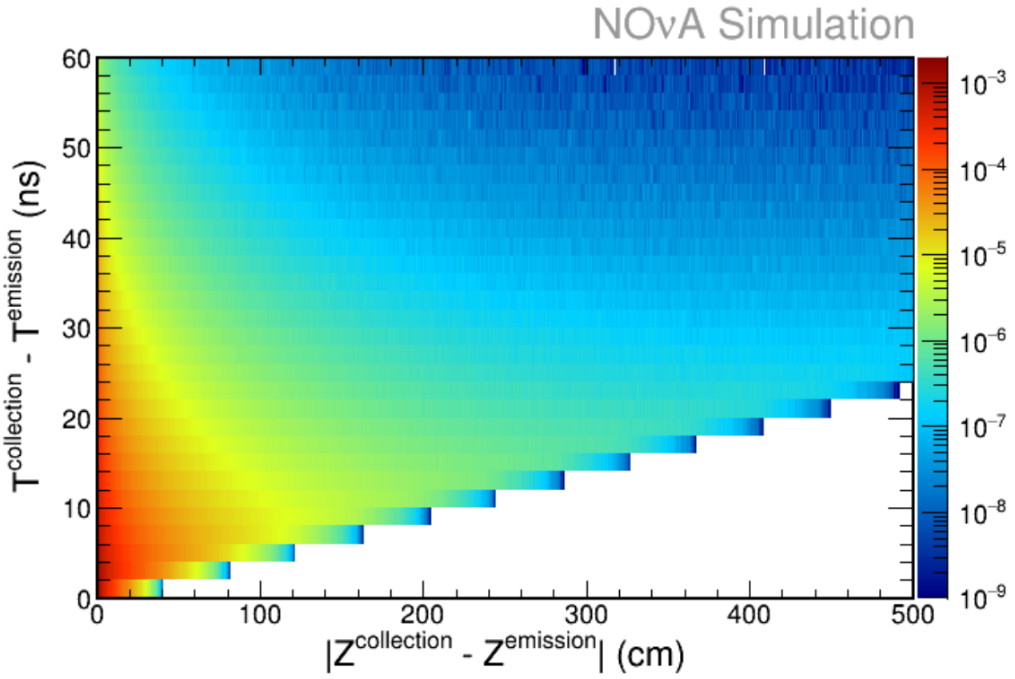


Figure 4.1: 2D histogram of photon collection rate as a function of time difference ΔT and projected location difference ΔZ [55].

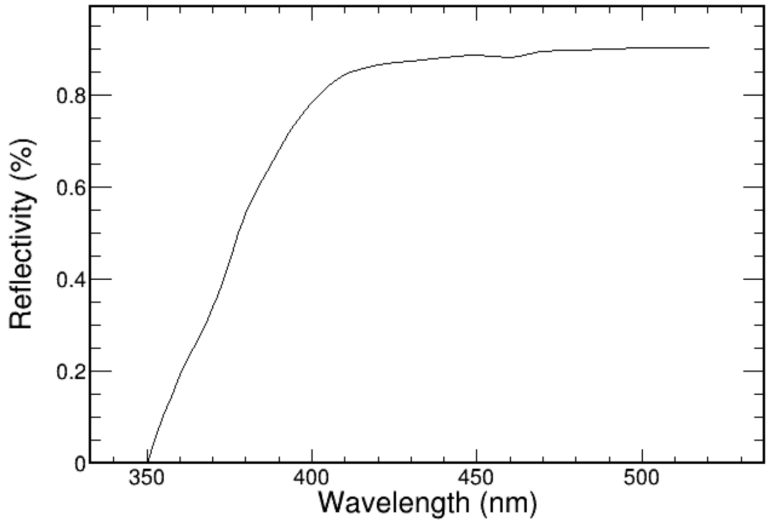


Figure 4.2: Wall reflectivity of readout cells as a function of wavelength [55].

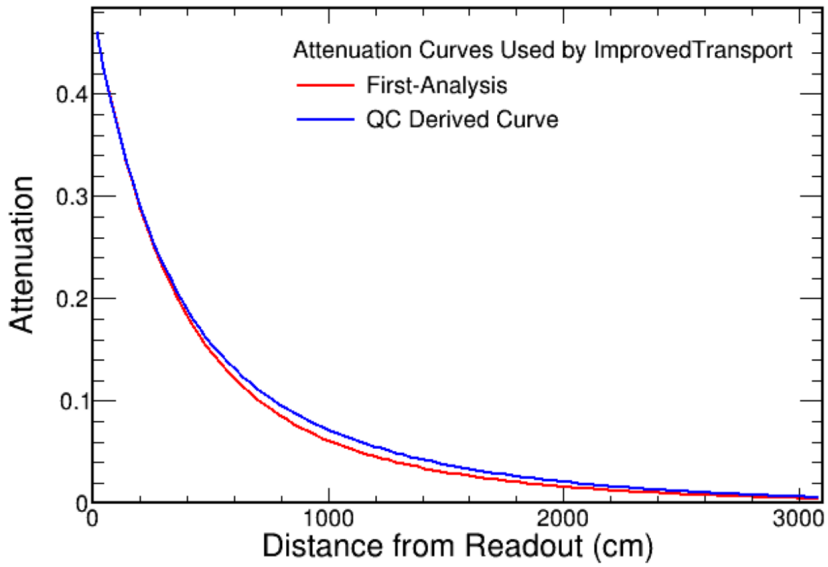


Figure 4.3: The light attenuation curve as a function of the distance to readout. There are two versions of curves. The red curve was used in the first analysis and the blue curve was an updated version which better matches data [55].

path while the other half will take the long path. An attenuation curve is used to describe the number of surviving photons that travels in the fiber as a function of the distance from the APD readout. In the first analysis, the curve was determined in bench tests. This curve was found to be inaccurate for long distance. The later simulation adopted a new curve derived from the average results of quality control (QC) tests conducted on multiple spools of wavelength-shifting fiber (see Figure 4.3).

4.3.4 Electronics response simulation

The next step in the simulation involves converting the number of photons reaching the APD into the number of photoelectrons (N_{pe}) absorbed by the APD. This

process takes into account the quantum efficiency of the APD and applies Poisson sampling with $\lambda = N_{pe}$. In addition, an excess noise factor (F) is introduced to expand the variance of signal distribution to FN_{pe} to match the real distribution.

The front-end boards contains three chips: 1. an ASIC that performs pulse shaping using a CR-RC circuit, 2. Analog to Digital Converter (ADC) that digitized the analog signals from ASIC into integers between 0 and 4095, 3. an FPGA that performs real-time zero suppression by applying a dynamic threshold to the dual correlated sampling (DCS) trace, identifying and retaining only significant signal peaks. All these three chips need to be simulated. We will not delve into the details here, as the process is significantly more complex.

The comparison between low-level quantities in the data and full simulation is conducted using cosmic ray data, with the signal regions kept blind. The results demonstrate a good agreement between the two.

4.4 Energy Calibration

NOνA performs independent energy calibrations in two detectors [56]. Cosmic ray muons are utilized in both detectors to derive the calibration parameters. Digital signals from ADC are converted into energy values during calibration.

4.4.1 Relative Calibration

The calibration process begins with a simple rescaling of the ADC counts to the approximate number of photoelectrons (PE). Following this, a relative calibration is carried out to eliminate the variations in detector response across different cells.

This takes the following factors into account.

Light Attenuation

The position of energy deposition matters. As photons propagate through the fibers, attenuation occurs, leading to a greater reduction in ADC counts for photons originating farther from the fiber ends. This effect is more obvious in Far Detector, as the cell is longer compared to the Near Detector.

Threshold Bias

The FEBs set a readout threshold that filters noise. Therefore, higher energy deposition is required to trigger the recording of signals. Due to attenuation, the average energy to trigger recording increases for energy deposition located farther from the ends of the fiber. As a result, a given energy deposition may not be detected if it is located far from the ends of the fibers, whereas it may be detected if it is located near the ends of the fibers.

Detector Shadowing

Since NOνA uses cosmic muon data for calibration, it is crucial to check whether the energy deposition per unit length along the track (dE/dx, as known as the stopping power) of cosmic muons remains consistent across the detector. Unfortunately, due to the large dimensions of the Far Detector, muons that travel through the entire detector to reach the bottom generally have a higher average energy than those that stop near the top, which results in a higher average dE/dx in the lower part of the detector.

Other Factors

- The fiber reflectivity varies across different cells, leading to differences in attenuation lengths.
- The scintillator is also different among the detector cells, and air bubbles within the scintillator create a non-uniform light propagation medium in each cell.
- For horizontal cells, gravity causes the fibers to sag.
- The ADC response to energy deposition at the edge of the cell differs significantly from that further away. Thus, a correction is needed to eliminate the difference.

To study how the ADC response varies with the energy deposition location, NOνA constructs a 2D histogram (see Figure 4.4) of the photoelectrons per unit length (PE/dx) along the track[3] as a function of the position W along the cell. Here, W is defined as the relative projected position with respect to the cell center.

NOνA fits the following attenuation profile to PE:

$$PE/dx = C + A(e^{\frac{W}{X}} + e^{-\frac{L+W}{X}}). \tag{4.3}$$

A, C are coefficients that vary cell-by-cell. C is referred to as the "background". X is the attenuation length which also vary in different cells. L is the length of cell. The attenuation fit does not perfectly match the data, with some discrepancies,

[3]The unit of track is cm. PE/dx is commonly expressed in PE per cm (PE/cm).

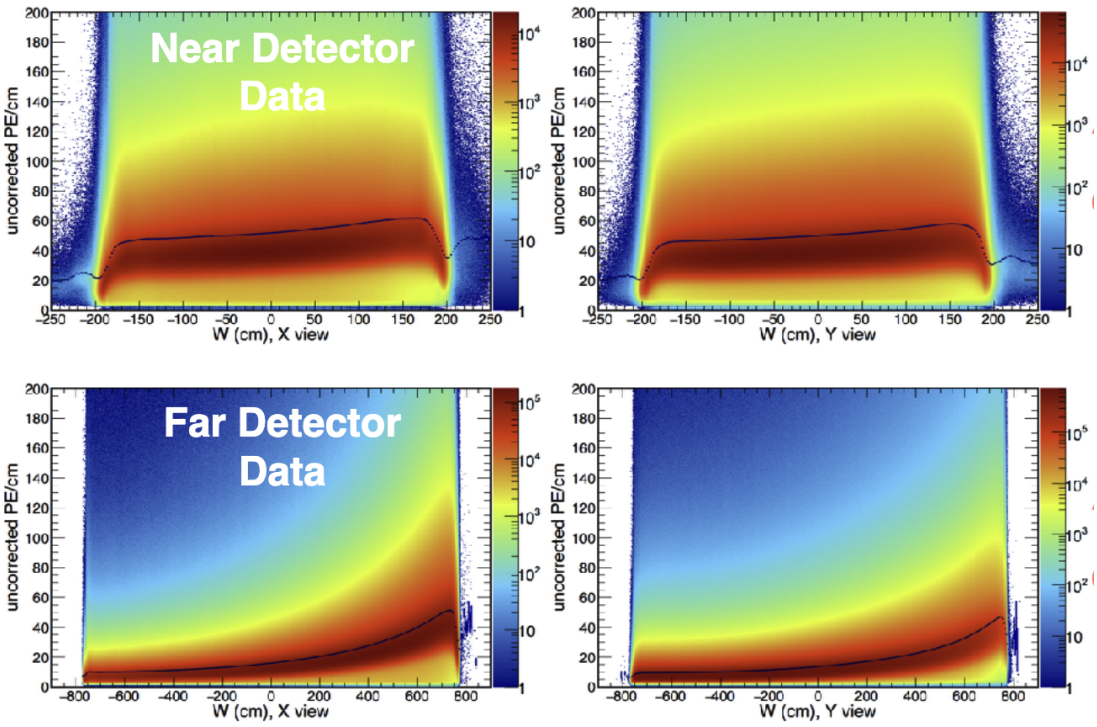


Figure 4.4: Uncorrected PE per unit length versus location W on the cell for FD and ND in X view and Y view. Black dots correspond to the mean values. Boundary effect is obvious for $|W| \gtrsim 200$ cm in ND and $|W| \gtrsim 800$ cm in FD. [57].

including points deviating from the fitted curve at the edges. To address this, locally weighted scatterplot smoothing (LOWESS) is used as the second fit upon the attenuation fit. Parameters in both fits are saved with the cell number[4] in a CSV file.

Using the parameters obtained from the fit in each cell, we can now normalize the ADC response to a corrected PE value, denoted as PECorr, which is independent of the position in the detector.

4.4.2 Drift Calibration

As detector materials and electronics age over time, the detector response undergoes gradual changes, commonly referred to as drift. The change of mean PECorr/dx in each cell is monitored. Assuming that the energy distribution of cosmic muons remains invariant over time, a linear fit is applied to derive a time-dependent correction for PECorr/dx. This correction ensures that the corrected PECorr/dx remains independent of time[5]. Drift calibration is conducted across different detector cells.

4.4.3 Absolute Calibration

The absolute calibration aims to convert the position- and time-independent PECorr into an absolute energy scale (e.g., in GeV). The conversion factor, sometimes referred as the calorimetric energy scale, cannot be obtained from data, since the true energy deposition in the detectors is not directly measurable. A

[4]Each cell has its unique number in the detector.
[5]This is a new PECorr, in contrast to the old PECorr, which lacks drift calibration.

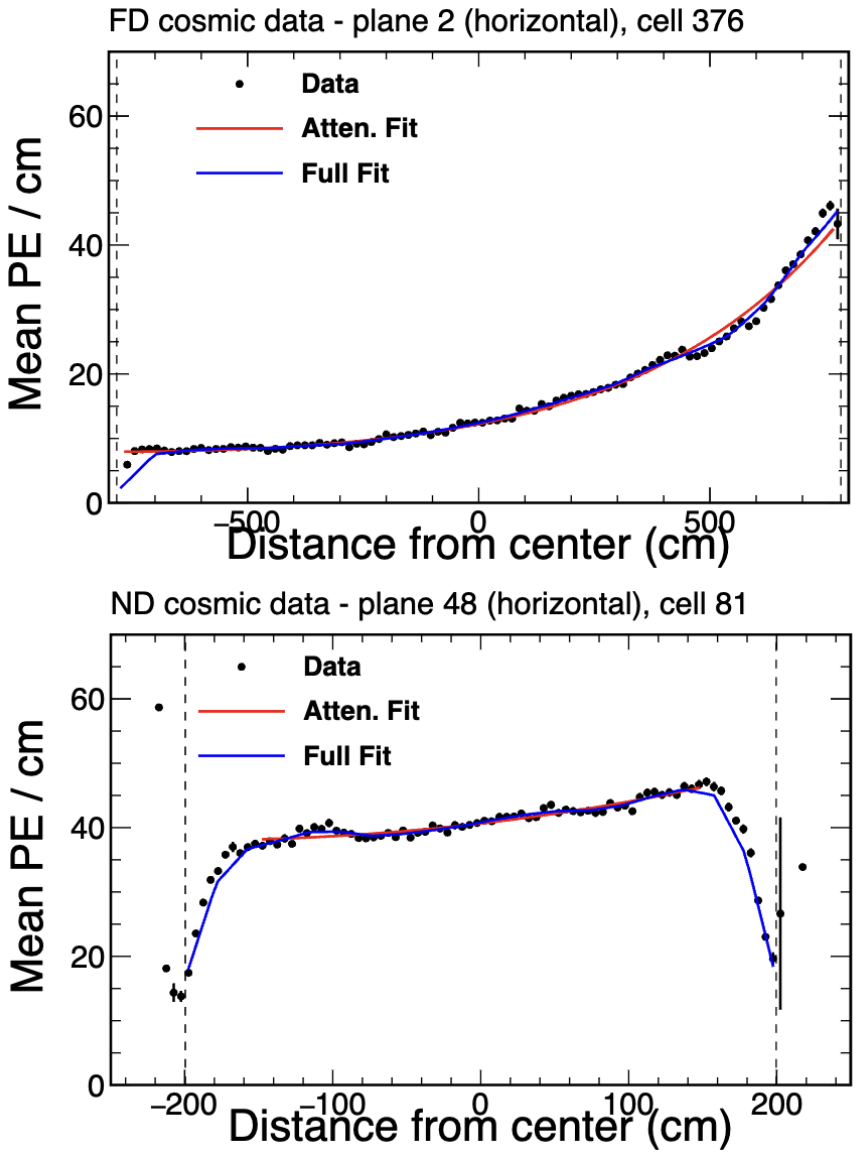


Figure 4.5: Attenuation fit (red curve) and full fit (attenuation fit + LOWESS) on FD cell 81 and ND cell 376 [57]. We can see that the full fit (blue curve) accurately captures the data pattern at the edges.

simulation is required to determine the conversion factor.

The energy deposition dE/dx of muons can be well described by the Bethe-Bloch equation. The absolute calibration relies on muons that stop within the detector, known as stopping muons. Muons primarily decay through the channel

$$\mu^- \rightarrow e^- + \overline{\nu}_e + \nu_\mu.$$

The electrons (as known as Michel electrons) generated at the end of muon tracks can be used to identify the stopping muons. Position reconstruction is performed using the detector hits generated by these muons to extract higher-level information, such as track length and endpoint position. A series of quality cuts (selection criteria) are then applied to filter out tracks with poorly reconstructed positions [58]. Hits between 100 cm and 200 cm (the track window) to the end of the track are selected. In this region, the cosmic muon energy loss per unit length is close to a constant.

With the selection strategy mentioned above, the following distributions are obtained:

- dE/dx distribution calculated using the Bethe-Bloch equation for simulation. See Figure 4.6.
- PECorr/dx distributions of simulation and data. See Figure 4.7.

The calorimetric energy scale is defined as the ratio of the means of simulated dE/dx to the means of simulated PECorr/dx. Assuming that the data and

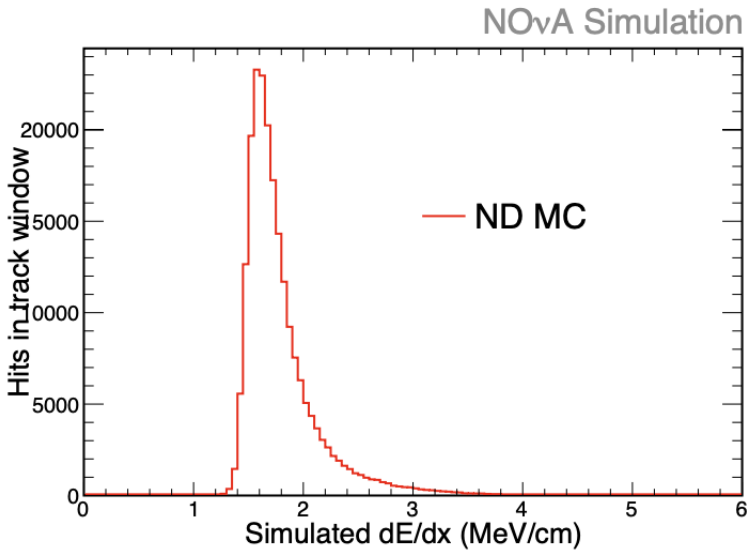


Figure 4.6: An example histogram of simulated dE/dx of ND MC in track window[58].

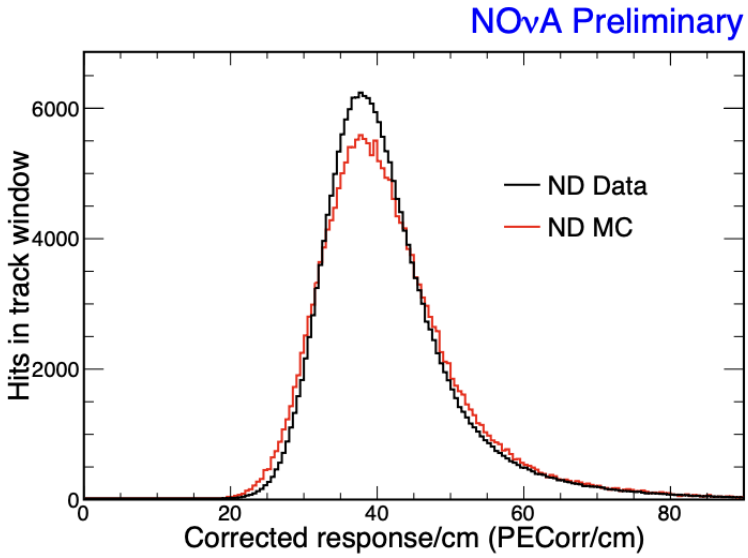


Figure 4.7: An example histogram of Near Detector PECorr/dx of data (black) and MC (red) in track window[58]. The two distributions closely match each other.

simulation share the same calorimetric energy scale, we have

$$(\text{calibrated dE/dx})_{\text{data}} = (\text{calorimetric energy scale}) \times (\text{PECorr/dx})_{\text{data}} \quad (4.4)$$

4.5 Timing Calibration

The NOνA experiment aims to detect neutrino interactions from the NuMI beam using both the Near and Far Detectors. The NuMI beam delivers a 10 μs spill every 1.3 seconds. The expected neutrino interaction rate in the Far Detector is only 1–2 events per day, whereas the background cosmic ray rate is approximately 1.2×10^5 Hz. This vast difference in rates makes an activity-based trigger impractical, as it cannot effectively distinguish neutrino signals from the overwhelming cosmic background. Instead, NOνA employs a fully software-based trigger along with a temporary buffer that stores data for several minutes before processing. When a NuMI spill is generated, both detectors are informed and data collected from the 500 μs window centered on the spill are read out. NOνA requires that all the readout channels are synchronized within 10 μs.

4.5.1 Timing System

A distributed system was built to synchronize all the channels in detectors [59]. Each detector has two redundant timing chains (one as the primary chain and the other as a hot spare) and each chain has one Master Timing Distribution Unit (MTDU) connected to a Global Positioning System (GPS) antenna. An MTDU accesses the receiver's 10 MHz reference clock and is linked to multiple Slave

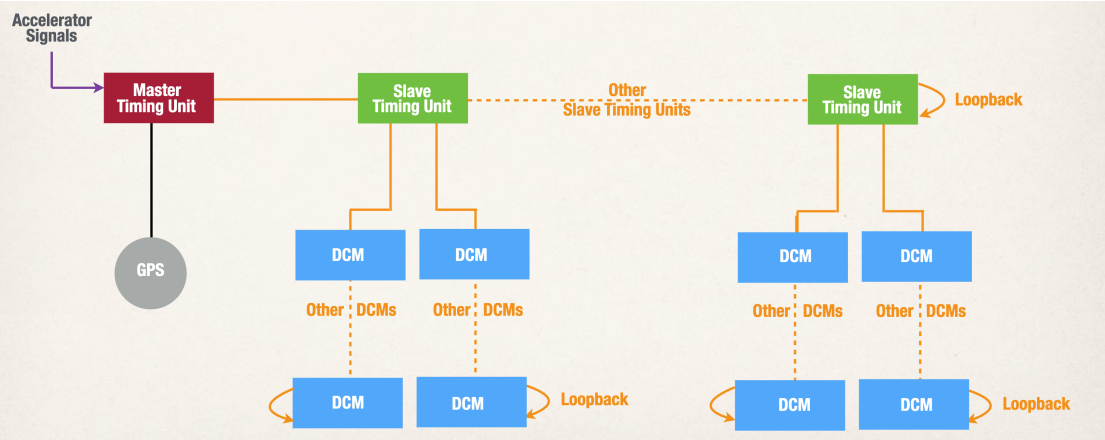


Figure 4.8: Schematic of a timing chain. All the STDUs in a chain form a loopback, and likewise, all the DCMs in a branch also establish a loopback.

Timing Distribution Units (STDUs). STDUs in the same chain form a loopback. Each STDU connects to several DCMs, with each DCM further connected to 64 FEBs. Each FEB serves 32 APDs. Finally, we want all DCMs and TDUs to be synchronized to the master clock.

To ensure that MTDUs are synchronized to accurate time, an independent system called the Timing Calibration Reference (TCR) was established. The TCR system ensures the accuracy of the Master Timing Units (MTDUs) by providing a reliable 1 Hz pulse via an independent GPS antenna and receiver. Each MTDU compares its clock phase with that of the TCR to detect potential inaccuracies caused by glitches or clock drift.

4.5.2 Online Timing Synchronization

To achieve high timing precision, the delays of signal propagation cannot be ignored. A calibration is performed to "learn" the delay of each device in the timing

chain. At the start of the calibration process, the MTDU broadcasts a special signal to all devices in the chain, including STDUs and DCMs. When the next synchronization signal is sent from the MTDU, all the STDUs start a counter, which stops upon receiving the return sync signal from the last STDU in the chain. Each STDU then calculates and stores half of the time-of-flight (TOF) value in its delay register, corresponding to the time it takes for a signal to travel from that STDU to the last STDU. Each STDU also maintains an independent counter for delay values along each DCM branch. Now each STDU knows how long to wait to start its own clock after the signal from MTDU arrives. NOνA imposes a requirement that all cables connecting a DCM to the FEBs must have the same length. A measured delay value between the DCM and FEB is configured in the hardware. The entire delay calibration process in the timing chain is performed periodically to monitor stability and the effects of fluctuations in ambient temperature.

4.5.3 Offline Timing Calibration

An additional calibration was implemented as a crosscheck of the online calibration. It utilizes muon events in the detectors. For the Far Detector, the abundant cosmic muon events are used, while for the Near Detector, muons generated by neutrino interactions with the surrounding rock are utilized. Strict quality cuts are applied to these events to select well-reconstructed muon tracks. The time differences between hits recorded by different DCMs along the same track are calculated. Corrections are applied by accounting for the time of flight (assuming

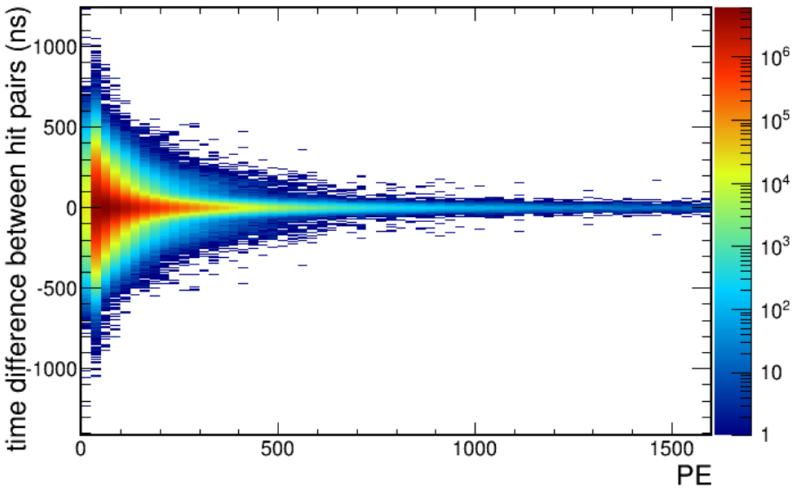


Figure 4.9: 2D histogram of time difference between cell hits in the same DCM vs PE [59].

muons travel at the speed of light) and the distance between the hits and the APD. If the online calibration functions properly, the relative time difference distribution between a pair of DCMs will align closely with the expected distribution determined by the detector operating conditions.

The timing resolution of a DCM can be calculated using the existing well-reconstructed tracks in data. A 2D histogram of the time difference between two cell hits in the same DCM versus the number of photoelectrons (PE) is made. The root-mean-square of time difference in each PE bin is the timing resolution (see Figure 4.9).

Chapter 5

Event Reconstruction

In the previous two chapters, we discussed how the detectors convert scintillator lights into electronic signals. The raw digits collected by the detector hardware are not easily interpretable for physics analysis. To facilitate a more effective understanding of particle interactions, we need to transform the raw data into more meaningful (user-friendly) physics quantities.

In NOνA, the reconstruction process plays a crucial role in converting raw detector data into higher-level information. Its primary goal is to extract key physical properties of the detected particles and address the following fundamental questions: 1. What is the identity of the particle that produced a given cluster of hits? 2. What is the energy of the particle? It is important to note that reconstruction does not provide the "true" information about the particles. Instead, it offers an estimate based on the available detector data and the applied algorithms. For example, the reconstructed energy of a particle may differ from the true energy, due to the limitations of performance of detectors, accuracy of

algorithms, statistical fluctuations, etc.

Reconstruction is a complex and large-scale project. NOνA employs the *ART* framework for reconstruction, a versatile ROOT-based framework for particle physics data processing and generation, designed to provide a unified platform for various modern neutrino experiments such as NOνA, DUNE and MicroBooNE. The entire reconstruction process is divided into multiple tasks, each implemented as a program named "module" in *ART*.

In this chapter, I will discuss NOνA's reconstruction chain, from raw hits to the fully reconstructed objects such as particle energy and particle identification (PID). By the end of this chapter, we will have a big picture of the entire reconstruction process.

5.1 Raw digits to cell hits

The CalHit module includes algorithms that convert raw digits from DAQ into cell hits, represented by a C++ class that encapsulates both calibrated information (e.g. Attenuation-corrected Photoelectrons) and limited raw data (e.g. cell number, ADC value and time in nanoseconds) associated with the hit. This module forms the basis of all the subsequent reconstruction tasks.

5.2 Slicer

A slicer is a module designed to identify hits resulting from a single physics interaction and group them into a "slice". In this context, a physics interaction

refers to either a neutrino interaction or a cosmic ray. For instance, a neutrino interacting with a nucleus in the detector may produce multiple charged particles, which are scattered outward. Each charged particle may generate a small group of hits (a sub-cluster). Together, all the hits from this interaction form a cluster.

During the 500 μs data collection window in the detectors, numerous hits occur due to the high rates of interaction. These hits can be further classified into two categories, signal and background. Ideally, we want to collect all the hits that belong to one interaction. However, true clusters often overlap in both space and time, necessitating the use of a dedicated algorithm for their separation.

Below, I will briefly introduce the two major slicing algorithms implemented in NOνA.

5.2.1 Slicer4D

The Slicer4D is a clustering algorithm based on Density Based Spatial Clustering of Applications with Noise (DBSCAN) [60]. DBSCAN requires two parameters to be set:

1. ϵ: the radius that defines the neighborhood of a point.

2. minPts: the minimum number of points required to form a dense region.

DBSCAN classifies points into three categories:

1. Core points: A point is considered a core point if there are at least minPts points (including itself) within a distance of ϵ.

2. Reachable points: A point A is directly reachable from a core point B if A is within a distance of ϵ from B. Furthermore, if A is reachable from B and B is reachable from C, then A is indirectly reachable from C.

3. Outliers (Noise points): Any point that is not reachable from any other points is classified as an outlier or noise point.

DBSCAN iterates over all points in the dataset. If a point is not a core point, it is temporarily marked as noise (though it may later be assigned to a cluster). If a point is identified as a core point, the algorithm begins constructing a cluster by adding the core point and its directly reachable points to the cluster. If any of these reachable points are also core points, their reachable points are recursively added to the cluster. This process continues until no more points can be added. Once a cluster is fully constructed, the algorithm selects a new unvisited point and repeats the process.

Slicer4D is a variant of DBSCAN. It utilizes a special function to calculate the "distance" between two points (hits), which incorporates a causality term, two penalty terms for large spatial separation, and an additional penalty for hits with low photoelectron (PE) counts. The distance function is defined as

$$\epsilon = \left(\frac{\Delta T - \Delta\vec{r}/c}{T_{\text{res}}}\right)^2 + \left(\frac{\Delta Z}{\text{D}_{\text{pen}}}\right)^2 + \left(\frac{\Delta XY}{\text{D}_{\text{pen}}}\right)^2 + \left(\frac{\text{PE}_{\text{pen}}}{\text{PE}}\right)^5. \tag{5.1}$$

T_{res} is the Euclidean norm[1] of the timing resolutions of these two hits. D_{pen} and PE_{pen} are distance penalty and PE penalty, respectively. ΔXY and ΔZ denote the projection distance between two hits on different axis. $\vec{r}$ is the position

[1]The Euclidean norm of two variables x and y is $\sqrt{x^2 + y^2}$

difference between two hits. Depending on whether the two hits are in the same view, the calculation differs. A detailed description of the calculation can be found in the original technical note of Slicer4D [61].

5.2.2 TDSlicer

The Time Density Slicer (TDSlicer) [62] serves as a replacement algorithm for Slicer4D in NOνA. It consists of multiple clustering steps. Unlike Slicer4D, which clusters hits from both views at the outset, TDSlicer first constructs clusters separately in each view and then merges them in a later stage. First, the following density function of hit i with respect to the surrounding hits is defined:

$$\rho_i = \sum_j e^{-d_{ij}^2}, \tag{5.2}$$

where $d_{ij} = \left| \frac{dt - \frac{\vec{r}}{c}}{\tau} \right|$, $\vec{r}$ and dt are the position difference and time difference between hits, respectively. τ is a tunable parameter named time scale. TDSlicer requires a function to calculate the "isolation score" of a hit:

$$\delta_i = \min_{j|\rho_j > \rho_i} (d_{ij}). \tag{5.3}$$

A threshold is applied to select hits with high density and isolation scores, which are designated as centroids. The threshold of isolation ensures that a centroid is sufficiently distant from other centroids. A preliminary time-based clustering is then performed, associating each hit with its nearest centroid. Subsequently, Prim's algorithm [63] is used to construct clusters in the XZ and YZ views. Finally,

the average z and t values of clusters from different views are compared to identify the best-matched cluster pairs and construct 3D clusters.

5.3 Hough Transform

NOνA uses a modified Hough transform to detect straight lines or "guidelines" in each view of a slice [64]. The Hesse normal form of lines in 2D space is used:

$$\rho = x\cos\theta + y\sin\theta, \tag{5.4}$$

where ρ represents the perpendicular distance from the origin to the nearest point on the line, and θ is the angle between the x-axis and the line segment connecting the origin to that nearest point.

The Hough accumulator matrix is a two-dimensional array where the two dimensions correspond to the discretized values of ρ and θ, respectively. Each element in the matrix represents a specific (ρ, θ) pair and stores the accumulated votes cast by hit pairs. The vote is calculated with the following expression:

$$v(\rho,\theta) = e^{-\frac{(\theta-\theta_0)^2}{2\sigma_\theta^2}} e^{-\frac{(\rho-\rho_0)^2}{2\sigma_\rho^2}}, \tag{5.5}$$

where θ_0 and ρ_0 are the parameters of the line that passes through the two hits, $\sigma_\rho = \frac{3}{\sqrt{12}}$ and $\sigma_\theta = \frac{3}{d\sqrt{6}}$. d denotes the distance between the two hits.

The algorithm iterates over all the selected hits pairs [2]. Once the accumulator

[2]There are some rules to select hit pairs. These rules help reduce the computation loads and prevent bias. For details of selection, see the technical note of modified Hough transform [64].

is constructed, a Gaussian blur is applied with a limited window size to smooth the accumulated votes. This step was proven to help improve the accuracy of algorithm.

It has been observed that picking all peaks above a threshold in the Hough accumulator results in an excessive number of lines in the final output. To address this, the modified Hough transform uses a multi-loop approach to identify legitimate lines. The highest peak in the accumulator is selected, and all points associated with this peak are removed. A new Hough accumulator is then created with the remaining points, and this process is repeated until no peaks remain above the threshold.

5.4 Vertex Reconstruction

The "elastic arm" algorithm (also known as "deformable templates") is one of the primary algorithms in NOνA used to reconstruct the neutrino interaction vertex and its associated arms.

An arm is defined as a finite-length directed ray originating from a vertex. Mathematically, a prong α can be expressed as:

$$x(s) = x_0 + s \sin\theta_\alpha \cos\phi_\alpha, \tag{5.6}$$

$$y(s) = y_0 + s \sin\theta_\alpha \sin\phi_\alpha, \tag{5.7}$$

$$z(s) = z_0 + s \cos\theta_\alpha \tag{5.8}$$

where θ_α and ϕ_α are angular coordinates that define the arm's direction, and s is

the distance from vertex (x_0, y_0, z_0). The number of arms, M, within a given slice is determined by the maximum number of lines identified by the Hough transform in either the XZ or YZ view. The following cost function is minimized to find the optimal vertex and arms [65]:

$$E = \sum_{i=1}^{N}\sum_{\alpha=1}^{M} V_{i\alpha} M_{i\alpha} + \lambda \sum_{i=1}^{N}\left(\sum_{\alpha=1}^{M} V_{i\alpha} - 1\right)^2 + \frac{2}{\lambda_v}\sum_{\alpha=1}^{M} D_\alpha \tag{5.9}$$

where $M_{i\alpha}$ is a measure of the distance between the hits and the arms, $V_{i\alpha}$ is the likelihood that a hit i is associated with an arm α, and D_α is the distance between the vertex and the first hit in the arm α. λ and λ_v are the coefficients that are manually set to control the strength of each term in Eq Equation (5.9). The first term measures the goodness of fit between the hits and the arms. The second term is a penalty applied to hits that do not belong to any arms. The final term imposes an additional penalty on arms whose first hits are significantly distant from the vertex.

$$M_{i\alpha} = \left(\frac{d_{i\alpha}}{\sigma_i}\right)^2, \tag{5.10}$$

where $d_{i\alpha}$ is the distance between hit i and arm α and σ_i is the spatial resolution of the detector. If hit i lies in the backward direction relative to arm α, we raise the exponent to 4.

$$V_{i\alpha} = \frac{e^{-\beta M_{i\alpha}}}{e^{-\beta\lambda} + \sum_{b=1}^{M} e^{-\beta M_{i\alpha}}}. \tag{5.11}$$

We assume that the probability of a hit being noise is a constant $e^{-\beta M_{i\alpha}}$, while the probability of a hit being associated with an arm is proportional to $e^{-\beta M_{i\alpha}}$. β

is a hyper-parameter that needs to be set manually.

5.5 Prong Reconstruction

The Elastic Arms algorithm has demonstrated high accuracy in predicting interaction vertices. However, it falls short in accurately determining the number of arms and establishing precise associations between hits and arms. A more robust algorithm is needed to effectively cluster hits into distinct prongs. A prong is defined as a cluster of hits corresponding to a single final-state particle in the interaction. NOνA developed a package named *FuzzyKVertex* [66] based on the Fuzzy K-Means Clustering (FKM) algorithm. The cost function of FKM is

$$J_{FKM}(\mu, a) = \sum_{i=1}^{c} \sum_{j=1}^{n} \mu_{ij}^{m} ||x_j - a_i||^2, \tag{5.12}$$

where m is a number larger than 1 and it represents the "degree of fuzziness", x_j is the position of the hit j, and a_i is the center of prong i. The algorithm will keep updating μ_{ij} and a_i with the following equations until J_{FKM} is minimized:

$$\mu_{ij} = \left(\sum_{k=1}^{c} \frac{||x_j - a_i||^{\frac{2}{m-1}}}{||x_j - a_k||^{\frac{2}{m-1}}} \right)^{-1}, \tag{5.13}$$

$$a_i = \frac{\sum_{j=1}^{n} \mu_{ij} x_j}{\sum_{j=1}^{n} \mu_{ij}}. \tag{5.14}$$

The denominator normalizes μ_{ij} so that

$$\sum_{i=1}^{c} \mu_{ij} = 1 \tag{5.15}$$

for hit j. Unfortunately, FKM has two issues that need to be addressed. First, the number of prongs in each slice is unknown. Second, some noise hits are incorrectly assigned to multiple prongs with equal likelihood. NOνA employs the Possibilistic Clustering Algorithm (PCA) [67] to address this issue. PCA defines μ_{ij} without requiring normalization, as follows:

$$d_{ij} = \left(\frac{x_j - a_i}{\sigma_j} \right)^2, \tag{5.16}$$

$$\mu_{ij} = e^{-\frac{m\sqrt{a}d_{ij}}{\beta}}, \tag{5.17}$$

where a denotes the number of prong centers, β is a hyperparameter, and σ_j represents the spatial resolution of the detector. The algorithm begins with a small c and gradually increases it until all hits are assigned to a prong or the maximum allowable number of prongs is reached.

Note that the algorithm operates on each view independently. To reconstruct 3D prongs, prongs from different views are matched based on their energy deposition and positions. The merged prongs are referred to as 3D prongs, while those that remain unmatched are termed 2D prongs.

5.6 Break Point Fitter

The Break Point Fitter [68, 69] is an algorithm developed to track particles in a detecter assuming that a charged particle experiences multiple Coulomb scattering at fixed intervals. The energy of a particle is estimated by walking back along the fitted straight line of a prong, cumulatively adding the expected energy loss to approximate the particle's energy at the beginning of the prong. The expected Coulomb scattering angle is given by the following expression:

$$\theta_{rms} = \frac{13.6 \text{ MeV}}{\sqrt{3}\beta p}\sqrt{s}(1 + 0.038\ln(s)), \tag{5.18}$$

where β is the ratio of v to c, p is the momentum of the particle, s is the distance that the particle travels. We begin at the starting point of the prong and insert a scattering plane if the expected scattering angle exceeds a predefined tolerance value. This procedure is repeated until the end of the prong is reached. The track of particle is assumed to follow the function:

$$\zeta_i = a + bz_i + \sum_{J=1}^{M} \alpha_J(z_i - Z_J)\Theta(z_i - Z_J), \tag{5.19}$$

where Z_J is the position of J'th scattering, a is the vertex of the prong, b is the initial track slope with respect to the z-axis, α_J is the J'th scattering angle, and Θ is the Heaviside step function[3]. The coordinate system is constructed with the z-axis aligned with the direction of the fitted prong line, while the x- and

[3] $\Theta(x) = \begin{cases} 1 & \text{if } x \geq 0, \\ 0 & \text{if } x < 0. \end{cases}$

y-axes are as parallel as possible to the original x- and y-axes in the detector. The optimization of parameters in Eq. Equation (5.19) is performed by minimizing a χ^2 of all the hits and scattering angles

$$\chi^2 = \sum_{i=1}^{n} \frac{(\zeta_i - \zeta_i')^2}{\sigma_i^2} + \sum_{j=1}^{N} \frac{\alpha_j^2}{\sigma_{\alpha j}^2}, \tag{5.20}$$

where ζ_i' is the measured position of hit i, ζ_i is the predicted position of hit i by Eq. Equation (5.19), σ_i is the uncertainty of position measurement of hit i, and $\sigma_{\alpha j}$ is the RMS of scattering angle by Eq. Equation (5.18).

Track fitting is performed under three different particle identity assumptions (muon, proton, and pion), resulting in three separate track fitting outcomes.

5.7 Kalman Track

Kalman Track [69] is another algorithm used to reconstruct the tracks of particles. It is based on Kalman filtering [70], a well-known algorithm that estimates the state of a dynamic system from noisy measurements. The filter predicts the system's next state using a linear model, then updates this prediction with new observations, accounting for uncertainties. Kalman Track begins with a seeding step, where a pair of hits within a specified distance in the xz or yz view are selected. A track is then reconstructed from these two hits, and the Kalman filter is used to predict the hit positions in subsequent cells. If a hit is sufficiently close to the predicted position, it is added to the track, and the track parameters are updated. This process continues until no more hits can be added. The track

is extended from the downstream to the upstream of the detector. Additionally, another track reconstruction is performed in the opposite direction to identify any potentially missed hits.

The 2D tracks in different views are matched based on the following metric:

$$S = \frac{\text{Start}_{\text{diff}} + \text{Stop}_{\text{diff}}}{\text{length of overlap in z-direction}}, \tag{5.21}$$

where $\text{Start}_{\text{diff}}$ and $\text{Stop}_{\text{diff}}$ represent the differences in the start and stop positions along the z-direction, respectively.

5.8 Event Classification

NOνA employs deep neural networks for event classification. Specifically, a computer vision model known as the Convolutional Visual Network (CVN)[4] was designed and trained for this purpose.

The CVN takes as input two pixel-mapped views of detector hits, corresponding to different projections of the event. Each view is processed independently through multiple convolutional layers, which extract spatial and feature information. The outputs from these convolutional stacks are then combined and passed through a series of fully connected layers, which further refine the classification decision.

[4]Also referred to as Convolutional Neural Networks (CNNs).

The final layer of the network produces a set of scores between 0 and 1, representing the probability of the event belonging to each class. These classes include ν_μ charged-current (CC), ν_e CC, ν_τ CC, neutral-current (NC) interactions, and cosmic-ray-induced backgrounds, among others. The model is trained using supervised learning, with labeled event samples generated from Monte Carlo simulations.

By leveraging deep learning, CVN provides high-accuracy classification, improving NOνA's ability to distinguish between different neutrino interactions and background events.

5.9 Particle Classification

Another CVN model, ProngCVN, was developed to classify individual particles (prongs) within an event. Similar to the event-level CVN, ProngCVN takes as input both event-level and prong-level pixel maps. The prong-level hits are identified using the FuzzyK algorithm, as described earlier.

The network processes these inputs through convolutional layers and outputs five scores, each representing the likelihood that the prong originates from one of the five particle types: electron, gamma, muon, pion, or proton. This approach enhances particle identification by leveraging both global event context and localized prong features.

5.10 Energy Reconstruction

The NOνA experiment, designed to study neutrino oscillations, relies on precise energy reconstruction to accurately determine neutrino properties. An energy estimator is an algorithm that estimates the energy of particles based on the input features of detected events. Various energy estimators have been developed for different types of neutrino interactions, including ν_μ CC, ν_e CC, and NC interactions. In this section, we will focus on the traditional energy estimators used in NOνA. These methods have been widely adopted in numerous analyses within the experiment and have been extensively validated. Machine learning-based energy estimators, which offer potential improvements, will be discussed in later chapters.

In energy reconstruction, several key questions are of interest, including but not limited to:

- Accuracy of the reconstructed energy – This is typically described by the energy resolution, which quantifies the relative error between the reconstructed and true energy. A better resolution indicates a more accurate energy estimation.

- Robustness of the energy estimator – This refers to how stable the reconstructed energy remains under systematic uncertainties. A robust estimator should exhibit minimal variation when systematic shifts are introduced.

Given the reconstructed energy E_{reco} and true energy E_{true}, the accuracy of

an energy estimator is described by the quantity known as energy resolution:

$$Resolution = \frac{E_{reco} - E_{true}}{E_{true}}. \tag{5.22}$$

The resolution is calculated event-by-event and we focus on the distribution of the energy resolution within each energy range, specifically its mean and root mean square (RMS). The resolution of an energy estimator is evaluated in Monte Carlo samples where the E_{true} values are available.

As we discussed previously, neutrinos interact with matter through two different interaction vertices[5]: charged current (CC) and neutral current (NC). In the final state, both a lepton and a hadronic shower are produced. The hadronic showers arise from the decay of the scattered nucleus.

5.10.1 ν_μ CC Spline Energy Estimator

The default energy estimator for ν_μ CC events is a spline-based estimator. In a ν_μ CC event, the outgoing lepton is a muon or anti-muon, which produces a characteristic long, clean track. The ν_μ CC spline energy estimator [71, 72, 73] consists of two separate energy estimators: one for the muon energy and one for the hadronic energy.

The reconstructed energy of the ν_μ event is the sum of the muon energy (E_μ) and the hadronic energy (E_{had}):

$$E_{\nu_\mu} = E_\mu + E_{\text{had}}. \tag{5.23}$$

[5]We do not consider gravity, as its effect is too weak.

Muon energy

The muon energy is estimated based on the length of 3D Kalman muon tracks determined by ReMId. ReMId (Reconstructed Muon Id) [74] is an algorithm that evaluates the log-likelihood of a track being a muon track. The track with the highest ReMId value in an event is selected as the muon track. To train the spline model, a 2D histogram of true muon energy versus Kalman Track length is made with Monte Carlo sample. The histogram is then divided into a number of bins[6] with respect to the track length. Inside each bin, a Gaussian fit is performed to get the Gaussian mean of true energy. A piece-wise linear function is fit to the Gaussian means by minimizing the energy resolution (see Figure 5.1). The resulting optimized spline function is the muon energy estimator.

Hadronic Energy

The hadronic energy estimator is constructed in a similar way to the muon energy estimator. The true hadronic energy is the difference between E_{ν_μ} and E_μ. Instead of the Kalman track length, the calorimetric energy (also known as the "visible" energy) of the shower is used to estimate the energy (see Figure 5.2). The visible energy is calculated by summing the calorimetric energy of hits that are not associated with the muon track.

[6] The bin width is much smaller, so that each linear segment of the spline function corresponds to many bins.

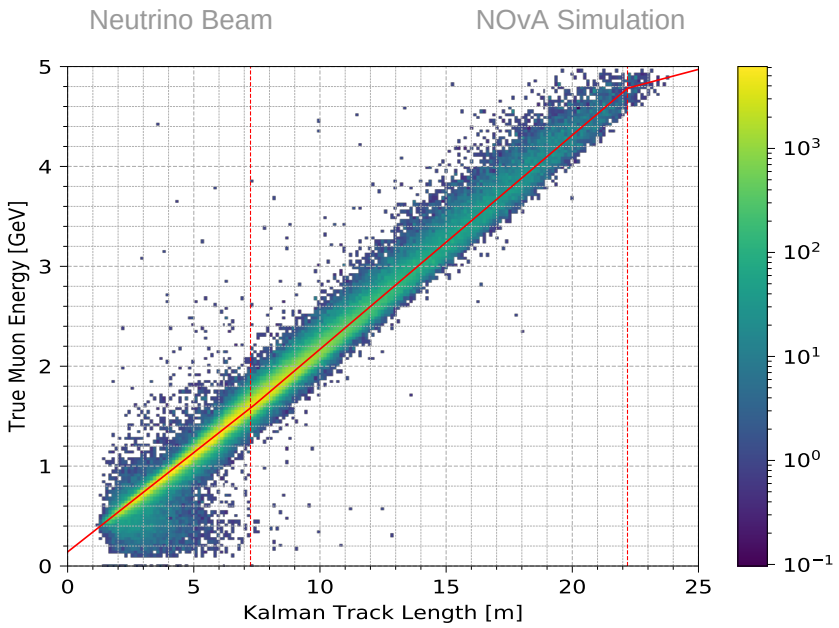


Figure 5.1: The 2D histogram of true muon energy versus Kalman Track length for FHC Monte Carlo sample. The solid red line represents the best-fit spline, while the dashed red lines represent the boundaries of the linear segments of the spline function.

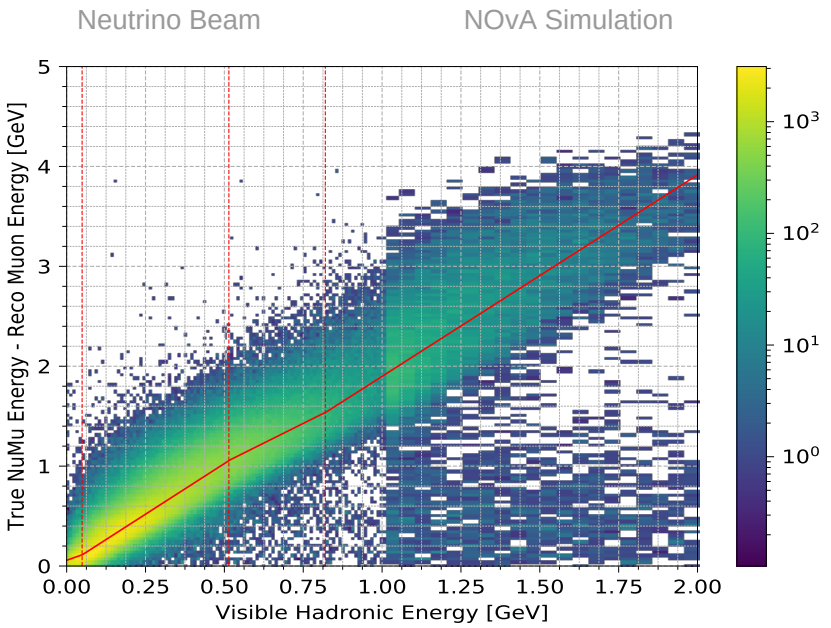


Figure 5.2: The 2D histogram of True Hadronic Energy vs Visible Energy for FHC Monte Carlo sample. The solid red line represents the best-fit spline, while the dashed red lines represent the boundaries of the linear segments of the spline function.

5.10.2 ν_e CC Energy Reconstruction

Unlike muons produced in ν_μ CC events, the outgoing electrons in ν_e CC events do not produce a straight and clean track in the detector. Instead, these electrons generate electromagnetic showers which spread out widely in the detector, making it difficult to estimate the energy using simple dimensional information.

The ν_e energy estimator was designed based on the observation that electromagnetic and hadronic depositions exhibit different detector responses [75]. The ProngCVN scores are used to classify a 3D Fuzzy-K prong as either an electromagnetic or hadronic deposition. Two new scores are defined as follows:

$$S_{\text{EM}} = S_e + S_\gamma + S_{\pi^0}, \tag{5.24}$$

$$S_{\text{Had}} = S_p + S_n + S_\pi + S_\mu + S_{\text{other}}, \tag{5.25}$$

where S_e is the electron ProngCVN score, with other symbols defined similarly. The electromagnetic (EM) shower energy of an event is defined as the total calorimetric energy of all prongs for which $S_{\text{EM}} > S_{\text{Had}}$. The hadronic energy is given by the total calorimetric energy of the slice minus the EM energy.

A polynomial function is used to fit the energy of ν_e:

$$E_{\nu_e} = a\ E_{\text{EM}} + b\ E_{\text{Had}} + c\ E_{\text{EM}}^2 + d\ E_{\text{Had}}^2. \tag{5.26}$$

The energy spectrum of E_{ν_e} is not flat. If the function is directly fitted to E_{ν_e}, the predicted energy tends to be biased toward the peak of the spectrum. To address this issue, each event is assigned a weight based on its E_{ν_e} value. Events in energy

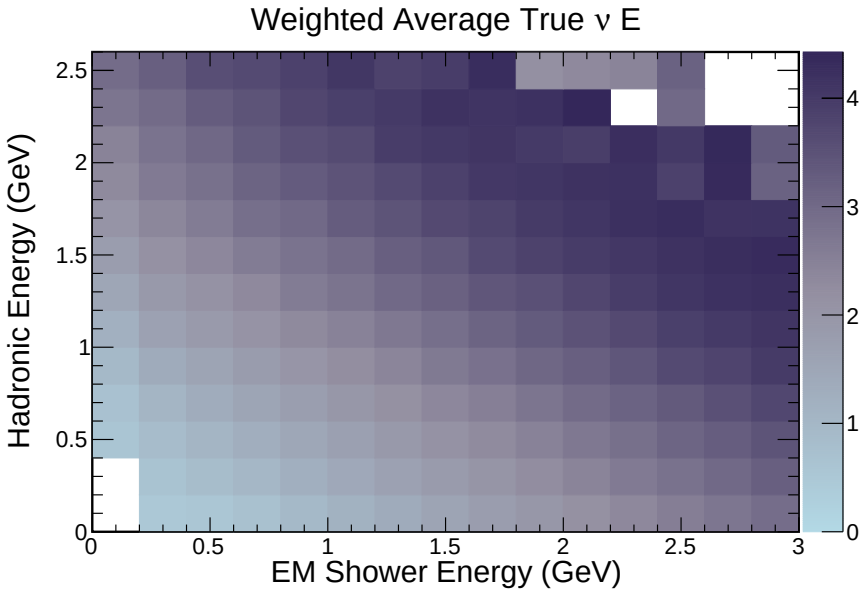


Figure 5.3: Reweighted average True E_{ν_e} vs Electron Shower Energy and Hadronic Energy.

ranges with a high event density are given smaller weights, while those in sparsely populated energy ranges receive larger weights. This weighting strategy helps balance the contribution of different energy regions in the fit. A 2D histogram (see Figure 5.3) was created with EM shower energy and hadronic energy as the axes, and the reweighted average true E_{ν_e} as the z-value in each bin. The function Equation (5.26) is fitted to the 2D histogram using ROOT to get the optimized 4 parameters. The resulting energy estimator achieves a mean resolution of 5.7%. To correct this bias, the predicted E_{ν_e} is multiplied by a scaling factor [75] of $\frac{1}{1+0.057}$ to ensure that the mean resolution is zero.

How scaling works

Assuming that the mean resolution of an energy estimator has

$$< \frac{E_{\text{reco}} - E_{\text{true}}}{E_{\text{true}}} >= \delta \Leftrightarrow < \frac{E_{\text{reco}}}{E_{\text{true}}} >= \delta + 1, \tag{5.27}$$

we want to make a corrected estimated energy that satisfy

$$< \frac{E'_{\text{reco}} - E_{\text{true}}}{E_{\text{true}}} >= 0 \Leftrightarrow < \frac{E'_{\text{reco}}}{E_{\text{true}}} >= 1. \tag{5.28}$$

It is evident that

$$E'_{\text{reco}} = \frac{E_{\text{reco}}}{1 + \delta} \tag{5.29}$$

is the sufficient condition of Eq Equation (5.28).

5.10.3 NC Energy Reconstruction

In a neutral current interaction, the energy of the incoming neutrino E_ν is shared between the outgoing neutrino and the hadronic shower. Since the energy of the outgoing neutrino is immeasurable, estimating E_ν directly is challenging. Instead, the current standard NC energy estimator used in sterile neutrino analysis focuses on estimating the energy of the hadronic shower (E_{dep}).

Various NC energy estimators were proposed [76] in the early stages of the analysis. Before discussing the standard energy estimator, it is necessary to review these candidate estimators to better understand the potential challenges in NC event energy reconstruction and the approaches used to address them. Note that

we construct energy estimators for FD and ND separately.

Calorimetric Estimator

Using calorimetric energy (E_{cal}) is a straightforward method to estimate the energy of the hadronic shower. However, neutrons in the hadronic shower do not directly ionize the scintillator; they only deposit energy when they occasionally collide with a nucleus in hadronic activity. Therefore, the calorimetric energy tends to be lower than the true hadronic shower energy.

Linear Estimator

To correct the bias of the calorimetric estimator, a simple linear fit was performed. The new energy estimator was constructed as:

$$E_{\text{dep(est)}} = a + b\ E_{\text{cal}}. \tag{5.30}$$

It was observed that the linear energy estimator responds differently to events dominated by electromagnetic activity compared to those dominated by hadronic activity on average, as part of the energy can be carried away by neutrons in hadronic interactions.

Orphaned Energy Estimator

The orphaned energy (E_{Orph}) is defined as the total calorimetric energy of all hits that are not reconstructed into any prongs. Hits produced by neutrons are more difficult to reconstruct into prongs than those produced by charged particles. The

orphaned energy estimator use the same equation (Equation (5.30)) as the linear estimator while it performs fit for events with $E_{\text{Orph}} > 0.01$ GeV and events with $E_{\text{Orph}} < 0.01$ GeV separately. The orphaned energy estimator achieves better resolution than the linear estimator, but the resolution is not symmetric around zero. This is due to the non-linear relationship between the average hadronic shower energy and calorimetric energy in ND, and the presence of electromagnetic activity in events with neutrons in the final state.

Quadratic Estimator

The quadratic energy estimator is similar to the ν_e CC energy estimator. It classifies prongs into hadronic prongs and electro-magnetic prongs and calculate E_{EM} and E_{Had} respectively. A cross term is added to the function:

$$E_{\text{dep(est)}} = \alpha\ E_{\text{EM}} + \beta\ E_{\text{had}}^2 + \gamma\ E_{\text{EM}}^2 + \delta\ E_{\text{had}}^2 + \epsilon\ E_{\text{EM}} \times E_{\text{had}}. \tag{5.31}$$

The estimator performs well in the E_{cal} range with sufficient events, but its performance deteriorates at higher E_{cal} where there are less events.

Scaling Estimator

The scaling energy estimator computes the estimated energy as the sum of the scaled E_{EM} and E_{Had}:

$$E_{\text{dep(est)}} = \frac{E_{\text{EM}}}{1+\alpha} + \frac{E_{\text{Had}}}{1+\beta}, \tag{5.32}$$

where α and β represent the biases in $\frac{E_{\text{cal}}-E_{\text{dep(true)}}}{E_{\text{dep(true)}}}$ for event samples dominated by electromagnetic and hadronic activity, respectively. These biases are determined by performing a Gaussian fit to the distribution of $\frac{E_{\text{cal}}-E_{\text{dep(true)}}}{E_{\text{dep(true)}}}$ in different event samples, where the bias is taken as the mean value of the fitted Gaussian distribution. However, the bias of E_{dep} still exist even if we provide correction to E_{EM} and E_{Had}.

Bias Corrected Estimators

To mitigate the bias in the above energy estimators (see Figure 5.4), we want to know how much the bias is in each E_{cal} range. An empirical function is introduced to model the bias as a function of E_{cal}:

$$f(E_{\text{cal}}) = P \times (E_{\text{cal}} + Q)^2 \times e^{-R \times E_{\text{cal}}} + S. \tag{5.33}$$

First, the events are divided into several bins with respect to E_{cal}. Within each bin, a Gaussian fit is performed around the peak of the $\frac{E_{\text{dep(est)}}-E_{\text{dep(true)}}}{E_{\text{dep(true)}}}$ distribution to determine the bias. The obtained bias as a function of E_{cal} is then fitted using Equation (5.33). The corrected energy estimator is then given by

$$E'_{\text{dep(est)}} = \frac{E_{\text{dep(est)}}}{1 + f(E_{\text{cal}})}. \tag{5.34}$$

This procedure is applied to all candidate energy estimators (see Figure 5.5) to obtain their bias-corrected versions. Among these, the bias-corrected scaling estimator demonstrates the least bias and was consequently chosen as the standard

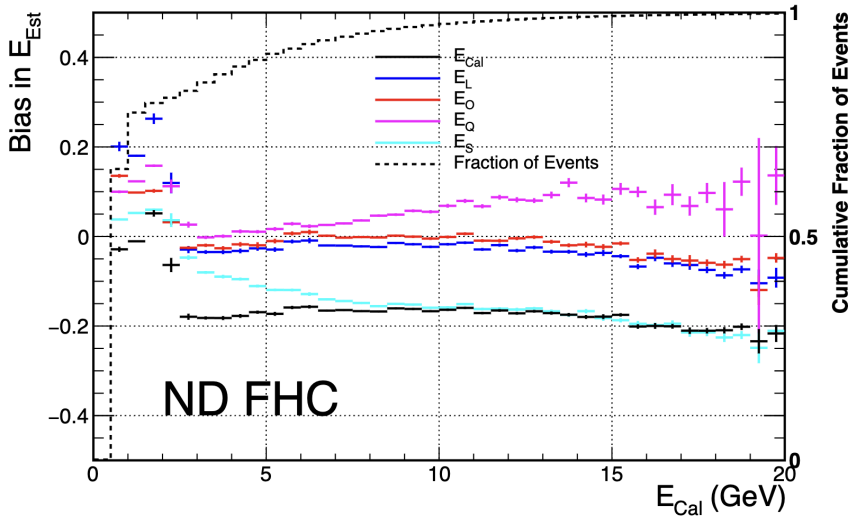


Figure 5.4: Biases as a function of E_{cal} for the calorimetric, linear, orphaned energy, quadratic, and scaling estimators for ND [76]. The dashed black line is the cumulative fraction of events below E_{cal}.

energy estimator for the sterile neutrino analysis in 2022. This energy estimator is called Nus20 energy estimator.

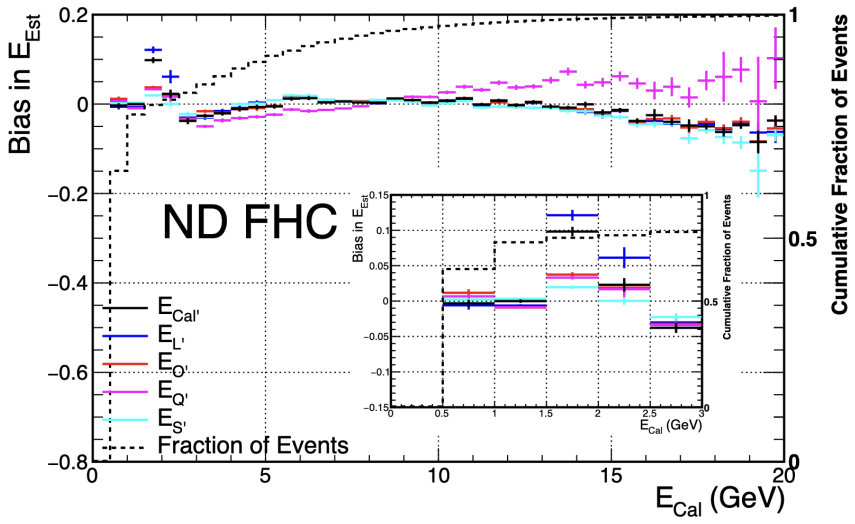


Figure 5.5: Corrected biases as a function of E_{cal} for the calorimetric, linear, orphaned energy, quadratic, and scaling estimators for ND [76]. The dashed black line is the cumulative fraction of events below E_{cal}. The small plot is a zoomed-in view for $E_{\text{cal}} < 3$ GeV.

Chapter 6

Sterile Analysis

In this chapter, I will outline the general process of the sterile neutrino analysis conducted in 2022. This includes event selection and binning, the implementation of systematic uncertainties, and the parameter fitting methodology.

The 2022 sterile neutrino analysis utilizes both ν_μ charged-current (CC) and neutral-current (NC) samples, with the ν_μ CC sample playing a crucial role in constraining atmospheric parameters.

6.1 Event Selection

Event selection is a crucial step in the analysis. Our goal is to identify high-quality events of interest (the signal) while effectively rejecting both poorly reconstructed events and background contamination. A common approach to event selection is the use of "cuts", which are specific criteria applied to the data to filter out unwanted events. These cuts can be based on various reconstructed quantities,

such as energy, momentum, particle identification, or low-level information. By carefully tuning these selection criteria, we aim to maximize signal purity while maintaining a sufficient event sample for statistical significance.

6.1.1 ν_μ CC Event Selection

The 2022 sterile neutrino analysis employs the same ν_μ CC selection cuts as the 2020 three-flavor analysis [77, 78]. Separate sets of cuts are implemented for the Far Detector (FD) and Near Detector (ND), with each set further divided into sub-categories listed below.

Quality Cuts

The quality cuts aim to remove events with low-quality reconstruction by enforcing a few minimal requirements. These requirements include: the number of hits in a slice must be at least 20, the reconstructed energy must be below 5 GeV, the slice must extend across more than 4 continuous planes (to remove slices that are nearly vertical to the beam direction), at least one cosmic track [1] is reconstructed, and at least one 3D Karman track must be successfully reconstructed with a valid RemID.

Containment Cuts

To achieve accurate energy reconstruction, it is essential that the final state particles remain fully contained within the detector. The energy carried by the escaping

[1] Cosmic tracks are the tracks reconstructed by an algorithm called "Cosmic Tracker". This algorithm processes all slices, not just those identified as cosmic slices.

particles will introduce a bias in the energy reconstruction. Additionally, to improve the quality of the analysis, it is important to reject background events such as cosmic rays in the Far Detector (FD) and rock muons—muons produced from interactions between neutrinos and the rock surrounding the Near Detector (ND). The containment cuts utilize reconstructed dimensional and positional information from showers, tracks, and slices to remove events that are too close to the edges of the detector. For example, we reject events with the slice starting within the first two planes. The parameters used in these cuts are carefully tuned through extensive studies, taking into account factors such as the signal-to-background ratio, the energy resolution of estimators after applying the cuts, and the fraction of energy lost due to escaping particles. Note that since the two detectors have different geometric properties, the containment cuts are implemented and optimized separately for each detector.

PID Cut

A Boosted Decision Tree (BDT) [79] was trained to classify events as either cosmic rays or ν_μ charged-current (CC) interactions. The BDT takes as input various reconstructed event features, including the direction, position, and length of the muon track, the fraction of muon hits relative to the total hits, and the event classification score from the Convolutional Visual Network (CVN). The output of the BDT is a score that quantifies the likelihood of an event being ν_μ CC. A Cosmic Rejection cut is implemented by defining a threshold on this score—events with scores above the threshold are classified as ν_μ CC.

The PID cut relies on three separate PIDs: Cosmic Rejection score, ReMId

and Event CVN muon score. The optimal threshold values for these cuts are determined through a grid search, maximizing the Figure of Merit (FoM):

$$FOM = \frac{S}{\sqrt{S+B}}, \tag{6.1}$$

where S is the signal event number and B is the background event number.

6.1.2 NC Event Selection

Similar to the ν_μ CC selection, the NC selection cuts can be broadly categorized into three types: (1) quality cuts, (2) containment cuts, and (3) signal selection cuts. Aspects already discussed in the ν_μ CC selection will not be repeated here.

Quality Cuts

To pass the quality cuts, each event is required to have a reconstructed vertex and have at least a reconstructed prong. In addition, the slice must extend across more than 2 continuous planes.

Containment Cuts

The NC containment cuts use positions of reconstructed vertex and distance from reconstructed prongs to the edges of detector to ensure that the interaction occurs well within the fiducial volume. For far detector, the vertex position variables are instead used by cosmic rejection BDT cuts.

Signal Selection Cuts

To distinguish signal from cosmic backgrounds in the Far Detector, a cosmic rejection BDT was trained with 20 input variables. These variables include, but are not limited to, the vertex position, the number of hits per slice, and the number of reconstructed showers.

A cut on the CVN score was applied to distinguish signal events from background. The threshold values of BDT score, FD CVN score and ND CVN score are tuned together using a special Figure of Merit defined below:

$$FOM = \frac{1}{N}\sum_{i}^{N}\frac{S_i}{\sqrt{S_i + B_i + \sigma_i^2}}, \tag{6.2}$$

where S_i and B_i denote the signal and background event counts in i'th bin of the spectrum, respectively, and σ_i represents the total systematics in that bin. The thresholds used in the selection cuts were determined by maximizing the FOM, resulting in a ND CVN score of 0.98, a FD CVN score of 0.1, and a BDT score of 0.85.

In addition to the three types of selections described above, an energy cut was applied, requiring events to have a reconstructed energy (as estimated by the standard energy estimator) between 0.5 GeV and 20 GeV.

6.1.3 Data Selection Summary

The data used in 2022 sterile analysis are summarized in Table 6.1.

Besides the experimental data, an idealized dataset called Asimov fake data is constructed using high-statistics Monte Carlo oscillation predictions scaled down

Sample	selected candidate event number	POT
ND ν_μ CC	2,826,066	11.0×10^{20}
FD ν_μ CC	209	13.6×10^{20}
ND NC	103,109	10.9×10^{20}
FD NC	469	13.6×10^{20}

Table 6.1: Summary of selected data.

to match the experimental data. This dataset has ignorable statistical fluctuations and no systematic shifts. The prediction is performed with standard 3-flavor oscillation parameters, meaning that Asimov fake data assumes no sterile neutrino oscillations.

6.2 Binning

6.2.1 ν_μ CC binning

There are 19 bins in the ν_μ CC spectra. The bin width is set to 0.1 GeV between 1 and 2 GeV, taking into account the relatively high statistics in this energy region, the performance of the energy estimator with a resolution of approximately 10%, and the fact that this is the energy range where the majority of 3-flavor neutrino oscillations occur in NOνA. Outside this region, wider bins are used to account for lower event statistics and reduce computational time. The bin edges are defined as follows [77, 80]: 0, 0.75, 1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.25, 2.5, 2.75, 3, 3.5, 4, and 5 GeV.

Resolution	Bin edges (unit: GeV)
30%	0.5, 0.65, 0.85, 1.10, 1.43, 1.86, 2.41, 3.14, 4.08, 5.30, 6.89, 8.96, 11.65, 15.14, 20
20%	0.5, 0.60, 0.72, 0.86, 1.04, 1.24, 1.49, 1.79, 2.15, 2.58, 3.10, 3.72, 4.46, 5.35, 6.42, 7.70, 9.24, 11.09, 13.31, 15.97, 20
10%	0.5, 0.55, 0.61, 0.67, 0.73, 0.81, 0.89, 0.97, 1.07, 1.18, 1.30, 1.43, 1.57, 1.73, 1.90, 2.09, 2.30, 2.53, 2.78, 3.06, 3.36, 3.70, 4.07, 4.48, 4.92, 5.42, 5.96, 6.55, 7.21, 7.93, 8.72, 9.60, 10.56, 11.61, 12.77, 14.05, 15.46, 17.00, 20
5%	0.5, 0.53, 0.55, 0.58, 0.61, 0.64, 0.67, 0.70, 0.74, 0.78, 0.81, 0.86, 0.90, 0.94, 0.99, 1.04, 1.09, 1.15, 1.20, 1.26, 1.33, 1.39, 1.46, 1.54, 1.61, 1.69, 1.78, 1.87, 1.96, 2.06, 2.16, 2.27, 2.38, 2.50, 2.63, 2.76, 2.90, 3.04, 3.19, 3.35, 3.52, 3.70, 3.88, 4.07, 4.28, 4.49, 4.72, 4.95, 5.20, 5.46, 5.73, 6.02, 6.32, 6.64, 6.97, 7.32, 7.68, 8.07, 8.47, 8.89, 9.34, 9.81, 10.30, 10.81, 11.35, 11.92, 12.52, 13.14, 13.80, 14.49, 15.21, 15.97, 16.77, 17.61, 18.49, 20

Table 6.2: NC bin edges for different resolutions

6.2.2 NC binning

The NC samples use a resolution-based binning scheme to divide the energy range from 0.5 to 20 GeV into bins, where the bin width is given by the following equation:

$$(\text{bin width}) = \text{resolution} \times (\text{bin's low edge}). \tag{6.3}$$

The "resolution" is a manually set parameter used to control the bin width. In the 2022 analysis, various binning resolutions ranging from 5% to 30% were implemented. However, smaller binning resolutions result in significantly more computational time. Therefore, the 30% resolution was chosen as the most feasible option. See Table 6.2 for a complete list of the resolutions used and their corresponding bins.

6.3 Systematics

A variety of systematics are considered in the sterile analysis, which can be grouped into seven categories: flux, detector, normalization, muon, cross section, MEC, and miscellaneous (uncertainty of neutrino flux from kaon decay, neutron uncertainty, tau cross section uncertainty). These systematics are represented by individual covariance matrices, which capture both the uncertainty in each bin of the spectra and the correlations of uncertainty across different bins and detectors.

6.3.1 Flux Systematics

Flux-related uncertainties come from mis-modeling of hadron production in the graphite target and subsequent beam transport with the latter being affected by factors such as target position, horn current, and magnetic fields in the decay pipe. These uncertainties are estimated using the PPFX package, which employs a multi-universe approach to generate 100 covariance matrices from 100 simulated "universes" with varied proton–target interaction cross-sections and beam transport parameters [81]. These matrices are averaged to form a single matrix. A Principal Component Analysis (PCA) is then applied to reduce this matrix to a small set of principle components, with the five most significant components used in the oscillation fit.

6.3.2 Detector Systematics

The detector systematics account for the uncertainties of calibration procedure and detector response. The calibration energy scale uncertainty is determined

from the energy scale discrepancy between data and Monte Carlo samples; The calibration shape uncertainty covers the differences between reconstructed energy and true energy for the events close to the edges of detector; The drift systematics models the aging effects in the detectors; The light level systematics takes into account the uncertainties of the number of scintillation protons collected by the fibers while the Cherenkov systematics takes care of the uncertainty of number of Cherenkov photons.

6.3.3 Normalization Systematics

The normalization systematics account for uncertainties in the protons on target (POT), detector mass, and selection efficiency. In the 2022 sterile analysis, two separate uncertainties are considered for selection efficiency: one for the ν_μ selection and another for the neutral current (NC) selection. Each of these reflects the discrepancy in selection efficiency between data and Monte Carlo (MC) simulations for the respective ν_μ or NC samples. Additionally, a dedicated systematic is included to account for a known bug in the simulation that causes leptons to be unintentionally dropped. Since this issue is still under investigation, the potential impact of the missing leptons is incorporated as a separate systematic uncertainty [77].

6.3.4 Muon Energy Systematics

The muon energy systematics address uncertainties [82] in the reconstructed muon energy derived from the muon track length. They are implemented for the Near

Detector (ND), Far Detector (FD), and the Muon Catcher. These systematics can be categorized into two types: uncorrelated uncertainties, which are primarily driven by detector mass accounting, and correlated uncertainties, which are dominated by uncertainties in physics simulation tools such as GEANT4.

6.3.5 Cross Section Systematics

There are two types of cross section systematics: the GENIE knobs and the NOνA custom systematics. In GENIE, systematic uncertainties are handled through a set of reweighting knobs that affect various aspects of neutrino interaction modeling. These knobs take care of the uncertainties of GENIE cross-section model parameters for axial and vector mass, form factors, decay branching ratios, etc. Among them, the QE and RES knobs have the most significant impact on modeling neutral current interactions (see Table 6.3). The reweighting knobs adjust the weights of individual neutrino events so that distributions (e.g., energy spectra) can be modified without regenerating the events. The NOνA custom systematics are implemented to cover the additional uncertainties including the Final State Interaction (FSI) systematics. In neutrino interactions, hadrons produced within the nucleus may undergo further interactions as they propagate through the nuclear medium. These secondary interactions are collectively referred to as Final State Interactions (FSI). FSIs are typically categorized into three topological channels: Absorption(ABS, no pions are observed in the final state), Charge Exchange(CX, a single π^0 is observed in the final state) and Quasi-Elastic(QE,

Knob	Description
MaNCEL	Axial mass in neutral current quasi-elastic (NCEL) scattering model, affects $\frac{d\sigma_{(\mathrm{NCEL})}}{dQ^2}$ both in shape and normalization
EtaNCEL	NCEL strange axial form factor eta, affects $\frac{d\sigma_{(\mathrm{NCEL})}}{dQ^2}$ both in shape and normalization
MaNCRES	Axial mass in neutral current resonance (RES) pion production model, affects $\frac{d\sigma_{(\mathrm{NCRES})}}{dW\,dQ^2}$ both in shape and normalization
MvNCRES	Vector mass in neutral current resonance (RES) pion production model, affects $\frac{d\sigma_{(\mathrm{NCRES})}}{dW\,dQ^2}$ both in shape and normalization
BR1gamma	*Resonance* $\rightarrow X+\gamma$ branching ratio, e.g. $\Delta^+_{(1232)} \rightarrow p+\gamma$
BR1eta	*Resonance* $\rightarrow X+\eta$ branching ratio, e.g. $N^+_{(1440)} \rightarrow p+\eta$
Theta_Delta2Npi	distort π angular distribution in $\Delta \rightarrow N+\pi$

Table 6.3: Selected GENIE QE and RES knobs.

single pion with same sign as initial pion beam is observed). There are four final state interaction(FSI) systematics defined in NOvA: khNFSISyst2020_EV1, khNFSISyst2020_EV2, khNFSISyst2020_EV3 and khNFSISyst2020_MTP. The first three (EV1, EV2, EV3) are derived by diagonalizing the covariance matrix associated with the three main FSI channels (ABS, CX, QE), resulting in uncorrelated eigenvectors representing variations in the underlying physics processes. The fourth systematic, MTP, is associated with the mean free path of hadrons inside the nucleus, and effectively scales the total FSI cross section inversely. These four systematics are defined to be mutually uncorrelated, allowing for independent treatment in uncertainty analysis. For a complete description of all cross-section systematics used in NOνA, see [83].

6.3.6 NC MEC Systematics

Meson Exchange Current (MEC) systematics were originally implemented only for charged-current (CC) interactions by default. For the 2022 sterile neutrino analysis, a dedicated neutral-current (NC) MEC systematic was introduced. However, subsequent studies indicated that the NC event selection includes only a negligible fraction of MEC events. As a result, the impact of the NC MEC systematic on the analysis is considered negligible.

6.4 PISCES

Parameter Inference with Systematic Covariance and Exact Statistics (PISCES) [84] is a hybrid test statistic developed for sterile analysis in NOνA. PISCES performs covariance matrix-based fits, treating systematic uncertainties as a multivariate Gaussian and statistical uncertainties using exact Poisson statistics. The Poisson likelihood is employed because it remains valid in the low-statistics regime, such as in the ν_e appearance channel where the total event count is small.

The likelihood with Gaussian form can be expressed as:

$$\mathcal{L} = \frac{1}{\sqrt{(2\pi)^N |V|}} \exp\left[-\frac{1}{2}\sum_{ij}(x_i - \mu_i)V_{ij}^{-1}(x_j - \mu_j)\right], \tag{6.4}$$

where V is the covariance matrix. x_i and μ_i are the data and oscillation prediction in i'th bin, respectively.

Use the relationship $\chi^2 = -2\log\mathcal{L}$ for Gaussian likelihood and remove constant terms, we have:

$$\chi^2_{\text{Gaussian}} = \sum_{ij}(x_i - \mu_i)V^{-1}_{ij}(x_j - \mu_j) + \log|V|. \tag{6.5}$$

PISCES separates the matrix term into statistical and systematic components and replaced the Gaussian form of statistical uncertainty with Poisson form. The new equation reads [84]:

$$\chi^2_{\text{Hybrid}} = 2\sum_i \left[m_i - x_i + x_i \log\left(\frac{x_i}{m_i}\right)\right] + \sum_{ij}(\mu_i - m_i)\Sigma^{-1}_{ij}(\mu_j - m_j) + \log|V|, \tag{6.6}$$

where $m_i = \mu_i + \sum_j^K \alpha_j s_{ji}$ and Σ_{ij} is the covariance matrix for systematics. Here α_j is the fit parameter for the j'th systematic uncertainty and s_{ji} is the reweight for j'th systematic in i'th bin. K is the total number of systematics types.

Considering the beam components (e.g. FD NC) and assigning independent binning schemes for them, the expression can be transformed to the following form:

$$\chi^2_{\text{Hybrid}} = 2\sum_i \left[\left(\sum_\alpha \mu_{\alpha i} f_{\alpha i}\right) - x_i + x_i \log\left(\frac{x_i}{\sum_\alpha \mu_{\alpha i} f_{\alpha i}}\right)\right] \tag{6.7}$$

$$+ \sum_{\alpha\beta}\sum_{ij}(f_{\alpha i} - 1)F^{-1}_{\alpha i \beta j}(f_{\beta j} - 1), \tag{6.8}$$

where $F_{\alpha i \beta j}$ is the component-wise covariance matrix and $f_{\alpha i}$ is the multiplicative systematic shift (systematic pull). $F_{\alpha i \beta j}$ is a fixed matrix, with elements normalized by the absolute values of bins i and j, and does not vary during the fit.

PISCES employs the profile likelihood technique. Two sterile neutrino oscillation parameters (Δm^2_{41} vs $\sin^2\theta_{24}$ or $\sin^2\theta_{34}$) are selected as parameters of interest (POI), and a 2D parameter grid is constructed. At each grid point, all remaining oscillation parameters and systematic pulls are allowed to float and are optimized to minimize the χ^2. At the end of the fit, the χ^2 is obtained as a function of the two parameters. In addition, an independent fit is performed in which all oscillation parameters, including the parameters of interest (POIs), are allowed to float. This provides the global best-fit point and the corresponding minimum chi-square value, χ^2_{min}. Confidence contours are obtained by evaluating $\Delta\chi^2 = \chi^2 - \chi^2_{\text{min}}$ and selecting the parameter regions that correspond to fixed confidence levels, using critical values derived from either Wilks' theorem or the profiled Feldman–Cousins method. This procedure evaluates the sensitivity of the experiment, showing the range of parameter space that can be excluded under the no-signal (null hypothesis) assumption.

Chapter 7

The LSTM energy estimator

Various deep neural networks have been developed for ν_μ charged-current (CC) energy estimation in NOνA, achieving neutrino energy resolution with an RMS below 10%. Despite the success of these deep learning approaches for ν_μ CC events, there is currently no neural network-based energy estimator deployed for neutral-current (NC) interactions.

In this chapter, I will discuss the design of neural networks for neutrino energy estimation, the training and evaluation of these models, and their integration into the sterile neutrino analysis.

7.1 Introduction to Neural Networks

7.1.1 A Brief history of Neural Networks

Neural networks, inspired by the structure and functioning of the human brain, have undergone significant evolution over the past few decades. The concept of

neural networks can be traced back to the 1940s, when Warren McCulloch and Walter Pitts proposed the first mathematical model of a neuron [85]. The model represented a simplified version of biological neurons, where each neuron was a binary [1] unit that received inputs, performed a weighted summation, and produced an output if the sum exceeded a certain threshold. The output was then passed on to other neurons, forming a network. This model laid the groundwork for later developments in artificial neural networks by demonstrating that complex logical functions could be performed through interconnected simple units, mirroring basic neural activity in the brain. However, it wasn't until the 1980s that neural networks gained significant attention with the development of the backpropagation algorithm [86], which allowed for effective training of multi-layer networks. Conceptually rooted in the chain rule of calculus, backpropagation computes the gradient of a loss function with respect to each weight (trainable parameters) in the network by recursively applying the chain rule from the output layer back to the input layer. This gradient information is then used to update the network parameters, thereby minimizing the error between the predicted and target outputs.

In the decades that followed, neural networks encountered periods of both enthusiasm and stagnation. However, the resurgence of interest in the 2000s, driven by the availability of large datasets and powerful computing resources, especially the General-Purpose Graphics Processing Unit (GPGPU), led to the advent of deep learning, a subfield of neural networks focused on training multi-layer architectures. In particular, deep neural networks (DNNs) have achieved

[1] The input of unit is either true or false.

groundbreaking successes in various applications, including computer vision [87], natural language processing [88], and speech recognition [89].

7.1.2 Basic Structure of Neural Networks

At a fundamental level, a neural network can be viewed as a parameterized function that maps input tensors to output tensors through a sequence of linear transformations and non-linear activations. Each layer of the network performs a tensor contraction—typically a matrix multiplication—between the input and a learnable weight tensor, followed by the addition of a bias term. Formally, for a given layer, the transformation can be written as

$$\mathbf{z} = \mathbf{W}\mathbf{x} + \mathbf{b}, \tag{7.1}$$

where $\mathbf{x} \in \mathbb{R}^n$ is the input vector (or tensor), $\mathbf{W} \in \mathbb{R}^{m \times n}$ is the weight matrix, and $\mathbf{b} \in \mathbb{R}^m$ is the bias.

To introduce non-linearity into the model—essential for representing complex, non-linear functions—the linear output $\mathbf{z}$ is passed through a non-linear activation function, applied element-wise:

$$\mathbf{a} = \sigma(\mathbf{z}), \tag{7.2}$$

where σ may be a ReLU (i.e. $\sigma(z) = \max(0, z)$), sigmoid, tanh, or other non-linear function. The role of the non-linear activation is crucial: without it, the entire network would collapse into a single linear transformation, regardless of depth. By interleaving nonlinearities with affine layers ($\mathbf{W}\mathbf{x} + \mathbf{b}$), we can build

a Multi-Layer Perceptron (MLP). The MLP is able to construct complex, highly non-linear mappings from inputs to outputs.

7.2 Design of Neural Networks for Neutrino Energy Estimation

Broadly, there are two categories of neural network-based energy estimators used in NOνA. The first category is based on convolutional neural networks (CNNs) [90]. Each neutrino event is represented by two 2D pixel (hit) maps corresponding to the XZ and YZ views of the NOvA detector [91]. These images are processed by two separate CNN branches—one for each view—with independent weights. After convolutional processing, the outputs of the two branches are combined with the reconstructed interaction vertex position (a 3D coordinate). This combined feature vector is then passed through a linear layer to predict a single positive scalar: the energy of the neutrino or electron (for electron shower energy estimation).

The other type of neural networks use the reconstructed prong and slice information as inputs. Each neutrino event contains a variable number of 3D prongs and 2D prongs, each ranging independently from 0 to 10. Since the number of prongs is not fixed across events, we adopt neural network architectures that can handle variable-length sequences. In practice, we use either Long Short-Term Memory (LSTM) [92] or Transformer [93] layers to process the prong features, as these models are well-suited for sequential data with dynamic lengths. The

slice information is represented as a fixed-size feature vector and is processed using an MLP[2]. After individual processing, the outputs are combined and passed through another MLP to produce the final outputs: the energy values of incoming neutrinos and outgoing leptons.

7.2.1 Input format of prong and slice

The 3D (or 2D) prongs of an neutrino event can be represented as a 2D tensor with shape (P, F_{prong}), where P denotes the number of prongs in the event and F_{prong} denotes the number of features associated with each prong. Each feature corresponds to a physical quantity, such as the reconstructed energy, the starting position, and other relevant attributes of the prong. The slice associated with the event is represented by a 1D tensor of shape (F_{slice}), where F_{slice} denotes the number of slice-level features. While the number of prongs P varies from event to event, the lists of prong features and slice features remain fixed across the dataset for a given energy estimator.

7.2.2 Neural Networks for Sequential Data

RNN

Recurrent Neural Networks (RNNs) are a class of neural architectures designed to model sequential data by maintaining a hidden state that is updated as new inputs from the sequence are processed. At each step, an RNN updates its hidden state

[2]For LSTM EE, there is no MLP that processes the slices. In other words, the MLP is replaced by an identity matrix.

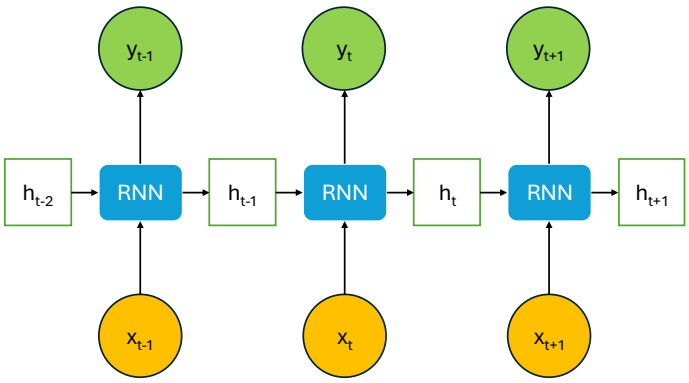


Figure 7.1: Unrolled structure of an RNN over multiple steps, where each hidden state h_t is updated based on the previous hidden state h_{t-1} and the current input x_t.

based on the current input and the previous hidden state, following the function

$$(y_t, h_t) = \text{RNN}(x_t, h_{t-1}) \tag{7.3}$$

where x_t is the input at step t, h_{t-1} is the previous hidden state, h_t is the updated hidden state, and y_t is the output. This recurrent structure makes RNNs naturally suited for a wide range of tasks involving sequential data.

Despite their conceptual advantages, standard RNNs often struggle to learn long-range dependencies. During training, the backpropagation of gradients through many sequential steps can result in vanishing or exploding gradients, making it difficult for the network to retain information across long sequences. Consequently,

standard RNNs tend to bias toward short-term patterns and fail to effectively capture distant relationships.

To overcome these limitations, Long Short-Term Memory (LSTM) networks were introduced. LSTMs extend the basic RNN architecture by incorporating a memory cell and a gating mechanism that regulate the flow of information. Specifically, the input, forget, and output gates control how information is added to, removed from, and extracted from the memory cell. This design allows LSTMs to maintain and manipulate information over long sequences, enabling the network to learn both short-range and long-range dependencies more effectively.

Due to these properties, LSTM networks have become a widely adopted architecture for sequential modeling tasks where retaining complex or distant relationships within the data is critical.

Transformer

The Self-Attention mechanism, introduced by Vaswani et al. in the Transformer architecture, is designed to enable a model to process and capture relationships between all elements in a sequence, regardless of their position. This approach is particularly useful for tasks where understanding dependencies over long ranges is crucial. Self-attention allows each element of a sequence to "attend" to every other element in the sequence when creating its representation. Unlike models like RNNs or CNNs, which process data sequentially or locally, self-attention calculates how much attention each part of the input should pay to every other part.

7.2.3 LSTM Energy Estimator

The LSTM Energy Estimator in NOνA was developed to estimate the energy of ν_μ charged-current (CC) events. It takes as input reconstructed 3D prong information, 2D prong projections, and slice-level data, and predicts the energies of both the incoming ν_μ neutrino and the outgoing muon. The architecture of the LSTM Energy Estimator is shown in Figure 7.2.

All 3D prongs of each event, represented as a tensor of shape $(P \times 33)$, where P is the number of prongs, are first sorted by prong length in descending order and then processed through interleaved Batch Normalization and Linear layers, resulting in a transformed tensor of shape $(P \times 128)$. These tensors are then fed into an LSTM layer, the output of which is a fixed-length vector of size 32. The 2D prongs are processed in an identical and independent group of Batch Normalization and Linear layers. The slice information is represented as a vector of length 6. These three vectors are concatenated into a single vector of size 70, which is subsequently passed through another sequence of interleaved Batch Normalization and Linear layers, resulting in a final output vector of length 2 representing ν_μ and muon energies. The list of input variables and their summaries are shown in Table 7.1.

While the model is trained on ν_μ CC events, it does not impose a strict requirement that the input be a true ν_μ CC event; it can be applied to any events[3]. In such cases, the model performs the energy prediction under the assumption that the event is a ν_μ CC interaction.

[3]If there is no prong in an event, a tensor of all zeros is used to represent the prong input

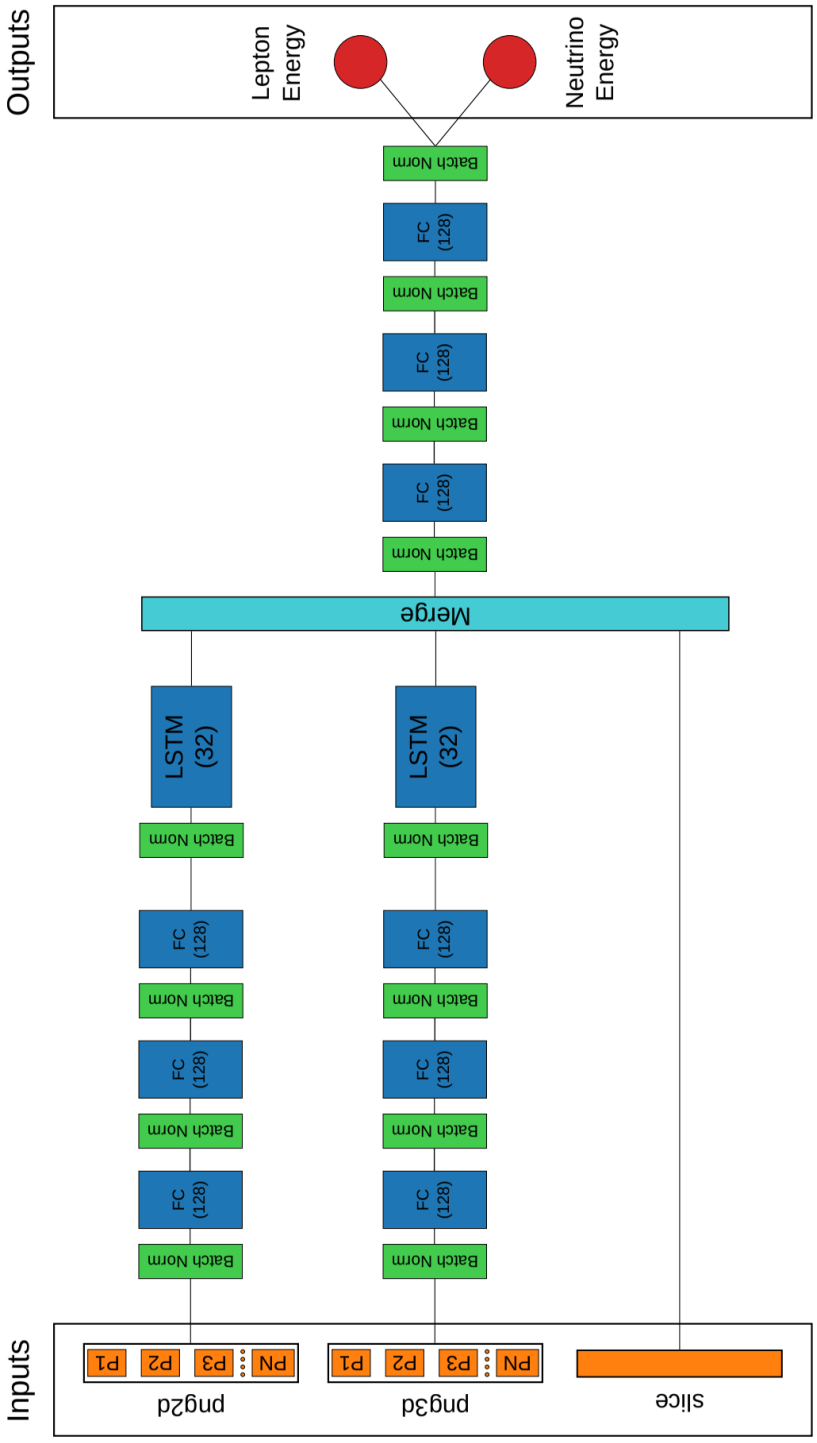


Figure 7.2: Schematic of ν_μ CC LSTM Energy Estimator.

Type	**Description**
3D prong	x, y, z components of the direction vector for 3D prong
	x, y, z components of the position vector for 3D prong
	ProngCVN output scores representing the likelihood that the prong was produced by a muon,electron,pion,proton,photon
	Estimated energy by Break Point Fitter under different particle assumptions (0: proton, 1: pion, 2: muon)
	BPF energy of the 3D prong shared with other prongs
	x, y, z components of the BPF momentum for 3D prong
	Prong length
	Number of hits in the 3D prong
	Number of hits of the 3D prong in XZ, YZ view
	Number of planes crossed by the 3D prong track
	Sum of the calibrated energy of all hits associated with the 3D prong. The full energy of hits shared among multiple prongs is included.
	Sum of the calibrated energy of all hits associated with the 3D prong. For hits shared among prongs, the energy is apportioned using the weights determined by FuzzyK.
2D prong	2D prong variables analogous to those of 3D prong
Slice	Sum of the calibrated energy of all hits in the slice
	calE - sum(png.calE)
	Number of hits in the slice
	Calorimetric energy of hits that don't appear in any FuzzyK prongs
	NOνA used coarse timing for FD in period 1. In all other periods, and for all ND data, fine timing is used[4].
	A boolean that indicated if data was collected during low gain run periods.

Table 7.1: Summary of input variables of LSTM Energy Estimator. There are 33 variables of 3D prong, 13 variables of 2D prong and 6 variables of slice.

[4]Periods are intervals of time characterized by uniform running conditions.

7.2.4 Transformer Energy Estimator

The Transformer Energy Estimator is a prong-based model developed for ν_μ CC events in NOνA. Originally designed as an alternative to LSTM models in the DUNE experiment, it was later adapted for use in NOνA. This model takes both 3D prong and slice-level information as input and is capable of predicting not only the energies of muon neutrinos and muons, but also other quantities.

The 3D prong features are processed through a stack of 12 transformer encoder layers, while the slice information is handled by an MLP. The outputs from the transformer encoders are summed over the prong dimension to form a vector of size F_{prong}. The output of the slice MLP is a vector of size 16. The two vectors are concatenated into a single vector of size $F_{\text{prong}} + 16$ and passed to another MLP to generate the final output. See Figure 7.3 for architecture of the Transformer Energy Estimator.

The final output of the transformer is invariant to the prong order, as sum pooling is applied over the encoder outputs across the prong dimension. Moreover, transformer models tend to perform better when the number of prongs is large, as RNNs including LSTMs are prone to forgetting long-range dependencies due to their sequential nature. In contrast, transformers process all prongs in parallel and capture global relationships more effectively, regardless of sequence length.

7.2.5 Choosing An Energy Estimator

The LSTM-based energy estimator is currently implemented in the NOvA analysis framework. The Transformer-based energy estimator, however, will only become

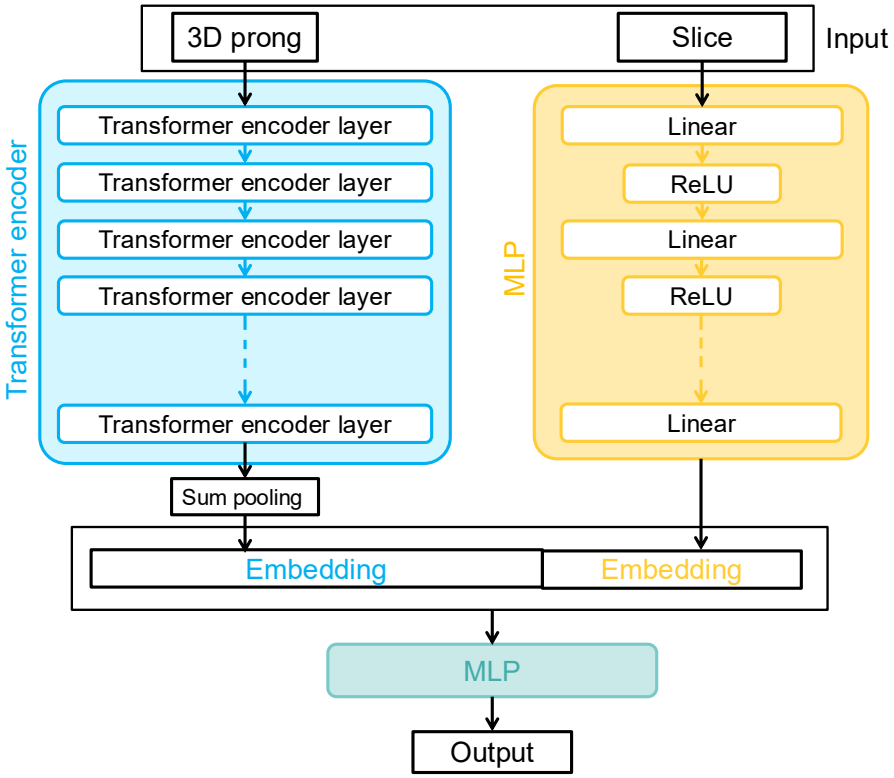


Figure 7.3: Schematic of ν_μ CC Transformer Energy Estimator. The embeddings are vectors that contains processed information of prongs and slices.

available in the next software iteration, which is still in the process of being produced. Moreover, the current implementation does not allow for the direct incorporation of systematic uncertainties into the Transformer's input. As a result, we choose to use the LSTM architecture for neutral current (NC) energy estimation.

7.3 Train LSTM Energy Estimator for NC Events

Although LSTM models have shown promising results on CC events, there has been no prior attempt, to our knowledge, to train a deep learning–based energy estimator directly on NC events in NOνA. To enable this, we need to prepare dedicated neutral current event datasets for training, adapt the existing training strategies, and modify the model and hyper-parameters accordingly.

7.3.1 Neutral Current Dataset

For this analysis, we use the same analysis framework and files that were used for the NOvA 2022 sterile neutrino results. For training, we randomly selected approximately 50% of FD events and 50% of ND events from the FHC NC nominal nonswap samples. These samples were generated assuming no neutrino oscillations, which consists primarily of muon neutrinos. Separate LSTM models are trained on the ND and FD datasets. The remaining events, referred to as the sterile analysis set, are reserved for use in the sterile neutrino analysis.

The events are filtered using the same selection criteria as in the 2022 sterile neutrino analysis. In addition, two further requirements are applied: the event must originate from a true neutrino interaction, and there must be at least one

primary particle produced at the interaction vertex. After event selection, a total of 303,628 NC events remain in the FD sample, and 92,907 NC events remain in the ND sample.

7.3.2 The modified version of LSTM for NC

Model outputs

As noted earlier, the ν_μ CC LSTM model yields two outputs: the reconstructed energies of the ν_μ and the corresponding outgoing muon. However, for neutral current (NC) events, there is no outgoing muon. One possible alternative is to repurpose the second output of the model to predict the energy of the outgoing neutrino instead. While this could, in principle, allow the model to retain a two-output structure, it introduces significant challenges. Unlike charged particles such as muons and protons, neutrinos typically do not leave direct signals in the detector. As a result, the energy of the outgoing neutrino cannot be measured directly, and any attempt to reconstruct it must rely solely on the visible remnants of the interaction. Moreover, the limited size of our dataset further complicates the reliable training of such a target. Therefore, for NC events, we choose a simpler approach by having the LSTM model predict only the energy of the incoming neutrino.

Loss functions and weights

The loss function is the mean absolute percentage error between the reconstructed neutrino energy by LSTM and the true neutrino energy:

$$Loss = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{E_{\nu,i}^{reco} - E_{\nu,i}^{true}}{E_{\nu,i}^{true}}\right|, \tag{7.4}$$

where $N = 1024$ is the batch size, i.e., the number of training examples processed by the model before each weight update. This loss function effectively corresponds to the average absolute value of the energy resolution.

Since the distribution of true neutrino energies is not uniform but instead concentrated around the energy of the neutrino beam, it is necessary to apply event-wise weighting during training to prevent bias toward overrepresented energy regions. We adopt the same weighting scheme used in the training of the ν_μ CC LSTM model. Specifically, we divide the energy range into bins and compute the number of events in each bin. The weight assigned to events in a given bin is set to be inversely proportional to the event count in that bin. These weights are then multiplied by the individual MAPE values to compute a weighted average loss.

An analysis of both the FD and ND datasets reveals that more than 95% of the events have true neutrino energies within the range of 0 to 20 GeV. Based on this observation, we define 100 uniform bins over the interval [0, 20] GeV for weight computation. To avoid excessively large weights that could destabilize training, we cap the maximum weight at 50 times the global minimum weight. Any computed weight exceeding this threshold is clipped to this maximum value.

Now Equation (7.4) becomes:

$$Loss_{\text{weighted}} = \frac{1}{N}\sum_{i=1}^{N}\left| w(E_{\nu,i}^{\text{true}}) \times \frac{E_{\nu,i}^{\text{reco}} - E_{\nu,i}^{\text{true}}}{E_{\nu,i}^{\text{true}}} \right|, \tag{7.5}$$

where weight w is a function of $E_{\nu,i}^{\text{true}}$.

Train and evaluation

We trained the models (one for the FD and one for the ND) for a maximum of 200 epochs. Early stopping was employed with a patience of 40 epochs by monitoring the validation loss during training. In details, if the validation loss did not show any improvement for 40 consecutive epochs, the training process was terminated early to prevent overfitting. Additionally, we applied a learning rate scheduler that reduced the learning rate by a factor of 0.5 if the validation loss did not improve for 5 consecutive epochs. The initial learning rate is 0.001.

Evaluations of the models were conducted on the sterile analysis set. In Figures 7.4 and 7.5, we present two-dimensional histograms of the LSTM-predicted energy versus the true neutrino energy. For comparison, the corresponding histograms for the Nus20 energy estimator are shown in Figures 7.6 and 7.7. It is important to note that the Nus20 energy estimator is designed to reconstruct the hadronic shower energy, rather than the energy of the incoming neutrino. While the Nus20 estimator effectively reconstructs the hadronic energy, the presence of invisible outgoing neutrinos introduces an intrinsic uncertainty in estimating the true neutrino energy. This poses a challenge for oscillation analyses, which rely

directly on accurate knowledge of the neutrino energy. In contrast, the LSTM-based model is trained to directly predict the energy of the incoming neutrino. To account for this difference, we also provide two-dimensional histograms of the Nus20 energy versus the true hadronic energy in Figures 7.8 and 7.9. In Figure 7.10, we present the distributions of LSTM energy resolution in the FD and the ND respectively.

We further provide the resolution versus true neutrino energy plots in Figure 7.11. We observe that the LSTM model exhibits suboptimal performance in reconstructing neutrino energy in the low-energy regime, particularly for events with true energies below approximately 1.5 GeV.

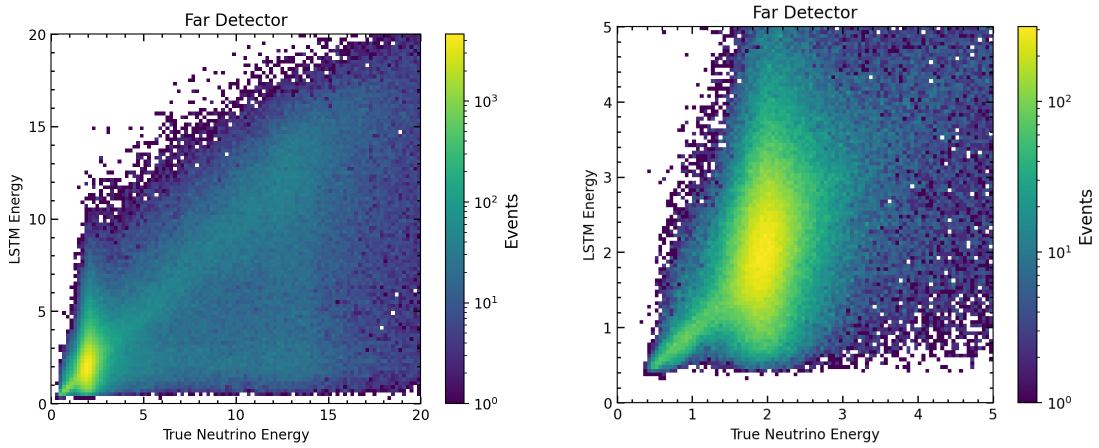


Figure 7.4: 2D histograms of LSTM-predicted energy vs. true neutrino energy for FD events in the sterile analysis set. Left: energy range from 0 to 20 GeV. Right: Zoomed-in view focusing on the 0 to 5 GeV region.

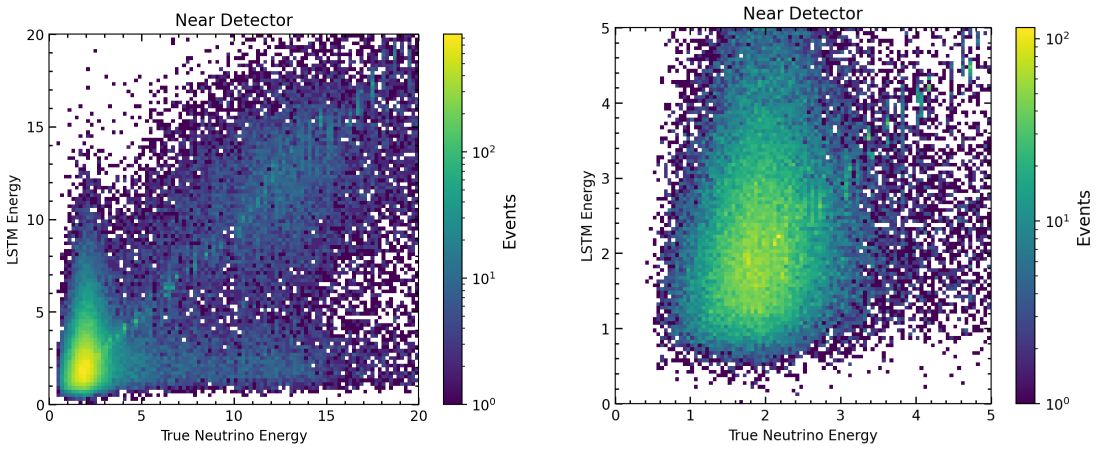


Figure 7.5: 2D histograms of LSTM-predicted energy vs. true neutrino energy for ND events in the sterile analysis set. Left: energy range from 0 to 20 GeV. Right: Zoomed-in view focusing on the 0 to 5 GeV region.

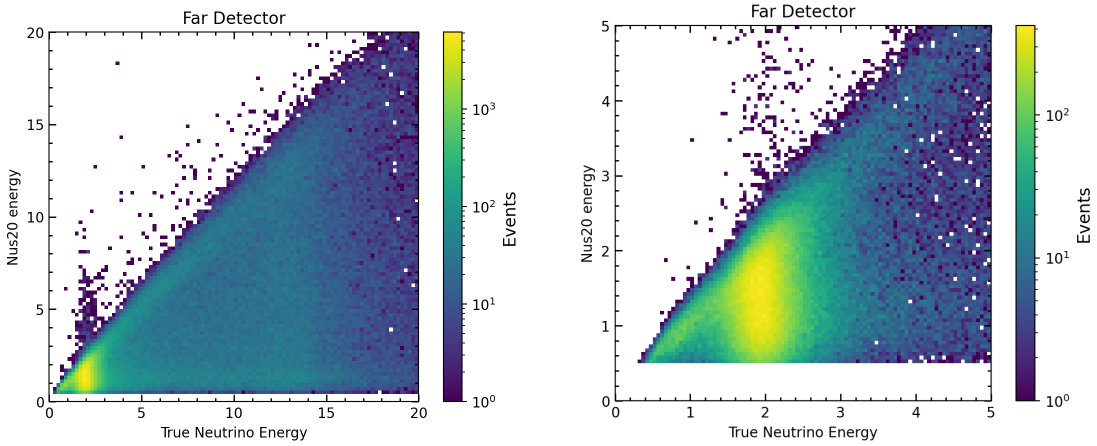


Figure 7.6: 2D histograms of Nus20 energy vs. true neutrino energy for FD events in the sterile analysis set. Left: energy range from 0 to 20 GeV. Right: Zoomed-in view focusing on the 0 to 5 GeV region.

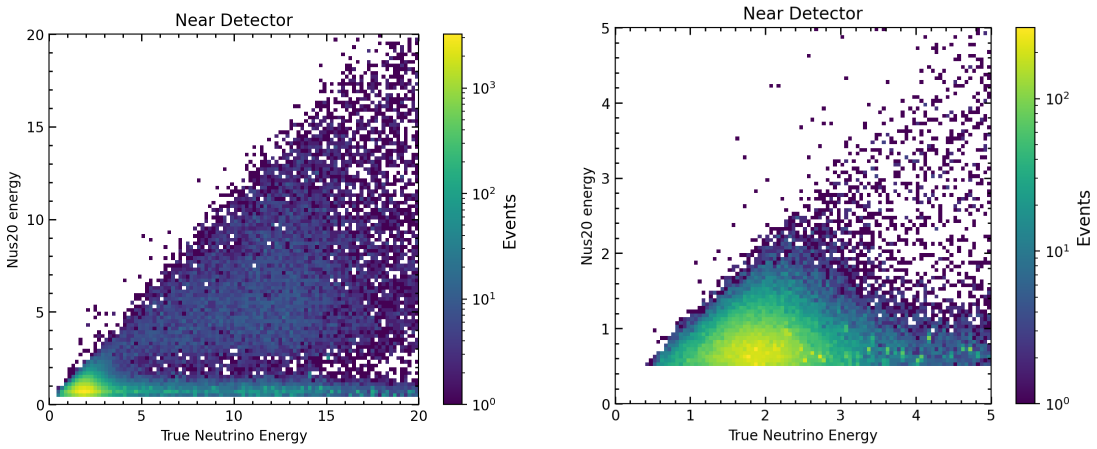


Figure 7.7: 2D histograms of Nus20 energy vs. true neutrino energy for ND events in the sterile analysis set. Left: energy range from 0 to 20 GeV. Right: Zoomed-in view focusing on the 0 to 5 GeV region.

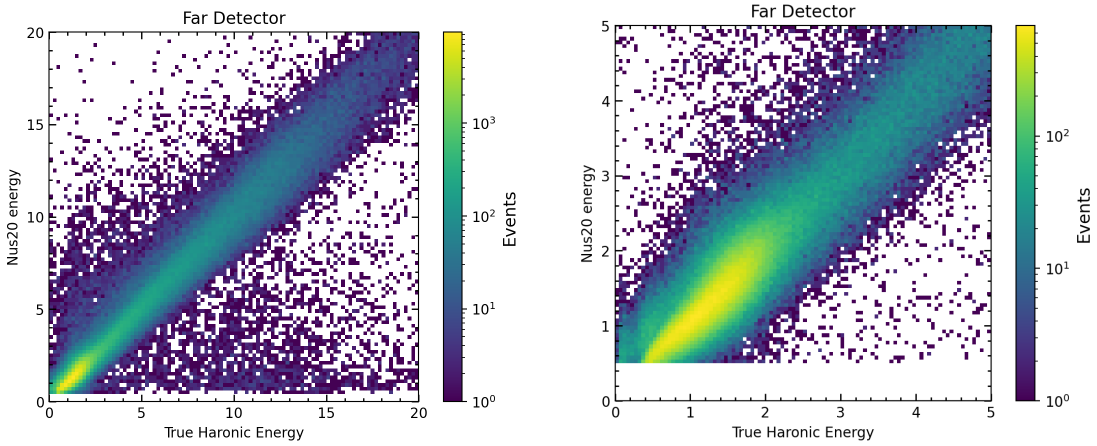


Figure 7.8: 2D histograms of Nus20 energy vs. true hadronic energy for FD events in the sterile analysis set. Left: energy range from 0 to 20 GeV. Right: Zoomed-in view focusing on the 0 to 5 GeV region.

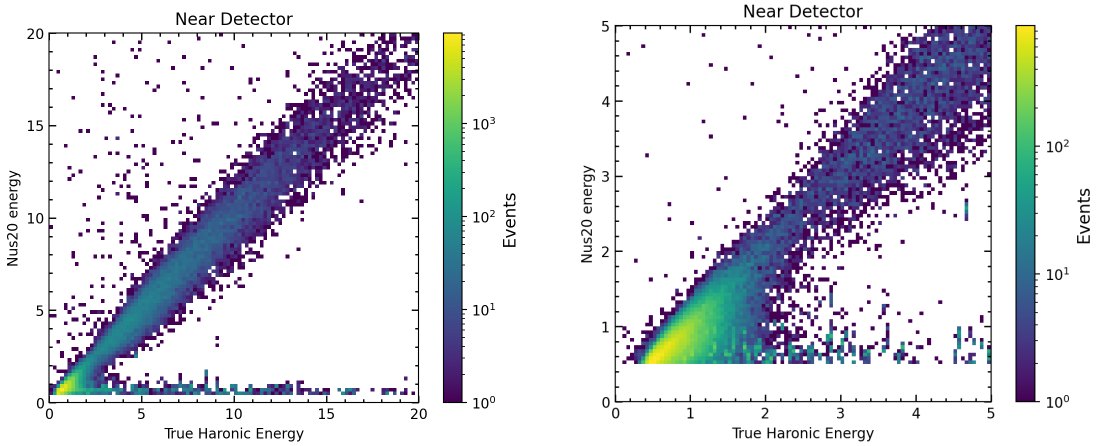


Figure 7.9: 2D histograms of Nus20 energy vs. true hadronic energy for ND events in the sterile analysis set. Left: energy range from 0 to 20 GeV. Right: Zoomed-in view focusing on the 0 to 5 GeV region.

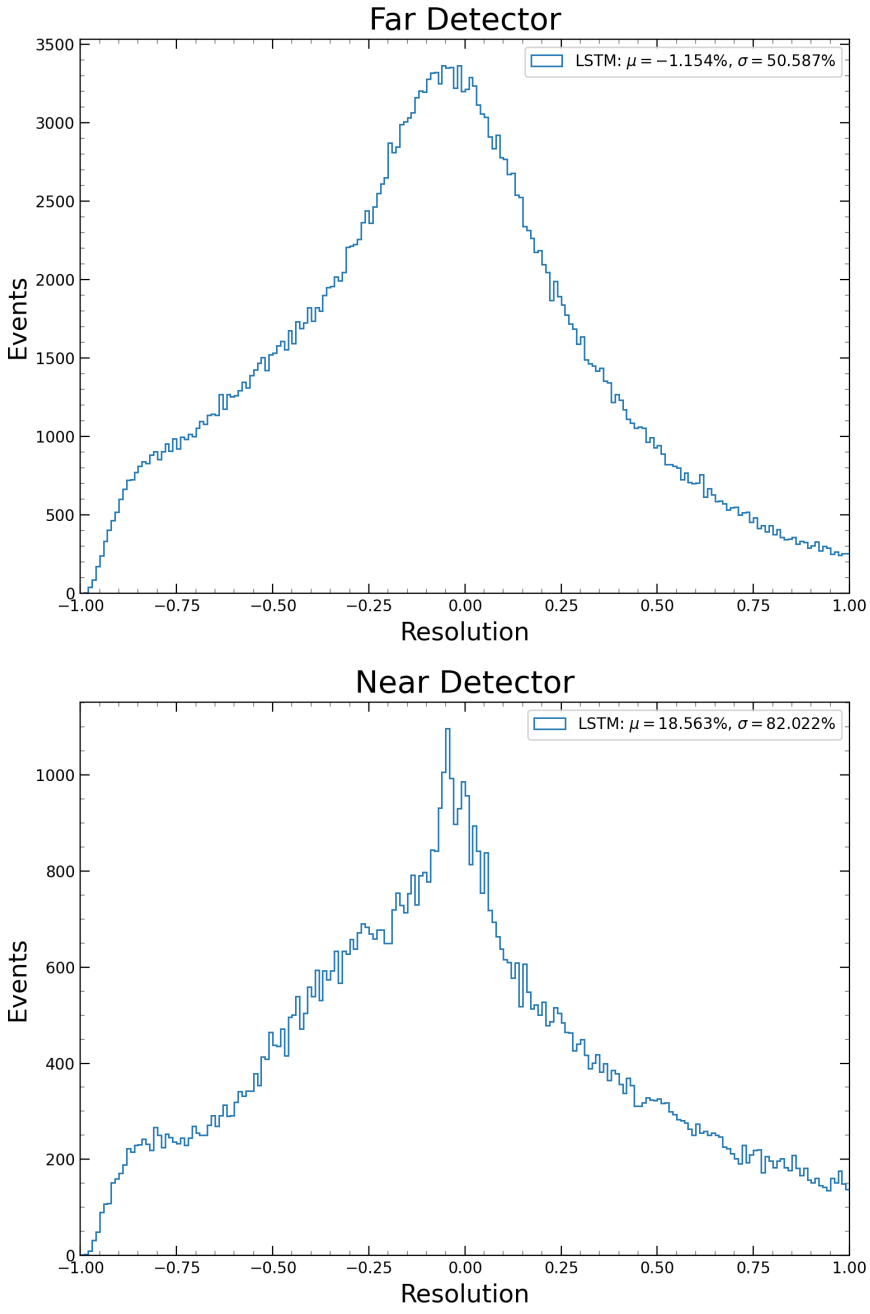


Figure 7.10: Histograms of LSTM energy resolution for FD (top) and ND (bottom) events in the sterile analysis set.

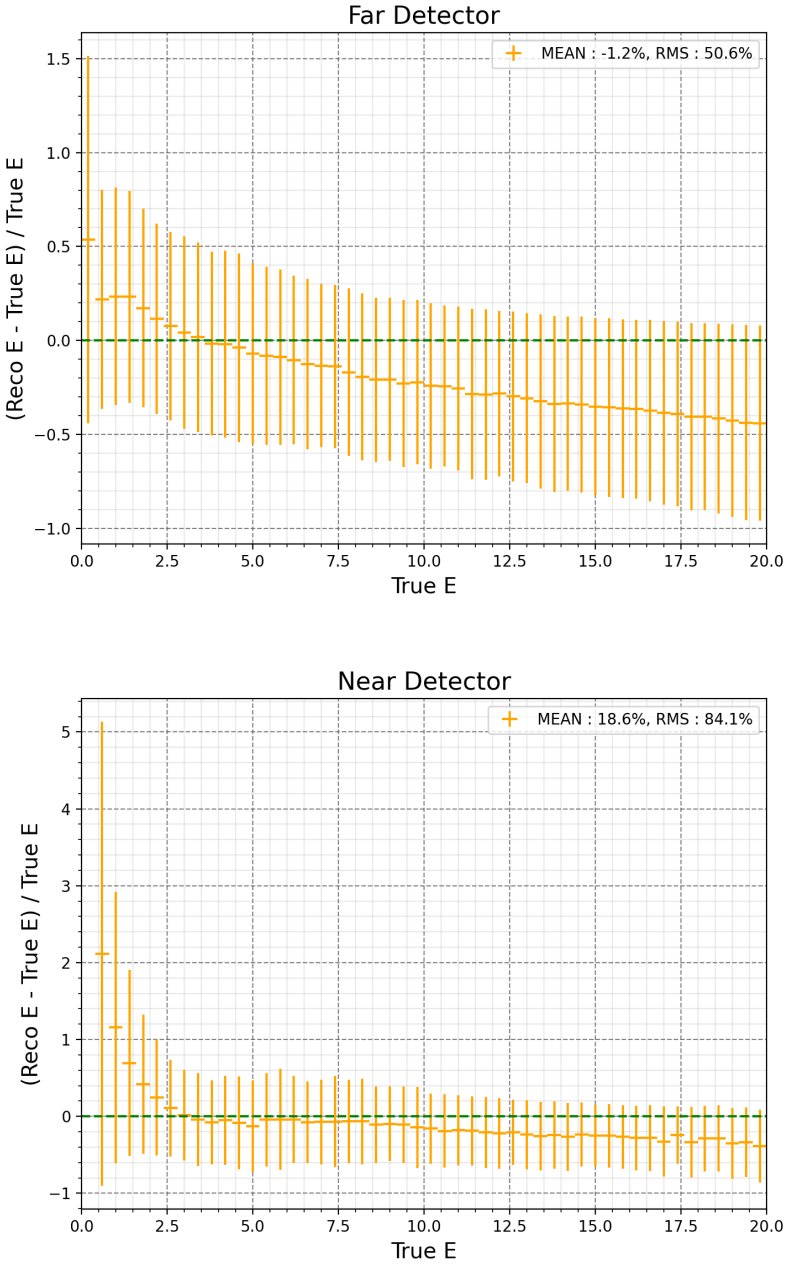


Figure 7.11: LSTM energy resolution as a function of true neutrino energy for FD (top) and ND (bottom) events in the sterile analysis set. In each bin of true neutrino energy, the mean and root-mean-square (RMS) of the resolution distribution are computed and shown as the central value and the corresponding error bar, respectively.

7.4 Evaluating the Robustness of LSTM Energy Estimator under Different Systematics

In this section, we evaluate the robustness of the LSTM Energy Estimator against various sources of systematic uncertainties. We apply $+1\sigma$ and -1σ shifts to the input variables of the LSTM and recompute the mean and root-mean-square of the resulting energy resolution distributions. In addition to the previously considered systematics, we introduce a new source of uncertainty associated with prong length.

7.4.1 RES and QE Systematics

Tables 7.2 and 7.3 summarize the impact on the LSTM energy prediction when each RES and QE cross-section systematic is applied. See Table 6.3 for descriptions of each systematic.

Systematic	Shift	mean of resolution	change of mean	RMS of resolution	change of RMS
None	0	-0.0259		0.466	
MaNCEL	+1 σ	-0.0294	-0.00349	0.466	-0.000393
	-1 σ	-0.0238	0.00211	0.467	0.000304
EtaNCEL	+1 σ	-0.026	-9.74e-05	0.466	-3.76e-06
	-1 σ	-0.0258	9.23e-05	0.466	3.64e-06
MaNCRES	+1 σ	-0.0328	-0.00689	0.468	0.00201
	-1 σ	-0.0195	0.00633	0.465	-0.00186
MvNCRES	+1 σ	-0.026	-0.00017	0.467	0.000356
	-1 σ	-0.0259	6.82e-06	0.466	-0.000208
BR1gamma	+1 σ	-0.0259	4.86e-06	0.466	-1.26e-05
	-1 σ	-0.0259	-4.87e-06	0.466	1.26e-05
BR1eta	+1 σ	-0.0249	0.000929	0.467	0.000218
	-1 σ	-0.0268	-0.000924	0.466	-0.000219
Theta_Delta2Npi	+1 σ	-0.0251	0.000783	0.467	9e-05
	-1 σ	-0.0253	0.000523	0.466	6.02e-05

Table 7.2: Effects of RES and QE Systematics on LSTM Energy Resolution Mean and RMS in FD.

Systematic	Shift	mean of resolution	change of mean	RMS of resolution	change of RMS
None	0	0.0841		0.574	
MaNCEL	+1 σ	0.0844	0.000302	0.571	-0.00331
	-1 σ	0.0838	-0.000382	0.577	0.0026
EtaNCEL	+1 σ	0.0841	-6.72e-05	0.574	-0.000269
	-1 σ	0.0842	6.57e-05	0.574	0.000266
MaNCRES	+1 σ	0.092	0.00784	0.581	0.00665
	-1 σ	0.0742	-0.00999	0.565	-0.0091
MvNCRES	+1 σ	0.0903	0.00612	0.577	0.00305
	-1 σ	0.0781	-0.00606	0.571	-0.003
BR1gamma	+1 σ	0.0841	-3.46e-05	0.574	-4.49e-05
	-1 σ	0.0842	3.46e-05	0.574	4.49e-05
BR1eta	+1 σ	0.0831	-0.00108	0.573	-0.000879
	-1 σ	0.0852	0.00106	0.575	0.000861
Theta_Delta2Npi	+1 σ	0.086	0.00188	0.575	0.000612
	-1 σ	0.0854	0.00125	0.575	0.000409

Table 7.3: Effects of RES and QE Systematics on LSTM Energy Resolution Mean and RMS in ND.

7.4.2 Final State Interaction Systematics

Tables 7.4 and 7.5 summarize the impact of Final state interaction (FSI) systematics, which directly influence the production and propagation of final-state particles in the nuclear medium.

Systematic	Shift	mean of resolution	change of mean	RMS of resolution	change of RMS
None	0	-0.0259		0.466	
MFP	+1 σ	-0.0273	-0.00138	0.466	-0.000389
	-1 σ	-0.0221	0.00379	0.468	0.00128
EV1	+1 σ	-0.0254	0.000458	0.467	0.000441
	-1 σ	-0.0243	0.0016	0.466	-3.26e-05
EV2	+1 σ	-0.0244	0.00145	0.466	6.46e-06
	-1 σ	-0.0246	0.00129	0.466	-0.000236
EV3	+1 σ	-0.0243	0.00155	0.466	-0.000106
	-1 σ	-0.024	0.00187	0.466	-0.000211

Table 7.4: Effects of Final State Interaction Systematics on LSTM Energy Resolution Mean and RMS in FD.

7.4.3 Prong Length Systematics

Prong length systematics are not included in the NOνA analysis framework by default. We implemented this effect by applying a +5% (-5%) shift in the prong length (both 3D and 2D) per standard deviation (σ). This systematic is included in the PISCES framework when we replace Nus20 energy with LSTM energy. See Tables 7.6 and 7.7 for impact of this systematic.

Systematic	Shift	mean of resolution	change of mean	RMS of resolution	change of RMS
None	0	0.0841		0.574	
MFP	+1 σ	0.0847	0.000601	0.575	0.000696
	-1 σ	0.085	0.000811	0.573	-0.000634
EV1	+1 σ	0.083	-0.00111	0.574	6.73e-05
	-1 σ	0.0847	0.000586	0.574	-0.000153
EV2	+1 σ	0.0849	0.000733	0.574	-0.000207
	-1 σ	0.0848	0.000627	0.574	-9.21e-05
EV3	+1 σ	0.0847	0.000589	0.574	-0.000501
	-1 σ	0.0856	0.00144	0.574	5.06e-05

Table 7.5: Effects of Final State Interaction Systematics on LSTM Energy Resolution Mean and RMS in ND.

Systematic	Shift	mean of resolution	change of mean	RMS of resolution	change of RMS
None	0	-0.0259		0.466	
Prong length	+1 σ	-0.0283	-0.00247	0.466	-0.000367
	-1 σ	-0.0232	0.00272	0.467	0.000541

Table 7.6: Effects of Prong Length Systematics on LSTM Energy Resolution Mean and RMS in FD.

Systematic	Shift	mean of resolution	change of mean	RMS of resolution	change of RMS
None	0	0.0841		0.574	
Prong length	+1 σ	0.0936	0.00946	0.578	0.00395
	-1 σ	0.0754	-0.0087	0.572	-0.00262

Table 7.7: Effects of Prong Length Systematics on LSTM Energy Resolution Mean and RMS in ND.

7.4.4 Calibration Systematics

Among the calibration-related systematics considered in the NOνA analysis, the calibration scale systematic has the most significant impact on energy reconstruction. To improve the robustness of the LSTM model against this uncertainty, discrete random noises were applied to the input variables directly associated with the calibration scale during model training. Specifically, approximately one-third of the events had their calibration-related variables scaled by a factor of 1.2, another one-third by a factor of 0.8, while the remaining third were left unchanged.

We implemented a simplified calibration scale systematics to evaluate the robustness of model. The systematic applies a +5% (-5%) shift on the variables mentioned above.

From the results summarized in the tables above, it can be preliminarily concluded that the LSTM model exhibits negligible sensitivity to the listed categories

Systematic	Shift	mean of resolution	change of mean	RMS of resolution	change of RMS
None	0	-0.0259		0.466	
Calibration Scale	+1 σ	-0.0181	0.00779	0.47	0.00333
	-1 σ	-0.0336	-0.00774	0.463	-0.00334

Table 7.8: Effects of Calibration Scale Systematics on LSTM Energy Resolution Mean and RMS in FD.

Systematic	Shift	mean of resolution	change of mean	RMS of resolution	change of RMS
None	0	0.0841		0.574	
Calibration Scale	+1 σ	0.0842	2.11e-05	0.578	0.00428
	-1 σ	0.0836	-0.000515	0.572	-0.00181

Table 7.9: Effects of Calibration Scale Systematics on LSTM Energy Resolution Mean and RMS in ND.

of systematics. The variations in the mean and standard deviation of the resolution are both within approximately 0.01, indicating a high degree of robustness.

Chapter 8

Analysis Result

In this chapter, I will discuss the sterile analysis result using alternative energy estimation algorithms. Specifically, the NC energy estimator is replaced with an improved version, while the muon neutrino (numu) energy estimator remains unchanged.

As is shown in Equation (2.64), the muon neutrino survival probability is independent of θ_{34}. Sensitivity to θ_{34} arises from the measurement of active neutrino survival probability in neutral current events while sensitivity to θ_{24} is primarily driven by observations of muon neutrino disappearance [94].

8.1 Using True Energy as Reconstructed Energy

Before replacing the Nus20 energy estimator with the LSTM model, it is informative to evaluate the analysis sensitivity using the true NC neutrino energy as

a stand-in for the reconstructed energy. This provides a best-case scenario, illustrating the maximum potential improvement that could result from enhancing the energy estimator in NC events.

In the 2022 sterile neutrino analysis, cosmic data were included in the Far Detector background. However, since the true neutrino energy is only available for Monte Carlo (MC) events, we exclude cosmic backgrounds when using true energy in this sensitivity study.

The sensitivity is expressed as a contour in a two-dimensional parameter space. Assuming that the statistic $\Delta\chi^2$, calculated via PISCES's profiled likelihood technique, asymptotically follows a χ^2 distribution, we construct a 90% confidence level contour. For two fit parameters, the corresponding critical value is $\Delta\chi^2 = 4.61$ [23].

In the NOvA experiment, confidence intervals are derived using the profiled Feldman–Cousins method [95], which ensures correct statistical coverage even in regions where the assumptions of Wilks' theorem break down. However, this approach requires extensive Monte Carlo simulations, resulting in a substantially higher computational demand relative to the chi-square approximation. Hence, the Feldman-Cousins method will not be applied for this thesis result.

The resulting contours are shown in Figures 8.1 and 8.2. For comparison, we also include the sensitivity obtained using the Nus20 energy estimator, with cosmic background excluded in both cases. Sensitivity calculations are performed using 5% and 10% NC energy binning schemes. As shown in the plots, replacing the reconstructed NC energy with the true neutrino energy shifts the contours to the left in the high Δm^2_{41} region. This shift indicates an improvement in sensitivity,

Resolution	Bin edges
20%	0.5, 0.70, 0.99, 1.26, 1.54, 1.79, 2.15, 2.58, 3.10, 3.72, 4.46, 5.35, 6.42, 7.70, 9.24, 11.09, 13.31, 15.97, 20
10%	0.5, 0.70, 0.99, 1.26, 1.54, 1.73, 1.90, 2.09, 2.30, 2.53, 2.78, 3.06, 3.36, 3.70, 4.07, 4.48, 4.92, 5.42, 5.96, 6.55, 7.21, 7.93, 8.72, 9.60, 10.56, 11.61, 12.77, 14.05, 15.46, 17.00, 20
5%	0.5, 0.70, 0.99, 1.26, 1.54, 1.61, 1.69, 1.78, 1.87, 1.96, 2.06, 2.16, 2.27, 2.38, 2.50, 2.63, 2.76, 2.90, 3.04, 3.19, 3.35, 3.52, 3.70, 3.88, 4.07, 4.28, 4.49, 4.72, 4.95, 5.20, 5.46, 5.73, 6.02, 6.32, 6.64, 6.97, 7.32, 7.68, 8.07, 8.47, 8.89, 9.34, 9.81, 10.30, 10.81, 11.35, 11.92, 12.52, 13.14, 13.80, 14.49, 15.21, 15.97, 16.77, 17.61, 18.49, 20

Table 8.1: NC bin edges for different resolutions

allowing a larger region of parameter space to be excluded.

8.2 Analysis Result Using LSTM as NC Energy Estimator

After replacing the Nus20 energy estimator with the LSTM-based estimator, we performed the 2022 sterile neutrino analysis within the PISCES framework. In addition to the existing systematics, we added prong length systematic while the new LSTM NC energy estimator is used. However, preliminary results indicated that the fitter failed to converge, as PISCES was unable to complete the parameter fitting within 5000 epochs. The convergence issue was traced to substantial prediction errors of the LSTM energy estimator in the sub-1.5 GeV energy range (see Figure 7.11). This necessitated the adoption of a binning scheme modified for the estimator's resolution in this region. We made adjustment to the 5%, 10%, 20% binning scheme so that we have wider bins under 1.5 GeV (see Table 8.1).

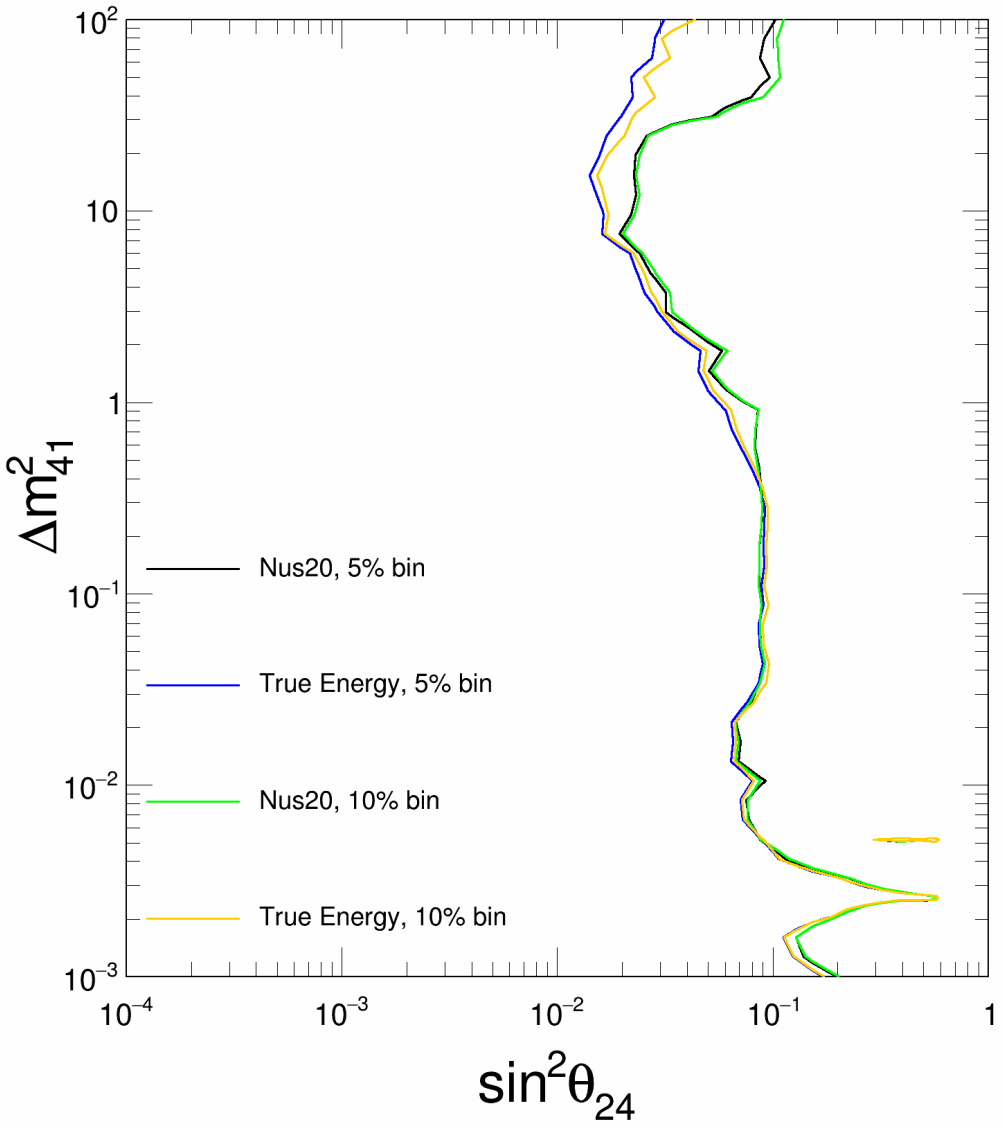


Figure 8.1: Sensitivity contour (90% confidence limits) in $\Delta m_{41} - \sin^2\theta_{24}$ space. For each contour, region to the right of the contour are excluded at 90% confidence level.

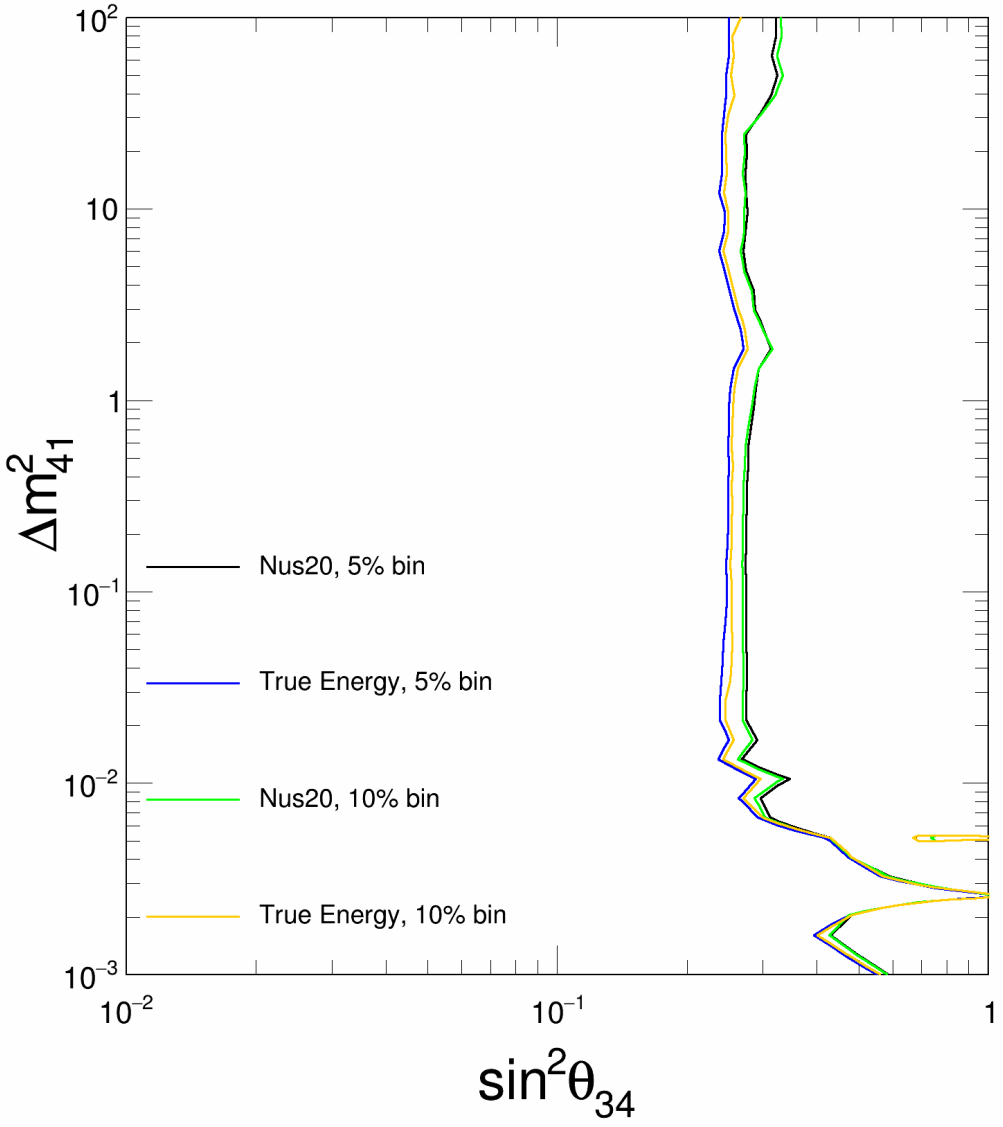


Figure 8.2: Sensitivity contour (90% confidence limits) in $\Delta m_{41} - \sin^2\theta_{34}$ space. For each contour, region to the right of the contour are excluded at 90% confidence level.

The resulting contours are shown in Figures 8.3 and 8.4. We also include the sensitivity obtained using the Nus20 energy estimator [1].

The new NC energy estimator results in ignorable changes in sensitivity of $\sin^2\theta_{24} - \Delta m^2_{41}$ and $\sin^2\theta_{34} - \Delta m^2_{41}$.

[1]The prong length systematic is not used as it doesn't impact Nus20 energy estimator.

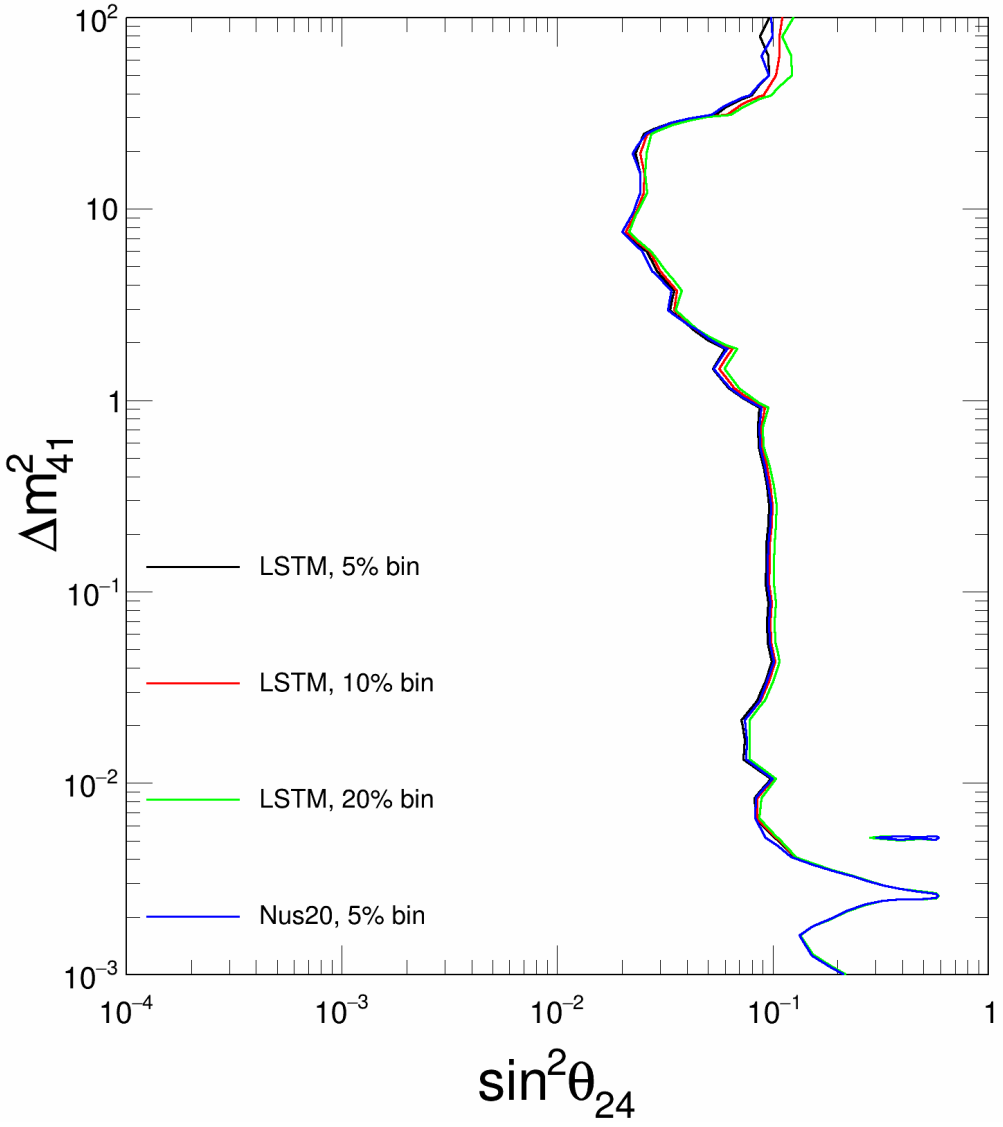


Figure 8.3: Sensitivity contour (90% confidence limits) in $\Delta m_{41} - \sin^2\theta_{24}$ space. For each contour, region to the right of the contour are excluded at 90% confidence level.

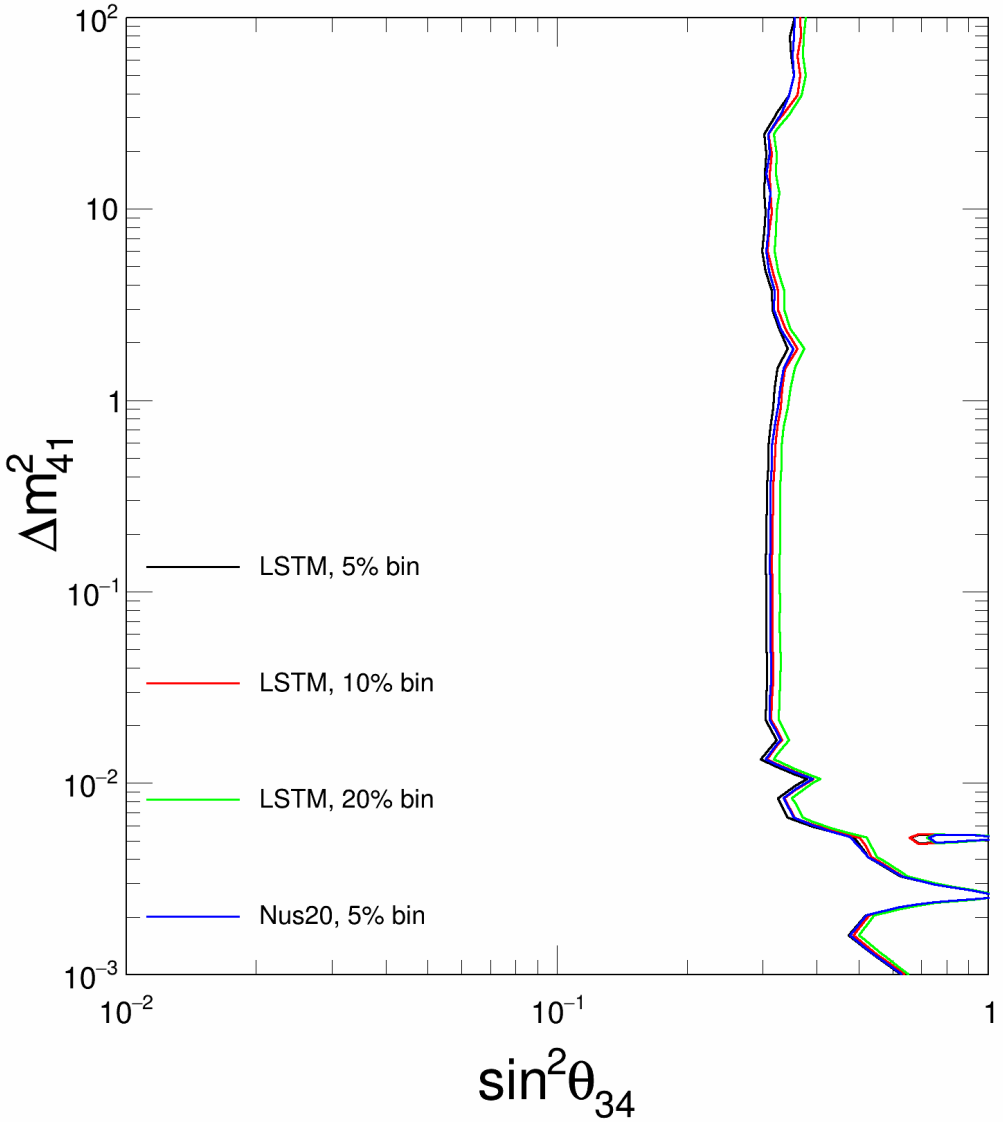


Figure 8.4: Sensitivity contour (90% confidence limits) in $\Delta m_{41} - \sin^2\theta_{34}$ space. For each contour, region to the right of the contour are excluded at 90% confidence level.

Chapter 9

Conclusion and Discussion

The results of the sterile neutrino analysis with the new LSTM energy estimator is presented. The current performance of LSTM models for NC events remains suboptimal, producing a sensitivity to sterile neutrino oscillations that is only marginally better than the results with the current visible energy estimator. As previously discussed, the absence of the outgoing neutrino energy in NC interactions poses a significant challenge for accurately reconstructing the incoming neutrino energy. We propose several potential strategies to improve the performance of the model:

- Mixed Energy Estimator: The LSTM exhibits poor performance in the low-energy regime. To address this, we propose a hybrid approach that combines the existing Nus20 energy estimator with the LSTM. Specifically, the traditional estimator would be used in the low-energy range, while the LSTM would be employed for higher energies. A critical challenge in this approach lies in determining which estimator to apply to a given event. One possible

solution is to develop a classifier that selects the appropriate estimator, or alternatively, computes a weighted average of the two estimators based on their reliability in different energy regimes.

- Final-State Specific Reconstruction: In certain final states—such as those involving a single outgoing proton—the neutrino energy can be more reliably reconstructed. We propose implementing a preliminary classification of events based on their final-state topology, followed by the application of tailored reconstruction strategies for each category. This targeted approach could lead to significant improvements in reconstruction accuracy for specific interaction channels.

- Alternative fitting framework: The current implementation of PISCES incurs substantial computational costs when using finer bin sizes. To mitigate this, we consider developing alternative fitting frameworks that can handle smaller bin sizes more efficiently. For example, a fitting framework implemented using GPU-based parallel computing.

- More MC: Our early experiments indicate a correlation between the performance of the LSTM model and the size of the dataset. Specifically, as the dataset grows, the performance of the LSTM model improves accordingly. Currently, the ND LSTM model underperforms compared to the FD LSTM model, primarily due to the limited size of the ND training dataset, which contains fewer than 100,000 events, in contrast to over 3,000,000 events in the FD dataset. In future work, we plan to enhance the model performance by generating larger MC datasets. Additionally, we may need to re-optimize

the event selection criteria for training, as the current criteria were originally designed for the Nus20 energy estimator.

- The Transformer model has demonstrated promising performance on the NOνA muon neutrino event data. In the future, a dedicated Transformer model can be trained on NC events to further improve the results in that domain.

References

[1] Heinz-Eberhard Mahnke. Lise Meitner, β-decay and non-radiative electromagnetic transitions. *Notes and Records: the Royal Society Journal of the History of Science*, 76(1):107–116, 2022.

[2] Wolfgang Pauli. Open letter to the group of radioactive people at the Gauverein meeting in Tübingen., 1930.

[3] Nota di Enrico Fermi. An attempt to a β rays theory. *IL NUOVO CIMENTO*, 1933.

[4] C. L. Cowan, F. Reines, F. B. Harrison, H. W. Kruse, and A. D. McGuire. Detection of the free neutrino: a confirmation. *Science*, 124(3212):103–104, 1956.

[5] G. Danby, J-M. Gaillard, K. Goulianos, L. M. Lederman, N. Mistry, M. Schwartz, and J. Steinberger. Observation of high-energy neutrino reactions and the existence of two kinds of neutrinos. *Phys. Rev. Lett.*, 9:36–44, 1962.

[6] K. Kodama, N. Ushida, C. Andreopoulos, N. Saoulidou, G. Tzanakos, P. Yager, B. Baller, D. Boehnlein, W. Freeman, B. Lundberg, J. Morfin, R. Rameika, J.C. Yun, J.S. Song, C.S. Yoon, S.H. Chung, P. Berghaus, M. Kubantsev, N.W. Reay, R. Sidwell, N. Stanton, S. Yoshida, S. Aoki, T. Hara, J.T. Rhee, D. Ciampa, C. Erickson, M. Graham, K. Heller, R. Rusack, R. Schwienhorst, J. Sielaff, J. Trammell, J. Wilcox, K. Hoshino, H. Jiko, M. Miyanishi, M. Komatsu, M. Nakamura, T. Nakano, K. Niwa, N. Nonaka, K. Okada, O. Sato, T. Akdogan, V. Paolone, C. Rosenfeld, A. Kulik, T. Kafka, W. Oliver, T. Patzak, and J. Schneps. Observation of tau neutrino interactions. *Physics Letters B*, 504(3):218–224, 2001.

[7] Carin Cain. Neutrino detectors for national security. `https://physics.aps.org/articles/v13/36`.

[8] Takaaki Kajita. Atmospheric neutrinos and discovery of neutrino oscillations. *Proceedings of the Japan Academy, Series B*, 86(4):303–321, 2010.

[9] M. Gell-Mann and A. Pais. Behavior of neutral particles under charge conjugation. *Phys. Rev.*, 97:1387–1389, 1955.

[10] B. Pontecorvo. Mesonium and anti-mesonium. *Sov. Phys. JETP*, 6:429, 1957.

[11] Bruno Pontecorvo. Neutrino experiments and the problem of conservation of leptonic charge. *Sov. Phys. JETP*, 26(984-988):165, 1968.

[12] V. Gribov and B. Pontecorvo. Neutrino astronomy and lepton charge. *Physics Letters B*, 28(7):493–496, 1969.

[13] Ziro Maki, Masami Nakagawa, and Shoichi Sakata. Remarks on the unified model of elementary particles. *Prog. Theor. Phys.*, 28:870–880, 1962.

[14] P. R. Demarque and J. R. Percy. A series of solar models. *American Astronomical Society*, 1964.

[15] John N. Bahcall. Solar neutrinos. I: Theoretical. *Phys. Rev. Lett.*, 12:300–302, 1964.

[16] Raymond Davis, Don S. Harmer, and Kenneth C. Hoffman. Search for neutrinos from the sun. *Phys. Rev. Lett.*, 20:1205–1209, 1968.

[17] K. Nakamura. Present status and future of Kamiokande. *Conf. Proc. C*, 890928:297, 1989.

[18] Christopher W. Walter. The super-kamiokande experiment. *arXiv preprint*, 0802.1041, 2008.

[19] Y. Fukuda, T. Hayakawa, E. Ichihara, K. Inoue, K. Ishihara, H. Ishino, Y. Itow, T. Kajita, J. Kameda, S. Kasuga, K. Kobayashi, Y. Kobayashi, Y. Koshio, M. Miura, M. Nakahata, S. Nakayama, A. Okada, K. Okumura, N. Sakurai, M. Shiozawa, Y. Suzuki, Y. Takeuchi, Y. Totsuka, S. Yamada, M. Earl, A. Habig, E. Kearns, M. D. Messier, K. Scholberg, J. L. Stone, L. R. Sulak, C. W. Walter, M. Goldhaber, T. Barszczxak, D. Casper, W. Gajewski, P. G. Halverson, J. Hsu, W. R. Kropp, L. R. Price, F. Reines, M. Smy, H. W. Sobel, M. R. Vagins, K. S. Ganezer, W. E. Keig, R. W. Ellsworth, S. Tasaka, J. W. Flanagan, A. Kibayashi, J. G. Learned, S. Matsuno, V. J.

Stenger, D. Takemori, T. Ishii, J. Kanzaki, T. Kobayashi, S. Mine, K. Nakamura, K. Nishikawa, Y. Oyama, A. Sakai, M. Sakuda, O. Sasaki, S. Echigo, M. Kohama, A. T. Suzuki, T. J. Haines, E. Blaufuss, B. K. Kim, R. Sanford, R. Svoboda, M. L. Chen, Z. Conner, J. A. Goodman, G. W. Sullivan, J. Hill, C. K. Jung, K. Martens, C. Mauger, C. McGrew, E. Sharkey, B. Viren, C. Yanagisawa, W. Doki, K. Miyano, H. Okazawa, C. Saji, M. Takahata, Y. Nagashima, M. Takita, T. Yamaguchi, M. Yoshida, S. B. Kim, M. Etoh, K. Fujita, A. Hasegawa, T. Hasegawa, S. Hatakeyama, T. Iwamoto, M. Koga, T. Maruyama, H. Ogawa, J. Shirai, A. Suzuki, F. Tsushima, M. Koshiba, M. Nemoto, K. Nishijima, T. Futagami, Y. Hayato, Y. Kanaya, K. Kaneyuki, Y. Watanabe, D. Kielczewska, R. A. Doyle, J. S. George, A. L. Stachyra, L. L. Wai, R. J. Wilkes, and K. K. Young. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81(8):1562–1567, 1998.

[20] Herbert H. Chen. Direct approach to resolve the solar-neutrino problem. *Phys. Rev. Lett.*, 55:1534–1536, 1985.

[21] Q Retal Ahmad, RC Allen, TC Andersen, JD Anglin, JC Barton, EW Beier, M Bercovitch, J Bigu, SD Biller, RA Black, et al. Direct evidence for neutrino flavor transformation from neutral-current interactions in the sudbury neutrino observatory. *Phys. Rev. Lett.*, 89(1):011301, 2002.

[22] Gareth Kafka. *A search for sterile neutrinos at the NOνA far detector*. PhD thesis, Harvard University, 2016.

[23] S. Navas et al. Review of particle physics. *Phys. Rev. D*, 110(3):030001, 2024.

[24] Steve Boyd. Lecture notes on neutrino oscillations. `https://warwick.ac.uk/fac/sci/physics/staff/academic/boyd/stuff/neutrinolectures/lec_oscillations.pdf`, 2020.

[25] E Kh Akhmedov. Neutrino physics. *arXiv preprint*, hep-ph/0001264, 2000.

[26] Stephen J Parke, Peter B Denton, and Hisakazu Minakata. Analytic neutrino oscillation probabilities in matter: Revisited. *arXiv preprint*, 1801.00752, 2018.

[27] Peter B Denton, Hisakazu Minakata, and Stephen J Parke. Compact perturbative expressions for neutrino oscillations in matter. *Journal of high energy physics*, 2016(6):1–23, 2016.

[28] D. Casper, R. Becker-Szendy, C. B. Bratton, D. R. Cady, R. Claus, S. T. Dye, W. Gajewski, M. Goldhaber, T. J. Haines, P. G. Halverson, T. W. Jones, D. Kielczewska, W. R. Kropp, J. G. Learned, J. M. LoSecco, C. McGrew, S. Matsuno, J. Matthews, M. S. Mudan, L. Price, F. Reines, J. Schultz, D. Sinclair, H. W. Sobel, J. L. Stone, L. R. Sulak, R. Svoboda, G. Thornton, and J. C. van der Velde. Measurement of atmospheric neutrino composition with the IMB-3 detector. *Phys. Rev. Lett.*, 66:2561–2564, 1991.

[29] A. Aguilar, L. B. Auerbach, R. L. Burman, D. O. Caldwell, E. D. Church, A. K. Cochran, J. B. Donahue, A. Fazely, G. T. Garvey, R. M. Gunasingha, R. Imlay, W. C. Louis, R. Majkic, A. Malik, W. Metcalf, G. B. Mills, V. Sandberg, D. Smith, I. Stancu, M. Sung, R. Tayloe, G. J. VanDalen, W. Vernon,

N. Wadia, D. H. White, and S. Yellin. Evidence for neutrino oscillations from the observation of $\overline{\nu}_e$ appearance in a $\overline{\nu}_\mu$ beam. *Phys. Rev. D*, 64:112007, 2001.

[30] A. A. Aguilar-Arevalo et al. A combined $\nu_\mu \to \nu_e$ and $\bar{\nu}_\mu \to \bar{\nu}_e$ oscillation analysis of the MiniBooNE excesses. *arXiv preprint*, 1207.4809, 2012.

[31] Mario A. Acero, Carlo Giunti, and Marco Laveder. Limits on ν_e and $\overline{\nu}_e$ disappearance from Gallium and reactor experiments. *Phys. Rev. D*, 78:073009, 2008.

[32] Adam Aurisano, Gavin S. Davies, and Brian Rebel. Executive summary: Sterile analysis - 2022. NOVA-doc-54138-v1, 2022.

[33] JJ Evans. The MINOS experiment: results and prospects. *Advances in High Energy Physics*, 2013(1):182537, 2013.

[34] Noel Stanton. Status of Fermilab E803 (COSMOS). *'95 electroweak interactions, unified theories. Proceedings, Leptonic Session of the 30th Rencontres de Moriond, Moriond Particle Physics Meetings*, pages 187–190, 1995.

[35] L. Fields, J. Chvojka, L. Aliaga, O. Altinok, B. Baldin, A. Baumbaugh, A. Bodek, D. Boehnlein, S. Boyd, R. Bradford, W. K. Brooks, H. Budd, A. Butkevich, D. A. Martinez Caicedo, C. M. Castromonte, M. E. Christy, H. Chung, M. Clark, H. da Motta, D. S. Damiani, I. Danko, M. Datta, M. Day, R. DeMaat, J. Devan, E. Draeger, S. A. Dytman, G. A. Díaz, B. Eberly, D. A. Edmondson, J. Felix, T. Fitzpatrick, G. A. Fiorentini, A. M. Gago, H. Gallagher, C. A. George, J. A. Gielata, C. Gingu, B. Gobbi,

R. Gran, N. Grossman, J. Hanson, D. A. Harris, J. Heaton, A. Higuera, I. J. Howley, K. Hurtado, M. Jerkins, T. Kafka, J. Kaisen, M. O. Kanter, C. E. Keppel, J. Kilmer, M. Kordosky, A. H. Krajeski, S. A. Kulagin, T. Le, H. Lee, A. G. Leister, G. Locke, G. Maggi, E. Maher, S. Manly, W. A. Mann, C. M. Marshall, K. S. McFarland, C. L. McGivern, A. M. McGowan, A. Mislivec, J. G. Morfín, J. Mousseau, D. Naples, J. K. Nelson, G. Niculescu, I. Niculescu, N. Ochoa, C. D. O'Connor, J. Olsen, B. Osmanov, J. Osta, J. L. Palomino, V. Paolone, J. Park, C. E. Patrick, G. N. Perdue, C. Peña, L. Rakotondravohitra, R. D. Ransome, H. Ray, L. Ren, P. A. Rodrigues, C. Rude, K. E. Sassin, H. Schellman, D. W. Schmitz, R. M. Schneider, E. C. Schulte, C. Simon, F. D. Snider, M. C. Snyder, J. T. Sobczyk, C. J. Solano Salinas, N. Tagg, W. Tan, B. G. Tice, G. Tzanakos, J. P. Velásquez, J. Walding, T. Walton, J. Wolcott, B. A. Wolthuis, N. Woodward, G. Zavala, H. B. Zeng, D. Zhang, L. Y. Zhu, and B. P. Ziemer. Measurement of muon antineutrino quasielastic scattering on a hydrocarbon target at $E_\nu \sim 3.5$ GeV. *Phys. Rev. Lett.*, 111:022501, 2013.

[36] Mitchell Soderberg. ArgoNeuT: A liquid argon time projection chamber test in the NuMI beamline. *arXiv preprint*, 0910.3433, 2009.

[37] D. S. Ayres et al. The NOvA technical design report. 2007.

[38] P. Adamson et al. The NuMI neutrino beam. *Nucl. Instrum. Meth. A*, 806:279–306, 2016.

[39] Robert Miles Zwaska. *Accelerator Systems and Instrumentation for the NuMI Neutrino Beam*. PhD thesis, University of Texas at Austin, 2005.

[40] The NuMI technical design handbook. `https://www-numi.fnal.gov/numwork/tdh/tdh_index.html`, 2002.

[41] Marco Del Tutto. Blessing package - beam simulation plots. NOVA-doc-13524-v6, 2015.

[42] Stuart Mufson. NOνA liquid scintillator production. In *Proceedings of FroST 2016 Workshop*, 2016.

[43] Christian Farnese. The ICARUS experiment. *Universe*, 5(2):49, 2019.

[44] Veera Mikola. *Improving the NOνA 3-flavour neutrino oscillation analysis*. PhD thesis, University College London, 2023.

[45] S Mufson, B Baugh, C Bower, TE Coan, J Cooper, L Corwin, JA Karty, P Mason, MD Messier, A Pla-Dalmau, et al. Liquid scintillator production for the NOνA experiment. *Nucl. Instrum. Meth. A*, 799:1–9, 2015.

[46] Gregory Pawlosky. The NOνA Experiment. *PoS*, NEUTEL2015:037, 2015.

[47] The FLUKA code: overview and new developments. *EPJ - Nuclear Sciences & Technologies*, 10, 2024.

[48] J. Allison, K. Amako, J. Apostolakis, P. Arce, M. Asai, T. Aso, E. Bagli, A. Bagulya, S. Banerjee, G. Barrand, B.R. Beck, A.G. Bogdanov, D. Brandt, J.M.C. Brown, H. Burkhardt, Ph. Canal, D. Cano-Ott, S. Chauvie, K. Cho,

G.A.P. Cirrone, G. Cooperman, M.A. Cortés-Giraldo, G. Cosmo, G. Cuttone, G. Depaola, L. Desorgher, X. Dong, A. Dotti, V.D. Elvira, G. Folger, Z. Francis, A. Galoyan, L. Garnier, M. Gayer, K.L. Genser, V.M. Grichine, S. Guatelli, P. Guèye, P. Gumplinger, A.S. Howard, I. Hřivnáčová, S. Hwang, S. Incerti, A. Ivanchenko, V.N. Ivanchenko, F.W. Jones, S.Y. Jun, P. Kaitaniemi, N. Karakatsanis, M. Karamitros, M. Kelsey, A. Kimura, T. Koi, H. Kurashige, A. Lechner, S.B. Lee, F. Longo, M. Maire, D. Mancusi, A. Mantero, E. Mendoza, B. Morgan, K. Murakami, T. Nikitina, L. Pandola, P. Paprocki, J. Perl, I. Petrović, M.G. Pia, W. Pokorski, J.M. Quesada, M. Raine, M.A. Reis, A. Ribon, A. Ristić Fira, F. Romano, G. Russo, G. Santin, T. Sasaki, D. Sawkey, J.I. Shin, I.I. Strakovsky, A. Taborda, S. Tanaka, B. Tomé, T. Toshito, H.N. Tran, P.R. Truscott, L. Urban, V. Uzhinsky, J.M. Verbeke, M. Verderi, B.L. Wendt, H. Wenzel, D.H. Wright, D.M. Wright, T. Yamashita, J. Yarba, and H. Yoshida. Recent developments in Geant4. *Nucl. Instrum. Meth. A*, 835:186–225, 2016.

[49] John Allison, Katsuya Amako, JEA Apostolakis, HAAH Araujo, P Arce Dubois, MAAM Asai, GABG Barrand, RACR Capra, SACS Chauvie, RACR Chytracek, et al. Geant4 developments and applications. *IEEE Transactions on nuclear science*, 53(1):270–278, 2006.

[50] S. Agostinelli, J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Arce, M. Asai, D. Axen, S. Banerjee, G. Barrand, F. Behner, L. Bellagamba, J. Boudreau, L. Broglia, A. Brunengo, H. Burkhardt, S. Chauvie, J. Chuma, R. Chytracek, G. Cooperman, G. Cosmo, P. Degtyarenko, A. Dell'Acqua,

G. Depaola, D. Dietrich, R. Enami, A. Feliciello, C. Ferguson, H. Fesefeldt, G. Folger, F. Foppiano, A. Forti, S. Garelli, S. Giani, R. Giannitrapani, D. Gibin, J.J. Gómez Cadenas, I. González, G. Gracia Abril, G. Greeniaus, W. Greiner, V. Grichine, A. Grossheim, S. Guatelli, P. Gumplinger, R. Hamatsu, K. Hashimoto, H. Hasui, A. Heikkinen, A. Howard, V. Ivanchenko, A. Johnson, F.W. Jones, J. Kallenbach, N. Kanaya, M. Kawabata, Y. Kawabata, M. Kawaguti, S. Kelner, P. Kent, A. Kimura, T. Kodama, R. Kokoulin, M. Kossov, H. Kurashige, E. Lamanna, T. Lampén, V. Lara, V. Lefebure, F. Lei, M. Liendl, W. Lockman, F. Longo, S. Magni, M. Maire, E. Medernach, K. Minamimoto, P. Mora de Freitas, Y. Morita, K. Murakami, M. Nagamatu, R. Nartallo, P. Nieminen, T. Nishimura, K. Ohtsubo, M. Okamura, S. O'Neale, Y. Oohata, K. Paech, J. Perl, A. Pfeiffer, M.G. Pia, F. Ranjard, A. Rybin, S. Sadilov, E. Di Salvo, G. Santin, T. Sasaki, N. Savvas, Y. Sawada, S. Scherer, S. Sei, V. Sirotenko, D. Smith, N. Starkov, H. Stoecker, J. Sulkimo, M. Takahata, S. Tanaka, E. Tcherniaev, E. Safai Tehrani, M. Tropeano, P. Truscott, H. Uno, L. Urban, P. Urban, M. Verderi, A. Walkden, W. Wander, H. Weber, J.P. Wellisch, T. Wenaus, D.C. Williams, D. Wright, T. Yamada, H. Yoshida, and D. Zschiesche. Geant4—a simulation toolkit. *Nucl. Instrum. Meth. A*, 506(3):250–303, 2003.

[51] Luke A. Corwin. Plan for Beam MC with FLUGG. NOVA-doc-3728-v1, 2009.

[52] C. Andreopoulos et al. The GENIE Neutrino Monte Carlo Generator. *Nucl.*

Instrum. Meth. A, 614:87–104, 2010, 0905.2517.

[53] Chris Hagmann, David Lange, Jerome Verbeke, and Doug Wright. Cosmic-ray shower library (CRY). Lawrence Livermore National Laboratory document UCRL-TM-229453, 2012.

[54] Matthew Strait. Geometry changes tech note. NOVA-doc-23132-v3, 2018.

[55] Adam Aurisano. The NOνA detector simulation. NOVA-doc-13577-v3, 2015.

[56] Adam Lister. Calibration: a one page summary. NOVA-doc-65198-v1, 2025.

[57] Tyler Alion. NOνA calibration: calorimetric energy scale in the NOνA detectors. IoP APP / HEPP Conference, 2018.

[58] D Méndez. Second analysis calorimetric energy scale calibration of the NOνA detectors. NOvA-doc-13579-v34, 2016.

[59] Evan Niner. Timing calibration technical note. NOVA-doc-12570-v2, 2015.

[60] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd*, volume 96, pages 226–231, 1996.

[61] Michael D Baird. A side by side comparison of Slicer, Cosmic Slicer, and Slicer4D. NOVA-doc-9195-v1, 2013.

[62] Dan Pershey, Junting Huang, and Matthew Judah. TDSlicer technote. NOVA-doc-27689-v5, 2020.

[63] R. C. Prim. Shortest connection networks and some generalizations. *The Bell System Technical Journal*, 36(6):1389–1401, 1957.

[64] Michael D. Baird. Global vertex reconstruction beginning with a modified hough transform. NOVA-doc-8241-v1, 2012.

[65] Mark Messier. Vertex reconstruction based on elastic arms. NOVA-doc-7530-v4, 2012.

[66] E Niner. Vertex clustering with possibilistic fuzzy-k means algorithm. NOVA-doc-7648-v2, 2012.

[67] R. Krishnapuram and J.M. Keller. A possibilistic approach to clustering. *IEEE Transactions on Fuzzy Systems*, 1(2):98–110, 1993.

[68] Michael Baird and Mark Messier. Break point fitter technical note. NOVA-doc-32455-v1, 2018.

[69] Biswaranjan Behera, Gavin Davies, and Fernanda Psihas. Event reconstruction in the NOνA experiment. *arXiv preprint*, 1710.03772, 2017.

[70] R. E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1):35–45, 1960.

[71] Susan M. Lein. ν_μ charged current energy estimators. NOVA-doc-15211-v1, 2016.

[72] Wei Mu. NuMu energy estimator technote for prod5 MC. NOVA-doc-43757-v1, 2020.

[73] D Torbunov. Numu energy estimator technote for Prod3 MC. NOVA-doc-23342-v2, 2017.

[74] Nicholas J Raddatz. ReMId technical note. NOVA-doc-11206-v1, 2014.

[75] Shiqi Yu, Fernanda Psihas, and Zelimir Djurcic. Energy reconstruction for 2017 ν_e appearance analysis. NOVA-doc-22502-v14, 2017.

[76] Harry Hausner and Brian Rebel. NC energy estimation. NOVA-doc-46448-v2, 2020.

[77] Michael Baird, Louise Suter, and Jeremy Wolcott. Summary of the 2020 FHC+RHC 3-flavor oscillation analysis. NOVA-doc-44422-v3, 2020.

[78] E. Catano-Mur, N. Nayak, A. Sutton, and K. Warburton. Event selection for the 2020 3-flavor analysis. NOVA-doc-44040-v2, 2020.

[79] Kirk Bays. NOνA cosmic rejection package and algorithms technical note. NOVA-doc-11205-v5, 2017.

[80] Michael Baird and Luke Vinton. Extrapolation technote for the numu third analysis. NOVA-doc-23390-v2, 2017.

[81] N Nayak. Flux systematics for the 2018 NOνA oscillation analyses. NOVA-doc-27884-v4, 2018.

[82] Matthew Strait. Muon energy scale systematic. NOVA-doc-20816-v26, 2023.

[83] Kirk Bays, Michael Dolce, Shane MacDonald, Mar´ıa Mart´ınez Casales, Aaron Mislivec, Greg Pawloski, and Jeremy Wolcott. Cross-section tuning 2020 tech note. NOVA-doc-43962-v3, 2020.

[84] V Hewes. Parameter Inference with Systematic Covariance and Exact Statistics. NOVA-doc-45977-v3, 2022.

[85] Warren S McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *Bulletin of mathematical biology*, 52:99–115, 1990.

[86] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. Learning representations by back-propagating errors. *Nature*, 323:533–536, 1986.

[87] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, 2012.

[88] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)*, pages 4171–4186, 2019.

[89] Geoffrey Hinton, Li Deng, Dong Yu, George E. Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick Nguyen, Tara N. Sainath, and Brian Kingsbury. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine*, 29(6):82–97, 2012.

[90] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

[91] Pierre Baldi, Jianming Bian, Lars Hertel, and Lingge Li. Improved energy reconstruction in NOνA with regression convolutional neural networks. *Phys. Rev. D*, 99(1):012011, 2019.

[92] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[93] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

[94] M. A. Acero, B. Acharya, P. Adamson, N. Anfimov, A. Antoshkin, E. Arrieta-Diaz, L. Asquith, A. Aurisano, A. Back, N. Balashov, P. Baldi, B. A. Bambah, E. F. Bannister, A. Barros, A. Bat, K. Bays, R. Bernstein, T. J. C. Bezerra, V. Bhatnagar, D. Bhattarai, B. Bhuyan, J. Bian, A. C. Booth, R. Bowles, B. Brahma, C. Bromberg, N. Buchanan, A. Butkevich, S. Calvez, T. J. Carroll, E. Catano-Mur, J. P. Cesar, A. Chatla, R. Chirco, B. C. Choudhary, A. Christensen, M. F. Cicala, T. E. Coan, A. Cooleybeck, C. Cortes-Parra, D. Coveyou, L. Cremonesi, G. S. Davies, P. F. Derwent, P. Ding, Z. Djurcic, K. Dobbs, M. Dolce, D. Doyle, D. Dueñas Tonguino, E. C. Dukes, A. Dye, R. Ehrlich, E. Ewart, P. Filip, M. J. Frank, H. R. Gallagher, F. Gao, A. Giri, R. A. Gomes, M. C. Goodman, M. Groh, R. Group, A. Habig,

F. Hakl, J. Hartnell, R. Hatcher, H. Hausner, M. He, K. Heller, V. Hewes, A. Himmel, T. Horoho, A. Ivanova, B. Jargowsky, J. Jarosz, M. Judah, I. Kakorin, A. Kalitkina, D. M. Kaplan, B. Kirezli-Ozdemir, J. Kleykamp, O. Klimov, L. W. Koerner, L. Kolupaeva, R. Kralik, A. Kumar, V. Kus, T. Lackey, K. Lang, J. Lesmeister, A. Lister, J. Liu, J. A. Lock, M. Lokajicek, M. MacMahon, S. Magill, W. A. Mann, M. T. Manoharan, M. Manrique Plata, M. L. Marshak, M. Martinez-Casales, V. Matveev, B. Mehta, M. D. Messier, H. Meyer, T. Miao, V. Mikola, W. H. Miller, S. Mishra, S. R. Mishra, A. Mislivec, R. Mohanta, A. Moren, A. Morozova, W. Mu, L. Mualem, M. Muether, D. Myers, D. Naples, A. Nath, S. Nelleri, J. K. Nelson, R. Nichol, E. Niner, A. Norman, A. Norrick, H. Oh, A. Olshevskiy, T. Olson, M. Ozkaynak, A. Pal, J. Paley, L. Panda, R. B. Patterson, G. Pawloski, R. Petti, R. K. Plunkett, L. R. Prais, M. Rabelhofer, A. Rafique, V. Raj, M. Rajaoalisoa, B. Ramson, B. Rebel, P. Roy, O. Samoylov, M. C. Sanchez, S. Sánchez Falero, P. Shanahan, P. Sharma, A. Sheshukov, A. Shmakov, Shivam, W. Shorrock, S. Shukla, D. K. Singha, I. Singh, P. Singh, V. Singh, E. Smith, J. Smolik, P. Snopok, N. Solomey, A. Sousa, K. Soustruznik, M. Strait, L. Suter, A. Sutton, K. Sutton, S. Swain, C. Sweeney, A. Sztuc, B. Tapia Oregui, N. Talukdar, P. Tas, T. Thakore, J. Thomas, E. Tiras, M. Titus, Y. Torun, D. Tran, J. Tripathi, J. Trokan-Tenorio, J. Urheim, P. Vahle, Z. Vallari, J. D. Villamil, K. J. Vockerodt, M. Wallbank, C. Weber, M. Wetstein, D. Whittington, D. A. Wickremasinghe, T. Wieber, J. Wolcott, M. Wrobel, S. Wu, W. Wu, W. Wu, Y. Xiao, B. Yaeggy, A. Yahaya, A. Yankelevich, K. Yonehara, S. Zadorozhnyy, J. Zalesak, and R. Zwaska. Dual-baseline search for

active-to-sterile neutrino oscillations in NOνA. *Phys. Rev. Lett.*, 134:081804, 2025.

[95] MA Acero, B Acharya, P Adamson, L Aliaga, N Anfimov, A Antoshkin, E Arrieta-Diaz, L Asquith, A Aurisano, A Back, et al. The profiled feldman-cousins technique for confidence interval construction in the presence of nuisance parameters. *arXiv preprint*, 2207.14353, 2022.

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

- **Cosmic-Ray Muon** (**CR** μ) – A muon coming from the abundant energetic particles originating outside of the Earth's atmosphere.

A.2 Acronyms

Table A.1: Acronyms

Acronym	Meaning
CC	Charged Current
NC	Neutral Current
MC	Monte Carlo
FD	Far Detector
ND	Near Detector
SSM	Standard Solar Model
T2K	"Tokai to Kamioka", a neutrino oscillation experiment
APD	Avalanche Photodiode
DAQ	Data Acquisition System
FEB	Front-End Board
TDU	Timing Distribution Unit
ADC	Analog to Digital Converter
TDC	Time to digital Converter
ASIC	Application-Specific Integrated Circuit
DCM	Data Concentrator Module
PISCES	Parameter Inference with Systematic Covariance and Exact Statistics