Observing muon neutrino to electron neutrino oscillations in the NOνA Experiment

by

Tian Xin

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY

Major: High Energy Physics

Program of Study Committee:
Mayly Sanchez, Major Professor
Kerry Whisnant
Amanda Weinstein
John Lajoie
Vivekananda Roy

Iowa State University
Ames, Iowa
2016

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my thanks to many people without whose help and support the thesis would not be possible. I want to first thank my advisor, Mayly Sanchez, for her support through my PhD researching and the writing of this thesis. She introduced neutrino physics to me 5 years ago. Neutrino was pretty new to me when I started to work in the group. She taught me all aspects of being a good physicist from detailed analysis skills to the appreciation of the big physics picture. Her guidance is invaluable not just in my PhD study but also in my future career. I am deeply grateful.

Thanks also to the other members of my graduate committee: Kerry Whisnant, Amanda Weinstein, John Lajoie and Vivekananda Roy not only for reviewing my thesis but also for the tutoring during my study in ISU.

I would like to thank all the colleagues in the NOνA collaboration. The ν_e appearance analysis would not exist without the hard work from all the members, especially the ones from the ν_e group. I want to first thank the group conveners: Mayly Sanchez and Patricia Vahle for their guidance. Thanks to the other members in the ν_e group, including Chris Backhouse, Evan Niner, Kanika Sachdev, Ji Liu, Erika Catano-Mur, Joe Lozier, Dan Pershey, Alex Radovic, Nitin Yadav and Hongyue Duyang. Thanks to Jianming Bian for the advice and discussion on cosmic rejection. In my work as APD expert, I received help from members in the operation and DCS group, including Rick Tesarek, Mat Muether, Leon Methuer, Peter Shanahan, Rob Putrick, Mark Messier, Keith and Phenanda. Also big thanks to Brian and Tom Coan for offering me the opportunity of working in the Calibration group and the mentoring during the work.

Thanks to my colleagues from ISU experimental neutrino group. First very big thanks to Gavin who has given so many advices and help on computing issues and physics questions since I entered the group. His help was particularly important to me when I initially started

in NOνA. Thanks to Jose, who is my colleague and roommate for the past 2 years during my stay at Fermilab. I really appreciate the discussions we constantly have on different interesting topics as well as the help he offered every time I was in some difficult time. Also thanks to Mat, Ioana for being such good postdocs to work with and all the good advices I received from them.

Finally I want to express my thanks to my family. Thanks to my parents for their selfless support and care not just during my PhD but also through my whole life. I hope I can spend more time with you after my graduation.

I would like to dedicate this thesis to the memory of my grandfather, Wenqing, with whom I grow up and spent most of my childhood. Thanks for teaching me to be curious and accurate. I appreciate the memory of assisting you craft cabinets or fix electrcal devices in the yard when I was a kid. I would not achieve this thesis without your great influence in my personality.

ABSTRACT

Neutrino oscillations offers an insight on new physics beyond the Standard Model. The three mixing angles (θ_{12}, θ_{13} and θ_{23}) and the two mass splittings (Δm^2_{12} and Δm^2_{23}) have been measured by different neutrino oscillation experiments. Some other parameters including the mass ordering of different neutrino mass eigenstates and the CP violation phase are still unknown.

NOνA is a long-baseline accelerator neutrino experiment, using neutrinos from the NuMI beam at Fermilab. The experiment is equipped with two functionally identical detectors about 810 kilometers apart and 14 mrad off the beam axis. In this configuration, the muon neutrinos from the NuMI beam reach the disappearance maximum in the far detector and a small fraction of that oscillates into electron neutrinos. The sensitivity to the mass ordering and CP violation phase determination is greately enhanced. This thesis presents the $\nu_e$appearance analysis using the neutrino data collected with the NOνA experiment between February 2014 and May 2015, which corresponds to 3.45 $\times 10^{20}$ protons-on-target (POT). The ν_e appearance analysis is performed by comparing the observed ν_e CC-like events to the estimated background at the far detector. The total background is predicted to be 0.95 events with 0.89 originated from beam events and 0.06 from cosmic ray events. The beam background is obtained by extrapolating near detector data through different oscillation channels, while the cosmic ray background is calculated based on out-of-time NuMI trigger data. A total of 6 electron neutrino candidates are observed in the end at the far detector which represents 3.3 σ excess over the predicted background. The NOνA result disfavors inverted mass hierarchy for $\delta_{cp} \in [0, 0.6\pi]$ at 90% C.L.

CHAPTER 1. INTRODUCTION

Neutrinos are so abundant in our universe and, yet, elusive to our understanding. These two facts come from one important feature: neutrinos very weakly interact with other forms of matter. Neutrinos keep puzzling us as well as challenging us since they were first postulated by Pauli in the 1930s and even more since neutrino oscillations were discovered in the 1980s, which provide a clue of new physics beyond the Standard Model of particle physics. This thesis presents the observation of muon neutrino to electron neutrino oscillation in the NOνA experiment using neutrinos from the NuMI beam at Fermilab, USA. The measurement is designed to address questions of θ_{13}, the neutrino mass hierarchy and CP violation phase.

Chapter 2 will discuss the background of neutrino physics, including the history of neutrino physics as well as the theory of the Standard Model and the physics beyond the Standard Model. The discussion will cover the big picture of how neutrinos fit into the Standard Model theory, how neutrinos interact with other particles through the weak interaction and the structure of electro-weak interaction theory. Then we will describe the phenomenology of neutrino mass and neutrino oscillation. In the end, the current status of neutrino oscillation science experiments will be summarized.

Chapter 3 will introduce the NOνA experiment, including the design and characteristics of NuMI neutrino beam, the technology, configuration and capabilities of the two NOνA detectors and, last but not the least, energy calibration of the detectors.

Chapter 4 will describe NOνA capability to observe electron neutrino appearance. The discussion starts with the theoretical derivation of the oscillation probability function for electron neutrino appearance measurement and the NOνA detector capability of detecting different types of interactions. The chapter will also introduce the analysis tools, including reconstruction and particle identification algorithms that have been developed. Chapter 5 will describe the data and simulation used in the first analysis.

Chapter 6 will discuss the event selection for the near and far detector. In this chapter, the cosmic ray background prediction by applying the FD event selection to out-of-time NuMI trigger data will also be detailed. The next chapter will describe the estimation of beam background by using near detector data in a ratio with MC to correct FD MC. Chapter 8 will enumerate the different types of systematic uncertainties of the FD background prediction.

In Chapter 9, we will see the result of the first electron neutrino appearance analysis of NOνA. Several side-bands are studied. The measurement agrees with the prediction within fluctuations in the side-band regions. The chapter will also discuss the Feldman-Cousin method and the conclusion implied by this result.

CHAPTER 2. NEUTRINO: HISTORY AND THEORY

This chapter will introduce the history and theories related to the neutrino oscillation physics. The status of neutrino oscillation experiments will be summarized at the end.

2.1 Neutrino History

The first hint of the existence of the neutrino came from the observation of the continuous spectrum of the outgoing anti-electrons in β decay. It was first observed in the Carbon-14 β decay by James Chadwick in 1914 [1]. In 1930 Wolfgang Pauli very innovatively postulated the existence of a light, neutral, half-spin particle to fulfill the energy loss in the β decay as well as conserve the spin in the reaction [2]. In 1932, Chadwick found proof of the existence of neutron, which is neutral and yet much heavier than the postulated neutrino [3]. Enrico Fermi suggested a model to describe the decay reaction and introduced the famous Fermi constant to label the strength of the interaction in 1933 [4]. Using the model, the β decay rate was calculated in the next year by Bethe and Peierls [5].

The first experimental observation of the neutrino particle was made by Frederick Reines and Clyde Cowan in 1956 through an "inverse" β decay reaction [6] (Eq.2.1):

$$\bar{\nu_e} + p \rightarrow e^+ + n \tag{2.1}$$

A nuclear reactor was used as a source of electron anti-neutrinos. The detector was placed a few meters away from the reactor, where the flux is about 1000 per cm^2 per second and was surrounded by thick layers of earth and metal to reduce non-neutrino particles from the reactor and most of the cosmic ray background. The interaction was identified by the few μs delayed coincidence: a 0.5 MeV photon from positron-electron annihilation followed by a more energetic photon emitted from neutron capture. The experiment observed more than 500 events where the specific coincidence occurred, among which 200 were expected to be background. It was a strong evidence of the existence of the neutrino and Reines and Cowan were awarded the Nobel Prize for this work.

The helicity of neutrinos was measured by Maurice Goldhaber and Lee Grodzins at Brookhaven National Laboratory in 1958 [7]. The experiment found that neutrinos tend to have left-helicity over right-helicity, which was at the time consistent with the prediction of V-A theory founded by Feynman and Gell-Mann [8], in which only only left-handed neutrino is allowed in order to preserve the masslessness of neutrino.

In the 1960s, progress in the theory of elementary particle physics was made by Sheldon Glashow [9], Abdus Salam [10], and Steven Weinberg [11]. They unified electromagnetic and weak interactions under the framework of the Standard Model. Based on this model, W and Z bosons were predicted as the mediators of weak charged current and neutral current interaction respectively. The W particle was first measured at CERN through the W decay: $W \to e + \bar{\nu}_e, \mu + \bar{\nu}_\mu$ [12]. The reactions were characterized as an electron or muon with missing energy due to the neutrinos. The measured mass of W boson was about 82 GeV/c^2. On the other hand, the weak neutral current interaction that was believed to be mediated by the Z boson was first observed by a bubble chamber experiment at CERN [13]. The cross section of the weak neutral current interaction was measured to be about one-fifth of the weak charged current interaction. The Z boson was first discovered at CERN in 1983 [14]. The measured mass of Z boson was 92 GeV/c^2. As Z boson decay was known better, in 1989 the Mark II experiment at SLC and the LEP experiments at CERN made precise measurement of resonance width of Z boson, which showed a strong evidence of the existance of only 3 light active neutrino flavors [15], [16].

In 1962, the AGS neutrino experiment at Brookhaven made the first detection of muon neutrino [17]. It was also the first time when a measurement using the neutrinos from an accelerator was performed, in which muon neutrinos are produced in the Alternating Gradient Synchrotron (AGS) and muons from the reaction $\nu_\mu + n \to p + \mu^-$ were observed as a long and narrow minimum ionizing track in the detector. The tau neutrino was first detected in 2000 in the DONUT experiment, more than 20 years after the tau was discovered [18], [19].

The neutrino oscillation evidence is a relatively new occurrance in neutrino physics. The first clear experimental evidence of the neutrino oscillation of muon neutrino to the other two flavors was published in 1998 by the SuperK experiment [20]. Since then, plenty of experiments have been designed and conducted to measure the oscillation parameters and our understanding of this phenomenon has improved significantly. But we are still facing some unresolved questions, for example the neutrino mass hierarchy, the ordering of different neutrino mass eigenstates, and CP violation phase, whether CP symmetry is violated in a leptonic weak interaction. NOνA was designed to look for answers to these questions. More details on the neutrino oscillation experiments are discussed in the later part of this chapter.

2.2 Standard Model

The Standard Model is a theory that describes electromagnetic, weak and strong interactions. In the Standard Model, there are three types of elementary particles: leptons, quarks and interaction mediators. There are six leptons that are characterized by their non-zero lepton numbers. The leptons are classified into three families based on their flavored lepton number and charges: $(e,\ \nu_e)$, $(\mu,\ \nu_\mu)$, $(\tau,\ \nu_\tau)$. In each family, we have a charged lepton and an associated neutrino. The leptons interact with other particles through gravitational, electromagnetic and weak interactions. For neutrinos, only left-handed component participate in the weak interactions with other particles. According to the Standard Model, for each leptonic flavor, we have a SU(2) doublet consisting of a left-handed lepton and a left-handed neutrino and a singlet with right-handed lepton, $(e_L,\ \nu_{eL})$ and e_R.

The Standard Model also suggests six quarks as elementary particles, which also fall into three families: $(u,\ d)$, $(c,\ s)$, $(t,\ b)$. The quarks carry non-zero baryon numbers and extra color numbers, and interact via all types of fundamental forces. Leptons and quarks are all fermions, which are defined to have half-integer spin. Each lepton or quark has their own anti-particle, which has opposite values for all the quantum numbers.

The forces, through which leptons and quarks interact, are mediated by different type of mediator bosons: γ, $W^{\pm}$, Z^0, g(gluon), H(Higgs). All mediators are bosons, the spin of which is an integer. Table.2.1 lists the characteristics of the four fundamental forces and the corresponding intermediators.

Table 2.1: The table shows the interaction range, relative strength and the mediators of the four fundamental forces.

Forces	Range (m)	Relative Strength	Mediator
Gravitational	∞	1	graviton (postulated)
Electromagnetic	∞	10^{40}	γ
Weak	10^{-15}	10^{29}	$W^{\pm}, Z^0$
Strong	10^{-31}	10^{43}	gluon

2.2.1 The GWS model and the Weak Interaction

The Standard Model theory is the combination of quantum electrodynamics, GWS electroweak theory and quantum chromodynamics, which is represented by the direct product of the groups $SU(3) \otimes SU(2) \otimes U(1)$. The Glashow-Weinberg-Salam (GWS) model describes the interactions in the electroweak $SU(2) \otimes U(1)$ sector, the unification of electromagnetic and weak interactions. $SU(2)$ belongs to the weak isospin carried by only left-handed leptons and $U(1)$ to the hypercharge, $Y = 2(Q - I_3)$, for all fermions. Neutrinos have only a left-handed component as predicted by the Standard Model as well as suggested by the experimental measurement. Eq.2.2 is the Lagrangian for the free Dirac field with the mass terms neglected, and presents only electron flavor as an example. The full equation requires an extension to all flavors, e, μ, τ. Eq.2.3 is the Lagrangian for gauge boson field. In order to preserve invariance of weak isospin and hypercharge, D_μ, $\vec{F_{\mu\nu}}$ and $G_{\mu\nu}$ are defined as Eq.2.4, in which $\vec{W_\mu}$ and B_μ are massless gauge boson fields, $\vec{T}$ and Y are weak isospin ($\vec{T} = \vec{\sigma}/2$) and hypercharge operators and g and $g\prime$ are gauge coupling constants in $SU(2)$ and $U(1)$.

$$L_{int} = (\bar{\nu}_{eL}(x), \bar{e}_L(x))\, i\gamma^\mu D_\mu \begin{pmatrix} \nu_{eL}(x) \\ e_L(x) \end{pmatrix} + \bar{e}_R(x) i\gamma^\mu D_\mu e_R(x) \tag{2.2}$$

$$L_{field} = -\frac{1}{4}\vec{F^{\mu\nu}}\vec{F_{\mu\nu}} - \frac{1}{4}G^{\mu\nu}G_{\mu\nu} \tag{2.3}$$

where

$$\begin{aligned}
D_\mu &= \partial_\mu - ig\vec{T}\cdot\vec{W_\mu} - ig'YB_\mu \\
\vec{F_{\mu\nu}} &= \partial_\mu\vec{W_\nu} - \partial_\nu\vec{W_\mu} + g(\vec{W_\mu}\times\vec{W_\nu}) \\
G_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu
\end{aligned} \tag{2.4}$$

Defining:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}\left(W_\mu^1 \mp W_\mu^2\right) \tag{2.5}$$

$$Z_\mu = \frac{1}{\sqrt{g^2 + g'^2}}\left(gW_\mu^3 - g'B_\mu\right) = \cos\theta_W W_\mu^3 - \sin\theta_W B_\mu \tag{2.6}$$

$$A_\mu = \frac{1}{\sqrt{g^2 + g'^2}}\left(g'W_\mu^3 + gB_\mu\right) = \sin\theta_W W_\mu^3 + \cos\theta_W B_\mu \tag{2.7}$$

where θ_W is Weinberg angle, calculated as Eq.2.8,

$$\sin\theta_W = \frac{g'}{\sqrt{g^2 + g'^2}}, \quad \cos\theta_W = \frac{g}{\sqrt{g^2 + g'^2}} \tag{2.8}$$

the interactional Lagrangian can be written as Eq.2.9.

$$L_{int} = eA_\mu J^\mu_{em} + gW_\mu^+ J^{\mu+} + gW_\mu^- J^{\mu-} + \frac{g}{\cos\theta_W}Z_\mu^0 J^\mu_{NC} \tag{2.9}$$

where

$$\begin{aligned}
J_{em} &= \bar{e_L}\gamma^\mu e_L + \bar{e_R}\gamma^\mu e_R = \bar{e}\gamma^\mu e \\
J^{\mu+} &= \frac{1}{\sqrt{2}}\bar{\nu_{eL}}\gamma^\mu e_L \\
J^{\mu-} &= \frac{1}{\sqrt{2}}\bar{e_L}\gamma^\mu \nu_{eL} \\
J^\mu_{NC} &= \frac{1}{2}\bar{\nu_{eL}}\gamma^\mu\nu_{eL} - \frac{1}{2}\bar{e_L}\gamma^\mu e_L + \sin\theta_W{}^2 J^\mu_{em}
\end{aligned} \tag{2.10}$$

The Lagrangian.2.9 consists of four parts. The first term represents the electromagnetic interaction with the coupling constant $e = gg'/\sqrt{g^2 + g'^2}$. The second and third terms correspond to the weak charged current interaction with coupling constant g. $J^{\mu\pm}$ are charged currents with one end the electron and the other the corresponding neutrino and couple to charged gauge boson $W_\mu^\pm$. The last term corresponds to the weak neutral current interaction. J_{NC}^μ is the neutral current with both ends being either lepton-antilepton or neutrino-antineutrino and is coupled to the neutral gauge boson Z_μ^0 in the weak interactions.

The gauge bosons are initially massless to guarantee gauge invariance. The process called spontaneous symmetry breaking is needed for the gauge bosons to gain mass. Eq.2.11 is the Lagrangian for a classical scalar field:

$$L_\phi = (D_\mu\phi)^\dagger(D^\mu\phi) - \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2 \tag{2.11}$$

For $SU(2)$ symmetry breaking, ϕ can be defined as a doublet of scalar fields as $\phi = (\phi^\dagger,\ \phi^0)^T$. The minimum point of the potential term, $\mu^2\Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2$, occurs at $|\Phi|^2 = v/\sqrt{2} = \sqrt{-\mu^2/2\lambda}$. To break the $SU(2)$ symmetry, we choose the ground state to be Eq.2.12, in which the upper charged component is chosen to be 0.

$$\phi_0 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v \end{pmatrix} \tag{2.12}$$

The perturbated field $\phi(x)$ can be expanded around the ground state as Eq.2.13

$$\phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \tag{2.13}$$

Based on Eq.2.4, the lagrangian (Eq.2.11) is expanded to be:

$$\begin{aligned} L_\phi =& \frac{1}{2}(\partial^\mu h)(\partial_\mu h) - \frac{1}{2}(2\lambda v^2)h^2 \\ &+ \frac{1}{2}(\frac{gv}{2})^2 W^{+\mu}W_\mu^{+\dagger} + \frac{1}{2}(\frac{gv}{2})^2 W^{+\mu}W_\mu^{+\dagger} + \frac{1}{2}(\frac{v\sqrt{g^2+g\prime^2}}{2})^2 Z^{0\mu}Z_\mu^0 + \cdots \end{aligned} \tag{2.14}$$

From Eq.2.14, the weak gauge bosons achieve masses through the interaction with the h field, higgs field, mass of which is $m_h = \sqrt{2\lambda v^2}$

$$\begin{aligned} m_W^2 &= \frac{g^2v^2}{4} \\ m_Z^2 &= \frac{(g^2+g\prime^2)v^2}{4} \end{aligned} \tag{2.15}$$

According to the Particle Data Book 2015, the mass of W and Z bosons are measured to be 80.39 GeV and 91.19 GeV respectively. The Fermi constant G_F, which is initially introduced by Fermi to measure the strength of weak interaction is:

$$G_F = \frac{e^2}{4\sqrt{2}m_W^2(1 - m_W^2/m_Z^2)} = 1.12 \times 10^{-5} GeV^{-2} \tag{2.16}$$

As discussed earlier, neutrinos interact with matter only through the weak interaction. There are two types of weak interactions, the charged current (CC) interaction in which a charged W boson is exchanged and the neutral current (NC) interaction in which a Z boson is exchanged. The calculation based on V-A theory shows that the weak CC interaction, $\nu_\mu + n \to \mu + p$, is about 10 times stronger than the NC counterpart, $\nu_\mu + n \to \nu_\mu + n$ [21],[22].

Based on the depth in which neutrinos interact with a targeted nucleus, both CC or NC can be further classified into a few types: quasi-elastic, resonance, deep inelastic and coherent interactions.

Elastic or Quasi-elastic: The neutrino scatters off the nucleon and ejects a lepton out of the target, $\nu_l + n \to l^- + p,\ \nu_l + p \to \nu_l + p$. This interaction dominates the region where neutrino energy is about 1 GeV.

Resonance The neutrino excites the nucleon to a resonance state which then decays to a nucleon and a pion: $\nu_l + p \to l^- + p + \pi^+,\ \nu_l + p \to \nu_l + p + \pi^0$.

Coherent The neutrino scatters coherently off the nucleus and produces an extra pion, $\nu_l + (A, Z) \to l^- + (A, Z) + \pi^+,\ \nu_l + (A, Z) \to \nu_l + (A, Z) + \pi^0$, in which A and Z are atomic and proton numbers of nucleus. The produced charged pions will decay into leptons and neutrinos, while neutral pions will decay into gammas.

Deep Inelastic The neutrino scatters deeply off the nucleon with an exchange of large amount of energy and produces large number of hadronic particles in the final states: $\nu_\mu + N \to \mu + X$, in which X represents a hadronic shower. This type of interaction dominates the high neutrino energy region.

NOνA measures neutrinos in the 1-3 GeV energy region, where a significant fraction of interactions are quasi-elastic and resonance interactions.

2.3 The Theory of Neutrino Oscillations

2.3.1 Neutrino Mass

According to the Standard Model, nature only has left-handed neutrinos and no right-handed ones. So neutrinos do not receive Dirac mass through Yukawa coupling of the fermion with the Higgs, which requires both left-handed and right-handed components as shown in Eq.2.17.

$$L_D = -m_D(\bar{\nu_L}\nu_R + \bar{\nu_R}\nu_L) \tag{2.17}$$

On the other hand, the results from neutrino oscillation meaurements suggest another picture in which active neutrinos that interact must have non-zero mass in order to make possible the mixing between the flavors. The neutrino mass has been studied within different experiments using different technologies. The best upper limit on neutrino mass is from beta decay measurement, in which the end point of the electron spectrum is very precisely measured in order to get the missing energy due to neutrino mass. The effective neutrino mass that is measured in the beta decay experiments is $\sqrt{\sum_i |U_{ei}|^2 m_{\nu_i}} < 2.0$ eV [23]. From cosmology measurements, the lower limit suggests that the sum of neutrino mass is not smaller than 0.2 eV [24]. All the experiments put a very small limits on the neutrino mass, which is much smaller than the other particles.

The “SeeSaw” mechanism was developed to explain not just neutrino mass but also the smallness of it. One of the most popular types of the seesaw model is to assume the existance of right-handed neutrino component and neutrinos achieving masses through the combination of Dirac and Majorana mass terms. Eq.2.18 implies that neutrinos have the properties of both Dirac, for which the particle is different from anti-particle, and Majorana, for which the particle is identical with the anti-particle.

$$L^{MR} = -\frac{1}{2}m_{MR}(\bar{\nu}_R\nu_R^c + \bar{\nu}_R^c\nu_R) \tag{2.18}$$

We combine Eq.2.17 and Eq.2.18 to get

$$L = -\frac{1}{2}\begin{pmatrix} \bar{\nu_L} & \bar{\nu_R}^c \end{pmatrix}\begin{pmatrix} 0 & m_D \\ m_D & m_{MR} \end{pmatrix}\begin{pmatrix} \nu_L^c \\ \nu_R \end{pmatrix} \tag{2.19}$$

Under the assumption of $m_{MR} \gg m_D$, the eigenvalues of the middle Dirac-Majorana mass matrix are

$$\begin{aligned} m_+ &= \frac{1}{2}m_{MR} + \frac{1}{2}\sqrt{{m_{MR}}^2 + 4m_D^2} \sim m_{MR} \\ m_- &= \frac{1}{2}m_{MR} - \frac{1}{2}\sqrt{{m_{MR}}^2 + 4m_D^2} \sim \frac{m_D^2}{m_{MR}} \end{aligned} \tag{2.20}$$

In a universe with $m_{MR} \sim 10^{15}$ GeV and $m_D \sim 100$ GeV, we have a light neutrino with mass around 10^{-2} eV, which is close to our current experimental limit on neutrino mass, as well as a very heavy neutrino with mass, $\sim 10^{15}$ GeV, which is too heavy for current technology to create or measure. This hypothetic heavy neutrino is intriguing and efforts are in progress that pose this super heavy neutral particle to explain profound questions, such as how the matter/antimatter asymmetry of the universe comes to be.

Because the seesaw model involves a Majorana mass term, support for the seesaw model would be the demonstration of the existance of Majorana particles, which is the main goal of neutrinoless double beta decay measurements. In ordinary double beta decay involving Dirac neutrinos, we expect to see two anti-neutrinos in the final product. If there are neutrinos that are Majorana particles, which are interchangable with their antiparticles, the two neutrinos annihilate and the outgoing electrons carry almost all the binding energy. The current experiments have pushed the sensitivity of the half-lifetime of neutrinoless double beta decay to more than 2×10^{25} yr [25]. So far no positive signal has been observed. On the other hand, under the assumption at neutrinos are Majorana particles, the rate of the neutrinoless double beta decay is calculated to be proportional to the effective mass, $|m_{\beta\beta}^2| = |\sqrt{\sum_i {U_{ei}}^2 \nu_i}|$, which can also put a limit on the absolute active neutrino mass.

2.3.2 Neutrino Oscillation

According to previous measurements [15], there are three active light neutrino flavors, ν_e, ν_μ, ν_τ. As neutrinos interact with other particles, they do so as flavor eigenstates. However, as shown in section 2.3.1, neutrinos can gain masses through a mechanism like the "See−Saw". Neutrinos also have mass eigenstates, ν_1, ν_2 and ν_3, as which they propagate. If the mass eigenstates are different from the flavor eigenstates, neutrinos can oscillate from one flavor to another after traveling the propagation distance.

2.3.3 Oscillation with Two Neutrino Flavors

In a world with only two neutrino flavors, the neutrino flavor and mass eigenstates mix with mixing matrix U as shown in Eq.2.21.

$$\begin{pmatrix} \nu_\alpha \\ \nu_\beta \end{pmatrix} = U \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix} \tag{2.21}$$

A neutrino starts its propagation initially with flavor α, which can be expressed as the linear combination of mass eigenstates:

$$|\Phi(0)\rangle = |\nu_\alpha\rangle = \cos\theta|\nu_1\rangle + \sin\theta|\nu_2\rangle \tag{2.22}$$

During the propagation, mass eigenstates, ν_1 and ν_2 evolve as:

$$|\nu_i(t)\rangle = |\nu_i(0)\rangle e^{i\bar{p}_i \cdot x - iE_i t} = |\nu_i(0)\rangle e^{-i\phi_i} \tag{2.23}$$

After a propagation distance L, the wave function becomes:

$$\begin{aligned} |\Phi(L)\rangle &= \cos\theta|\nu_1\rangle e^{-i\phi_1} + \sin\theta|\nu_2\rangle e^{-i\phi_2} \\ &= \left(\cos^2\theta e^{-i\phi_1} + \sin^2\theta e^{-i\phi_2}\right)|\nu_\alpha\rangle + (\cos\theta\sin\theta)\left(e^{-i\phi_2} - e^{-i\phi_1}\right)|\nu_\beta\rangle \end{aligned} \tag{2.24}$$

The probability of ν_α to ν_β oscillation is:

$$P(\nu_\alpha \to \nu_\beta) = |\langle\nu_\beta|\Phi(L)\rangle| = \sin^2(2\theta)\sin^2\left(\frac{\phi_1 - \phi_2}{2}\right) \tag{2.25}$$

The phase difference $(\phi_2 - \phi_1)$ can be derived under the assumption, $L \sim ct$ (in natural units, $L \sim t$), $p_1 = p_2 = p$ and $p \gg m_{1,2}$.

$$\begin{aligned}
\phi_2 - \phi_1 &= (E_2 t - p_2 L) - (E_1 t - p_1 L) \\
&= (E_2 - E_1)L - (p_2 - p_1)L \\
&= \left[p\left(1 + \frac{m_2^2}{p^2}\right)^{\frac{1}{2}} - p\left(1 + \frac{m_1^2}{p^2}\right)^{\frac{1}{2}} \right] L - (p_2 - p_1)L \\
&\sim \frac{m_2^2 - m_1^2}{2p} L \sim \frac{\Delta m_{12}^2}{E} L
\end{aligned} \tag{2.26}$$

So the probability function can be finally expressed as:

$$P(\nu_\alpha \to \nu_\beta) = \sin^2(2\theta) \sin^2\left(\frac{\Delta m_{12}^2}{E} L\right) \tag{2.27}$$

The probability is affected mainly by four factors, neutrino energy E, propagation length L, mass splitting Δm_{12}^2 and mixing angle θ. E and the mass splitting between the two mass eigenvalues determine how fast the oscillation is. The smaller E and the greater mass splitting is, the faster the oscillation occurs. E and L can be controlled when the experiment is designed, while the mass splitting is one of the properties of neutrinos. The mixing angle decides the amplitude of the oscillation.

2.3.4 Oscillation with Three Neutrino Flavors and PMNS Matrix

In the picture with three active neutrino flavors, the mixing matrix between flavor and mass eigenstates is a 3 × 3 PontecorvoMakiNakagawaSakata matrix (PMNS) matrix, which is a unitary matrix and can be parameterized in the following way [26]:

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \tag{2.28}$$

where

$$\begin{aligned}U &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}\end{aligned} \tag{2.29}$$

where $s_{ij} = \sin\theta_{ij}, c_{ij} = \cos\theta_{ij}$, θ_{12}, θ_{23} and θ_{13} are the mixing angles and δ is the phase factor, which carries the information of the CP violation. If neutrinos are Majorana particles, two extra phases factors are needed to fulfill the extra degrees of freedom. With the definition of the mixing matrix, the oscillation probability is written as:

$$\begin{aligned}P_{\nu_\alpha \to \nu_\beta}(E, L) =& \delta_{\alpha\beta} - 4\sum_{i>j} Re(U_{\alpha i}U^*_{\beta i}U^*_{\alpha j}U_{\beta j}) \sin^2\left(\frac{\Delta m^2_{ij}L}{4E}\right) \\ &+ 4\sum_{i>j} Im(U_{\alpha i}U^*_{\beta i}U^*_{\alpha j}U_{\beta j}) \sin\left(\frac{\Delta m^2_{ij}L}{4E}\right) \cos\left(\frac{\Delta m^2_{ij}L}{4E}\right)\end{aligned} \tag{2.30}$$

The θ_{12} and Δm^2_{12} are known as solar sector parameters, which were historically first measured using solar neutrinos. Because solar neutrino oscillation experiments tend to have large L/E, it is easy to see the oscillation with mass splitting Δm^2_{12} as small as $\sim 10^{-5}$ eV^2. Similarly, θ_{23} is known as the atmospheric mixing angle. It was first measured in atmospheric neutrino experiments, such as Super-Kamiokande, which has small L/E and is suitable for the measurement of oscillation with large mass splitting $\Delta m^2_{23} \sim 10^{-3} eV^2$. In the 13 sector, the mass splitting Δm^2_{13} is approximately equal to the atmospheric mass splitting, $\Delta m^2_{13} \sim \Delta m^2_{23} = \Delta m^2_{atm}$. The mixing angle θ_{13} is already measured to be non-zero. Since the mixing angle determines the amplitude of the oscillations, a relatively large θ_{13} makes further measurements such as CP violation phase and mass hierarchy possible for experiments like NOνA. Notice that due to the imaginary term in the PMNS matrix, a non-zero δ will lead to different probability values for $\nu_\alpha \to \nu_\beta$ and its CP conjugated process, $\bar{\nu}_\alpha \to \bar{\nu}_\beta$. So δ is used as a variable to probe the CP violation behavior in the leptonic sector and is usually called CP violation phase.

The phenomenon discussed above describes the neutrino oscillation probability in vacuum, while neutrinos can travel through matter, such as the core of the sun or the earth. As neutrinos propagate through matter, the electron neutrinos undergo a coherent charged current forward scattering with electrons from the matter, which causes a signifiant modification in the oscillation probability. This effect is known as the Mikheyev-Smirnov-Wolfenstein (MSW) effect [29, 28, 27]. Here the effect is derived assuming only two flavors, ν_e and ν_μ, which is good enough to first order to understand the effect on the probability of ν_μ to ν_e oscillation in experiments such as NOνA. The Hamiltonian for propagating ν_e and ν_μ in the flavor representation is:

$$\begin{aligned} H_M^\alpha &= U \begin{pmatrix} \frac{m_1^2}{2E} & 0 \\ 0 & \frac{m_2^2}{2E} \end{pmatrix} U^\dagger + \begin{pmatrix} V_{cc} & 0 \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} -\Delta m^2 \cos 2\theta + 4EV_{cc} & \Delta m^2 \sin 2\theta \\ \Delta m^2 \sin 2\theta & \Delta m^2 \cos 2\theta \end{pmatrix} \end{aligned} \tag{2.31}$$

In the first row, the first term represents the kinematic term of free neutrinos, while the second term is representing the change in potential due to the charged current interaction between the electron neutrinos and the electrons in matter.

$$V_{cc} = \pm\sqrt{2} G_F N_e \tag{2.32}$$

In Eq.2.32, the plus sign is for neutrino, while the minus is for anti-neutrinos, which implies that the matter effect can amplify the difference between matter and antimatter. G_F is the Fermi constant and N_e is the density of electrons in the medium. The effective mixing angle, θ_M and effective mass splitting, Δm_M^2 with the matter effect taken into account can be achieved by diagonalizing the H_M^i in the mass eigenstate representation, $U_\dagger H_M^\alpha U$:

$$\begin{aligned} \sin 2\theta_M &= \frac{\sin 2\theta}{\sqrt{\left(2EV_{cc}/\Delta m^2 - \cos 2\theta\right)^2 + \sin^2 2\theta}} \\ \Delta m_M^2 &= \Delta m^2 \sqrt{\left(2EV_{cc}/\Delta m^2 - \cos 2\theta\right)^2 + \sin^2 2\theta} \end{aligned} \tag{2.33}$$

If $\sin 2\theta_M \sim 1$, the maximum oscillation occurs no matter how small θ is.

2.4 Neutrino Oscillation Experiments

Since the observation of neutrino oscillation, various experiments have been built to measure the PMNS parameters as well as mass splitting information. There are different types of experiments with different baselines L and neutrino energies E designed to measure parameters in different oscillation sectors. The experiments are usually categorized based on the source of neutrinos: solar, atmospheric, reactor and accelerator. In this section, the status of neutrino oscillation measurements will be discussed.

2.4.1 The 12 sector measurement

There are mainly two types of experiments that contribute to the measurement of oscillation in the 12 sector: solar neutrino and reactor neutrino experiments.

The sun provides energy to the earth through the interior nuclear fusion [31]. Electron neutrinos are generated as a side product in these reactions. Fig.2.1 shows the predicted solar neutrino flux from different processes based on the standard solar model. The total solar neutrino flux is 6.5×10^{10} neutrinos per cm^2 per second [33]. The neutrinos from the primary p-p process, $p + p \to D + e^+ + \nu_e$, responsible for about 91% of the total neutrino flux, peak at around 0.3 MeV, while the ones from p-e-p process are mono-energetic at 1.44 MeV but much less intense than the p-p process. There are also some other main processes, like ${}^3He + p$, ${}^7Be$ and ${}^8B$. The shaded areas in Fig.2.1 show the threshold for different experiments to detect neutrinos.

As discussed above, the solar neutrino energy is at the MeV level and the distance of propagation of electron neutrinos from the sun to earth is 1×10^8 km, which makes the solar neutrino experiments have the sensitivity to oscillation with Δm^2 as small as about 10^{-9} eV^2. The neutrino oscillation in the 12 sector with relatively small mass splitting was first clearly observed in the Homestake Solar Neutrino Experiment in the Homestake Gold Mine in South Dakota [30]. The experiment observed only one-third of the prediction made by the standard solar model. The result was then supported by the observation from other solar experiments, such as GALLEX[34], SAGE[35] and GNO[36]. The definite confirmation of solar neutrino

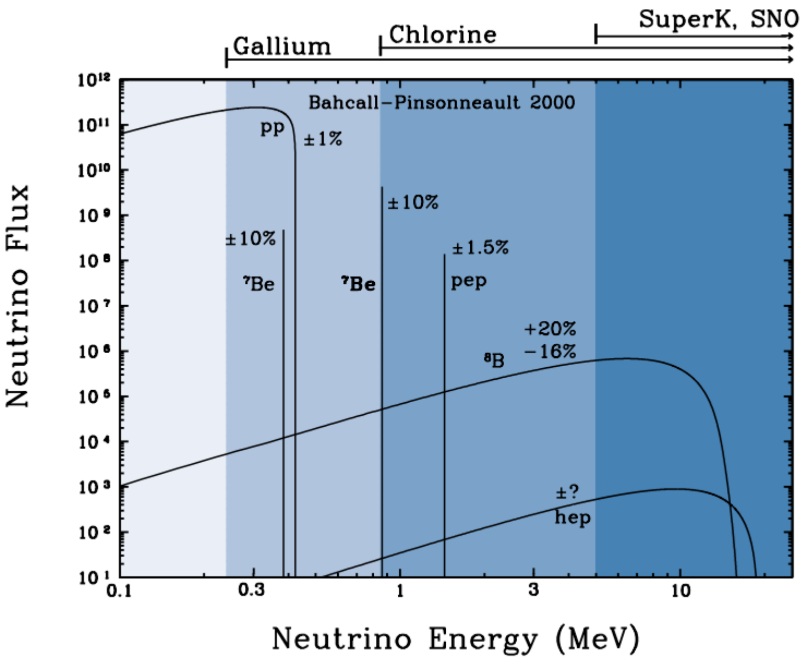


Figure 2.1: Solar neutrino flux predicted by the standard solar model [33].

deficit with respect to the standard solar model came from Sudbury Neutrino Observatory (SNO) [37]. The design of the SNO experiment enabled it to check the ν_e and $\nu_\mu + \nu_\tau$ flux in three different channels, CC($\nu_e + n \rightarrow e + p$), NC($\nu_\alpha + n(p) \rightarrow \nu_\alpha + n(p)$) and ES($\nu_\alpha + e \rightarrow \nu_\alpha + e$). Fig.2.2 shows SNO's latest result, the flux of $\nu_\mu + \nu_\tau$ versus ν_e in the three channels [38]. The intercept of the bands suggests the non-ν_e flux, $\phi_{\nu_\mu + \nu_\tau}$, to be $3.26 \pm 0.25^{+0.40}_{-0.35} \times 10^6$ cm^{-2}s^{-1}, The non-zero non-ν_e active neutrino flux is a strong evidence that solar neutrinos oscillate from the electron flavor to the other two flavors.

The precise measurement of the solar sector parameters was completed by the reactor neutrino experiment, KamLAND [40]. Kamland is a one-kiloton liquid scintillator neutrino detector observing $\bar{\nu_e} + p \rightarrow n + e^+$ with detection threshold 1.8 MeV. By precisely measuring the electron antineutrino disappearance, it has the sensitivity to determine θ_{12} and Δm^2_{12}.

$$P_{\bar{\nu_e} \rightarrow \bar{\nu_e}} \sim 1 - \sin^2 2\theta_{12} \sin^2(1.27 \Delta m^2_{12} L/E) \tag{2.34}$$

Fig.2.3 shows the latest KamLAND collaboration oscillation parameter result that was published in 2013. The result was a combined three-flavor analysis of solar and KamLAND data with θ_{13} constrained by accelerator and short-baseline reactor neutrino oscillation experiment

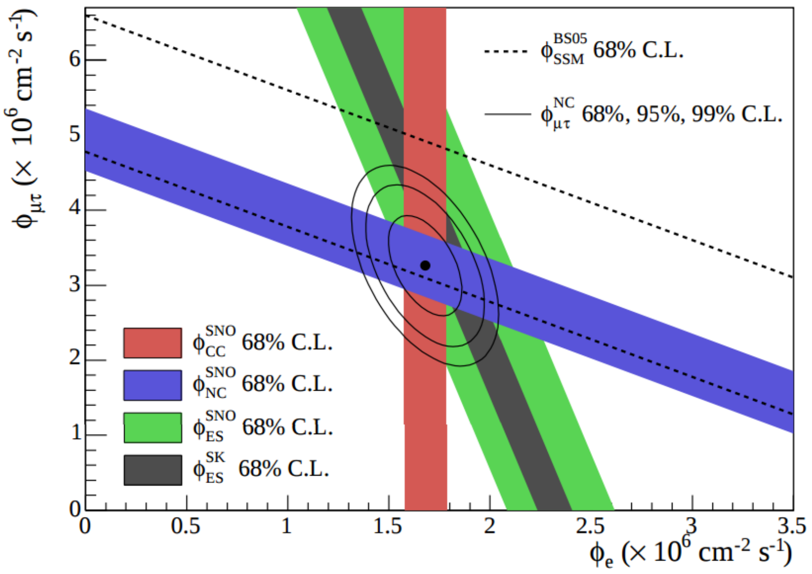


Figure 2.2: The result of $\phi_{\nu_\mu+\nu_\tau}$ versus ϕ_{ν_e} measurement by SNO [38]. The flux of $\nu_\mu + \nu_\tau$ is measured to be $3.26 \pm 0.25^{+0.40}_{-0.35} 10^6 cm^{-2} s^{-1}$.

result. The combined result puts a narrow constraint on both θ_{12} ($tan^2\theta_{12} = 0.436^{+0.029}_{-0.025}$) and Δm^2_{12} ($\Delta m^2_{12} = 7.53^{+0.18}_{-0.18} \times 10^{-5} eV^2$) [155].

2.4.2 The 23 sector measurement

The oscillation in the 23 sector was first seen in the atmospheric neutrino experiments. The protons and gammas from cosmic rays interact with protons from atmospheric molecules. The interaction produces pions, which further decay into leptons and neutrinos. The main channel is:

$$\begin{aligned} \pi^{\pm} &\rightarrow \mu^{\pm} + \nu_\mu(\nu_\mu) \\ &\rightarrow e^{\pm} + \bar{\nu_\mu}(\bar{\nu_\mu}) + \nu_e(\nu_e) + \nu_\mu(\nu_\mu) \end{aligned} \tag{2.35}$$

Based on this model, the number of muon (anti)neutrino measured on earth should be around twice of the electron (anti)neutrinos. Yet in reality, as people measured the ratio of $(\nu_\mu + \bar{\nu_\mu})/(\nu_e + \bar{\nu_e})$, a deficit was seen by experiments like Super-Kamiokande [20], IMB [41] and Soudan [42]. SuperK is a 50-kiloton water cherenkov multi-functional detector designed mainly to measure proton decay and to observe solar, atmospheric and supernova neutrinos. Fig.2.4 shows the zenith angle distributions of μ-like and e-like neutrinos updated by SuperK with

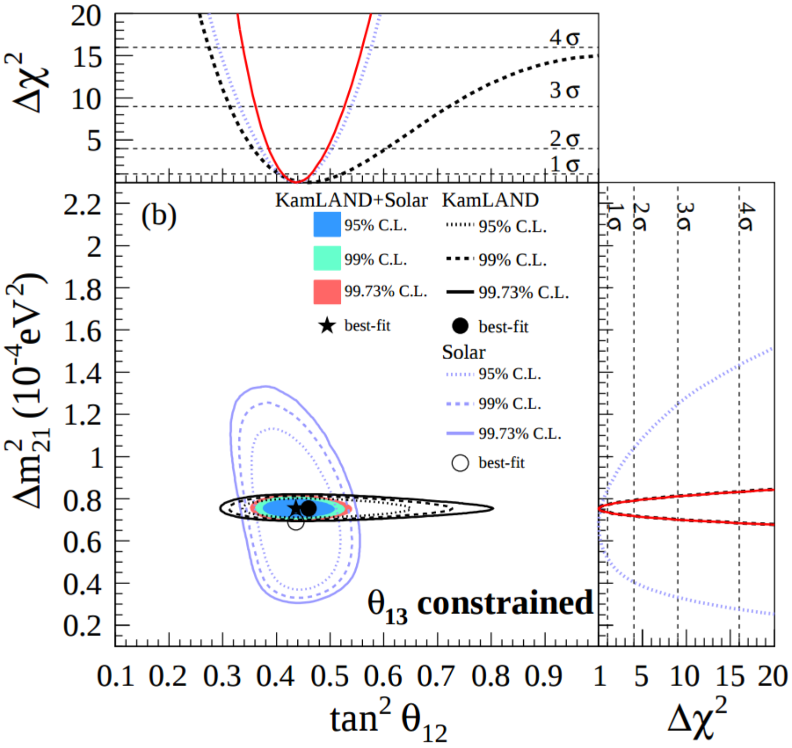


Figure 2.3: Allowed region for oscillation parameters, θ_{12} and Δm^2_{12} from Kamland and solar experiments. The combined result shows $tan^2\theta_{12} = 0.47^{+0.06}_{-0.05}$ and $\Delta m^2_{12} = 7.59^{+0.21}_{-0.21} \times 10^{-5} eV^2$ at 4σ.

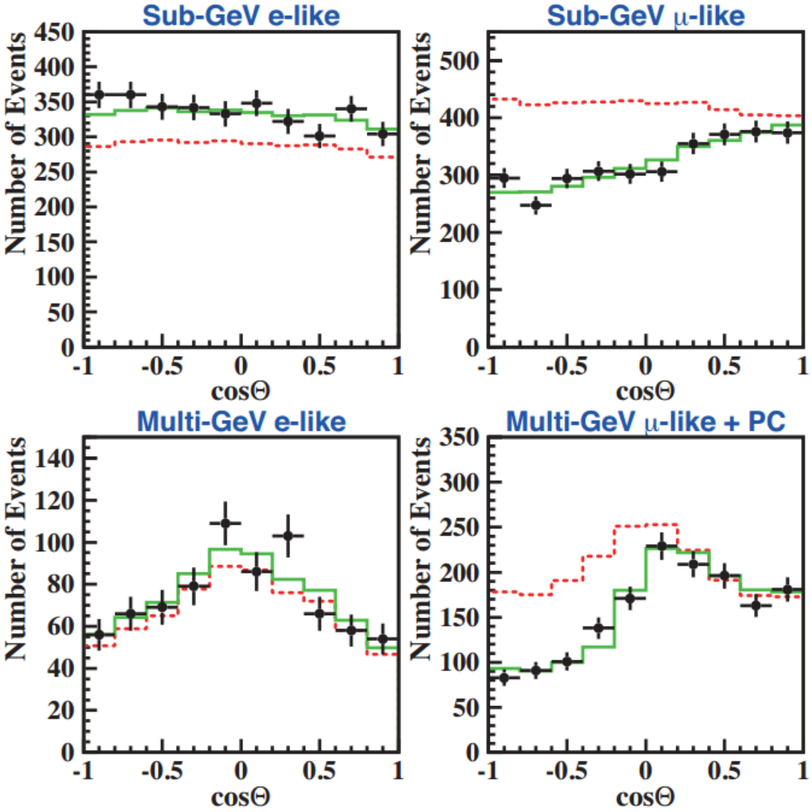


Figure 2.4: The zenith angle distribution of e-like and μ-like events with visible energy $<$ 1.33 eV and $>$ 1.33 GeV from SuperK experiment [156]. Dashed red line is the non-oscillated prediction, while the green line is oscillated scenario. A clear deficit in upward μ-like events can be seen.

higher statistics [156]. An obvious deficit with respect to the non-oscillated scenario can be seen for the μ-like neutrinos that come from the direction of $\cos\theta < 0$ (upward), the direction from which neutrinos travel through the earth. On the other hand, for e-like neutrinos, the measured flux is consistent with the non-oscillated prediction over the whole zenith angle region. The observation support the hypothesized conversion, $\nu_\mu \rightarrow \nu_\tau$ oscillation presented in Eq.2.36.

$$P_{\nu_\mu \rightarrow \nu_\mu} = 1 - \sin^2 2\theta_{23} \sin^2(1.27\Delta m^2_{23} L/E) \tag{2.36}$$

The result of SuperK was supported by long baseline accelerator neutrino experiments, such as K2K (KEK-to-Kamioka) and MINOS (Main Injector Neutrino Oscillation Search). Such experiments receive GeV neutrinos from an accelerator neutrino beam and search for dis-

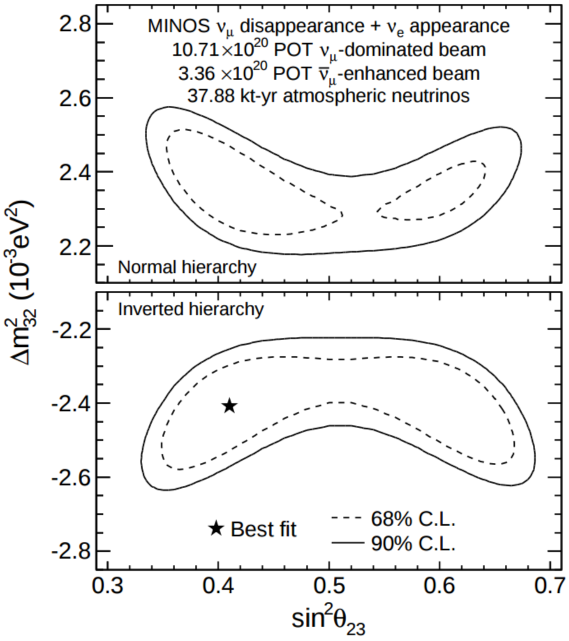


Figure 2.5: The MINOS result of θ_{23} and Δm^2_{23} result in a combined analysis of muon neutrino appearance and electron neutrino appearance using the complete set of atmospheric and accelerator neutrino data at 68% (dash) and 90% (solid) CLs. Star represents the best fit.

appearance of muon (anti)neutrino at a distance of hundreds of kilometers. Fig.2.5 shows the result of θ_{23} and Δm^2_{23} based on a combined analysis of muon neutrino appearance and electron neutrino appearance using the complete set of atmospheric and accelerator neutrino data [157]. The result shows $|\Delta m^2_{23}| = (2.28 - 2.46) \times 10^{-3} eV^2 (68\% CL), \sin^2(\theta^2_{23}) = 0.35 - 0.65 (90\% CL)$ for normal mass hierarchy and $|\Delta m^2_{23}| = (2.32 - 2.53) \times 10^{-3} eV^2 (68\% CL), \sin^2(\theta^2_{23}) = 0.34 - 0.67 (90\% CL)$ for inverted mass hierarchy. One thing to note is that because the mass splitting term is presented as proportional to $\sin^2(\Delta m^2 L/E)$ in the muon neutrino disappearance probability function, determining the sign of the splitting would be impossible for this method. The subject is known as neutrino mass hierarchy (Fig.2.6). When Δm^2 is positive, it is called normal mass hierarchy, where ν_3 is the heaviest mass eigenstate. In the inverted mass hierarchy case, ν_3 is the lightest neutrino and the sign of Δm^2 is negative. Knowing the mass hierarchy does not just provide an insight to the neutrino mass mechanism but also helps to guide the next generation of neutrinoless double beta measurements. For example, as shown in Fig.2.7, if the inverted mass hierarchy is true, not seeing any positive signal of neutrinoless double beta decay in an experiment with resolution at 0.01 eV means that the hypothesis of a

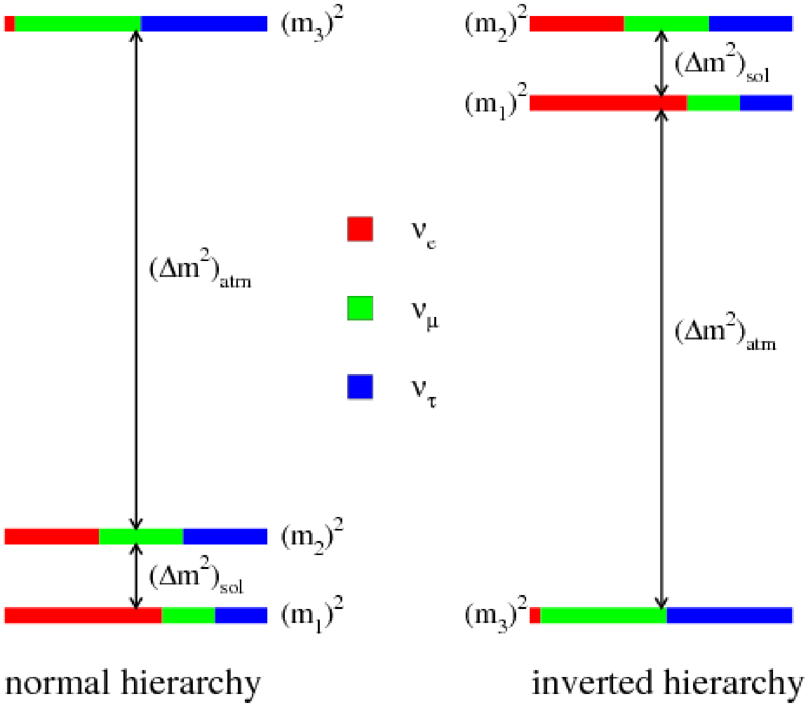


Figure 2.6: Neutrino mass hierarchy.

majorana neutrino is rejected at 90% confidence level [49]. Another unresolved question in the 23 sector is whether the mixing angle θ_{23} is above, below or equal to 45°, which is known as maximum mixing. According to the latest MINOS result shown in Fig.2.5, the best fit favors the scenario of $\theta_{23} < 0.5$ under the inverted mass hierarchy. NOνA also has the sensitivity to answer the θ_{23} question in a combined analysis of electron neutrino appearance and muon neutrino disappearance channels.

2.4.3 The 13 sector measurement

In the 13 sector, $\Delta m^2_{13} \sim \Delta m^2_{23} = \Delta m^2_{atm}$. Yet also in this case, the sign of Δm^2_{13} is unknown. The mixing angle θ_{13} was tested by reactor neutrino experiments CHOOZ in 1999 [45]. A precise measurement on the mixing angle was recently made by the Daya Bay Nuclear Neutrino experiment [46]. The Daya Bay Neutrino experiment uses electron anti-neutrinos from the reactors in Daya Bay. The setting of the experiment makes it sensitive to θ_{13} measurement through the electron antineutrino disappearance channel. Fig.2.8 shows the

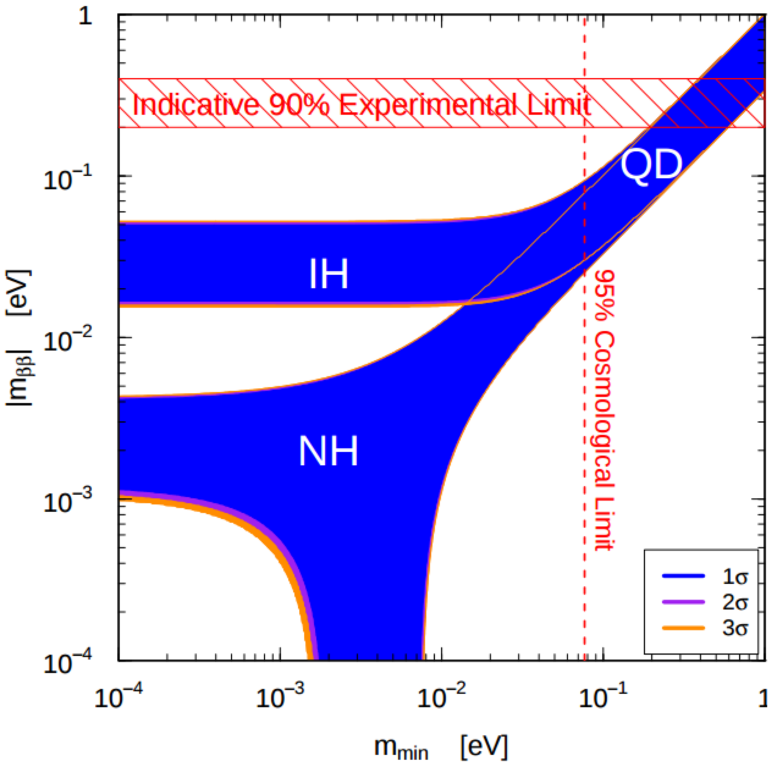


Figure 2.7: Effective majorana neutrino mass vs the mass of the lightest neutrino for normal hierarchy and inverted hierarchy [49].

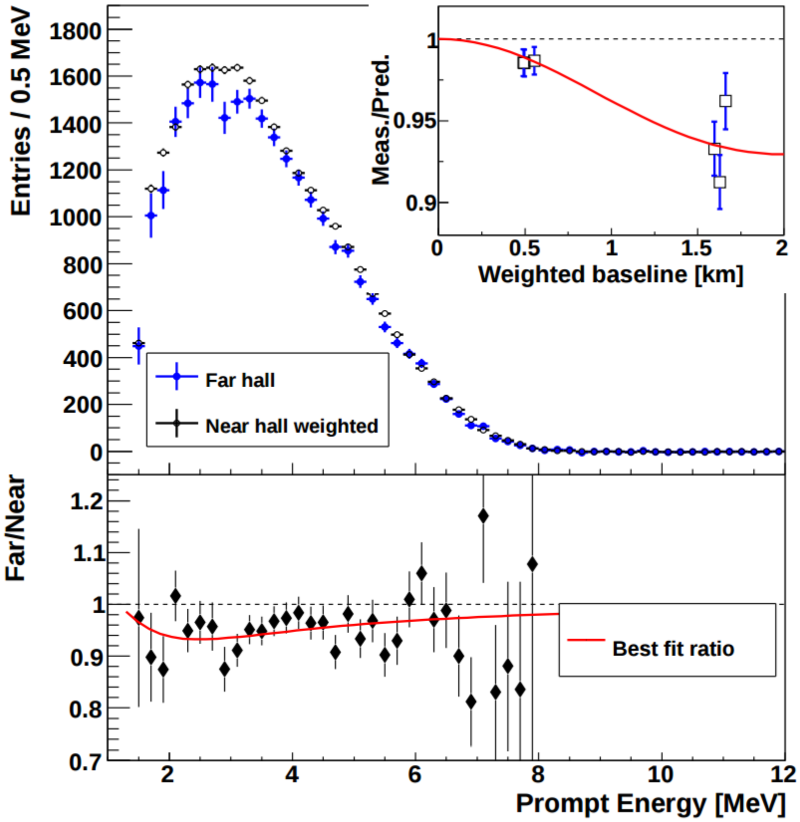


Figure 2.8: The updated esult of θ_{13} measurement by Daya Bay experiment in 2014. It shows the comparison of the measured event in the far hall (blue) and the prediction made using near hall data (Black circle). The fit (red) on the far to near ratio distribution suggests $\sin^2 2\theta_{13} = 0.089 \pm 0.008$.

updated θ_{13} result that the Daya Bay collaboration published in 2014 [158]. The graph shows the comparison in the number of measured electron antineutrinos to the prediction made using near hall data with zero θ_{13} assumption. The disappearance of the electron antineutrino is clear and according to the inset, the phenomenon of electron anti-neutrino disappearance becomes greater as the baseline is longer. One of the advantages that Daya Bay is equiped with six identical detectors, which enable relative measurements between near and far at two different baselines [159]. The measurement suggests $\sin^2 2\theta_{13} = 0.089 \pm 0.008$ and strongly proves a non-zero θ_{13}.

The long baseline accelerator neutrino experiments also proved to have the capability of measuring θ_{13} through $\nu_\mu \to \nu_e$ channel. T2K will be discussed briefly here and the electron neutrino appearance observation in NOνA is the main focus of this thesis. T2K is an off-

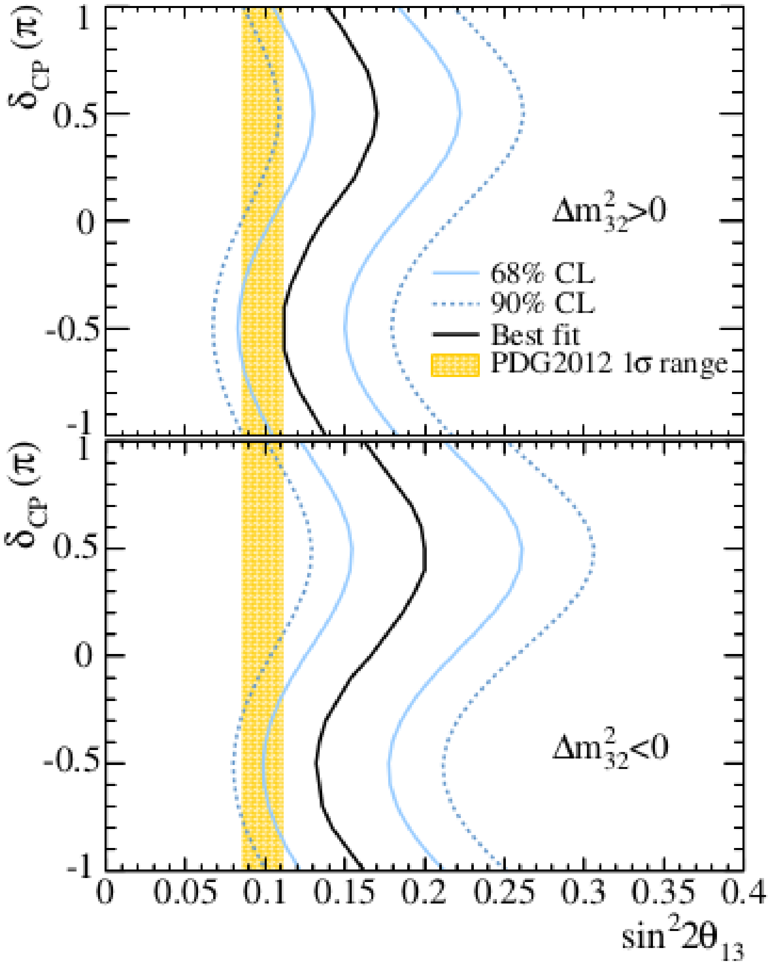


Figure 2.9: The best fit (black), 90% (solid blue) and 68% (dash blue) confidence level contour for δ_{cp} vs $\sin^2 2\theta_{13}$ from T2K experiment published in 2014.

axis accelerator neutrino experiment with neutrino energy peak at 0.6 GeV and baseline equal to 295 km. After about 3 years of data collecting, T2K accumulated 6.6×10^{20} POT muon neutrino data. The experiment observed 28 electron-like events over 4.92 ± 0.55 background events and proved a non-zero θ_{13} with 7.3σ. Fig.2.9 shows the best fit, 90% and 68% confidence level contour for δ_{cp} vs $\sin^2 2\theta_{13}$ [47]. From the picture, if we assume $|\Delta m^2_{23}| = 2.4 \times 10^{-3} eV^2$, maximum mixing for θ_{23} and normal mass hierarchy, $\Delta m^2_{23} > 0$ and zero CP violation phase, the best fit value of $\sin^2 2\theta_{13} = 0.140^{+0.045}_{-0.037}$ is obtained.

Another important parameter in the 13 sector is the CP violation phase. The CP violation phase has a profound influence in other area of physics because it provides a clue on the matter/antimatter asymmetry in the current universe. As the CP violation found in the baryonic sector is not big enough to explain today's asymmetry in the universe, theorists are building models to explain the lepton asymmetry first, which can then be converted to baryon asymmetry. Currently, the knowledge of δ_{cp} is limited by the resolution of our detection technologies. So far the best attempt to examine it was made by T2K [48]. At 90% confidence level, the region excluded by T2K is: (0.15,0.83)π for normal hierarchy and (-0.08,1.09)π for inverted hierarchy. So the normal hierarchy is slightly favored by T2K over inverted hierarchy. The NOνA experiment is designed to have the capability to examine the CP violation more precisely.

2.4.4 $\nu_\mu \to \nu_e$ Masurement in the NOνA Experiment

One of the main goal of NOνA is to observe electron (anti)neutrino appearance, through which NOνA has sensitivity to most of the remaining questions in neutrino oscillations. Equation.2.37

$$\begin{aligned} P(\nu_\mu(\bar{\nu_\mu}) \to \nu_e(\bar{\nu_e})) = \; & \sin^2 2\theta_{13} \sin^2 \theta_{23} \frac{\sin^2(A-1)\Delta}{(A-1)^2} \\ & \mp 2\alpha \sin\theta_{13} \sin\delta_{cp} \sin 2\theta_{12} \sin 2\theta_{23} \frac{\sin A\Delta}{A} \frac{\sin(A-1)\Delta}{A-1} \sin\Delta \\ & + 2\alpha \sin\theta_{13} \cos\delta_{cp} \sin 2\theta_{12} \sin 2\theta_{23} \frac{\sin A\Delta}{A} \frac{\sin(A-1)\Delta}{A-1} \cos\Delta \end{aligned} \tag{2.37}$$

shows the $\nu_\mu(\bar{\nu}_\mu)$ to $\nu_e(\bar{\nu_e})$ oscillation probability for 3 flavors with matter effect included. In the equation,

$$\alpha = \Delta m^2_{21}/\Delta m^2_{31} \tag{2.38}$$

$$\Delta = \frac{\pi}{2hc} * \frac{\Delta m^2_{31} * L}{E} = 1.27 * \frac{\Delta m^2_{13}/[eV^2] * L/[km]}{E/[GeV]} \tag{2.39}$$

$$A = \pm G_f N_e L/\sqrt{2}\Delta = \pm 7.56 \times 10^{-5} * \frac{\rho/[g/cm^{-3}] * E/[GeV]}{\Delta m^2_{13}/[eV^2]} \tag{2.40}$$

in which ρ is the density of the crust, ~ 3 g/cm^{-3}, and Δm^2_{13} is the mass splitting between ν_1 and ν_3, $m^2_{\nu_3} - m^2_{\nu_1}$, which is positive for normal mass hierarchy and negative for inverted mass hierarchy. The sign in the front of A is positive when it is in the neutrino mode and negative for antineutrino mode. The first term has $\sin^2 2\theta_{13}$ and $\sin^2\theta_{23}$, which means that the electron neutrino appearance channel has the sensitivity to the θ_{13} measurement and the θ_{23} octant question. Although the influence of the mass hierarchy and the CP violation is mixed in the oscillation probability, matter effect provides some power to decouple the two effects through A. The longer neutrinos travel, the greater the matter effect is and the greater the influence of the two effects is separated. For neutrinos at 2 GeV, assuming $\sin\delta_{cp} = 1$, the oscillation probability is enhanced by the matter effect by 13% for NH and suppressed by $\sim 50\%$ for IH scenario compared to the case without matter effect. In the anti-neutrino mode, the matter effect contributes in the opposite way that the probability is suppressed by 18% for NH and is enhanced by $\sim 100\%$ for IH scenario. The sensitivity to CP violation comes from $\sin\delta_{cp}$ and $\cos\delta_{cp}$ in the α terms. Therefore δ_{cp} can be tested by comparing the probabilities of $\nu_\mu \to \nu_e$ and $\bar{\nu}_\mu \to \bar{\nu}_e$.

Using Equation.2.37, Fig.2.10 shows the contour of probability of $\nu_\mu \to \nu_e$ vs probability of $\bar{\nu_\mu} \to \bar{\nu_e}$ under the assumption of $\theta_{23} = 45°$ and $\sin^2 2\theta_{13} = 0.095$ after 3 years of neutrino and 3 years of antineutrino running. The red ellipse represents different CP violation cases in the inverted mass hierarchy senario and the blue one is for normal mass hierarchy. The plot shows that the two different mass ordering senarios are decoupled by the matter effect and enable us to determine both mass hierarchy and CP violation phase. For example, if both $P(\nu_e)$ and $P(\bar{\nu_e})$ are measured to be at black star, all the inverted hierarchy cases will be rejected at 2σ confidence level. Fig.2.11 and Fig.2.12 show more details in the sensitivity to either of the two problems. Fig.2.11 shows that, at maximum θ_{23}, for about one-third of the possible δ_{cp} values, NOνA has higher than 2σ sensitivity to determining the mass hierarchy. Fig.2.12 suggests that, for about one-tenth of the possible δ_{cp} values, NOνA has higher than 1.5σ sensitivity to the CP violation measurement.

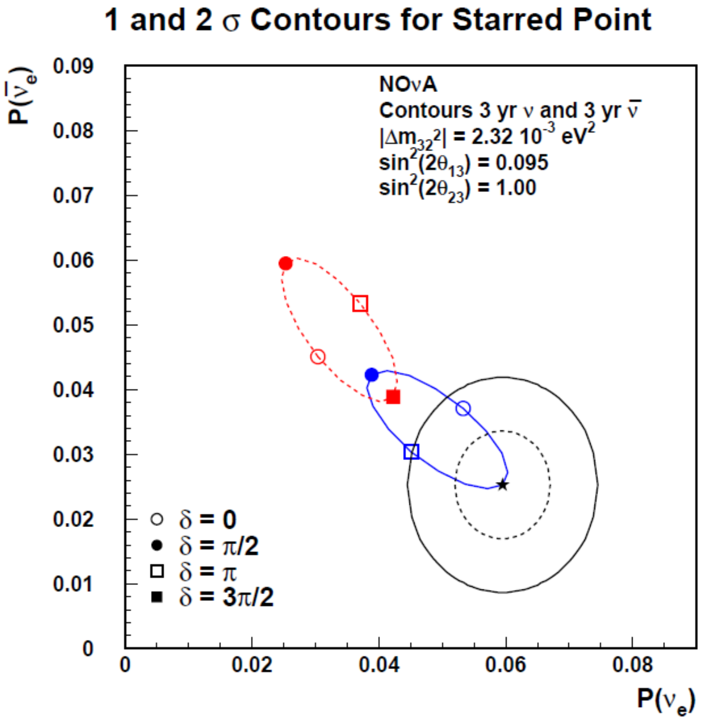


Figure 2.10: $P(\nu_\mu \to \nu_e)$ vs $P(\bar{\nu_\mu} \to \bar{\nu_e})$ assuming $\theta_{23} = 45^\circ$, $\sin^2 2\theta_{13} = 0.095$ after 3yr ν + 3yr $\bar{\nu}$ run. Red ellipse represents inverted mass hierarchy, while blue is for normal mass hierarchy. Points along the circles are corresponding to different CP violation angles.

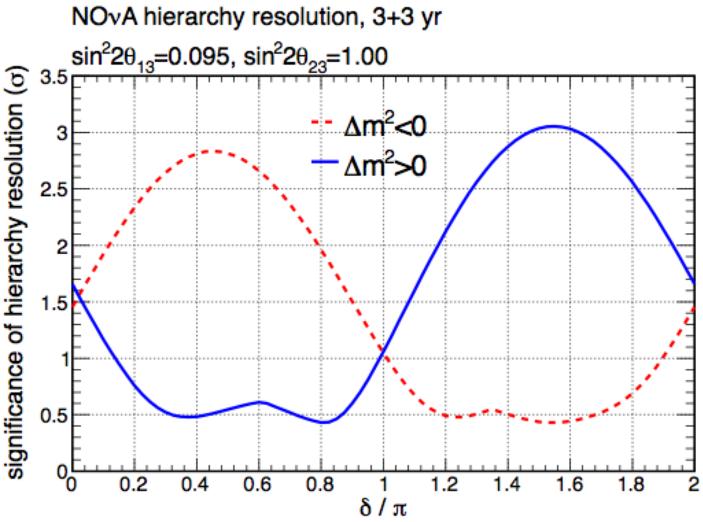


Figure 2.11: Significance of mass hierarchy resolution as a function of delta for the two hierarchy cases, normal (blue) and inverted (red), after 3yr ν + 3yr $\bar{\nu}$ run. At maximum θ_{23}, for about one-third of the possible δ_{cp} values, NOνA has higher than 2σ sensitivity to the mass hierarchy.

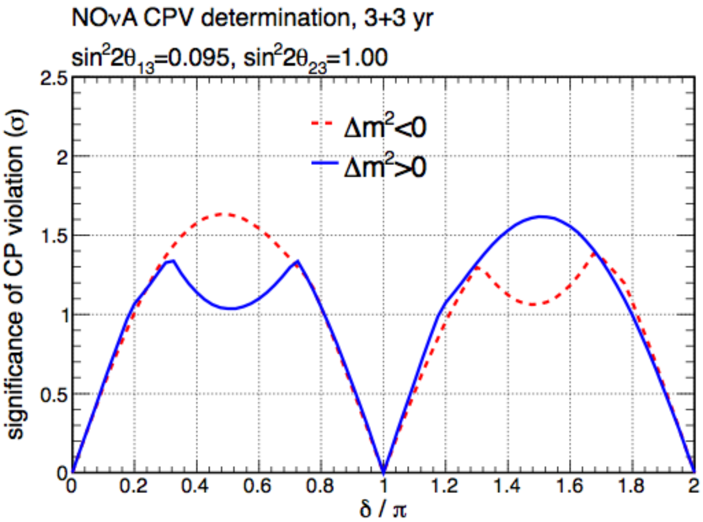


Figure 2.12: Significance of CP violation angle resolution as a function of δ_{cp} for the two hierarchy cases, normal (blue) and inverted (red). At maximum θ_{23}, for about one-tenth of the possible δ_{cp} values, NOνA has higher than 1.5σ sensitivity to the CP violation measurement.

2.4.5 Summary

A global fit including all neutrino oscillation data up to 2014 has been reported [50]. The global fit result (Table.2.2) shows that, especially for the unsolved questions, the picture of θ_{23} below maximum mixing is favored, however, with low significance. For the mass hierarchy question, the global fit does not show a strong preference to either hierarchies. Independent on the mass hierarchies of neutrinos, $\delta \sim (1.12, 1.60)\pi$ is favored at 1σ and $\delta \sim (0.20, 0.70)\pi$ is disfavored at 2σ.

Table 2.2: Global fit result for 3 ν oscillation in terms of the best fit, allowed 1σ and 3σ range. Δm^2 is defined as $m_3^2 - (m_1^2 + m_2^2)/2$, so $\Delta m^2 > 0$ is for NH and $\Delta m^2 < 0$ for IH.

Parameters	Best fit	1σ range	3σ range
$\Delta m_{12}^2/10^{-5} eV^2$	7.54	7.32 - 7.80	6.99 - 8.18
$\sin^2\theta_{12}/10^{-1}$	3.08	2.91 - 3.25	2.59 - 3.59
$\Delta m^2/10^{-3} eV^2$ (NH)	2.43	2.37 - 2.49	2.23 - 2.61
$\Delta m^2/10^{-3} eV^2$ (IH)	2.38	2.32 - 2.44	2.19 - 2.56
$\sin^2\theta_{13}/10^{-2}$ (NH)	2.34	2.15 - 2.54	1.76 - 2.95
$\sin^2\theta_{13}/10^{-2}$ (IH)	2.40	2.18 - 2.59	1.78 - 2.98
$\sin^2\theta_{23}/10^{-1}$ (NH)	4.37	4.14 - 4.70	3.74 - 6.26
$\sin^2\theta_{23}/10^{-1}$ (IH)	4.55	4.24 - 5.94	3.80 - 6.40
δ/π (NH)	1.39	1.12 - 1.77	**
δ/π (IH)	1.31	0.98 - 1.60	**

CHAPTER 3. THE NOνA EXPERIMENT

NOνA is a long baseline accelerator neutrino experiment, using muon neutrinos from NuMI beam at Fermilab. The NOνA experiment has two functionally identical detectors located as shown in Fig.3.1. The Far Detector is at Ash River, MN, about 810 km away from the neutrino source. The Near Detector is about 1 km in front of the source and is used to understand the initial state of the neutrinos from the beam. The two detectors are both at 14 mrad off the NuMI beam axis where the muon neutrinos peak at 2 GeV, the energy of which most of the muon neutrinos disappearance maximum in the far detector and a small fraction of them oscillate into electron neutrinos as shown in Fig.3.2.

Figure 3.1: The location of NOνA near and far detectors.

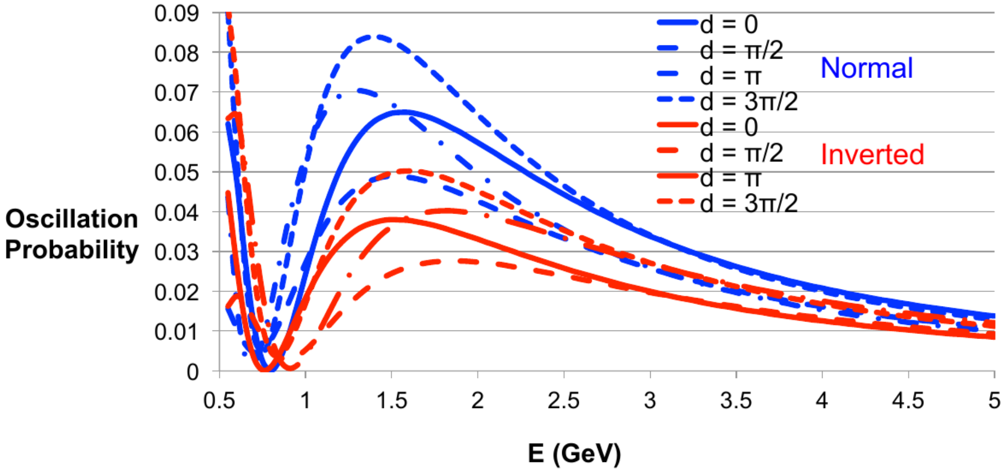


Figure 3.2: The ν_μ to ν_e oscillation probability for 4 possible δ_{cp} values in both normal and inverted mass hierarchy pictures under the assumption of $sin^2(2\theta_{23}) = 1$, $sin^2(2\theta_{13}) = 0.1$.

3.1 NuMI Beam

NuMI (Neutrinos at the Main Injector) is a neutrino beamline at Fermilab, IL USA. It was first designed to provide neutrinos for the MINOS experiment, which was also a long baseline neutrino oscillation experiment and performed a significant measurement of Δm^2_{23}. The beam line is also used by the MINERVA and NOνA experiments.

The neutrino-generating process starts from the Linear acceleractor, in which ~750 keV hydrogen ions are accelerated to 400 MeV [51]. The 400 MeV H^- ions are then sent to the Booster, which is a sychrotron facility and in which electrons are extracted from H^- and then the remaining proton is accelerated to 8 GeV. The protons are then batched, 4×10^{12} protons per batch, and transferred to the Main Injector. In the Main Injector, the protons are accelerated to an even higher energy, 120 GeV. The protons are then delivered to the NuMI beam line to produce neutrinos. The accelerator complex at Fermilab is undergoing an upgrade to produce protons with the intensity of 350 kW to 700 kW. The upgrade comes mainly from two aspects: higher efficiency in proton acceleration in the main injector and more protons per batch. Firstly, the frequency of the main injector is increased from 2.2 sec per circle to 1.3 sec per circle, resulting in an increase of the power by 70%. Secondly, Fermilab made use of the storage ring, the Recycler, which was previously used by the Tevatron facility. The technology

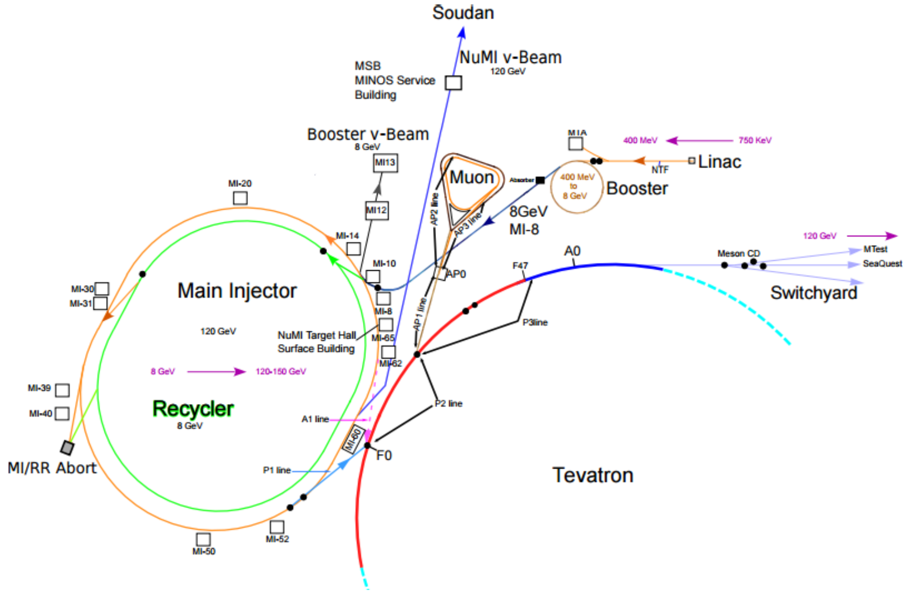


Figure 3.3: The Fermilab accelerator complex includes the linac, booster, recycler and main injector. The accelerated protons then propagate into NuMI beam facilities to produce neutrinos.

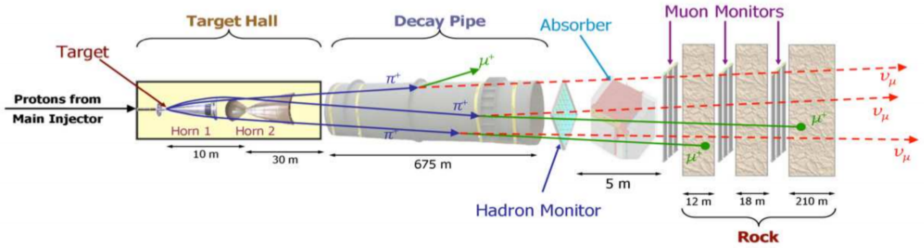


Figure 3.4: NuMI beam facility includes target, horns, decay pipe, hadron absorber and muon shield.

called slip-stacking us used to double the number of protons per batch, before the protons are sent to the main injector. Now after the Booster, the protons are first transferred to the Recycler, which has two RF cavities and thus is able to store protons with two different cycling frequencies. In the seven slots in the Recycler, six slow batches are cycling first. The faster batches are then added in through the seventh slot one by one and align with the first slow batches azimuthally on the ring. The six batches of slip-stacked protons are extracted to the main injector with about $4 - 5 \times 10^{13}$ protons per spill.

As the accelerator facilities are horizontal, the extracted protons in the NuMI beamline is bent by the magnets to about 58 mrad vertically downward in order to let the produced neutrinos pass through the earth. Fig.3.4 is a diagram of the NuMI beam facility including target, horns, decay pipe, hadron absorber and muon shield. The protons first arrive at the target hall and strike a graphite target with density of about 1.78 g/cm^3. The target system is designed to withstand the collision of 400 kW of protons with a stable supply of outgoing hadrons mainly consisting of pions and kaons. Charged pions and kaons are then focused by the magnetic horn. Using different horn configurations, the flux as well as the spectrum of the produced neutrinos can be modified. The magnetic field generated by forward current in the horn bends positive mesons inward which later decay to neutrinos. Therefore neutrinos are focused with forward horn current configuration, while reversed horn current focusing negative mesons is to produce anti-neutrinos. The strength of the horn current and the relative position of the horns have a big impact on the neutrino spectrum as well.

The focused mesons decay mainly through the decay chains, $\pi^{\pm} \rightarrow \mu^{\pm} + \nu_{\mu}(\bar{\nu}_{\mu}), K^{\pm} \rightarrow e^{\pm} + \nu_{e}(\bar{\nu}_{e}), \mu^{\pm} + \nu_{\mu}(\bar{\nu}_{\mu})$ in the decay pipe, which allows mesons to decay with little interaction with other matter. The decay pipe used in the NuMI beam is about ∼675 m in length, which corresponds to the decay length of a 10 GeV pion, which is a typical energy configuration of the outgoing pions seen in NuMI target.

A massive aluminum, steel and concrete structure is placed at the downstream of the decay pipe, working as an absorber, reducing the remaining particles like uninteracted protons, undecayed mesons, neutrons or electrons generated during the upstream interactions. The absorber also reduces the radiation of different particles. Some muons and neutrinos whose energy deposition rate is not large enough are pass through the absorber. The muon shield is an array of blocks of solid dolomite rock with 240 m in length in total, which is responsible for reducing the remaining muons. The flux of hadrons and muons are monitored by hadron and muon monitors, a set of ionization chambers, at the end of decay pipe and of the absorber respectively. As the NuMI beam is supplied by the Main Injector facility, the neutrinos are produced spill by spill which is synchronized with the proton spills in the Main Injector.

In the rest frame, the pion or kaon decay is isotropic so that the spectrum of outgoing neutrino does not have a dependence on the angle. In the lab frame, as the pions and kaons travel relativistically, the spectrum depends on the angle with respect to the beam direction.

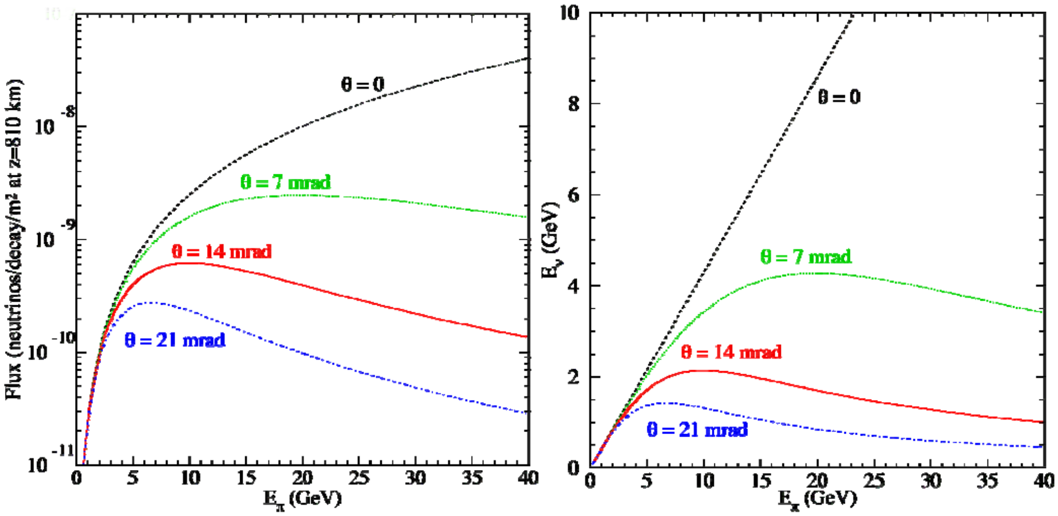


Figure 3.5: The distribution of the neutrino flux (left) and neutrino energy (right) as a function of decaying pion for different outgoing angles.

The following equations show the neutrino energy and neutrino flux as a function of the angle.

$$\begin{aligned} E_\nu &= \frac{1-\frac{m_\nu^2}{m_{meson}^2}}{1+\gamma^2\theta^2} E_{meson} \\ Flux_\nu &= \left(\frac{2\gamma}{1+\gamma^2\theta^2}\right)^2 \frac{A}{4\pi l^2} \end{aligned} \tag{3.1}$$

where E_{meson} represents the energy of the decaying mesons, which are pions and kaons in this case. γ is the boost of the mesons and θ is the angle of the outgoing neutrinos with respect to the beam direction. With these equations, the flux and neutrino energy distribution as a function of pion energy can be described in Fig.3.5. By combining the two sets of distributions above, we can obtain the neutrino spectra at different outgoing angles as shown in Fig.3.6. NOνA is located at 14 mrad off the NuMI beam axis, which leads to a muon neutrino spectrum with a narrow peak at 2 GeV, the energy at which $\nu_\mu \to \nu_e$ oscillation reaches the maximum in the Far Detector. Another big benefit that results from being off-axis is that as the spectrum peaks narrowly at around 2 GeV. In consequence, the ν_e signal is much less influenced by the background components that come from the interaction of the high energy muon neutrinos than the on-axis configuration. One considerable component of the background is the gammas from decaying π^0 which are generated in the neutral current interactions of the neutrinos. For an on-axis setting, the resulting gamma will significantly overlap with the ν_e signal and

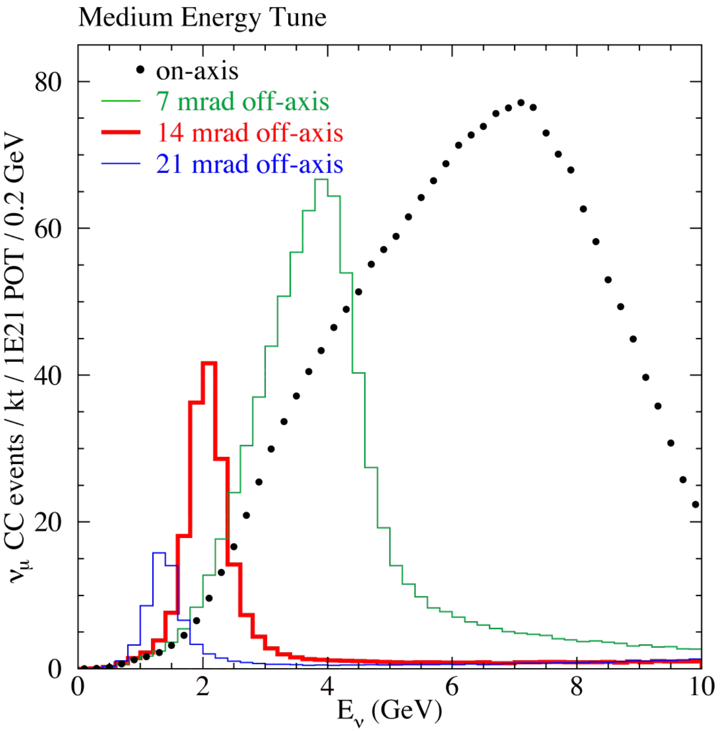


Figure 3.6: Muon neutrino spectra at different off-axis angles of the NuMI neutrino beam.

make the ν_e identification more challenging. Because only non-neutrino part of the neutral current interaction can be detected, neutral current interactions with high incident neutrino energy always cause a contamination in the signal region. At 14 mrad off-axis, as there are few neutrinos in the high energy region, the contamination from NC interactions can be greatly reduced according to Fig.3.7. Based on FLUKA/Flugg simulation, for the Forward Horn Current (FHC) mode and in the region of neutrino energy between 1 GeV and 3 GeV, the muon neutrinos are about 98% of the total neutrino flux in the far detector with 1.1% being muon antineutrinos and the rest 0.4% being electron neutrinos and electron antineutrinos [54]. In the Reverse Horn Current (RHC) mode, the muon antineutrinos are also at the level of 98% of the total neutrino flux (Fig.3.8).

3.2 NOνA Detectors

The two NOνA detectors are both finely-segmented liquid scintillator detectors. As a charged particle, such as electron or muon, passes through the scintillator material, the particle

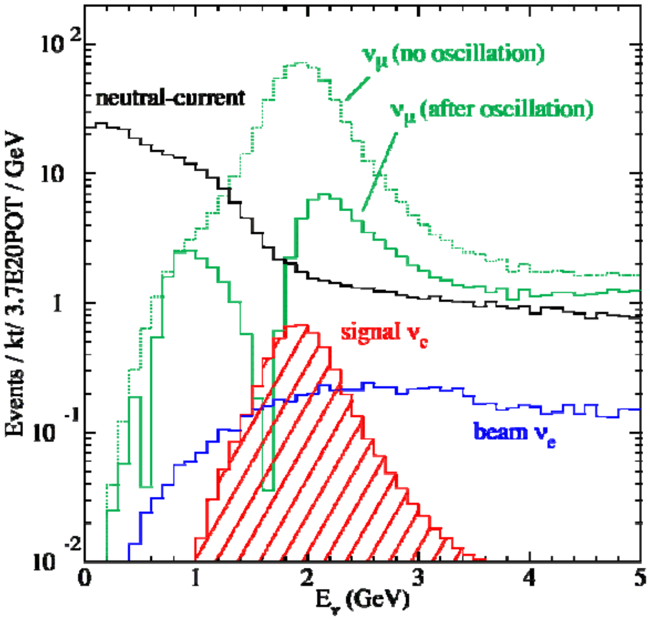


Figure 3.7: Spectra of different components, ν_eCC signal, ν_μCC, NC and intrisic ν_e, in the NOνA far detector based on the simulation with $m^2_{23} = 2.5 \times 10^{-3} eV^2, sin^2 2\theta_{23} = 1$, and $sin^2 2\theta_{13} = 0.10$.

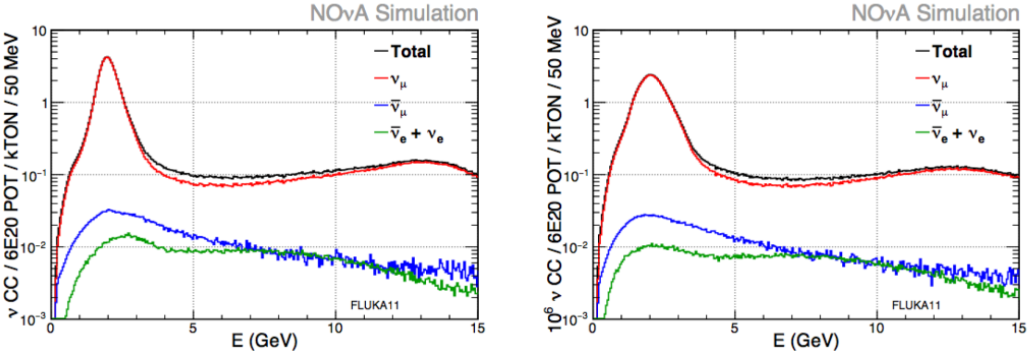


Figure 3.8: The FLUKA/Flugg simulated flux of different type of NuMI beam components, including ν_μ, $\bar{\nu_\mu}$, ν_e and $\bar{\nu_e}$ for the FHC mode in FD (left) and ND (right).

will interact with scintillator atoms through ionization, emitting low-energy photons that can be detected by photo-sensors (Fig.3.9a). As the scintillator is finely divided into different segments in the NOνA detectors, we are able to reconstruct the position of each energy deposition by the incident charged particle.

In this detection technology, liquid scintillator is used to absorb the energy of the charged particles as well as to convert the energy deposition into detectable light. As the incident particles interact with and deposit energy in the scintillation material, free electrons are excited. When they return to the ground state or intermediate exiton stage, light is emitted. The scintillator used by the NOνA detector has a composition of 95% of mineral oil to mainly absorb particle energy, 4% of scintillator material, pseudocumene, to transfer deposited energy into light and minor PPO (2,5-Diphenyloxazole) and bis-MSB ((1,4-bis-(o-methyl-styryl-)-benzene) to shift the wavelength of the emitted light to a specific range so that the light can be transferred more efficiently. The typical wavelength of the photons emitted by pseudocumene is 360-390 nm. After being shifted by PPO and bis-MSB, the wavelength of the photons ranges between 400-450 nm. Some other materials are also added in the scintillating liquid: an anti-static agent at 3 parts per million to reduce its conductivity and Vitamin E to fulfill the anti-oxidant requirement.

(a)

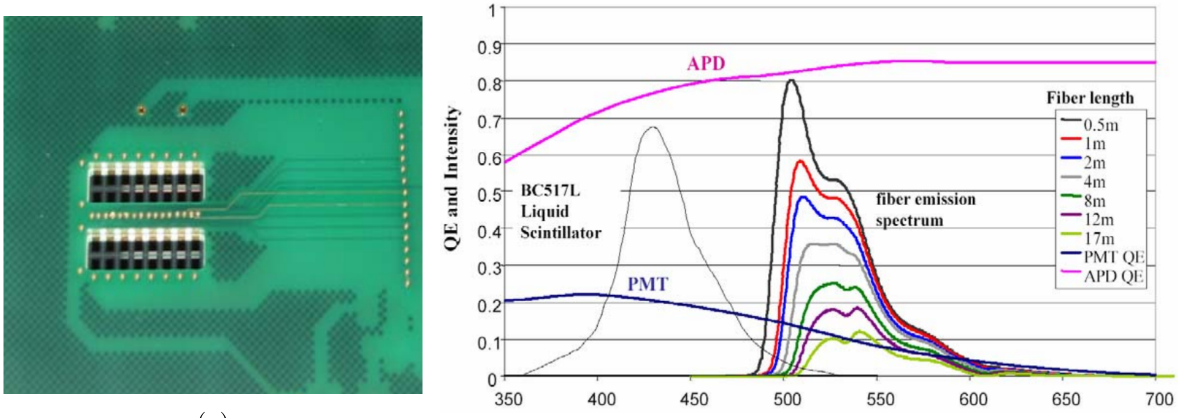
(b)

Figure 3.9: Structure of NOνA cells. The left shows the detection in one unit of NOνA detector. The right is the picture of one PVC extrusion module.

The light is collected and transported by the wavelength-shifting (WLS) fiber. The fiber is 0.7 mm in diameter and is made of polystyrene mixed with R27 dye to obtain wavelength shift. The total reflection is assured by two layers of clapping materials with lower refractive index. The clapping layers are both 3% of the fiber diameter. After the propagation in the WLS fiber, the wavelength of the incident photons is shifted to 490-550 nm. The fiber is looped in each cell. The pulses of light traveling to the two ends of the figer are then collected by a pixel on the photo-sensor. As the light propagates in the fiber, it attenuates. Since in a NOνA cell, an energy deposition is measured by two pulses of light with two different propagation distances, the overall attenuation effect is modeled by double-exponential function and is corrected in the calibration step, which will be discussed later in this chapter.

The NOνA detectors are segmented by PVC cells. The cross section of each PVC cell is 3.9 cm in width and 6.6 cm in depth [160]. The structure of the PVC extrusion is designed with two purposes: First, the wall must withstand the high pressure around the bottom of the detectors, especially the far detector, at the bottom of which the pressure from liquid scintillator is as high as 19 psi. The walls of the extrusion are carefully built so that the detector is prevented from creeping to failure. Secondly, the interior of the cells must have high reflective efficiency so that the emitted light still retain most of the energy after bouncing several times off the wall. In order to achieve high reflectivity, the PVC material used in the NOνA detectors is mixed with 15% of $\mathrm{TiO_2}$, which is demonstrated to preserve 90% of the energy of light at 430 nm. Within the detectors, the cells in the neighbouring planes are arranged orthorgonally so that all three dimensions of the energy deposition can be measured with a desired spatial resolution of approximately 5 cm. Two PVC extrusions side by side form a extrusion module, which is 1.3 m in width and 6.6 cm in depth (Fig.3.9b). On one extrusion module, there are 64 fiber ends, which are connected to a single unit of photo-sensors.

The photo-sensors used by NOνA are Avalanche Photo-Diode (APD). The photo-sensor transfers the light into photo-electrons (PE) and the PE signal is multiplied through avalanche multiplification for later processing. Fig.3.10a shows a picture of the front face of an APD chip, on which there are 32 pixels measuring light from 32 cells, each measuring the two ends of a fiber. One advantage of APD over the traditional photomultiplier tube (PMT) is that the APD performs at a much higher quantum efficiency (QE) than the PMT over the whole visible light region. The comparison between APD and PMT can be seen in Fig.3.10b. In the wavelength region of the WLS fiber, the QE of APD is approximately eight times higher than PMT, which results in a much higher detection efficiency in a detector with APDs than the one with PMTs. In order to reduce the thermal noise from the photoelectron generation, the working temperature of the APD is set at -15°C, which is achieved by thermoelectric cooller (TEC). The working gain is optimized to be at least 100 in order to achieve high signal to noise ratio (App.A). To reach the gain of 100, the operating voltage of APD is in the vicinity of 425 Volt. With this configuration, the typical value of signal-to-noise ratio is 10:1 at the far end of the cells in the far detector.



(a)

(b)

Figure 3.10: Left: Picture of the front face of one APD chip. Right: Distribution of quantum efficiency as a function of wavelength for APD (pink) vs PMT (blue).

The trace consisting of both the signal detected in APDs and noise are then further amplified and processed by the front-end board (FEB). A shaping function is firstly applied to the trace in application-specific integrated circuit (ASIC) to remove the low-frequency and high-frequency components. A narrow pulse will become broader with, for the far detector, a rising edge of 385 ns and a falling edge of 700 ns, and make it easier for later pulse height and pulse time extraction [55]. Then each analog point on the trace is digitized by analogy-to-digit (ADC) unit. The digitized trace is then sent through a process called dual correlated sampling (DCS), $DCS_i = ADC_i - ADC_{i-3}$, in which i represents each sample point on the trace. Fig.3.11

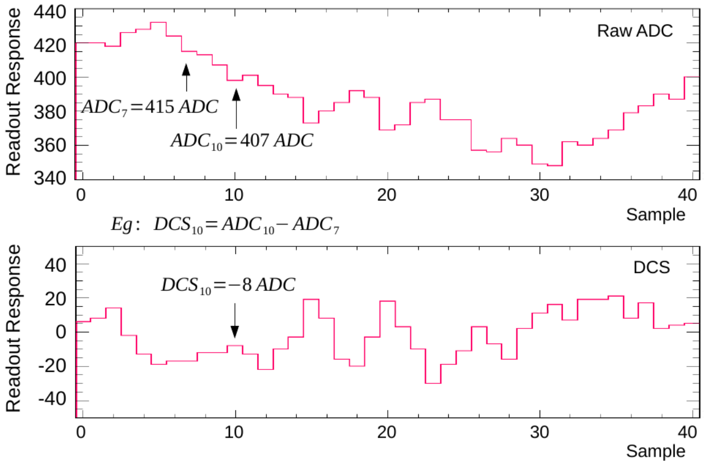


Figure 3.11: DCS process of raw ADC trace, $DCS_i = ADC_i - ADC_{i-3}$. An example is given in the figure that $DCS_1$0(-8) = $ADC_1$0(407) - ADC_7(415).

showns an example of how one sample point is processed through DCS process. The DCS process shifts the baseline to zero and reduces the low-frequency wandering of the baseline.

Then a threshold which is measured in a pedestal run of each channel is applied to the DCSed trace to extract signal from noise. The pulse with height above threshold are expected to come from the energy deposition from particles transversing the detector. The time and ADC value of the selected pulse is recorded as the time and strength of the energy deposition, which, combined with the location of the cell where the deposition occurs, packed into an analysis object called hit in FEB. Hit information from 64 FEBs is sent to one common Data Concentrator Module (DCM). The data is accumulated in the DCM for 5 millisecond and is propagated to a buffer node, which collects data from different DCMs at the same time and pack it into subruns, which are then stored into datadisks.

The process above can only tell the time of a energy deposition in the FEB time frame. In order to achieve the hit time in the detector frame, the FEB needs to be synchronized to a time distribution unit (TDU), which is synchronized to the main time distribution unit (MTDU) of the detector. The time of detector is further synchronized to an external GPS. As NuMI facility is also synchronized to a GPS, the detector is in the same time frame as the NuMI beam. For each spill of neutrinos delivered by the NuMI beam, the detector is able to open a 550 μs window that is centered at the neutrino spill to recorded neutrino interaction signal.

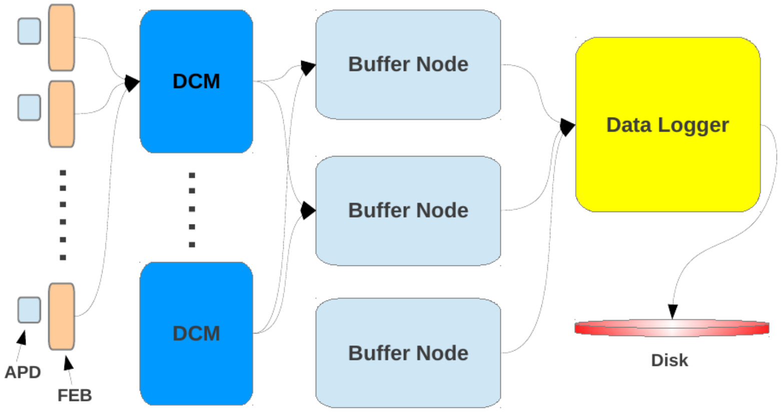


Figure 3.12: (a)APDs detect light from scintillator. (b)FEBs process signals from the APDs, extracts relevant data, formates them and transmit it to DCMs. (c)For 5ms, all data from all DCMs are transmitted to one single buffer node. For the next 5ms, all data goes to the next buffer node. (d)All data for a triggered event goes to one Data Logger and then written into disk.

The timing peaks of the NuMI neutrinos are later checked in both near and far detectors as part of detector commissioning and monitoring.

The NOνA far detector (FD) is located on the surface at Ash River, MN USA. It is $\sim$ 60 m long in the direction of NuMI beam and the cross section is 15.5×15.5 m^2. The scintillator represents $\sim$ 65% of total detector mass and PVC is approximately 30%. The detector has 344064 channels in total. The detector consists of 28 blocks lining up along the beam direction. Each block is then divided to 32 planes with each having 384 cells. The orientation of the cells switches in horizontal and vertical directions alternatively plane by plane. Electronically, as shown in Fig.3.13, there are 14 diblocks (DB), each being formed by two blocks. To control the data flow from all channels in the diblock, 12 DCMs are used for each diblock with DCM 0-6 for horizontal channels and DCM 7-12 for vertical channels One DCM is connecteed to 64 FEB/APD with each having 32 channels. The construction of the far detector started in January 2013. Each diblock was installed and commissioned independently which allows for data collection with partial detectors in parallel partitions. The NOνA far detector construction was finished in August 2014.

The overburden was built to reduce the cosmic ray contamination, especially cosmic gammas and neutrons, in the detector. This overburden is composed of 2.5 feet precast concrete plank,

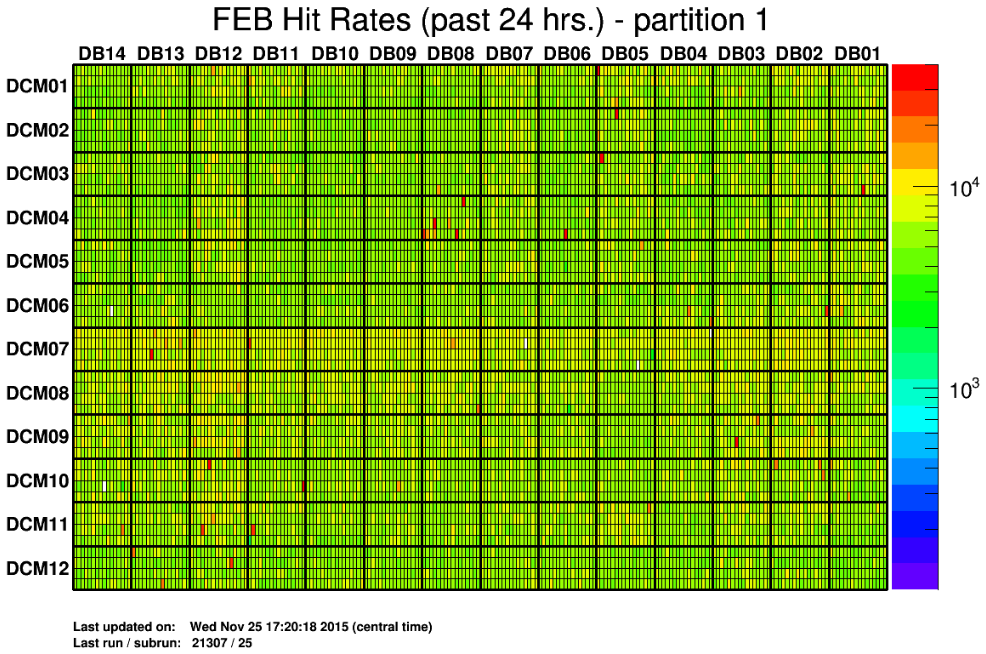


Figure 3.13: The map of the hit rate of FEBs for the far detector. The FEBs are grouped electronically into the DCMs that they connected to. Partition 1 is used to collect data for all the commissioned channels.

1.5 feet cast-in-place concrete and 0.5 feet loose barite concrete as shown in Fig.3.14a. The overburden in total presents a shield of 14 photon radiation length. Similarly, concrete and rocks are piled on the side wall of the detector hall in order to block the cosmic ray events with large zenith angles.

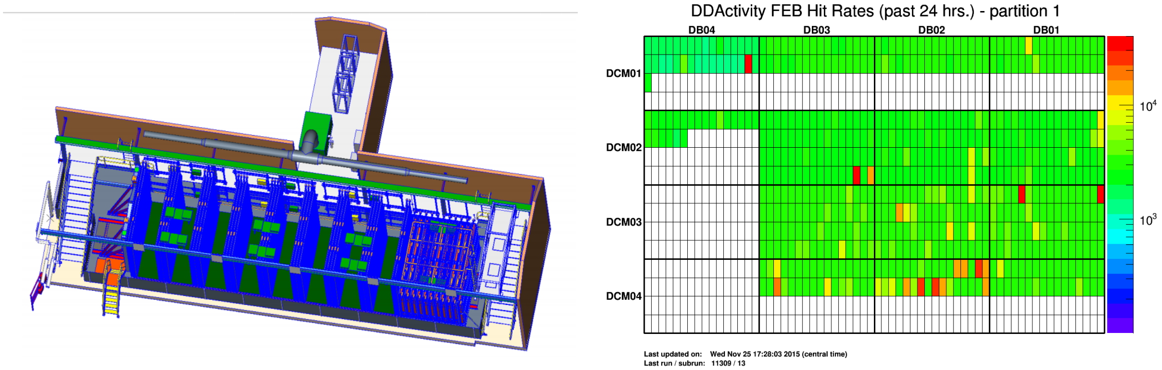
The near detector is a smaller version of the far detector and is 13 m in depth and has a $2.9 \times 4.1 m^2$ cross section. It is composed of 6 blocks, each of which has 31 planes. The X-view planes have 64 cells (2 modules) and Y-view planes have 96 cells (3 modules). The detector is followed by a muon catcher that consists of 22 scintillator planes and 10 steel plates to range muons from ν_μ CC interactions. A diagram of the top view of the near detector is shown in Fig.3.15a and the distribution of electronics is presented in Fig.3.15b. The construction of the near detector finished in August 2014.

The near detector is built underground at Fermi National Accelerator Laboratory, Batavia, IL USA, at 1 kilometer in front of the NuMI target. Thus the near detector has an approximately one million times higher intensity of neutrino induced interactions than in the far detector, as calculated by inverse square law. When beam intensity is approximately 500 kW,

(a)

(b)

Figure 3.14: Left: Diagram of different layers of overburden on the top of the far detector hall. Right: Picture of the electronics installed on the top face of the far detector.



(a)

(b)

Figure 3.15: Left: The diagram of the top view of the near detector. The NuMI neutrinos come from the left, passing the detector body and muon catcher. Right: The map of the hit rate of FEBs for the near detector.

within the vicinity of the NuMI spill window, there are on average 6 interactions [161, 162], including the neutrino interactions occuring inside the detector as well as interactions of rock muons, which is as the particles produced in the interactions of beam neutrinos with atoms in the rock. In order to distinguish between multiple interactions, the readout system of the near detector has a better timing resolution than the far detector: The rising and falling edges of the shaping function in ASIC are set at 140 ns and 4500 ns respectively for ND [53]. Since more energy deposition is expected to be detected in ND than in FD, ND requires 4 times more multiplexers in ASIC unit to process signal.

3.3 Detector Calibration

The detector response must be calibrated for any further useful analysis. The calibration has two main purposes: detector response calibration and absolute energy calibration. During detector response calibration, the detector responce at any temporal and spatial points is normalized to a common standard, which in this case is the average detector response. The detector response calibration includes attenuation calibration which corrects for in-cell difference, cell-to-cell calibration for the variation between different cells and drift calibration for aging and seasonal effects. Because the temporal variance in the average detector response is measured to be small, 0.5% for FD and 0.4% for ND [163], drift calibration is not applied in the first analysis. The absolute energy calibration then convert the "common standard", average detector response, to the energy in GeV by using the energy deposition by minimizing ionization particles [56].

The NOνA detector calibration is performed in several steps: Firstly the pulse height in ADC is converted to photo-electrons (PE) by multiplying 0.5, which roughly reflects the effect of FEB electronics. In-cell and cell-to-cell variations are corrected simultaneously in the attenuation calibration, in which a double-exponential function as Eq.3.3 is fitted to the cell responce as a function of cell depth W. In order to make the cell response match the attenuation fit function better, two effects: threshold and shadow effects are corrected. Due to the shadowing effect of the large far detector volume, the spectrum of the energy deposition in the cell at the bottom of the detector might be different from the one at the top. If we normalize the average energy in the cell to MIP energy, the PE value at the bottom might need to be adjusted higher than the PE value at the top. The threshold correction adjusts the effect of threshold cut on PE distribution that as the threshold cuts the lower part of the PE distribution, the mean of the measured PE will be higher than the actual energy deposition. A correction needs to be applied to scale down the measured PE to the actual amount of energy deposition. The deeper the energy deposition occurs, the larger the effect is. The threshold and shadow correction applied in the first analysis is based on the simulation and the value is calculated using Eq.3.2. A data-driven correction method is under development (See App.B).

$$Corr_{Thresh} * Corr_{Shadow} = \frac{PE_{Exp}}{\lambda_{Poisson}} * \frac{GeV_{True}}{MIP_{Exp}} \tag{3.2}$$

The threshold and shadow adjusted cell response is then corrected by attenuation calibration, in which Eq.3.3 is fitted to the distribution of energy deposition per path length, in the unit of PE/cm, as a function of W for each cell individually [56]. In Eq.3.3, y is the sum of the response at the two ends of fibers, L is the length of cells, X is attenuation length and W is the reconstructed depth of the energy deposition in the cell. A, C and X are the correction constants that are solved for. X is considered as attenuation length, which characterizes the properties of in-cell difference. C is a shift constant which is purely for cell-to-cell variation. A has effect on both in-cell and cell-to-cell difference.

$$y = C + A\left(e^{\frac{W}{X}} + e^{-\frac{L+W}{X}}\right) \tag{3.3}$$

Fig.3.16 shows the fit result for the cells in the near detector (left) and the far detector

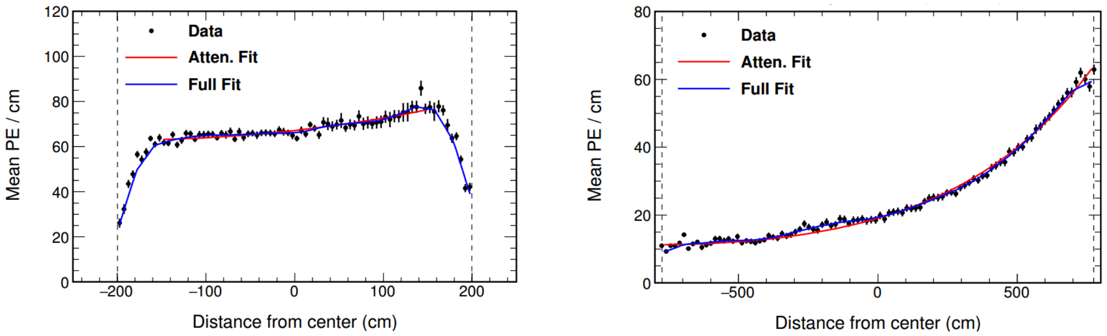
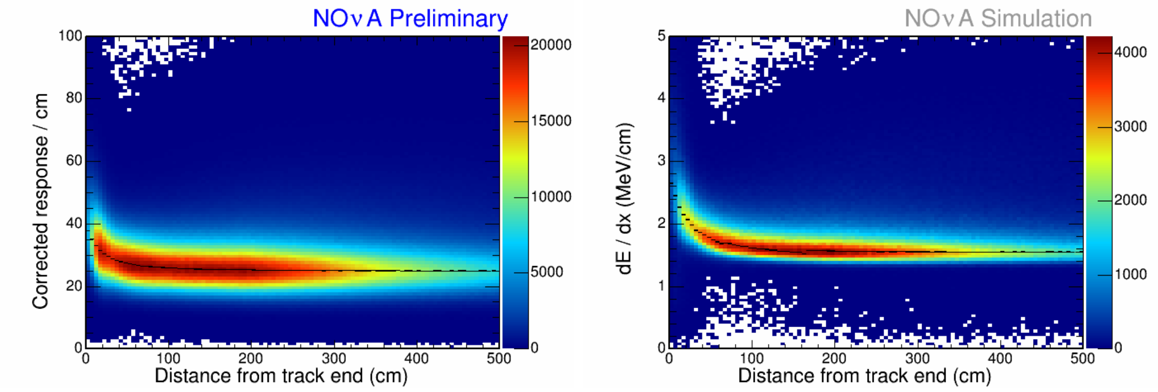


Figure 3.16: The attenuation correction as a function of W for ND (left) and FD (right). Data (black) shows a good agreement with the double exponential function (red) in the middle region. An improved function (blue) build the agreement in the full range.

(right) [57]. The double exponential function which is represented by the red line fits data well in the middle region. On the edge of cells, the agreement is improved by adding two extra functions, rolloff and LOWESS functions, to the double exponential function to take into account the roll-off effect on the edge and the uncertainty of the varying fiber position within the cell. Function with the LOWESS and roll-off terms added shows better agreement with the distribution than the attenuation function along. The near-to-far difference, the difference in

PE between the near and far ends of a cell, is much greater for the far detector, ∼600%, than it for the near detector, which is less than 110%. After the attenuation calibration, The response at any points inside the detector should have small variance and is in the of "corrected PE" (PECorr).

The last step, the absolute calibration, is made to convert corrected detector response PECorr to actual energy unit MeV by comparing to the response of hits from minimum ionizing particles (MIP). Fig.3.17a shows the distribution of distribution of corrected detector response as a function of the distance from track end of stopping muons. By comparing it to Fig.3.17b, the corrected detector response from data has a similar shape to the Bethe Bloch distribution based on true information of the simulation The absolute calibration constant is calculated to be ∼0.063 MeV/PECorr as the ratio between the y values of the MIP hits [58], the hits within 100-200 cm away from the track end, in the two distributions.



(a) (b)

Figure 3.17: The distribution of the corrected detector response from FD data (left) and true MeV from the simulation (right) as a function of the distance from track end.

CHAPTER 4. ν_e APPEARANCE ANALYSIS IN NOνA

The electron neutrino appearance measurement in NOνA is done by comparing the number of observed ν_e-like events to the estimated ν_e-like events in the far detector (FD). The estimation of the background ν_e-like events is done by extrapolating the observation in the near detector ND to the FD. In order to select ν_e-like events in both detectors, we need algorithms to first group the measured energy deposition, reconstruct interaction vertices and prongs and eventually evaluate the likelihood of the reconstructed pattern being ν_e events, which is usually called particle idenfication (PID) algorithm. These steps will be described in detail in this chapter. In the far detector, the selection that are based on the reconstruction and PID variables are designed to select ν_e-like and reject other event topologies, including cosmic ray events. In the near detector, the ν_e selection is applied to construct the sample used to predict the background channels in the far detector and a ν_μ selection is used to obtain the oscillated ν_e CC prediction in the FD. The ND selected sample is decomposed into different type of interactions, which are then extrapolated separately into the FD. The excess of the measured ν_e-like events over the extrapolated background is taken to determine the significance of rejecting $\theta_{13} = 0$ hypothesis.

4.1 Interactions in the Detectors

In NOνA, we can identify three main types of interactions: ν_e charged current (ν_e CC), ν_μ charged current (ν_μ CC) and neutral current (NC) interactions, which correspond to different topologies (Fig.4.1).

The electron neutrino signal in the electron neutrino appearance analysis is identified as an electromagnetic (EM) shower corresponding to a ν_e CC interaction. First an electron neutrino

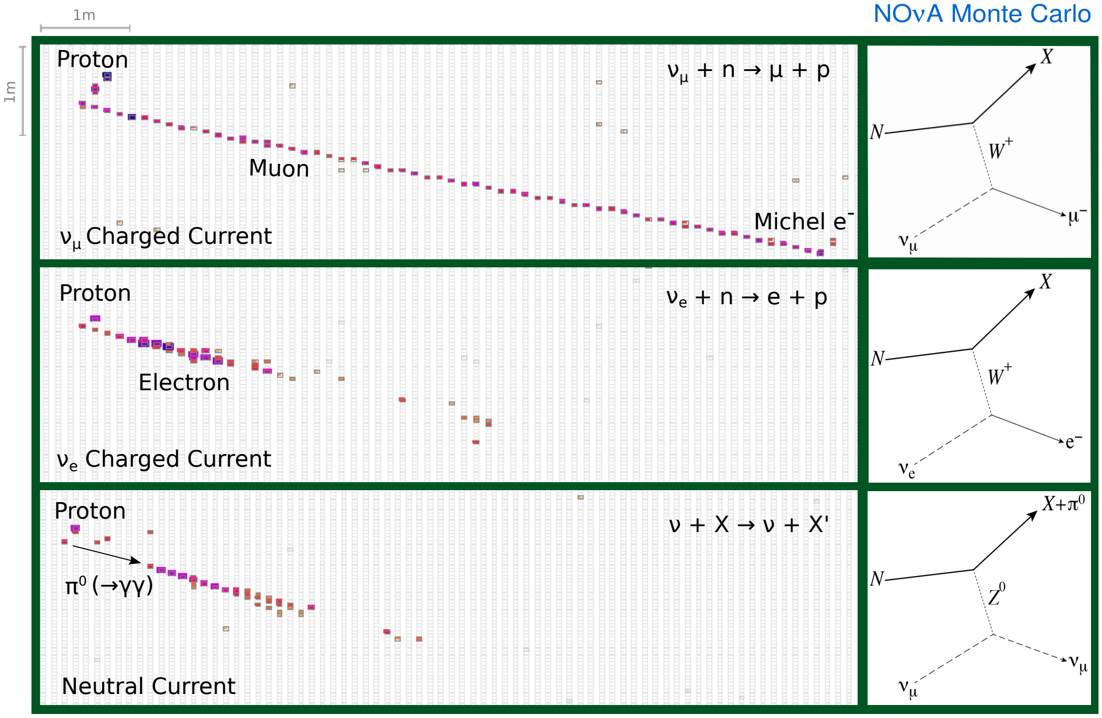


Figure 4.1: Event topology of ν_μ CC, ν_e CCand NC interactions based on NOνA simulation.

interacts with matter by exchanging a W boson, creating an electron and a proton or other hadrons depending on how deeply neutrinos interact with the nucleus. The electron initially loses energy through Bremsstrahlung radiation. The radiated photon further decays into an electron-positron pair, process known as pair production. The processes of radiation and pair production continue until either the electron or the photon reaches the critical energy. In the end, the energy of all created photons and electons is absorbed through ionization loss. The whole process forms an electromagnetic (EM) shower in the medium, which is taken as the signature of the ν_e CC interaction. One of the important parameters used to characterize an EM shower is the radiation length X_0. For a medium with atomic number greater than 13, the radiation length can be approximated as shown in Eq.4.1.

$$X_0[cm] \sim 180A/(Z^2\rho) \tag{4.1}$$

For the NOνA detectors, the X_0 is approximately 50 cm, equivalent to the depth of 10 planes. The interaction length of pair production is approximatly as $7/9X_0$. Another key parameter is the critical energy, E_c, which is defined as the energy below which the pair production process stops and is approximately $E_c[MeV] \sim 550/Z$. For an electron with initial energy E, the maximum number of particles generated in an EM shower is $N_{max} = E/E_c$ and the number of radiation lengths that the shower can extend to is $t_{max} = log_2(N_{max})$. For NOνA detectors, the critical energy is about 65 MeV. As EM shower developsin the longitudinal direction, the position at which energy deposition rate reaches the maximum is $X_0[ln(E/E_c) - 1]$ and the position at which half of the initial energy has been deposited is $X_0[ln(E/E_c) + 0.4]$ [61]. A more precise description was suggested by Longo and Sestili and it has form of Eq.4.2.

$$\begin{gathered} \frac{dE}{dt} = \frac{b^{a+1}t^a e^{-bt}}{\Gamma(a+1)}E \\ with\ a = 1.985 + 0.430ln(E[GeV]) \\ b = 0.467 - 0.021ln(E) \end{gathered} \tag{4.2}$$

t is the depth in radiation length and, for particle with energy less than 5 GeV. The shower grows wider and then narrows down with the increase in the depth of its propagation. In the transverse direction, an EM shower expands mainly due to electron multiple scattering via the Coulumb force. In this model, the shower radius is equal to Moliere radius R_m, $R_m[cm] = X_0E_s/(E_c\rho)$, in which $E_s = 21MeV$, E_c is critical energy. However, in practice, the transverse extension is a function of energy and the properties of medium.

As the oscillated ν_e CC signal event is expect to contain an EM shower, some other types of interactions also generate EM showers and thus become background components of the electron neutrino appearance analysis. In the electron neutrino appearance analysis, there are three main beam background channels: intrinsic beam ν_e CC, ν_μ CC and NC interactions.

Intrinsic Beam ν_e CC Background It is originally generated in the NuMI beam and is one of the main background components. Beam ν_e CC also leaves EM shower produced through ν_e CC weak interaction. Thus intrinsic ν_e is an irreducible background component against the oscillated ν_e that can only be partially separated by energy.

NC Background The bottom topology in Fig.4.1 represents a neutrino neutral current (NC) interaction with the production of a π^0, which is one of the most difficult background components to distinguish in the electron neutrino appearance analysis. A π^0 decays into

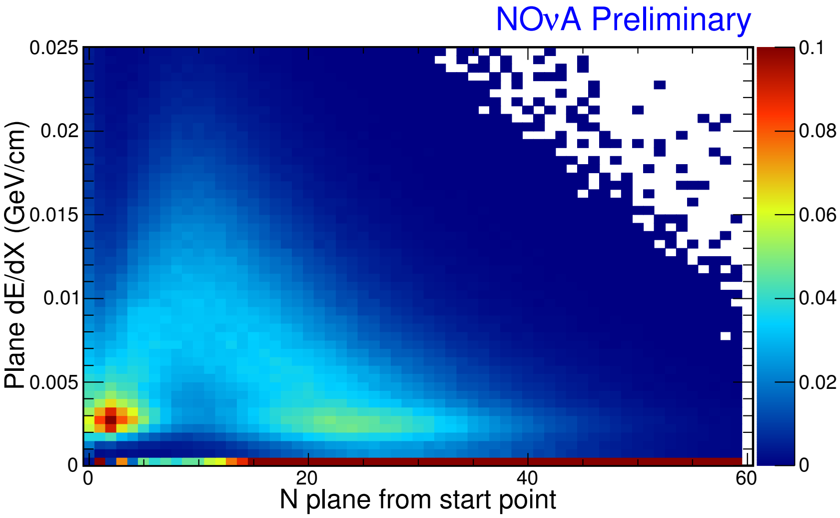


Figure 4.2: Distribution of energy deposition rate, dE/dx, as a function of the depth in the longitudinal direction based on NOνA simulation [140].

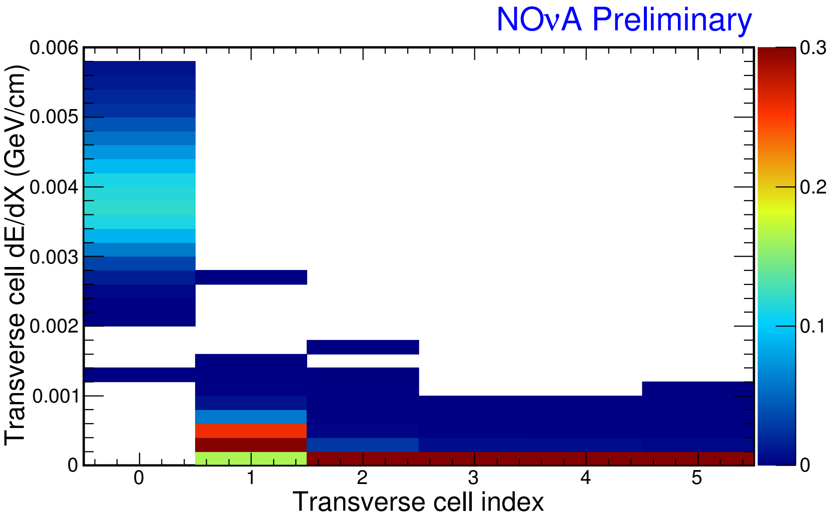


Figure 4.3: Distribution of energy deposition rate, dE/dx, as a function of the extension in the transverse direction based on NOνA simulation.

two γs in 99% of the cases. The produced photons then form two EM showers similarily to the process of EM shower produced by electrons. Based on this model, ideally in a π^0 event, we expect two EM prongs and the vertices of the two prongs to be on average about 50 cm downstream of the main interaction point. We can use this difference to distinguish between π^0 and ν_e. However, in certain cases, rejecting this background is harder. For example, when one of the photons from π^0 decay is much smaller than the other, the event can be misidentified as a ν_e CC event. On the other hand, for the photon-induced EM shower, a gap is expected between the prong and the main vertex, which can be used in tagging π^0 events out against the ν_e CC. But in some cases the gap is sometimes too small. If a π^0 travels fast and the two outgoing gammas have small angle between or even overlap with each other, only one prong might be reconstructed. In order to separate it from standard EM shower, its energy deposition per longitudinal extension tends to be twice of the one in a ν_e CC interaction.

ν_μ **CC Background** In which the muon neutrino interacts with matter through a $W^\pm$ and produces a muon in the detector. The muon is minimum ionizing particle, which loses energy constantly along the propagation with a minimum rate, $1.5/\rho[g/cm^3]$ MeV/cm. The topology of such event is a long track in the detector as shown in the top event display in Fig.4.1. In the NOνA detector, the energy deposition rate is 1.76 MeV/cm. For a 1 GeV muon, the distance it travels is about 6 m, which penetrates about 100 planes in the NOνA detector. However in some cases, the output muon does not carry most of the total energy and the hadronic part might contain one EM showers in the detector. The type of interaction adds into background events for the electron neutrino appearance analysis. One useful feature of the produced muon is that it tends to decay and creates an electron which is usually known as Michel electron, the energy of which is derived to have an upper limit at 53 MeV. The Michel electrons are usually used as a tool to check the detector energy calibration performance.

In additional to the beam background channels, since the NOνA far detector is located 300 m above the sea level, a large number of cosmic ray events are present in the far detector and are also background in the ν_e analysis. Primary cosmic rays originate from astrophysics processes. They consist of free protons (80%), alpha particle (15%) and some heavier nuclei as the rest. [101] As the primary cosmic rays enter atmosphere and interact with the molecules in the air, they produce secondary cosmic rays, which are mainly muons, electrons and positrons, photons, protons and neutrons and a small amount of mesons (Fig. 4.4). The cosmic ray events in the far detector are mostly secondary cosmic ray particles and the cosmic ray background contains all the topologies discussed, including EM shower, muon track as well as hadronic showers. For example, cosmic ray events are dominated by cosmic ray muons, which are mostly from the decay of hadrons, such as pions and kaons that are generated in the interaction of cosmic ray particles with atmospheric particles. The cosmic ray muon forms a track-like topology similar to a ν_μ CC interaction, which can be mostly rejected based on geometry, incident angle and some other properties. Cosmic ray photons induce EM showers which will also cause challenge of measuring the oscillated ν_e CC interactions in the detector. In addition, the EM shower caused by the electron in the cosmic neutrons decay: $n \rightarrow p + e^- + \bar{\nu}_e$ is also a background to the analysis. The cosmic ray background can be separated from a beam event based on some features in the event topology, such as event position and incident angle. More details of the cosmic ray background rejection procedure can be found in Sec.6.3.

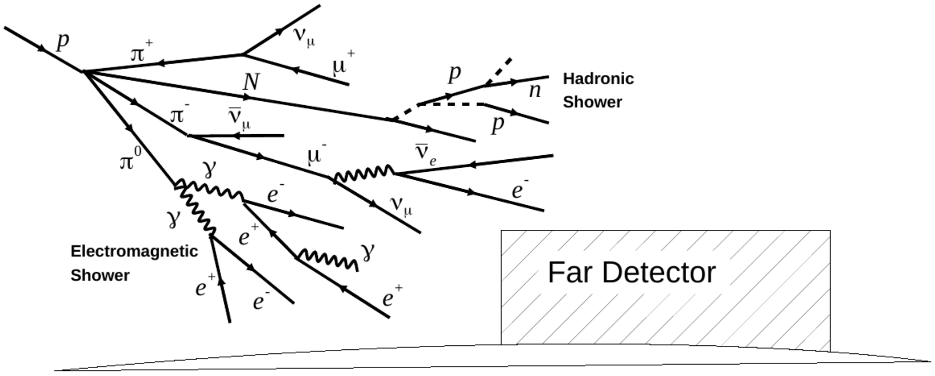


Figure 4.4: Primary cosmic rays interact with air molecules, producing secondary cosmic rays.

One different topology that can be commonly spotted in different types of interactions is the hadronic shower: It could be a byproduct in ν_e CC or ν_μ CC interaction, only detectable component in a NC interaction or a product by a cosmic ray event. The shape of the hadronic shower varies dramatically according to the initial particle and its energy [62]. Assuming the nuclear absorption length is λ_0, in the longitudinal direction, the shower maximum occurs at $X_{max} = \lambda_0(0.7 + 0.2ln(E[GeV]))$. The shower attenuation length is $\lambda_{atten} = \lambda_0(E[GeV])^{\frac{1}{8}}$. The full containment length calculated based on the previous two variables is approximated as $X_{max} + 2.5\lambda_{atten}$. In the transverse direction, the shower width increases linearly with depth [63], the description of the transverse extension of a hadronic shower requires to superimpose multiple Gaussian-like showers, each of which is characterized by Full Width at Half Maximum to be 0.28λ. The parameters for hadronic showers by GeV particles is still not well understood. Understanding the hadronic component of interactions is important. Firstly, for interactions that the produced hadrons carries large fraction of total energy and extend wide, the shape of hadronic shower will have great impact on the interaction reconstruction as well as particle identification. Secondly, a complete reconstruction of neutrino energy requires a robust count of hadronic energies. In the first ν_e appearance of NOνA, hadronic energy reconstruction is calculated as the total calorimetric energy of the hadronic part in the event. Finally, in some cases, various background components, such as NC and cosmic ray events, include hadronic shower, in which the secondary muons, photons or neutrons can further decay into EM showers. A better model of hadronic showers can provide a more accurate estimate of the total background of the analysis.

4.2 Reconstruction of interactions in the far detector

The reconstruction step is necessary to convert the raw observed energy depositions into variables such as interaction vertex position and energy. As discussed in the previous chapter, when a charged particle deposits a certain amount of energy in a cell, the following information is recorded: the amount of energy deposited (ADC), time the deposition occurs (TDC) and the location of the deposition (cell index). The first reconstruction step is to convert the raw information to variables with physics meanings: energy deposition in GeV, 3D position in cm and the time in ns with respect to the trigger window. The energy in GeV is obtained by scaling the ADC value by the calibration procedures which are detailed in the previous chapter. The position of a hit is stored in the format of X, Y and Z coordinates: The origin is defined as the center of the front face of the first plane. Z is the line passing the center of all the planes and pointing in the same direction as NuMI neutrino beam. Y is defined to be perpendicular to the Z direction and pointing vertically upward. X is then the vector product of vectors $\hat{Y}$ and $\hat{Z}$. The fact that the NOνA detectors are constructed with alternating planes in X and Y directions is used to obtain the spatial information of each energy deposition. For example, for a deposition in X-view cells, the X and Z position is translated according to the cell index and the Y position can be inferred according to the position of the depositions that occur in the neighbouring Y-positioning cells. From the design of the NOνA cells, the spatial resolution of interaction is expected to be 5 cm. The information is conveniently packed into a reconstructed object defined as hit. The reconstructed hits within specific spatial and timing windows are preliminarily grouped into different slices [65], which represent different primary interactions. The slice object is characterized based on its time, position, number of hits and total calorimetric energy. Then a hough transformation [67] is applied to a slice in order to recognize any geometric lines in it and to provide seeds to the vertex finding procedure. The vertex of the main interaction is reconstructed using the Elastic Arms algorithm [68]. Using the information of main vertex, the fuzzy K-mean [69] method is used to cluster the hits of the slice into different prongs, which represent different products of the interaction, for example EM showers, muon tracks or hadronic showers. The Fuzzy K-mean method was designed to handle the fuzzy edge of showers and solve the situation where the showers overlap. The reconstructed results in XZ and YZ views are matched based on Z position and form three dimentional prongs. The prongs obtained by the Fuzzy K-mean methed include information of starting position, direction and longitudinal extension. The prongs form a more precise representation of different outgoing particles of the interaction. At each step, the reconstruction is performed in both 2D X-view and Y-view planes, which then merge based on their spatial and timing coincidence to form the 3D object. More details of each reconstruction steps will be described in the following sections.

4.2.1 Slicer4D

The reconstruction chain starts with slicer4D which clusters hits that are close in space and time into the same group. The closeness is scored by Eq.4.3 [65, 66].

$$\epsilon = \left(\frac{\Delta T - \frac{\Delta \vec{r}}{c}}{T_{res}}\right)^2 + \left(\frac{\Delta Z}{D_{pen}}\right)^2 + \left(\frac{\Delta XY}{D_{pen}}\right)^2 + \left(\frac{PE_{pen}}{PEN}\right)^5 \tag{4.3}$$

The first term evaluates the timing residual between the two hits assuming particles traveling at the speed of light. The residual is designed to check the causality between the two hits and is counted in terms of the number of timing resolution, T_{res}, which is the quadratic sum of the timing resolution of the two hits. The rest are penalty terms designed to prevent extreme situations like two hits with too large spatial separation or a core hit with too small PE value. For hits in the same view, $\Delta \vec{r}$ is the two dimensional distance of the two points, $\sqrt{(\Delta XY)^2 + (\Delta Z)^2}$. If the hits are not in the same view, assume ΔXY is equal to 0 and contribute nothing in the residual. If the closeness score of a pair of hits is higher than a cutoff value ϵ_0, the two hits are defined as neighbouring hits.

The clustering procedure is then designed as follows:

1. Begin with a non-clustered hit. Loop and calculate its score with all the other non-clustered hits. If the score of a hit is lower than ϵ_0, tag it as neighbouring hit and group it in the same cluster.
2. Check the score of one of the grouped hit with all the rest hits and include hits with low ϵ score into the cluster.
3. Loop step 2 until no new hit can be included and a cluster is finished.
4. Start from step 1 and loop over step 1-2-3 to search for a new cluster.

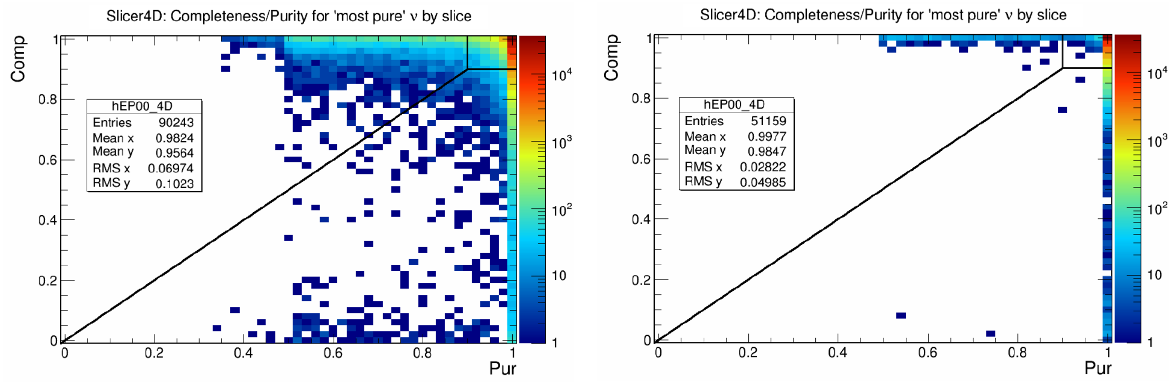
Based on this method, in each group there are two type of hits: core hit and broader hit. The core hit is the one with number of neighbouring hits higher than a criteria, $MinPts$. For the broader hit, the number of hits is lower than the criteria. The core hit tends to be closer to the core of a cluster defined as the average position of a cluster, than the broader hit. After the clustering procedure, a clean up method is applied to merge some small slices into nearby large ones.

A good slicing metric should have the flexibility to cluster events with large extension as well as enough resolution to separate between two close events. As discussed previously, in the NOνA detectors, for a typical 1 GeV beam muon neutrino, the longitudinal extension of the producing muon is about 500 cm. If the particles are assumed to travel at the speed of light, the timing difference is about 10 ns. For a shower-like particle like an electron or a photon, the extension is about 200 cm in longitudinal and about 20 cm in the transverse direction. Although the range of the extension for beam interactions seems controlable and not to vary greatly, for cosmic events whose size grows almost linearly with the energy, the extension could be a few times greater than the beam events. Fortunately, one of the advantage of this slicing method is that its performance does not depend on the maximum extension of the event but on the relative distance between the hits. In the far detector, the rate of beam events is low. On the other hand, the interaction due to the cosmic ray particles occur at a rate of about 148 kHz, which is translated to about 70 events in one 550 μs trigger window. Since the time span of a trigger window is much greater than the dimension of the detector, on average the 4 dimensional distance between the neighboring cosmic interactions is about 7 μs, in which distance in converted to time by being divided by the speed of light. In the near detector, even though most of the cosmic ray particles are blocked by a hundred meter of rock, one 550 μs spill trigger window consists of three to four rock events, which are due to the particles generated in the interaction of beam neutrinos with the surrounding rock. Assuming 4 events spread within 50 μs window around the beam spill, on average the 4D distance between two interactions in the near detector is more than 10 μs. Both 7 μs for FD and 10 μs for ND are greater than the 4D hit resolution of FD and ND, 4 μs and 1 μs respectively, which implies the capability of Slicer4D to separate multiple interactions.

The Slicer4D parameters are optimized using FD cosmic data and ND rock files. A special figure of merit is used to evaluate the tunning result. The performance of a clustering method can be evaluated from two aspects: completeness, the fraction of hits of a true interaction that are reconstructed in the corresponding slice, and purity, the fraction of hits of a reconstructed

Table 4.1: The optimized Slicer4D parameters for FD and ND.

Detectors	ϵ_0	PE_{pen}	D_{pen} [cm]	Completeness	Purity
Near Detector	0.6	40	400	98.5%	94.4%
Far Detector	9	20	50	99.3%	99.3%



(a)

(b)

Figure 4.5: Completeness vs purity distribution of the tuned Slicer4D for the near detector (left) and far detector (right). Slicer4D shows an average completeness of 98.5% and average purity of 94.4% for the near detector and both of 99.3% for the far detector.

slice that are from the corresponding true interaction (Eq.4.4).

$$
\begin{aligned}
Completeness &= \frac{E(hits\ reconstructed\ in\ slice \mid hits\ in\ slice\ by\ truth)}{E(hits\ in\ slice\ by\ truth)} \\
Purity &= \frac{E(hits\ in\ slice\ by\ truth \mid hits\ reconstructed\ in\ slice)}{E(hits\ reconstructed\ in\ slice)}
\end{aligned}
\tag{4.4}
$$

The figure of merit used in the optimization is calculated as the total number of events that have high values in both completeness and purity. The parameters suggested by the tuning are listed in Table.4.1 and Fig.4.5, showing an average completeness of 98.5% and average purity of 94.4% for the near detector and 99.3% for both completeness and purity for the far detector. Fig.4.6 and Fig.4.7 show the far detector in one cosmic trigger before and after Slice4D, which show that Slice4D is able to pick out most of the slices and forms preliminarily track/prong shapes.

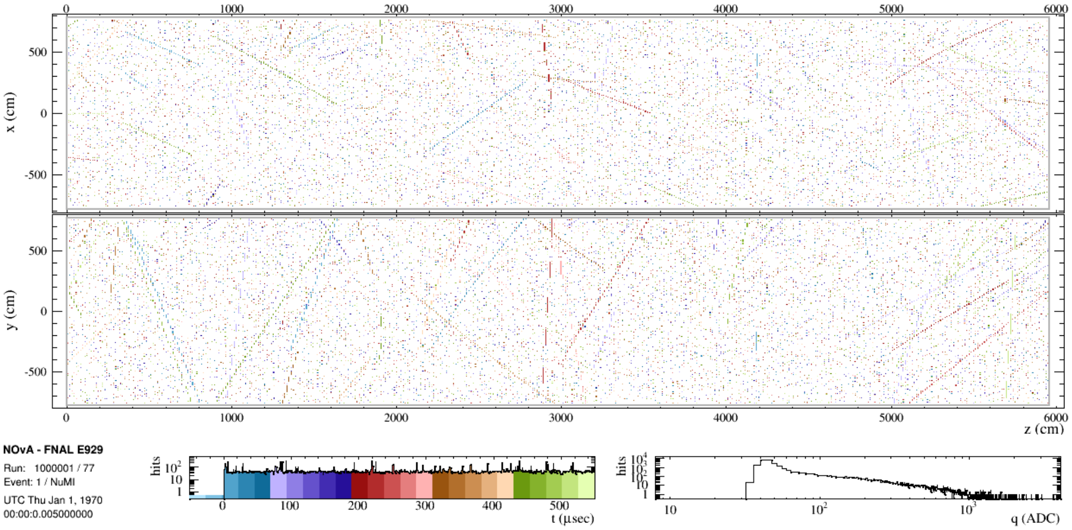


Figure 4.6: One simulated cosmic trigger of the NOνA far detector before any reconstruction procedure. The interaction is measured in two views, XZ and YZ view. In each view, energy depositions are represented as small colored squares.

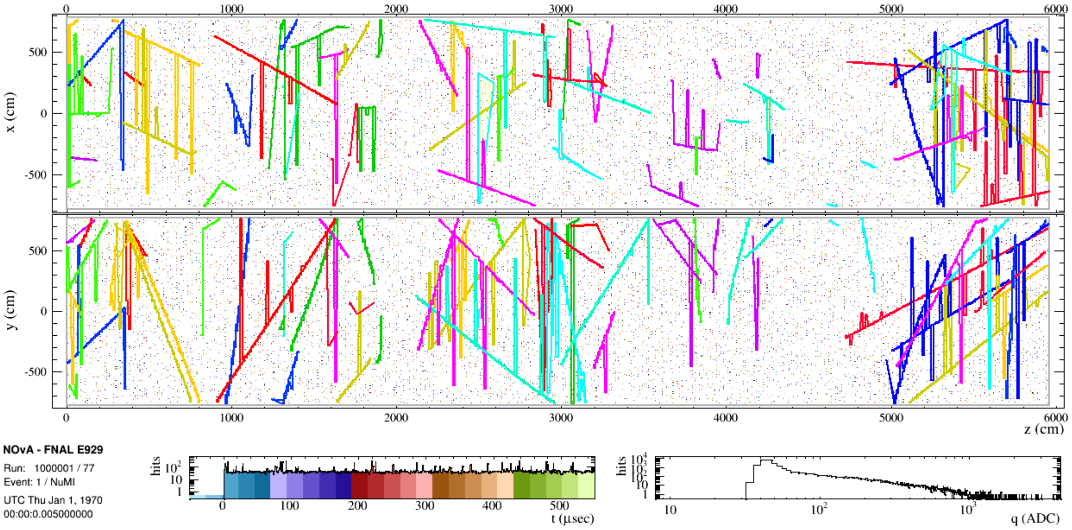


Figure 4.7: One simulated cosmic trigger of the NOνA far detector after Slicer4D.

4.2.2 Hough Transform

The clusters of hits selected by Slicer4D are then processed through the Hough Transform [67] method to look for linear patterns. The Hough transform is a pattern recognition method mostly used to search for patterns that can be parameterized. In NOνA, the pattern that is looked for is a line corresponding to linear tracks. In each view of the detector, a line can be described by two parameters, ρ and θ, where ρ is the perpendicular distance from the origin point to the line and θ is the angle between the line and the horizontal axis. In a standard Hough transform algorithm, the $(\rho,\ \theta)$ is calculated for any pair of hits and is plotted on a $(\rho,\ \theta)$ map. After looping over all pairs, it finds the peaks on the $(\rho,\ \theta)$ map, and the $(\rho,\ \theta)$ of the selected peaks represents a reconstructed linear pattern. In order to make the algorithm suitable for the NOνA experiment, some modifications were made [67]:

1. A limit is set such that the $(\rho,\ \theta)$ values are calculated only if the two hits that form the line are within an optimized distance. With the limit, the amount of calculations required is reduced and the algorithm becomes faster.

2. Since the strength of different particles in an interaction could vary widely, the peaks on the $(\rho,\ \theta)$ map have various heights, which makes it difficult to set an efficient threshold to select all the peaks. To avoid the problem, a multi-hough loop is applied with each loop selecting only the tallest peak and leave the rest hits into the next loop, which then search for the next tallest peaks. Two criteria are set to end the loop: either no peak can be found above a minimum threshold or the maximum number of lines is reached.

3. In each loop, after the map is created, a smoothing procedure is applied to avoid the risk of misidentifying peaks that are caused by the fluctuations mostly due to low statistics, which are usually observed in some shower-like interactions. For each bin, a Gaussian smoothing weight is applied to the ± 3 bins in both ρ and θ axis.

4. After finding the tallest peak in each iteration, a refining method is applied to find a more precise estimation of ρ and θ values. The refined value is an average over a 7×7 square of bins which is centered at the peak bin. The weight of each bin is proportional the height

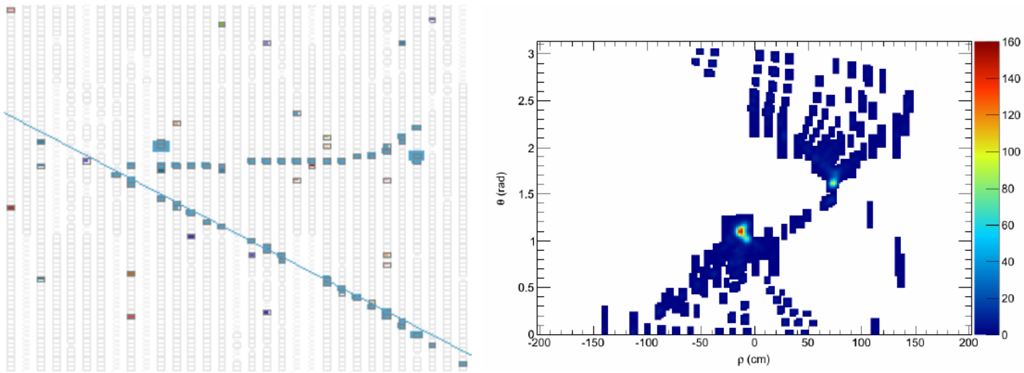


Figure 4.8: The first iteration of the hough transform, in which two clear peaks can be observed, in the (x, z) or (y, z) space (left) a in (ρ, θ) space (right).

of the bin and inversely proportional to the distance to the peak bin as Eq.4.5. For the peak bin, the distance is multiplied by $\sqrt{2}/2$ and carries more weight in the calculation.

$$\begin{aligned}\bar{\rho} &= Norm * \sum_{-3\leq i,j\leq 3}\left[\rho_{ij}\frac{height_{ij}}{\sqrt{i^2+j^2}}\right]\\ \bar{\theta} &= Norm * \sum_{-3\leq i,j\leq 3}\left[\theta_{ij}\frac{height_{ij}}{\sqrt{i^2+j^2}}\right]\end{aligned} \tag{4.5}$$

So with the applied modifications, the modified Hough transform can be described as:

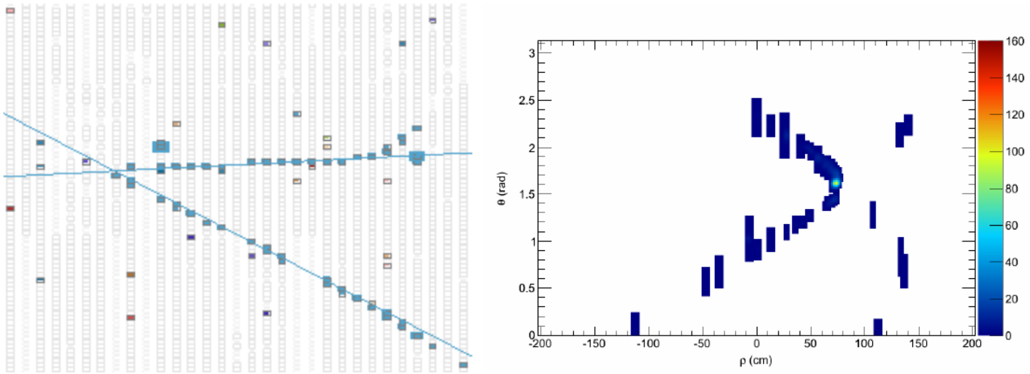


Figure 4.9: The second iteration of the hough transform, in which the hits from the first hough line are turned off and then the secondary peak is selected.

1. Run Hough transform analysis and make a smooth (ρ, θ) map. Pick the tallest peak and run the refining procedure to obtain a legitimate line.

2. Remove the hits within 6 cm from the new found line and use the remaining hits in the next iteration.
3. Repeat step 1 and 2 until either of the two criterion are fulfilled.
4. Merge the Hough results from the two views together.

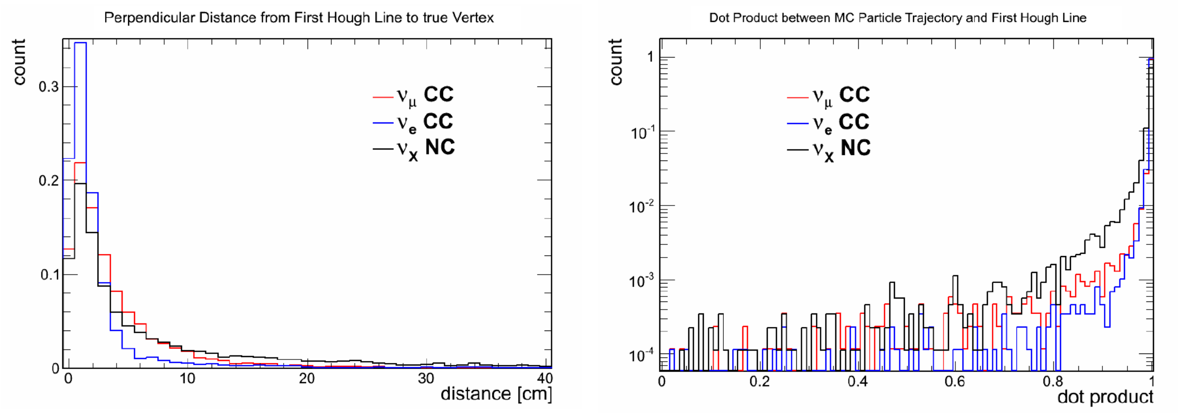
Fig.4.8 and Fig.4.9 shows the event display and the $(\rho,\ \theta)$ map for the first two iterations of one multi-hough procedure: The right plot of Fig.4.8 shows that in the first iteration, two clear peaks are formed in the map. The taller peak on the lower left corner of the map is selected in the first iteration. In the second iteration, the energy associated to the peak is removed and the next tallest peak stands out as shown in the right plot of Fig.4.9.

The performance of the Hough transform algorithm is evaluated by two quantities: perpendicular distance to the true vertex and dot product of the direction of hough line and the true direction. The former one can be understood as a measurement of the lateral displacement of the reconstructed line to the true one, while the latter is of the angular displacement. According to the validation for 8000-9000 of ν_μ CC, ν_e CC and NC interactions from far detector simulation files as shown in Fig.4.10, the mean perpendicular distance to the true vertex is 4 cm for ν_μ CC, 2.7 cm for ν_e CC and 6.7 cm for NC. The dot product quantity is 0.991 (ν_μCC), 0.996 (ν_eCC) and 0.98 (NC), which translates into 7.7° (ν_μCC), 5.1° (ν_eCC) and 11.5° (NC) angle difference between the reconstructed direction and true direction. Fig.4.11 is the display of the same cosmic trigger of NOνA far detector as in Fig.4.7 after Hough transform, in which the colored lines represent the reconstructed tracks.

4.2.3 Elastic Arms Vertex Finder

In the NOνA detectors, the event topology can be described as a combination of prongs that originate from a common vertex point. The mathematical description of the prongs is shown in Eq.4.6,

$$\begin{aligned} x(s) &= x_0 + ssin\theta_a cos\phi_a \\ y(s) &= y_0 + ssin\theta_a sin\phi_a \\ z(s) &= z_0 + scos\theta_a \end{aligned} \tag{4.6}$$



(a)

(b)

Figure 4.10: The perpendicular distance to the true vertex (left) and dot product of the hough line and the true line (right) for different type of interactions in the first iteration of multi-hough transform.

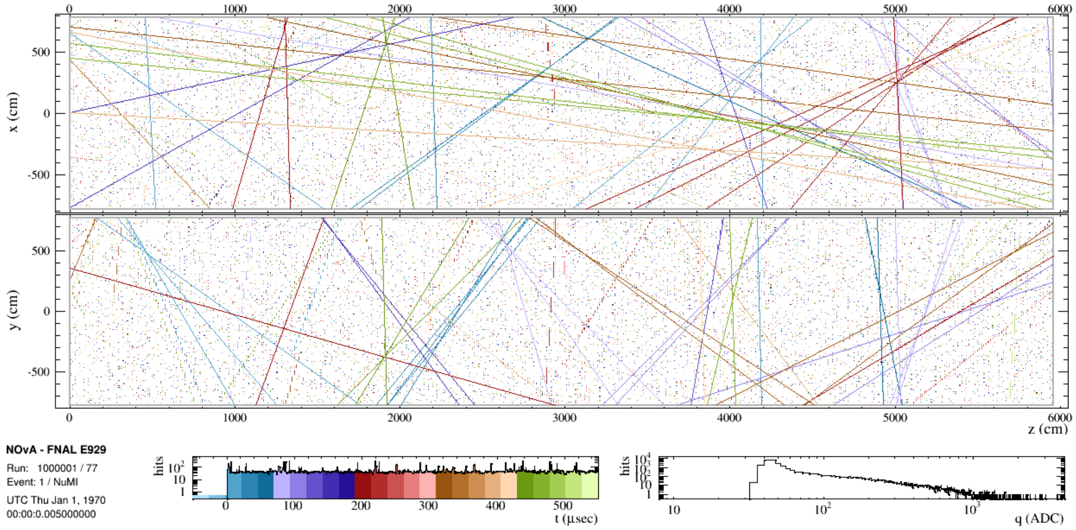


Figure 4.11: One simulated cosmic trigger of the NOνA far detector after hough transform.

where (x_0, y_0, z_0) is the location of the main interaction vertex, while θ_a, ϕ_a and s are zenith, azimuth angles and the extension of prong "a". The goal of the elastic arm algorithm is to fit for $(x_0, y_0, z_0, \vec{\theta}, \vec{\phi})$ which represents the minimum cost-function E value (Eq.4.7) [68].

$$E = \sum_{i=1}^{N}\sum_{a=1}^{M} V_{ia}M_{ia} + \lambda\sum_{i=1}^{N}(\sum_{a=1}^{M} V_{ia} - 1)^2 + \frac{2}{\lambda_\nu}\sum_{a=1}^{M} D_a \tag{4.7}$$

The cost function describes the situation with N hits in total which form M arms. M_{ia} is the perpendicular distance from the hit i to the arm a, while V_{ia} is the strength of the association of the hit i with the arm a. The first term measures the goodness of N hits matching M reconstructed arms. The distance in terms of spatial resolution is calculated in each view using:

$$M_{ia} = \left(\frac{d_{ia}^{perp}}{\sigma_i}\right)^2 \tag{4.8}$$

in which σ_i is set to be $3/\sqrt{12} = 0.9\ cm$, When the hit i is in the upstream with respect to the starting point of the arm a, the distance is calculated as:

$$\begin{aligned} M_{ia} &= \left(\frac{d_{ia}^{vtx}}{\sigma_i}\right)^2, \ if\ \frac{d_{ia}^{vtx}}{\sigma_i} \leq 1 \\ M_{ia} &= \left(\frac{d_{ia}^{vtx}}{\sigma_i}\right)^4, \ if\ \frac{d_{ia}^{vtx}}{\sigma_i} > 1 \end{aligned} \tag{4.9}$$

if the hit is within one σ from the starting point, the distance is equal to $(d_{ia}^{vtx}/\sigma_i)^2$, while when the hit locates greater than one σ awary from the vertex, the factor is $(d_{ia}^{vtx}/\sigma_i)^4$. With this modification, the lower the hits are located in the upstream of the vertex, the lower the energy cost value is. So the Elastic Arm method tend to reconstruct vertices to be close to the beam direction. The hit-track association term in Eq.4.9 is calculated by Eq.4.10.

$$V_{ia} = \frac{e^{-\beta M_{ia}}}{e^{-\beta\lambda} + \sum_{b=1}^{M} e^{-\beta M_{ib}}} \tag{4.10}$$

The closer the point is to the arms, the smaller M value is and consequently the greater the weight is assigned to the point. The first term in the denominator of Eq.4.10 is the overall weight of the noise hits, which is characterized by the parameter λ. The smaller λ is, the greater the weight is, which means that the noise hits are more likely to be labelled into arms. The β parameter in each term means the range of the association between hits and arms. The greater the β is, the faster the strength of the association decays. Thus the fit of minimizing the energy cost starts with low β value, which weakens the boundary between arms and reduces the fluctuation of the energy cost value among different trials. Beginning with low β value enables the algorithm to narrow the search region down relatively smoothly. The fine search, although, utilizes large β value, which enhances the edge of different arms and enables a precise fit for the main vertex. The third term in Eq.4.7 corresponds to the penalty for the displaced track, which is important in some situations such as when only one particle is recorded in the event. Ideally for one particle, the vertex is expected to be placed at the beginning of the particle prong, although the cost function with only the first two terms does not have the capability of telling the difference along one track. The idea of the term comes from the decay of a photon in space. The survival probability follows $exp(-d/\lambda_\nu)$, in which λ_ν is defined to be equal to the pair production length of a photon and d is the distance that the photon propagates. In the cost function, $d(D_a)$ is equal to sum of the distance of the vertex to the first hit from the arms in both XZ and YZ views, $D_a = d_a^{xz} + d_a^{yz}$. The contribution to the energy cost is calculated by:

$$\chi^2 = -2lnP = -2ln(\prod_a P_a) = -\frac{2}{\lambda_\nu}\sum_a D_a \tag{4.11}$$

In order to achieve good fit result, seeding is crucial. The number of arms is estimated as the maximum of the number of Hough lines found in xz and yz views based on the Multi-Hough Transform result, which is described in the previous section. In addition to the number of arms, a list of candidates for the arm vertex $\vec{v}$ and directions $\vec{d}$ is computed. The seeds for the vertex come from three sources: first, a minimum bias sample of the positions at the first 2%, 5%, ... 50% of the total hits in z direction; secondly, the positions of the first hit from the three most prominent hough lines; thirdly, the seed of vertex fit is generated as all the intersecting points of the five most prominent hough lines. Similarly, the seed for the arm directions has two sources: a minimum bias sample of 20 directions which are the vertices of a dodecahedron and the direction of hough lines. The hough lines from xz and yz views are matches according to the closeness in the height in the hough space:

$$-0.8 < \frac{H_{xz} - H_{yz}}{H_{xz} + H_{yz}} < +0.8 \tag{4.12}$$

The optimal seed is then selected by iterating the energy function among all the candidate vertices and directions until the minimum is identified. A good seed of vertex and directions can reduce the computational time as well as increase the success of finding the global minimum.

The parameters are optimized using MC files. Two variables are used as the measurement of the goodness of different candidates, the ratio of energy cost to the degrees of freedom, $E/(N-3-2M)$, and the sum of the association strength to the total number of hits, $\sum_{ia} V_{ia}/N$. The performance of the algorithm is shown in Fig.4.12 for ν_μCC, Fig.4.13 for ν_eCC and Fig.4.14 for NC interaction. The mean and RMS values of the difference between reconstructed and true vertices for different types of interactions are listed in Table.4.2. The vertex resolutions in the three dimensions are better than 20 cm (~ 4 cell-width) and worse than 10 cm (2 cell-width) for different types of interactions. The Elastic Arms method performs the best for ν_μ CC interactions, a bit less well for ν_e CC and the least well for NC interactions. Fig.4.15 shows one cosmic trigger of the NOνA far detector after elastic arm method, in which the crosses representing the elastic arm vertices and solid lines are the elastic arms found in the method.

Table 4.2: Mean and RMS of the difference between ElasticArm fitted and true vertices for ν_μ CC, ν_e CC and NC events.

	ν_μCC			ν_eCC			NC		
	X	Y	Z	X	Y	Z	X	Y	Z
Mean[cm]	0.0	0.4	1.5	-0.2	0.2	1.0	0.1	0.4	1.3
RMS[cm]	10.4	10.4	12.9	12.2	11.7	14.0	17.1	16.5	17.9

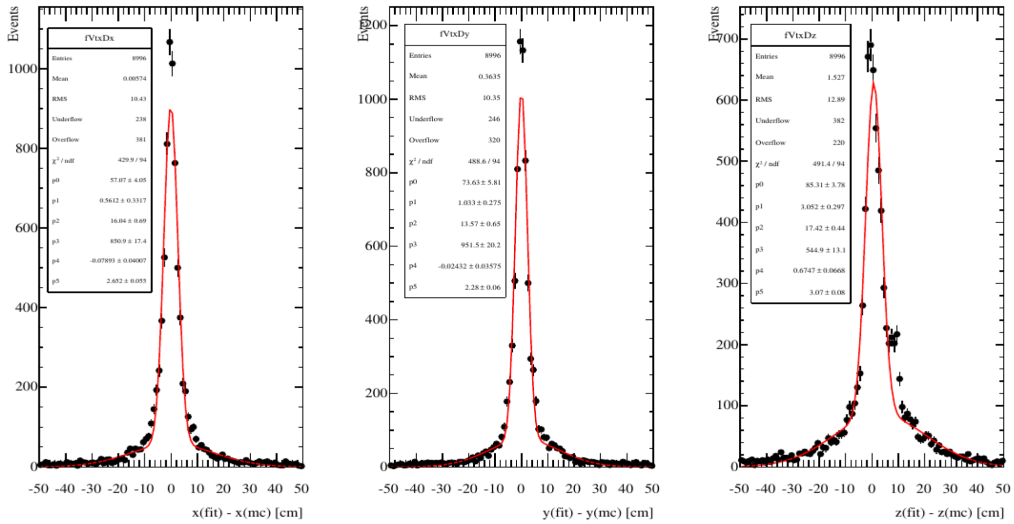


Figure 4.12: Difference between reconstructed vertex and true vertex position in X, Y and Z coordinates for ν_μCC interactions.

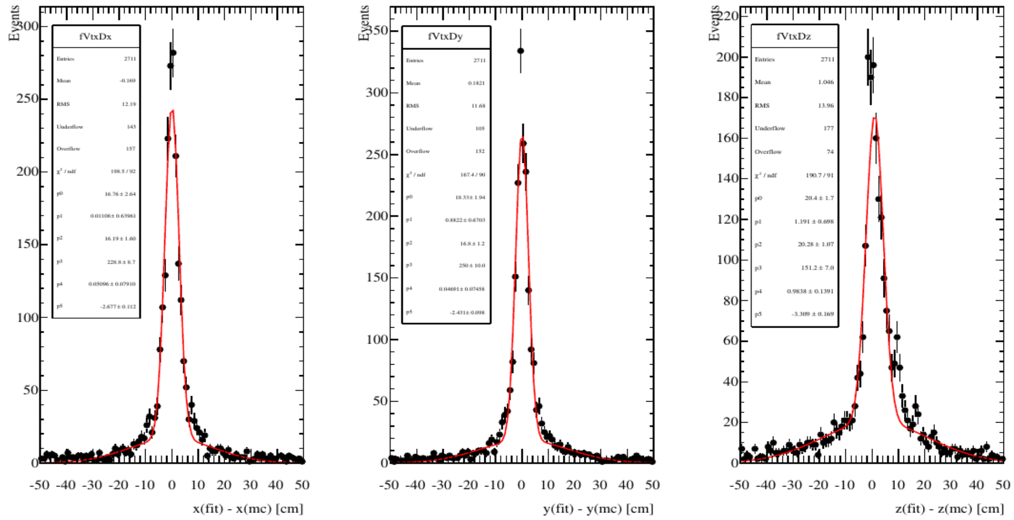


Figure 4.13: Difference between reconstructed vertex and true vertex position in X, Y and Z coordinates for ν_eCC interactions.

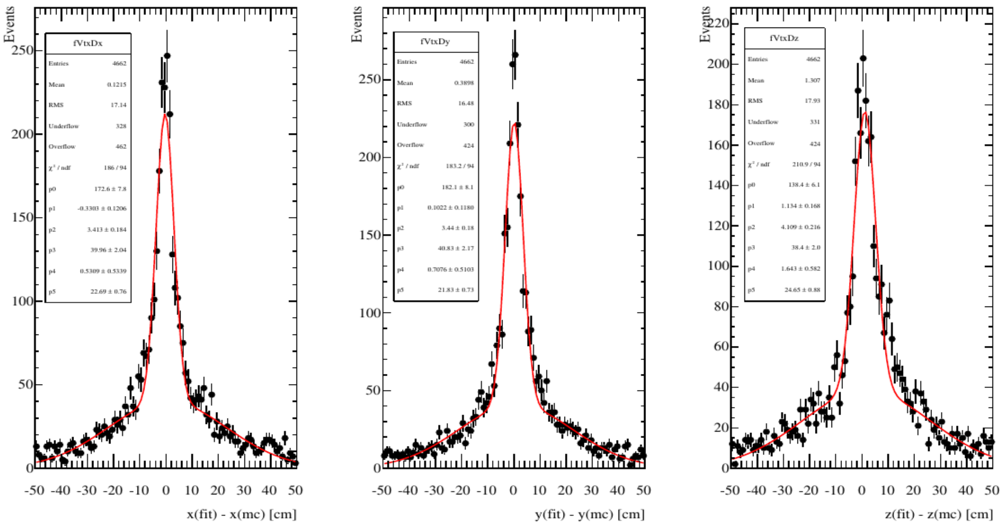


Figure 4.14: Difference between reconstructed vertex and true vertex position in X, Y and Z coordinates for NC interactions.

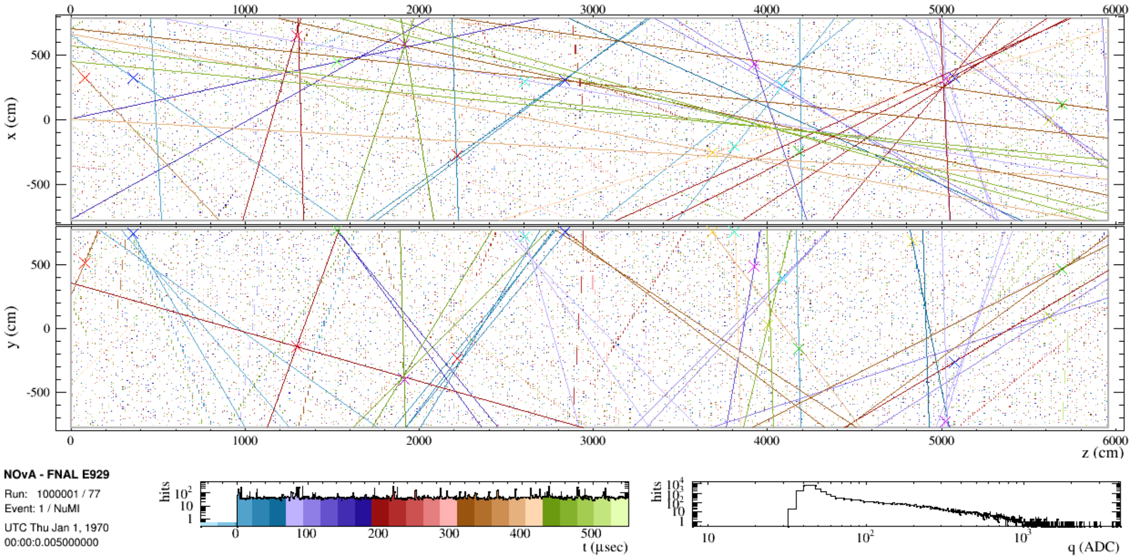


Figure 4.15: One simulated cosmic trigger of the NOνA far detector after elastic arm method.

4.2.4 Fuzzy K-Mean

The output of the previous Elastic Arms method is further processed by the Fuzzy K-Mean (FuzzyK) method, which optimizes the reconstruction result and is designed specifically to handle the fuzzy edges of showers. The optimization is done in XZ and YZ views separately and then the results in the two 2D views are merged into 3D reconstructed prongs. In each 2D view, FuzzyK first groups hits and the hits are then mapped on the distribution of the angle of hit from the Elastic Arm vertex, on which forward Z direction is corresponding to 0 radian. The range of the angle is from $-\pi$ to π. The angle distribution is divided into 360 bins, the content of which is calculated by Eq.4.13 [69].

$$w_k = \sum_{j=1}^{n} e^{-\left(\frac{\theta_k-\theta_j}{\sigma_j}\right)^2} \tag{4.13}$$

As distance of a hit from the angle bin increases, the association between the two is weakened at $\exp -(\theta_k - \theta_j)^2/\sigma_j^2$. The uncertainty, σ_j, is calculated by Eq.4.14, which is based on the simulation of 1 and 2 GeV muons in the detector. The uncertainty is large at low distance due to the 1/d term and decreases as the distance increases. After reaching the minimum, the uncertainty increases again gradually due to the increase in the number of multiple scatterings on the path. In Eq.4.14, d is the distance of hit_j from cell_k.

$$\sigma_{ij} = \frac{1.745}{d_{ij}} + 0.0204 + 0.000173 * d_{ij} \tag{4.14}$$

The seeds of prong centers are found as the peaks on the angle map where is considered to have the densest cell activity. For a track-like event, for example a muon, the transverse area of the event is narrow. So the peak of the prong center is thin and tall. For a shower-like event, the peak would be broader.

After the seeds are obtained, the centers are updated with higher precision. In each iteration, the center of the prong i is refined through the calculation of Eq.4.15, Eq.4.16 and Eq.4.17.

$$d_{ij} = \left(\frac{\theta_j - \theta_i}{\sigma_{ij}}\right)^2, \quad -\pi \leq (\theta_j - \theta_i) \leq \pi \tag{4.15}$$

$$\mu_{ij} = e^{-\frac{m\sqrt{c}d_{ij}}{\beta}} \tag{4.16}$$

$$\theta_i' = \theta_i + \frac{\sum_{i=1}^{c}\sum_{j=1}^{n}\frac{\mu_{ij}^{m}(\theta_j - \theta_i)}{\sigma_{ij}^2}}{\sum_{i=1}^{c}\sum_{j=1}^{n}\frac{\mu_{ij}^{m}}{\sigma_{ij}^2}} \tag{4.17}$$

The essential concept in the calculation of the membership of hit_j in $prong_i$, which as shown in Eq.4.16 decreases exponentially as the distance increases. The variable c is the total number of prongs, which makes all membership values slightly smaller as more prongs are added into test. The variable β is the normalization term, which is set to be 4, 2σ away from the prong centers. The factor m in Eq.4.17 can be understood as the degree of event fuzzyness that is expected in the detector. The larger m is, the fuzzier the clusters are expected. The value of m is set to be 2 in this method. After iterations, two main additional checks are performed. It is possible that multiple seeds converge to the same prong centers after a few iteractions. So the first check is to find the duplicates and remove one of them from the prong center list. Once a set of stable prong centers are formed, hit_j with membership value μ_{ij} higher than 1% is associated with $prong_i$. The second check is to revisit the angle distribution of just the unclustered hits to see whether any new peak can be formed. Then new iterations need to be performed, after the new seed is added to the stable prong centers found previously. These checks continue until either all hits are associated to prongs or the maximum number of prongs are reached.

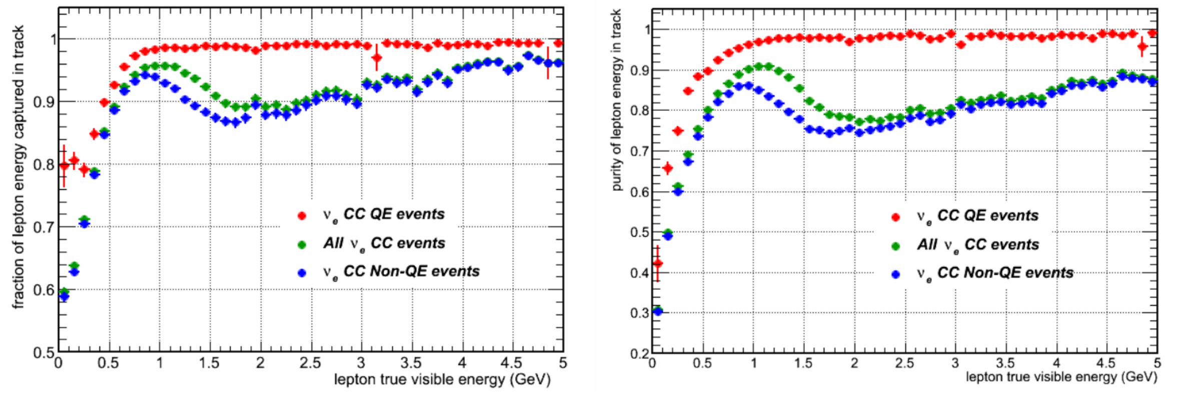


Figure 4.16: Efficiency (left) and Purity (right) as a function of true neutrino energy for different types of ν_e interactions.

The algorithm is tuned and tested on separate MC samples. The performance is tested in terms of purity ($P(True|Reco)$) and efficiency ($E(Reco|True)$) of the reconstruction result (Fig.4.16 and Fig.4.17 [70]), where Reco refers to the amount of energy associated to a prong by reconstruction, while True means the energy of a prong based on true information of the simulation. Efficiency and purity are tested as functions of neutrino visible energy. The efficiency for both ν_e and ν_μ events are above 85% in the energy region that we are most interested in, 1-3 GeV. The purity is around 80% for ν_e and ν_μ events. Since quasi-elastic (QE) interactions always have high efficiency and purity, non-QE events, especially deep inelastic, dominate the drop in efficiency and purity due to the large fuzziness of hadronic components produced in the interactions. Fig.4.18 shows one cosmic trigger of the NOνA far detector after fuzzy k-mean method with dotted lines representing fuzzy k-mean prongs and crosses for the elastic arm vertex from the previous step.

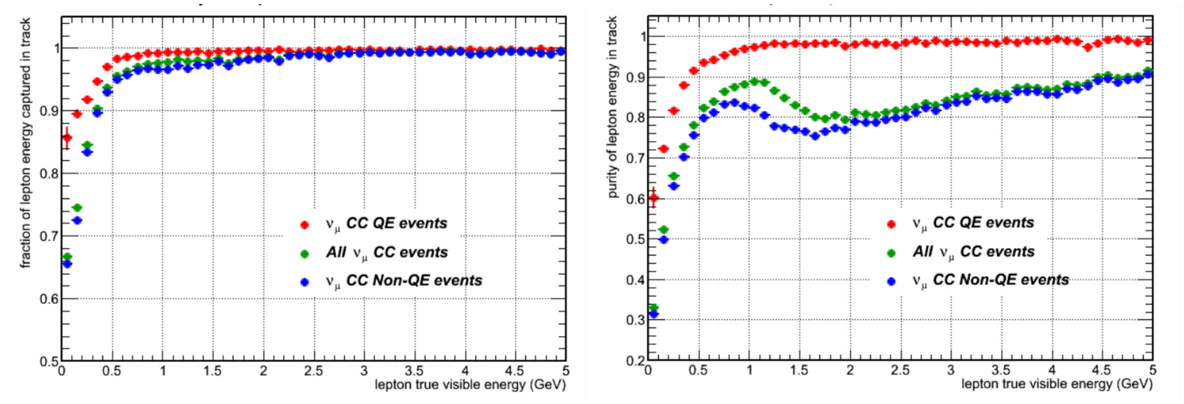


Figure 4.17: Efficiency (left) and Purity (right) as a function of true neutrino energy for different types of ν_μ interactions.

4.2.5 Event Energy Reconstruction

Once the events are reconstructed, the energy of them is calculated as the event total calorimetric energy. The calorimetric energy is then scaled by factors which take into account dead material and detector efficiency [126, 125]:

Dead material: The detector is comprised of mostly mineral oil, plastic material and glue, of which only the region with oil can be used to record particle energy deposition and is

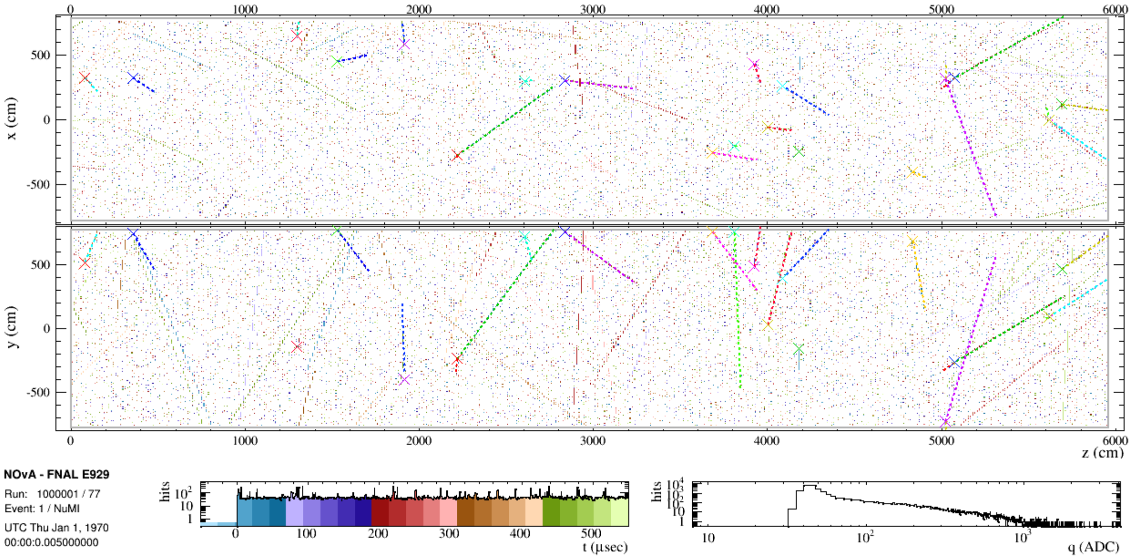


Figure 4.18: One simulated cosmic trigger of the NOνA far detector after fuzzy k-mean method.

known as active area. Materials such as plastic and glue are defined as inactive material in the energy measurement. So the calorimetric energy measured in an event is only a fraction of the actual neutrino energy. The fraction factor is usually known as sampling fraction and is calculated as the ratio of the total energy loss in active area and the sum of the energy loss in active and inactive materials.

Detector efficiency: Even though a particle deposits energy in the oil, there are multiple factors that will cause the incomplete measurement of the deposited energy. Some effects such as the conversion efficiency of scintillator material and absorption of PVC will be corrected in the calibration. One main effect, the energy loss of the hits that are removed by the threshold cut, is corrected by single hit calibration. The detector efficiency correction is applied as a factor scaling the total visible energy of an event up by some amount that is equal to the missing energy of filtered hits. When occuring deeper in cell, an energy deposition experiences greater attenuation effect and is more likely filtered by threshold cut. Although, the efficiency correction is usually a function of the depth in cell as shown in Fig.4.19, the factor used in the first ν_e apparance analysis is independent of the event position and is calculated using events near the center of detectors. This value is expected to be sufficiently accurate because the ν_e CC signal events are expected to be

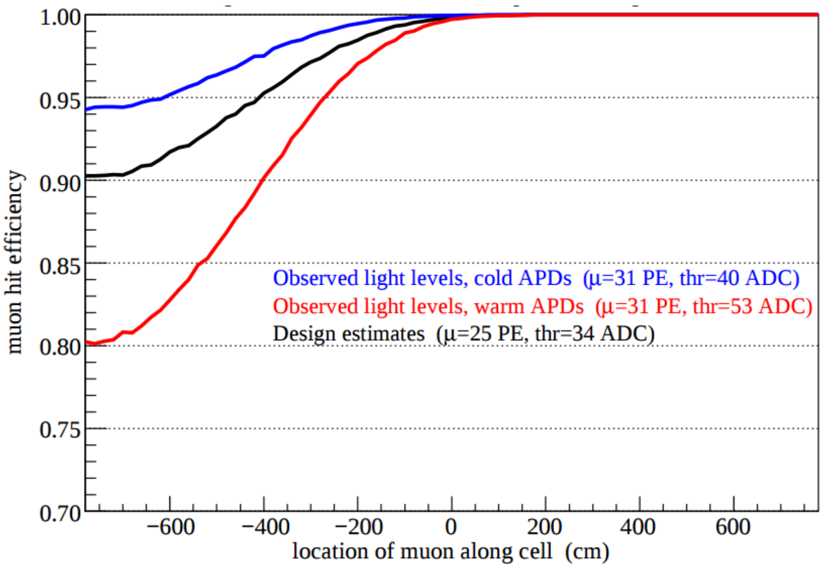


Figure 4.19: Detector efficiency as a function the depth in cell, W for warm (red) and cold (cold) APD configurations.

near the center of cells where position dependence is not strong.

The scaling factor including both dead material and detector efficiency is studied to be 1.78 for both FD and ND as the slope of true energy vs slice calorimetric energy profile. (Fig.4.20) According to the study based on MC, the event energy resolution for FD is 12.5% and 9.1% for ND. (Fig.4.21)

4.3 Electron neutrino identification Tools

Particle identification (PID) algorithms using physics variables, such as prong energy, direction and position, are applied to the reconstructed event. In the ν_e appearnce analysis, PID is designed to distinguish the ν_e CC interactions from other events topologies. NOνA has two independent PID algorithms, Likelihood-based Identifier (LID) and Library Event Matching (LEM). They utilize different approaches for identifying ν_e CC interaction and achieve almost equivalent performance. After a comprehensive evaluation of factors including physics performance, ease of explanation and systematic uncertainties, It was decided that LID would be main PID algorithm and LEM is used as a cross check of the main result [154]. The details of the two PID algorithms will be described in the next two subsections.

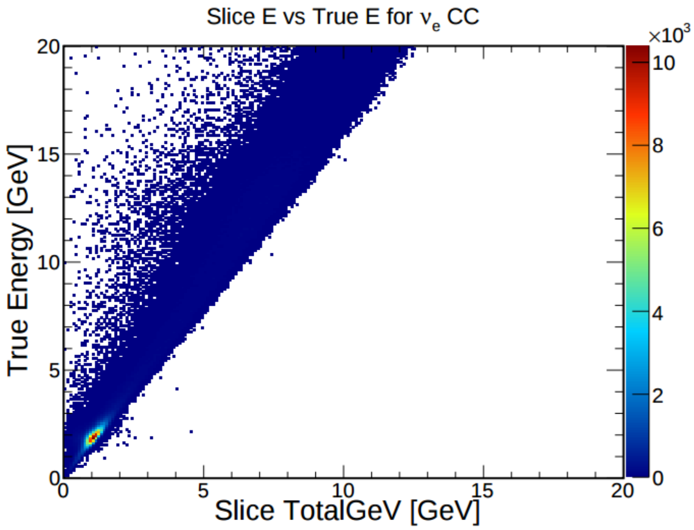


Figure 4.20: Event true energy vs the corresponding slice calorimetric energy. The scaling factor is calculated as the slope of the distribution.

4.3.1 LID

As the name indicated, LID is a particle identification method based on the likelihood of the unknown event being a certain type of event. It uses an artificial neural network (ANN) to score the probability of the most energetic prong in an event being an electron particle against other types of particle hypotheses, such as muon, proton, neutron, pion and gamma. The pattern of energy deposition rate is a good tool to distinguish among different type of particles, each having a unique energy propagation characteristics in both transverse and longitudinal directions. The pattern of a reconstructed particle is calculated based on the FuzzyK result.

LID uses the following variables:

Shower Core the reconstructed line connecting the start and the stop points of a shower.

Transverse Index the number of cells between the associated hit and the shower core.

Longitudinal dE/dx longitudinal energy deposition rate is calculated plane by plane as being the total energy deposited in the plane divided by the path length in the plane.

Transverse dE/dx transverse energy deposition rate is calculated in a similar way but cell by cell and in transverse direction as a function of transverse index.

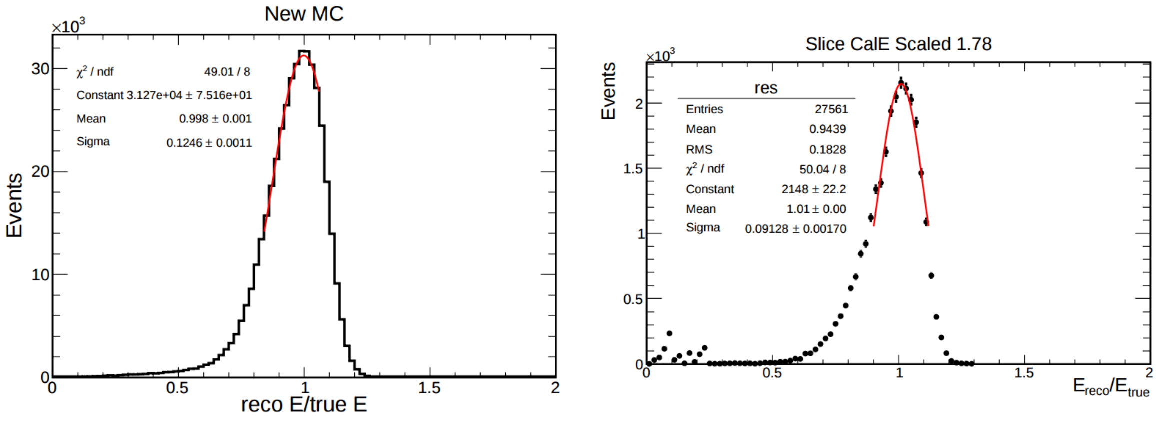


Figure 4.21: The Distributions of the ratio of scaled reconstruced energy to true energy in FD (left) and ND (right). The distributions are fitted using Gaussian function (red). The reconstructed energy resolution is fitted to be 12.5% for FD and 9.1% for ND.

As discussed in the previous chapter, for a typical electron, in the vicinity of the vertex, dE/dx is close to minimum ionizing and the spread in the transverse direction is small. As the electron propagates, the EM shower starts and develops, and thus the prong and dE/dx distribution becomes wider. The template distributions of energy deposition probability for different type of particles are needed to compute the likelihood are obtained from simulation. Fig.4.22 shows the comparison of the longitudinal dE/dx between electron and muon at two different depths in the particle propagation [71], [72]. The reconstructed vertex is defined to be in the plane 0 in the distribution. At plane 2, the deposition rate of the electron is close to the muon, which is a minimum-ionizing particle, while at plane 10, electron is no longer minimum ionizing and the dE/dx distribution is wider than the muon. The probability density distributions are constructed for different type of particles at different longitudinal and transverse depths. They are all normalized to 1 to match the condition of a probability density function. The hits within 2.5 cell-widths of the shower core for the first 8 planes and 20 cellwidths for the rest of the planes are used to calculated likelihoods.

As energy deposition is critical in the calculation, a refining method is implemented in order to correct the energy that is duplicated in the situation of a cell belonging to multiple showers.

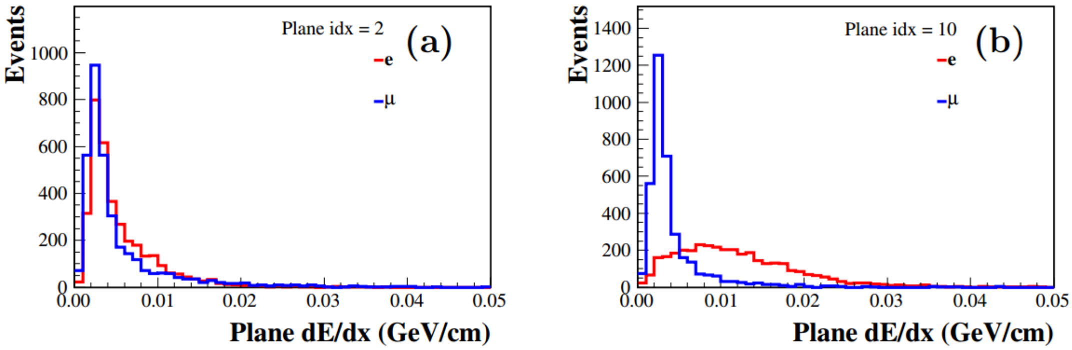


Figure 4.22: Comparison of longitudinal dE/dx distributions between electron and muon at two different depth in the particle propagation, plane index 4 (left) and index 10 (right).

The corrected energy is calculated as Eq.4.18,

$$E_i^{cell} = \frac{PE_i^{cell}}{atten_i} \frac{PE_i^{shower} e^{-D_i/\lambda}}{\sum_j PE_j^{shower} e^{-D_j/\lambda}} \tag{4.18}$$

where a_i is the attenuation correction factor of a particular cell, PE_i^{shower} is the total number of photoelectrons (PE) of the shower and D_i is the distance of the cell from the shower core. According to the simulation, the transverse energy deposition rate per shower energy follows $exp(-D_i/\lambda)$, as shown in Fig.4.23, which represents the energy deposition rate per shower energy as a function of transverse extension, which fits well with the fitted exponential function. So $PE_i^{shower} e^{-D_i/\lambda}$ represents the expected PE deposited in the cell and the ratio is the fraction of PE from the i^{th} shower. Once the probability of each hit with deposition rate dE/dx is calculated using the template distribution, the likelihood is obtained through Eq.4.19. The difference between the likelihood of being an electron and being another type of particle (muon, gamma, charged or neutral pion, neutron and proton), $LL(e^-) - LL(x)$, are used as inputs of the ANN. Fig.4.24 is an example of the likelihood difference between electron and muon hypotheses in both transverse and longitudinal directions. The distribution shows a clear separation between electron and muon in the longitudinal likelihood values. On the other hand, Fig.4.25 shows an example of the electron and gamma likelihood difference, which, due to the closeness in the shower-like propagation pattern, the separation is not as obvious as in the electron to muon comparison [72].

$$LL(x) = \sum_i^N log(P(x))/N \tag{4.19}$$

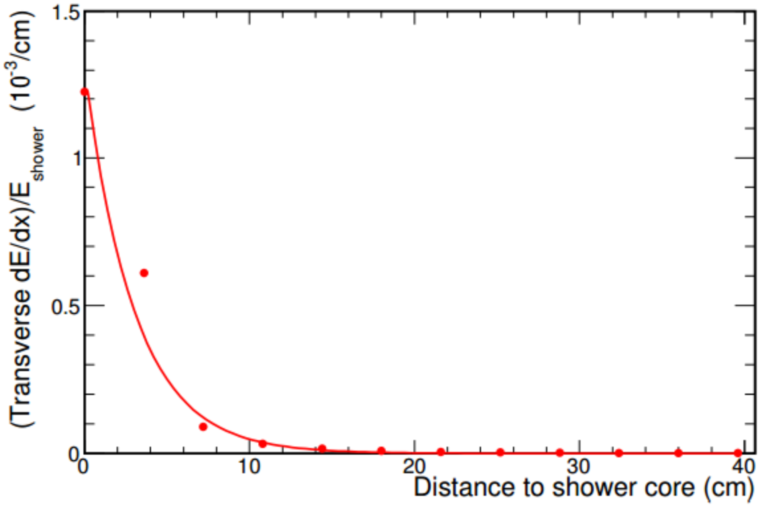


Figure 4.23: Energy deposition rate per electron energy at different transverse depth based on simulation, which is supposed to follow exponential relation (red).

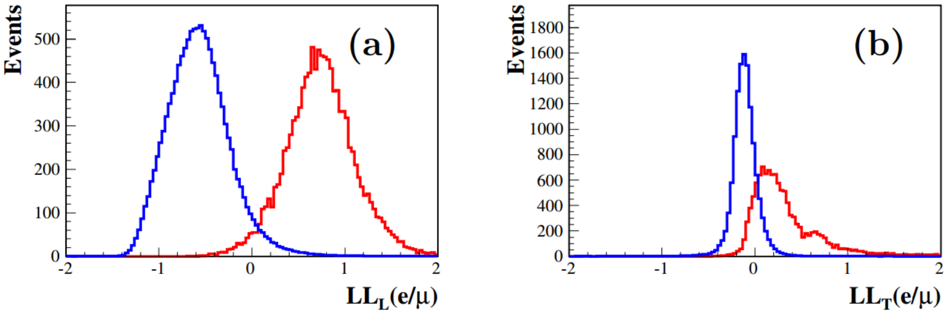


Figure 4.24: $LL(e) - LL(\mu)$ in longitudinal (left) and transverse (right) directions, when the true particle is electron (red) and muon (blue).

The likelihood of nine particle hypotheses, including electron, are considered and they form a total 16 of likelihood differences. In addition to the 16 likelihood difference variables, some other basic reconstruction variables have the power of separating electron from other events and are included in ANN training:

Shower energy fraction: the fraction of energy in the most energetic shower to total event energy. Because of the limit in the detector resolution, the recorded inelastic interaction is expected to have a smaller shower energy fraction than the quasi-elastic interaction.

π^0 **mass:** invariant mass of the most energetic shower combined with all other showers. The variable is used to label and reject potentially the π^0 events, which is one of the main

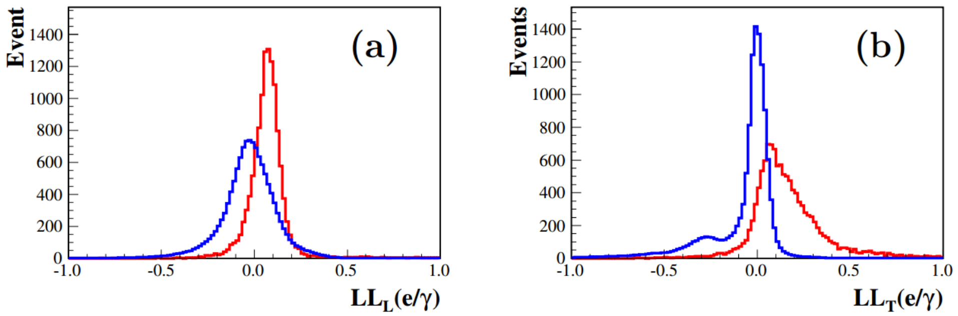


Figure 4.25: $LL(e) - LL(\gamma)$ in longitudinal (left) and transverse (right) directions, when the true particle is electron (red) and gamma (blue).

products in neutrino neutral current interactions.

Vertex energy: total energy in ±8 planes from the interaction vertex. The inelastic interaction tends to produce hadronic shower which deposits larger amount of energy at a relatively short propagating distance than the EM shower or muon track.

Gap: distance between the start point of the most energetic shower and the main vertex. Charged particles deposit energy immediately after they are produced in the detector, while neutral particle do not and tend to have a gap between the main vertex and the shower.

cosθ: cosine of the angle of the most energetic shower with respect to the beam direction. The cosmic ray events enter the detector mainly at large angle with respect to the horizontal direction, and hence the variable is used to reject cosmic ray events. It also has the power of separating ν_e CC from NC interaction, the product of which is expected to be scattered at a larger angle.

The method is tuned and tested on two independent simulation samples with $sin^2 2\theta_{13} = 0.1$, maximum θ_{23} mixing and without matter effect or CP violation in the oscillation weights. Fig.4.26 shows the LID distribution for the main interactions: ν_e CC, ν_μ CC, NC and beam ν_e CC, in the NOνA far detector. While oscillated ν_e CC interaction peaks at 1, all the other interactions concentrate at the low end of the LID axis. The arrows show the potential LID cuts to select oscillated ν_e CC events. In NOνA's first ν_e appearance analysis the LID cut is set to be: $LID > 0.95$, which corresponds to the maximum value of $s/\sqrt{b}$, in which s is the

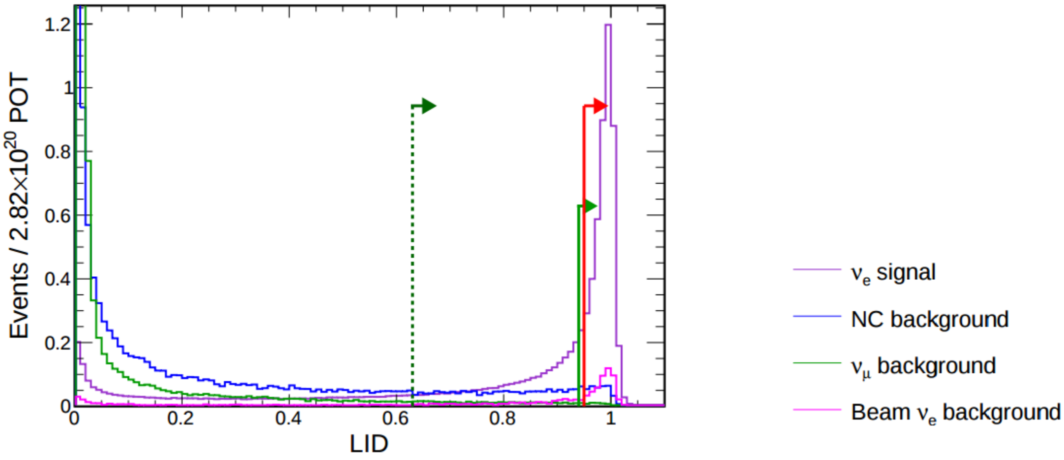


Figure 4.26: LID distribution for different types of interactions, ν_e CC, ν_μ CC, NC and beam ν_e CC. Dash line shows the optimal cut tuned using $s/\sqrt{s+b}$, while solid line represents the cut by $s/\sqrt{b}$. Red line is the LID cut used by the first analysis.

total number of oscillation ν_e CC signal events and b is the total number of background events. After all the other selection cuts are applied [73], the efficiency and purity of the LID cut for oscillated ν_e CC is 47.68% and 83.69% respectively. The efficiencies of $LID > 0.95$ for the other interactions are 3.50% for total beam background and 0.17% for cosmic ray background.

4.3.2 LEM

Another ν_e CC interaction idenfication algorithm used in NOνA is Library Event Matching (LEM). The principle of LEM is to compare the input trial event to a large number of known library events, which consists of both signal and background interactions. The final classification decision is made based on the collective properties of the matched library events. The comparison is made based on the basic cell energy depositions and the matching function is motivated from electrostatics:

$$E = E_A + E_B + E_{AB} \tag{4.20}$$

in which E_A and E_B represent the ”self energy” of the trial and library event, respectively. E_{AB} is the interactive energy of the trial and the library events. The calculation for each term is shown in Eq.4.21 [74],

$$
\begin{aligned}
E_A &= \frac{1}{2}\sum_{ij}^{cells} a_i^{\beta} T_{ij} a_j^{\beta} \\
E_B &= \frac{1}{2}\sum_{ij}^{cells} b_i^{\beta} T_{ij} b_j^{\beta} \\
E_{AB} &= -\sum_{ij}^{cells} a_i^{\beta} T_{ij} b_j^{\beta} \\
T_{ij} &= \left(\frac{\Delta p_{ij}^2}{\sigma_p^2} + \frac{\Delta c_{ij}^2}{\sigma_c^2}\right)^{-\alpha/2}
\end{aligned}
\tag{4.21}
$$

which can be understood as a modified version of the electrostatic energy. The cell energy deposition in the trial (a_i) and library events (b_i) perform similar to the charges in the energy calculation, in which β power is a tunable parameter in the method. T_{ij} represents the distance between a_i and b_i in terms of the number of cells, Δc_{ij}, and planes, Δp_{ij}. The falloff with distance is also generalized to $r^{-\alpha}$ to achieve a better sensitivity. The parameters are optimized to be: $\sigma_p = 0.286,\ \sigma_c = 0.095,\ \alpha = 0.25,\ \beta = 0.5$. With the idea of matching energy, the better two events match, the lower total "electrostatic" potential energy will be. Fig.4.27 shows an example of LEM energy match for a ν_e CC interaction in XZ and YZ view. The color in the z scale of the middle plot represents the "electrostatic" energy between the trial and library events. The warmer the color is, the higher the energy is. As both the trial and library events have two prongs, the energy distribution also shows highlights in the two particular directions. If the library has infinite number of template events, the fraction of the total matched events that are truly ν_e CC events is a good measurement of how close the trial event to the signal. However, in practice, we only select the best matched 1000 library events. In order to increase the sensitivity of ν_e CC identification, a decision tree is used, to which six variables describing the average properties of the 1000 matched events are fed to make the final score for the trial event. These are:

Weighted fraction of signal matches: each matched event is weighted according to its energy and the oscillation probability (Eq.4.22).

$$
w'_n = P_n^{cos} exp\left(-\lambda \left(\frac{E_n}{E_{1000}}\right)^{\gamma}\right) \tag{4.22}
$$

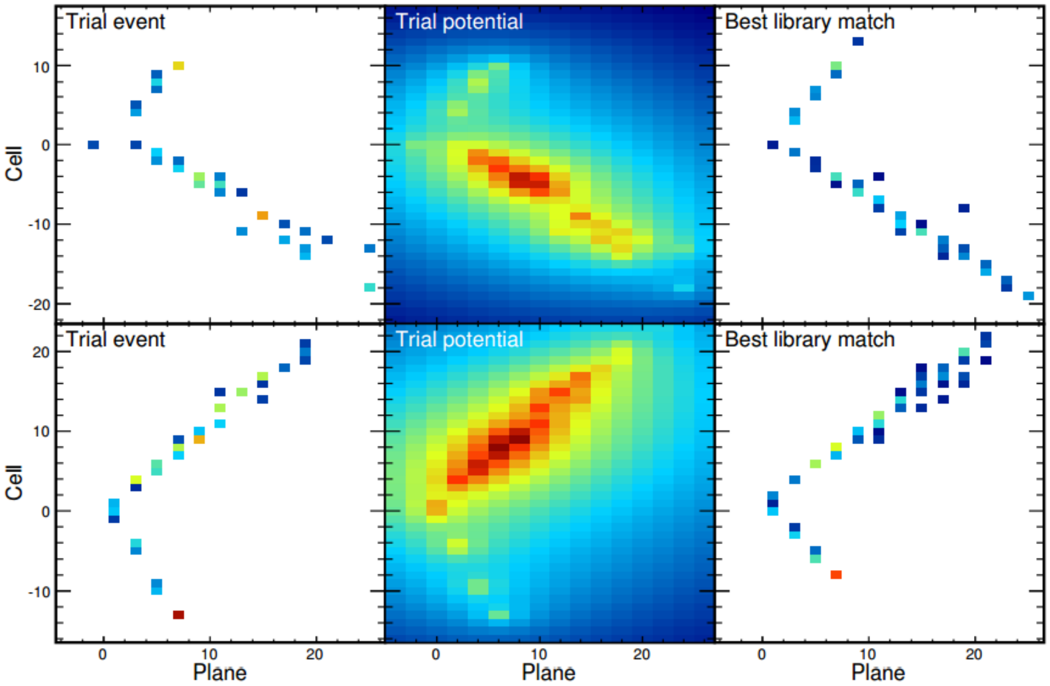


Figure 4.27: energy matching example of a ν_e CC interaction in XZ (upper) and YZ (lower) views. Left side is the trial event, the electrostatic potential field of which is in the middle. The right hand side is the matched library event.

where $\lambda = 6.67$, $\gamma = 10$ are optimized for the ν_e CC classification for NOνA. E_n and E_{1000} are the "electrostatic" energy of n^{th} and 1000^{th} best matched events. So the lower the electrostatic energy the matched event has, the better it matches and the greater weight it will be assigned to. The weighted signal fraction is calculated in Eq.4.23:

$$f_{sig} = \frac{\sum_{n,sig} w_n}{\sum_{n,tot} w_n} \tag{4.23}$$

Mean hadronic y: in a neutrino interaction, the outgoing electron is accompanied by some hadronic components. A good variable to distinguish ν_e CC from the other interactions is the the ratio of kinematic energy of the producing hadrons (K_h) to the kinatic energy of the incident neutrino (K_ν), $y = K_h/K_\nu$. The variable used in the decision tree regarding this effect is:

$$\langle y \rangle = \frac{\sum_{n,tot} w_n y_n}{\sum_{n,tot} w_n} \tag{4.24}$$

Mean matched charge fraction: the fraction of charge in cells that exist in both trial and library event. The more the trial and library event overlap, the greater the value of the average matched charge fraction will be.

$$\begin{aligned} f_Q &= \frac{2\sum_i^{cells} min(a_i, b_i)}{\sum_i^{cells}(a_i + b_i)} \\ \langle f_Q \rangle &= \frac{\sum_{n,tot} w_n f_{Q,n}}{\sum_{n,tot} w_n} \end{aligned} \tag{4.25}$$

Matched energy difference: the difference between the electrostatic energy of the true signal and the true background.

Enriched fraction: the fraction of the signal in a π^0-enriched library. Because π^0 from neutrino NC interactions is an important background for the ν_e CC identification, this variable has enhanced the power of removing π_0 background events.

Total calorimetric energy: the sum of all the energy deposited in the trial event, $E_{cal} = \sum_n^{cells} a_i$. NC interaction tends to have a low E_{cal} due to the outgoing neutrino, while E_{cal} of the intrinsic ν_e CC is usually high in the visible energy. So the calorimetric energy variable can be used to separate the oscillated ν_e CC apart from some background interactions.

Fig.4.28 shows the LEM distributions for ν_e CC, ν_μ CC, NC and beam ν_e CC interactions. The cut represented by the red solid line corresponds to the maximum $s/\sqrt{b}$ and is used by the first analysis. With the cut, $LEM > 0.8$, the efficiency and purity of the signal is 46% and 83%. The selection efficiency for background events is 3.1% for total beam background and 0.1% for cosmic ray background.

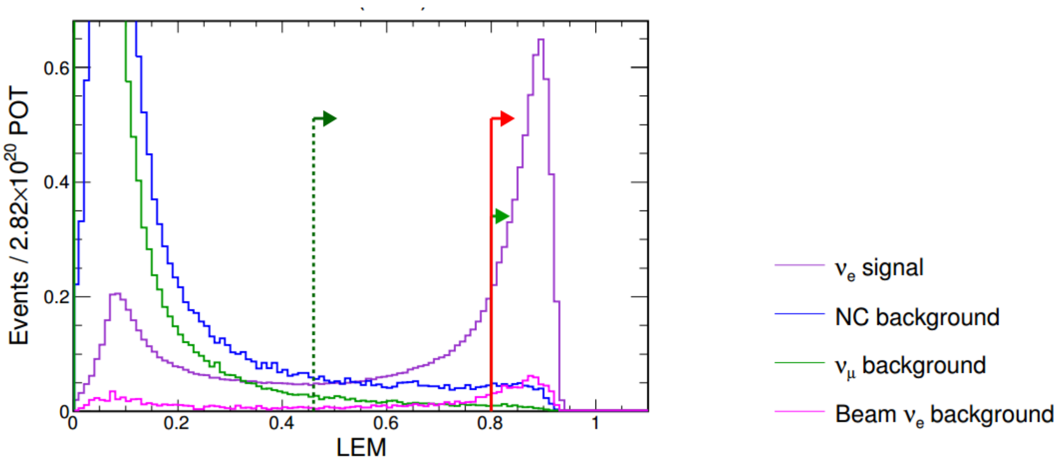


Figure 4.28: LEM distribution for different types of interactions, ν_e CC, ν_μ CC, NC and beam ν_e CC. Dash line shows the optimal cut when tuned using $s/\sqrt{s+b}$, while solid line represents the cut tuned using $s/\sqrt{b}$. Red line is the LEM cut used by the first analysis.

CHAPTER 5. NOνA DATA AND SIMULATION

5.1 The NOνA Data

In the NOνA experiment data acquisition system there are several type of triggers, each designed for specific purposes. For ν_e appearance analysis, two triggers are mainly used: NuMI beam spill trigger and cosmic trigger.

The **NuMI beam spill trigger** is a 550 μs window centered around the NuMI beam spill which spans 12 μs in time. The 30 μs window centered around the beam spill was defined as “in-time” window, while the region outside the 30 μs is considered as out-of-time window. A neutrino in the in-time window has a high probability of being observed. The in-time window is usually blind to the analysis until the box opening. The out-of-time data is used as side-band check and background prediction for the in-time data. Fig.5.1 shows the NuMI beam induced timing peak observed in both far and near detector (FD and ND) in the NuMI beam spill trigger data [76], [77]. Selection cuts are applied to the far detector data in order to reduce the impact of the cosmic ray background events. As the far detector is much further away from the NuMI beam than the near detector, the statistics in the timing peak are much lower in the far detector than in the near detector. For both far and near detector, the main spill peaks are well positioned in the nominal 208 - 238 μs in-time region.

The **Cosmic trigger** is a periodic 550 μs window, which collects mostly cosmic ray data outside of the spill trigger window. Because the far detector is on the surface, the cosmic ray events are abundant in the FD cosmic trigger data and are used to understand the performance of the far detector, including detector calibration, tuning of detector response simulation and studying cosmic ray background rejection. The counterpart to the cosmic trigger in the near detector is the “all activity” trigger, which records all activities in a periodic 500 μs trigger

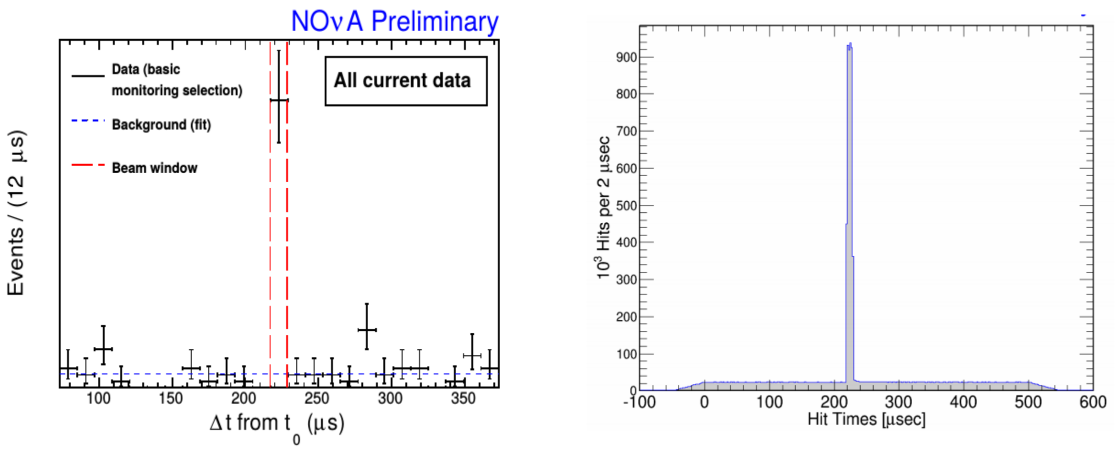


Figure 5.1: NuMI beam induced timing peak observed in FD (left) and ND (right) spill trigger data.

window. The all activity trigger selects mostly rock events in ND. Due to the high intensity of the beam events in the near detector, the all activity trigger provides a rich sample to study the detector performance.

The construction of the NOνA far detector started at the beginning of 2014 and the first diblock was completed in April, 2014. The segmentation design of the detector allowed the data collection during the construction. The data collected between February 6, 2014 and May 15, 2015 with at least four contiguous diblocks, is used in the first ν_e appearance analysis to ensure the completeness of the neutrino events recorded. The total accumulative exposure is 3.45×10^{20} protons on target (POT), which is equivalent to 2.74×10^{20} POT with full detector [75]. Fig.5.2 shows the cumulative POT exposure increases as a function of date. The flat region from September 5 2014 to October 24, 2014 corresponds to the period when the NuMI beam was shut down for upgrade and maintenance and only cosmic trigger data is recorded. The dataset is split by the NuMI shutdown activity as pre-shutdown and post-shutdown period. Due to a known but undetectable miscalibration in the timing system before NuMI shutdown, the far detector pre-shutdown data (23 % of the total first analysis data) requires an additional spill window 64 μs after the nominal spill window. The effect can be seen in Fig.5.1 where a smaller peak appears at the bin around 285 μs. The cumulative livetime exposure is 227.5 s and it is equivalent to 158.1 s in the full 14 kton detector.

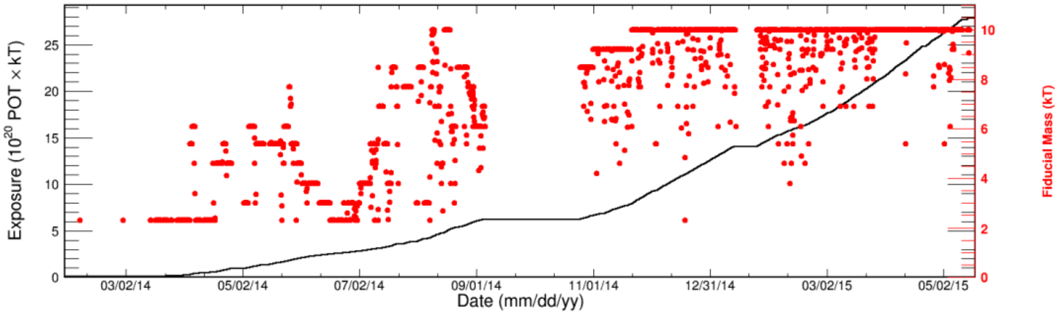


Figure 5.2: Cumulative POT (black) and detector fiducial mass (mass) as functions of date.

5.2 The NOνA Monte Carlo Simulation

Monte Carlo (MC) simulation is needed to help us understand the picture of particle interactions in the NOνA detectors. The detector performance such as detection efficiency can be examined by comparing data to the prediction by MC. As in a blind analysis, analysis tools such as reconstruction and particle identification algorithms can be developed and optimized. In addition the rate of beam events, including ν_e CCsignal and beam background events are predicted using MC samples. The NOνA simulation package has mainly two parts: the estimate of number of different types of events before entering the detector and the detector response after the particles entering the detector.

Before particles entering the detector, for beam events, the composition of neutrino flux at the NuMI beam is firstly predicted using the FLUGG software [79][80], which requires FLUKA for beam behavior simulation, such as including the hadron production in the target, the horn focusing and the decay of pions and kaons in the decay pipe, and the 4.9.6.p03c release of GEANT4 [81] to build the geometry interface. The output of this step is a spectrum of outgoing neutrinos and the fraction of different neutrino flavors. The challenge of achieving an accurate beam neutrino estimate come from two main categories: hadron production which refers to mainly the pion and kaon prediction after protons hitting the graphite target and beam transport which refers to the accuracy of modeling the condition of specific NuMI components. (See Sec.7.3.1) The interaction of the beam neutrinos with surrounding materials outside the NOνA detectors is then simulated using GENIE (version 2.8.0i) [82], which based on the the kinematics of simulated beam neutrinos determines whether a neutrino interaction can occurs, which type of interactions, QE, resonance or DIS, a neutrino will interact through and the 4-vector kinematic and position information of the secondary particles. On the other hand, the cosmic ray events before entering the NOνA detectors are simulated with CRY (version 1.7b) [83]. the output of which is the flux of different types of cosmic ray induced particles: including muons, photons, neutrons and other hadrons, at the level of the far detector.

After the simulated beam and cosmic ray particles enter the detectors, GEANT4 [84][85] is used to simulate the propagation of the particles inside the NOνA detectors and to calculate the amount of energy loss in each of the NOνA cells. For each energy deposition in NOνA cell, the downstream scintillation, light transport, collection as well as the readout system of each cell is modeled by a parameterized front-end simulation package. The rest of this section will focus on the simulation in the last step which is specifically formulated and parameterized for the NOνA detector system:

Scintillator Light: Ideally, the yield of scintillator light is proportional to the energy deposited by the charged particles. Although, in reality, especially for organic scintillator due to the effect of recombination, the actual light yield is lower than the ideal one. NOνA simulation corrects this effect to the second order, which is known as the Birks-Chou model as shown in Eq.5.1,

$$\frac{dL}{dx} = \frac{L_0 \frac{dE}{dx}}{1 + k_B \frac{dE}{dx} + k_C \left(\frac{dE}{dx}\right)^2} \tag{5.1}$$

where L_0 is the ideal light yield per unit of deposited energy. The constants k_B and k_C are studied to be 0.040 $gMeV_{-1}$ and -0.0005 cm^2MeV^{-2}, respectively, this combination of parameters giving the best match of the dE/dx distributions as a function of planes between data and MC for the selected ν_μ CCevents [86].

Photon Transport: After the light is emitted in the scintillator, it will drift in the scintillator or reflect off the PVC walls before being collected by fiber. Fig.5.3 shows the collection rate (colored z scale) as a function of traveled time (x axis) and distance along the cell (y axis), which is used in the first analysis as a template to simulate the light capture process, for which a complete simulation is very time-consuming. Bins with short distance and short time (bottom left corner) are corresponding to the situation in which

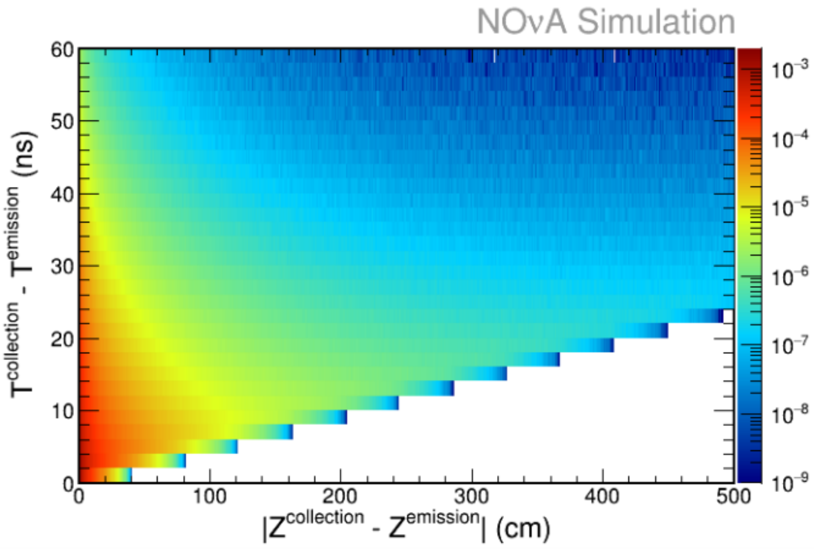


Figure 5.3: Collection rate as a function of traveled time and distance along the cell.

light is collected immediately, while the photons with short distance and long time (top left corner) on the other hand may bounce off the PVC walls several times before the collection. The empty bottom right corner is showing the limit of the speed of light in the fiber. Each emission is convoluted with this 2D distribution to predict the collection position and time of the emitted photon. After collection, photons travel to the readout with some level of attenuation which is formulated as a double-exponential function of the distance from the readout.

Photon Detection: Due to the attenuation in the fiber, a fraction of photons are lost before reaching the APD photodetector. It is assumed that the number of photoelectrons (PE) generated by the captured photons is N_{pe}, which includes the effect of quantum efficiency and poisson flucutation of the APD. The excess noise of the APD is modeled by a log-normal distribution with mean 1 and variance $(F-1)/N_{pe}$, in which F is the noise factor. Another effect, the “Sag” effect, is also found to be noticable in the analysis however it is well modeled in the simulation. Due to the fact that all pixels in one APD share the same power supply, as one pixel receives large amount of energy deposition and large power is driven from the power supply, a drop in baseline will appear in all the other pixels. The drop sometimes is large enough, ~ 0.0186 times the size of pulse due to particle energy deposition. It will be reconstructed as an energy deposit and thus cause additional hits in an event. In the first ν_e analysis of NOνA, in order to reduce the processing time, instead of a complete modelling, the "Sag" effect is only simulated in the pixels that contain an energy deposit within the simulation time window.

Electronics Simulation: The pulse shaping function in the ASIC is modeled by Eq.5.2,

$$f(t) = N_{pe} \frac{F}{F-R} (e^{-(t-t_0)/F} - e^{-(t-t_0)/R}) \tag{5.2}$$

in which R and F are the rising and falling time of the shaping RC-CR circuit. The electronics noise is simulated by the sum of two Gaussian Markov variables and is added to the trace measured by the APD. After having the full trace including both signal and noise, an ADC conversion is applied at each sample point through a dual correlated sampling (DCS) process, $DCS_i = ADC_i - ADC_{i-3}$, in which i denotes each sample point. (See Sec.3.2) Finally, the peaks above the thresholds that are measured based on real data are selected as indications of real energy depositions. After selecting pulses, the process of MC will be the same as that of data, including hit and event reconstruction, particle identification and event selections.

5.3 Data vs MC Comparison

5.3.1 ND Data vs MC Comparison

The comparison between data and MC was made to understand and improve the performance of the detectors as well as to tune the simulation to match the real detector conditions. A good data/MC agreement is important in the success of the analysis. Fig.5.4 through Fig.5.9 show the ND data versus total mc distributions for basic reconstruction variables for all the ND events that pass data quality selections. Fig.5.4 shows a good data/MC agreement in the selected region, $20 < nhit < 200$. Similarly, data and MC agrees well for the calorimetric energy in the interesting region, 1 - 3 GeV. The discrepancy in the calorimetric energy above 3 GeV is caused by the missing Birk effect for through-going muons in MC simulation, the impact of which is evaluated in Sec.7.3.2. The agreement is also demonstrated to be good for the length of longest prong (Fig.5.6), vertex position in the X, Y and Z coordinates (Fig.5.8, Fig.5.9 and Fig.5.7). Fig.5.10 and Fig.5.11 are the comparison in LID and LEM for the pre-selected events and data and MC also prove to agree well for both PIDs.

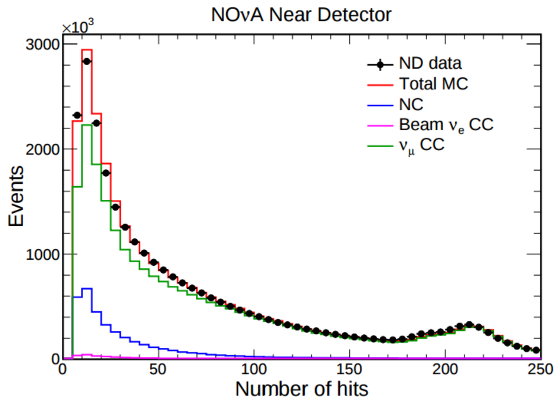


Figure 5.4: The ND data(black) versus total mc(red) distributions of the number of hits in slice for all events.

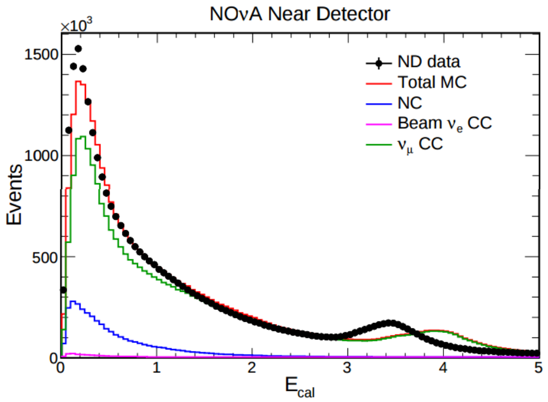


Figure 5.5: The ND data(black) versus total mc(red) distributions of the calorimetric energy of slice for all events.

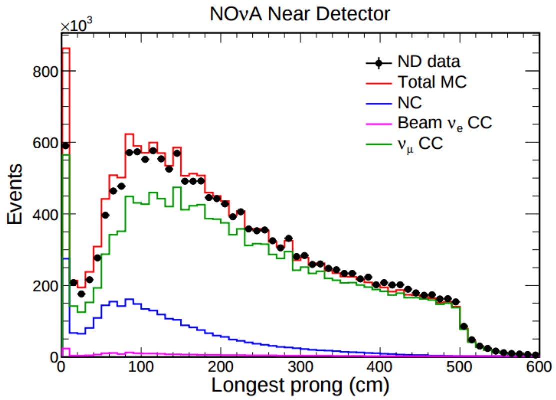


Figure 5.6: The ND data(black) versus total mc(red) distributions of the length of the longest prong in slice for all events.

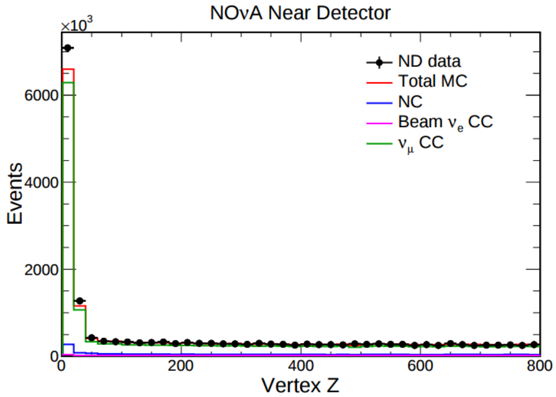


Figure 5.7: The ND data(black) versus total mc(red) distributions of vertex in z coordinate for all events.

5.3.2 FD Data vs MC Comparison

Because the beam events in FD is blinded, during the optimization of all analysis tools, the FD beam event prediction is based on the simulation. The data vs MC comparison is also checked in the far detector using cosmic ray events from FD cosmic trigger data. Fig.5.12 shows the comparison for the angle of the leading shower with respect to the beam direction

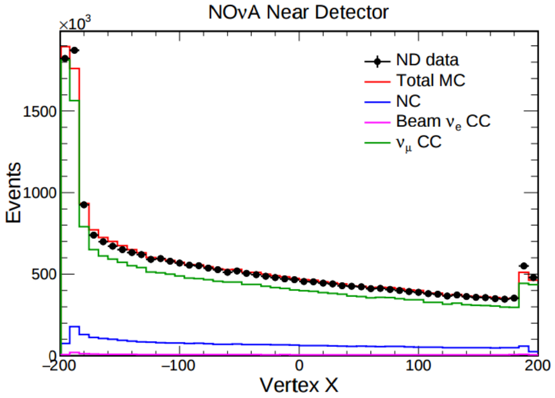


Figure 5.8: The ND data(black) versus total mc(red) distributions of vertex in x coordinate for all events.

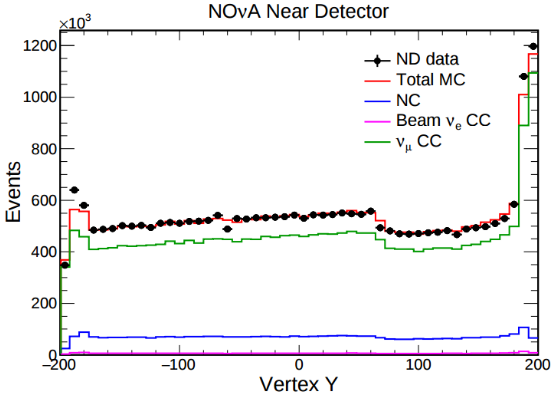


Figure 5.9: The ND data(black) versus total mc(red) distributions of vertex in z coordinate for all events.

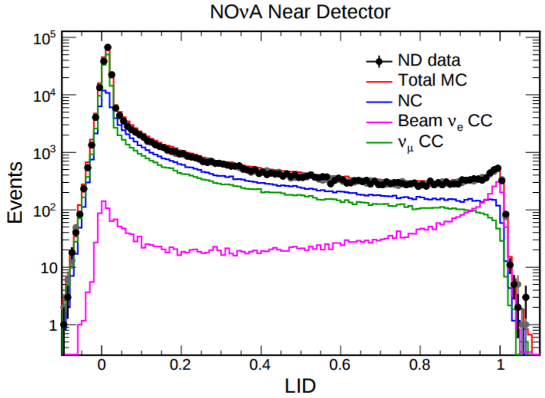


Figure 5.10: The ND data(black) versus total mc(red) distributions of LID for pre-selected events.

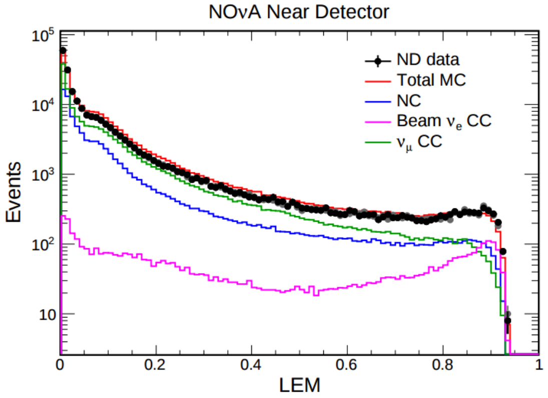


Figure 5.11: The ND data(black) versus total mc(red) distributions of LEM for pre-selected events.

for events passing the data quality cuts (See Sec.6.3.1), which are designed to remove some reconstruction failures. The agreement is good for the small angle range, $|cos\theta| > 0.5$, where the signal events are typically found. Fig.5.13 through Fig.5.17 show the distributions for basic reconstructed variables, number of hits, calorimetric energy of slice and vertex position in x, y and z coordinates The data and MC agree well for these variables for the events with small angle with respect to the beam direction as well as with vertices contained in the detector. Fig.5.18, Fig.5.19 and Fig.5.20 are showing the comparison for transverse and longitudinal energy deposition rate as well as the number of planes in shower, the variables that are used in the calculation of LID. Fig.5.21 shows the distributions for the output LID.

They all demonstrate to have a good data versus MC agreement that the difference between data and MC is consistently lower than 5% of the data for the interesting events that has small angle between the most energetic shower with respect to the beam direction is small and the vertex is contained.

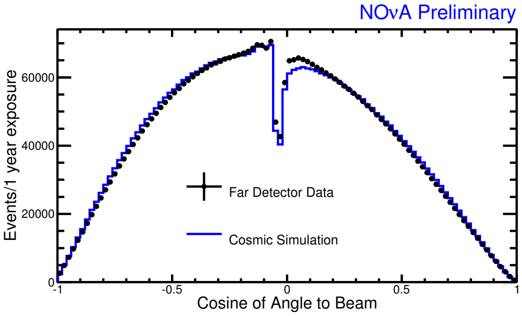


Figure 5.12: The FD data(black) versus mc(blue) distributions of the angle of the leading shower with respect to beam direction.

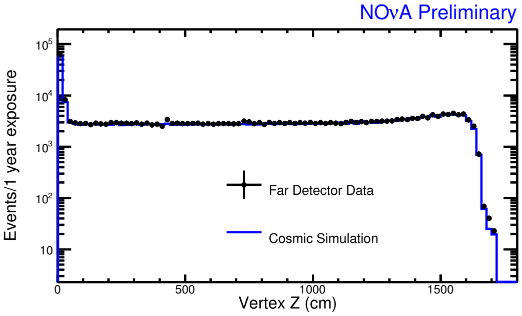


Figure 5.13: The FD data(black) versus mc(blue) distributions of vertex in z coordinate.

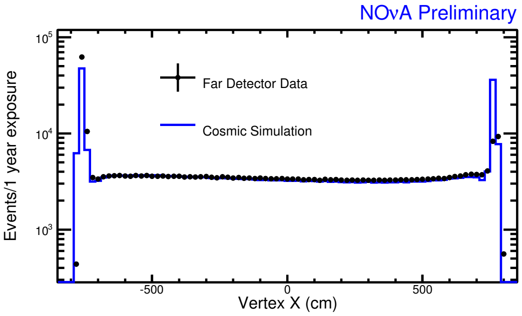


Figure 5.14: The FD data(black) versus mc(blue) distributions of vertex in x coordinate.

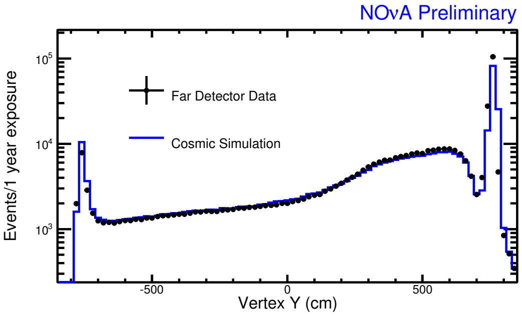


Figure 5.15: The FD data(black) versus mc(blue) distributions of vertex in y coordinate.

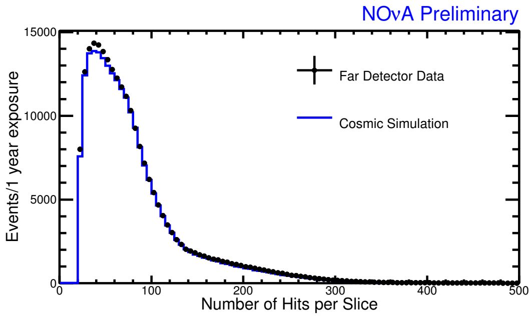


Figure 5.16: The FD data(black) versus mc(blue) distributions of the number of hits in slice.

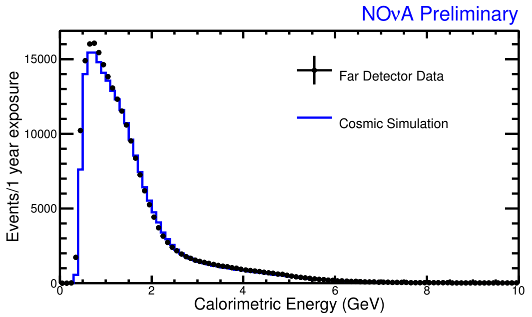


Figure 5.17: The FD data(black) versus mc(blue) distributions of total calorimetric energy of slice.

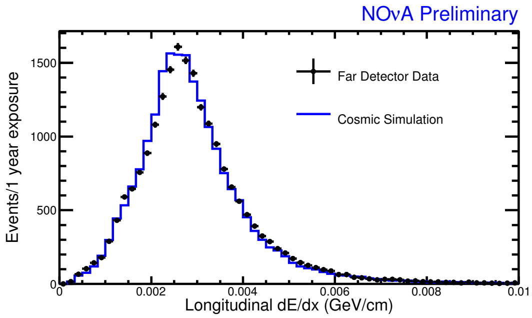


Figure 5.18: The FD data(black) versus mc(blue) distributions of longitudinal dE/dx.

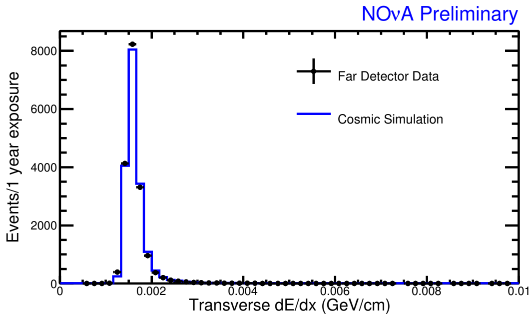


Figure 5.19: The FD data(black) versus mc(blue) distributions of transverse dE/dx.

Since the signal of the analysis is ν_e CC interaction, for which the signature is an EM shower induced by the outgoing electron, a more sophisticated comparison is performed to check specifically EM shower model using cosmic muon-induced EM showers. After high-energy cosmic muons enter the detector, they produce gammas through Bremsstrahlung radiation and also decay into electrons. Both the gammas and electrons undergo EM showers during the propagation. The study is divided into four steps: muon track selection, shower finding, muon removal and shower reconstruction and PID [103].

Muon track selection: The cosmic muon selected should fulfill three conditions: have both start and stop points outside the detector; present a not very large angle with respect to the beam direction so that the remaining prongs after the muon removal process will be close to the beam direction; penetrate at least 30 planes.

Shower finding: Once the cosmic muon is found, a shower finding algorithm is performed to

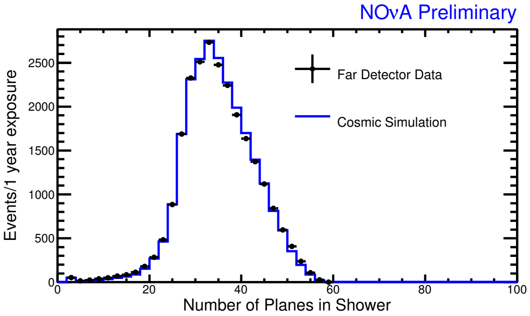


Figure 5.20: The FD data(black) versus mc(blue) distributions of number of planes in slice.

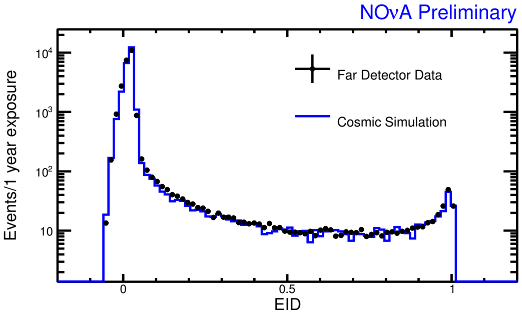


Figure 5.21: The FD data(black) versus mc(blue) distributions of LID.

select muon with showers attached. The search is done by checking the energy deposition in each plane. A consistant excess in the energy deposition rate beyond MIP rate suggests the existance of a EM shower, which is tagged for next step.

Muon removal: Muon track is removed from the events leaving only shower in the event at the raw hit level [104]. In the region where muon track and shower overlap, instead of removing the entire hit, the PE value of the hit is reweighted so that only the energy deposition from muon is removed. Fig.5.22 and Fig.5.23 shows a cosmic muon event before and after the muon removal algorithm. by which the EM shower is extracted.

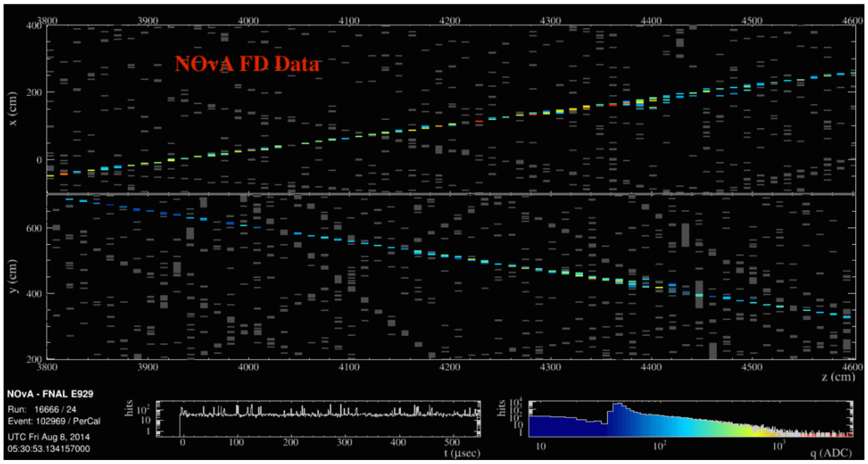


Figure 5.22: Event display of a selected cosmic muon event in the FD cosmic trigger data before muon removal.

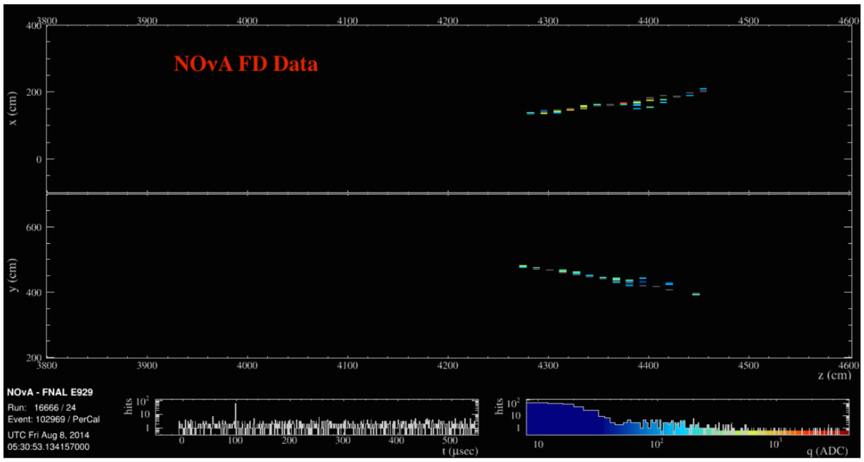


Figure 5.23: Event display of a selected cosmic muon event in the FD cosmic trigger data after muon removal.

Shower reconstruction and PID: The event after muon removal is processed by standard reconstruction and PID and data/MC comparison is performed for the resulting reconstruction and PID variables. Both the muon removed cosmic data and MC samples are reweighted based on shower energy and shower angle so that the muon removal sample has the same the leading shower energy and angle distributions as the ν_e CCsignal sample (See Fig.5.24 for shower energy and Fig.5.25 for cosine of the angle between the leading shower and the beam direction). Fig.5.26 through Fig.5.29 show the data/MC comparison for some key reconstructed shower variables, including shower width, shower length, the number of planes in shower and the number of hits in shower. All these shower variables show generally good agreement that the difference between data and MC is consistently lower than 5%. Fig.5.30 is the distribution of LID variable, according to which, for the examined EM showers, LID peaks near 1 and MC agrees with data well. Fig.5.31, Fig.5.32 and Fig.5.33 present the reconstruction efficiency as a function of reconstructed vertex position in X, Y and Z for showers passing LID > 0.7. The plots demonstrate a fairly stable reconstruction efficiency across the detector and a good agreement between data and MC in the efficiency.

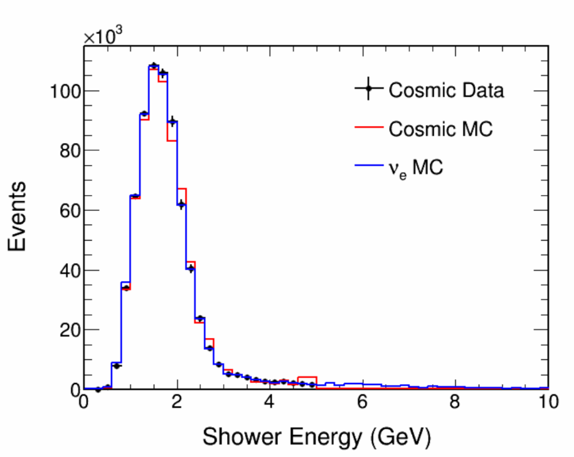


Figure 5.24: The data (black) vs MC (red) comparison in the reconstructed shower energy.

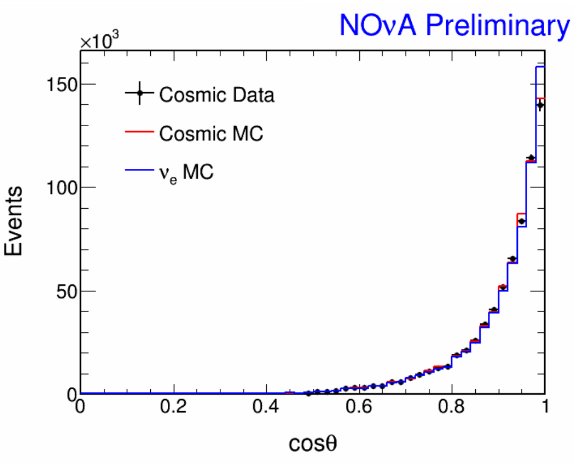


Figure 5.25: The data (black) vs MC (red) comparison in the cosine of the angle of the EM shower with respect to the beam direction.

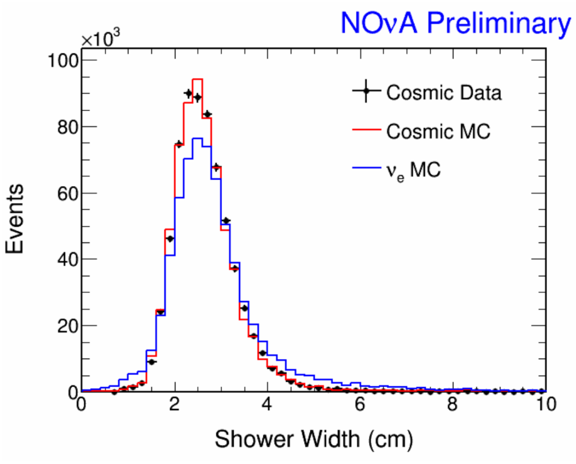


Figure 5.26: The data (black) vs MC (red) comparison in the reconstructed shower radius.

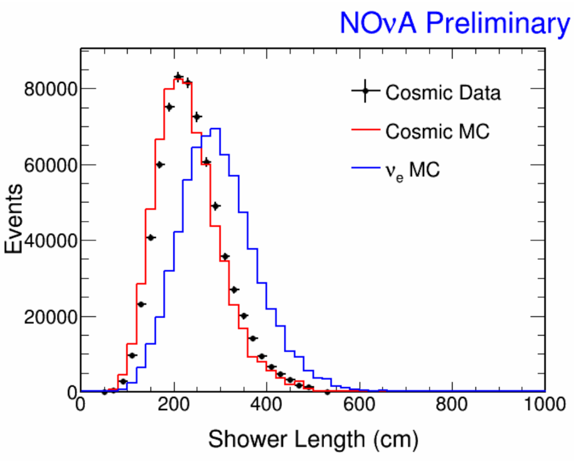


Figure 5.27: The data (black) vs MC (red) comparison in the reconstructed shower length.

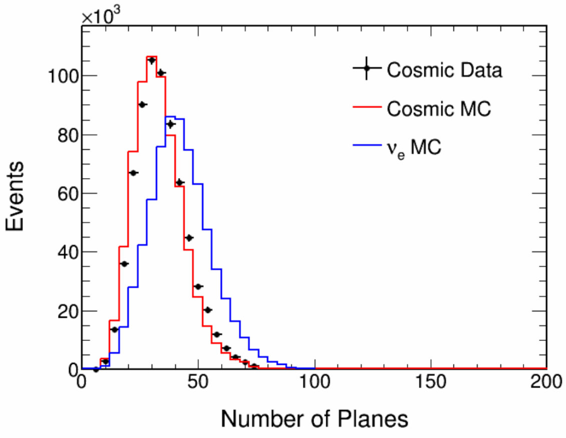


Figure 5.28: The data (black) vs MC (red) comparison in the number of planes in shower.

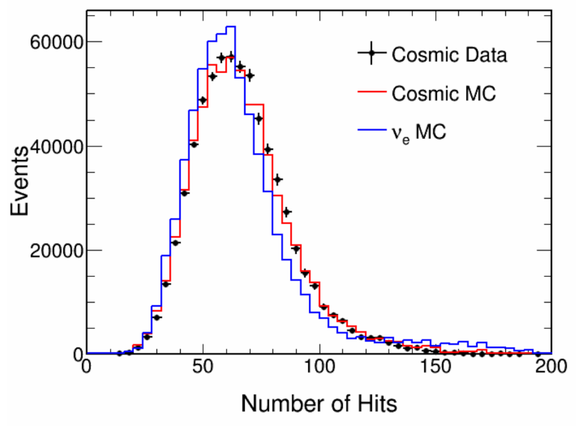


Figure 5.29: The data (black) vs MC (red) comparison in the number of hits in shower.

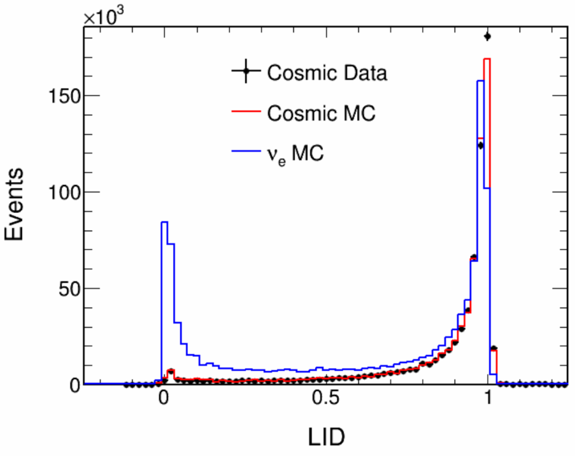


Figure 5.30: The data (black) vs MC (red) comparison in the number of hits in shower.

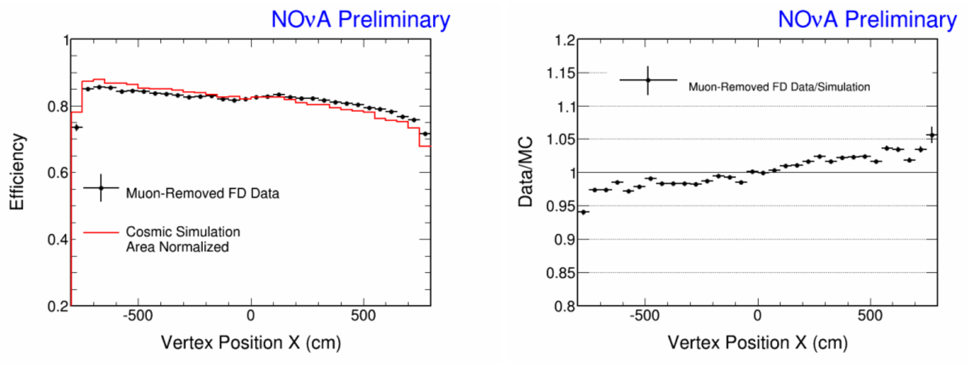


Figure 5.31: The data (black) vs MC (red) comparison in the reconstruction efficiency as a function of vertex position in X for events passing LID > 0.7.

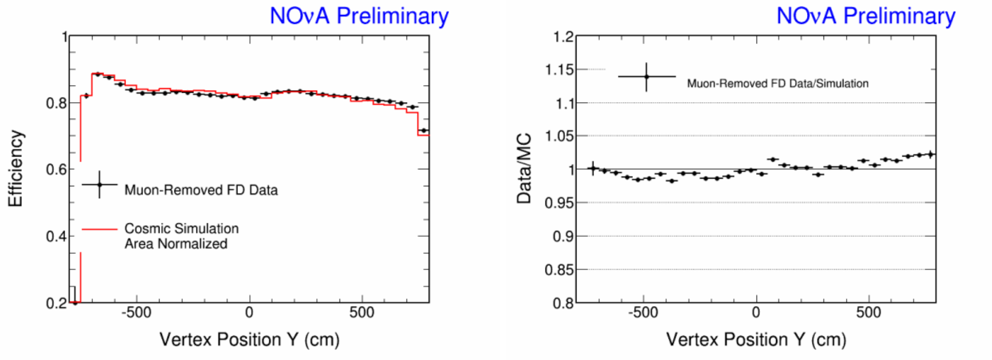


Figure 5.32: The data (black) vs MC (red) comparison in the reconstruction efficiency as a function of vertex position in Y for events passing LID > 0.7.

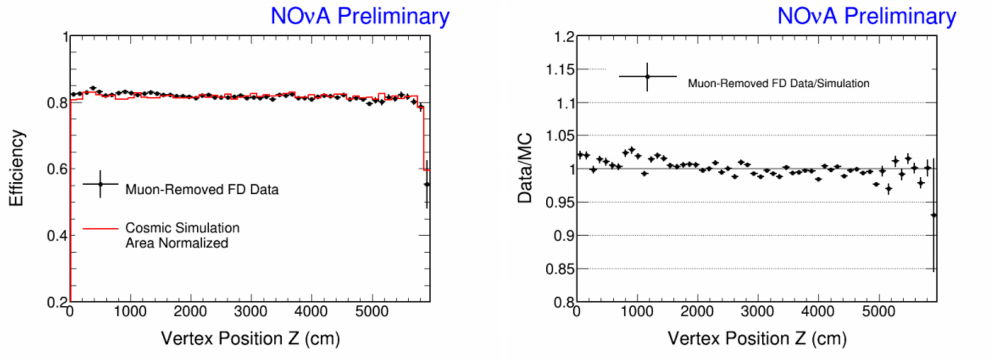


Figure 5.33: The data (black) vs MC (red) comparison in the reconstruction efficiency as a function of vertex position in Z for events passing LID > 0.7.

CHAPTER 6. FD AND ND ν_e SELECTIONS

Different levels of selections are applied in the first analysis: subrun level selections, spill level selections and analysis specific selections to select ν_e CCevents.

6.1 Good Run Selection

Good run selections are designed to make sure that the detectors are in a good running condition for each subrun. The cuts are based on basic variables, such as event time, hit rate, the number of slices and the number of 3D reconstructed tracks.

6.1.1 Far Detector

For the far detector, the typical duration time of a subrun is 2-3 minutes. Due to the large size of the far detector and the fact that its size varies during construction, the goodness of the detector is tested diblock by diblock. The bad behavioring diblocks are masked. In consequence, each subrun will have its own configuration, depending on the diblock masks. In this way, the loss of the far detector data can be greatly reduced. In order to select data with good detector conditions, the good run selection for the far detector has 8 steps [90]:

Number of non-empty triggers > 0: to exclude the case when no real triggers are recorded in the subrun, which might be caused by some unexpected data acquisition problem.

Timestamp of first event $<$ timestamp of last event: to ensure that the events are ordered properly in time in each subrun.

Timestamp of first and last events are later than Jan. 1st 2013: usually when the time stamp of a file is corrupted, the time will be set to Jan 1 2010. Check this variable to rule out the time stamp corruption problem.

subrun duration time > 1.0 sec: to ensure that number of events recorded in the subrun is large enough so that the preformance of the subrun can be evaluated with sufficient number of hits, especially the one within the signal ADC region.

13 Hz < median rate of signal ADC hits < 23 Hz: the signal ADC region is defined to be $ADC \in (175, 3200)$, in which the ADC values of most of the physics hits present [90]. The MIP hits are used as a standard reference to test the response of different pixels of the detector. The cut requires that the major portion of the detector operates in the nominal region.

Number of good consecutive diblocks > 1: to require at least 2 good consecutive diblocks. The good diblocks are defined following the thresholds below:

- A good diblock must contain at least 12 (100%) good DCMs.
- A good DCM must contain at least 56 (87.5%) good FEBs.
- A good FEB must contain at least 26 (81.25%) good pixels.

A good pixel requires that both overall hit rate and signal hit rate are within proper ranges, which are listed below. The nominal window for the overall hit rate is defined to be ($10^{0.5}$, $10^{3.5}$) Hz, in which the hit rates of most of the channels present [92]. Due to the large number of energy depositions from non-minimum ionizing particles, the upper limit is much higher than the one for signal hit rate.

Because cosmic trigger data is used to tag good pixels, pixels near the top of the detector tend to present higher hit rate than the ones near the bottom. It is necessary to set different hit rate thresholds for different DCMs. DCMs 01-06 reads signals from vertical cells, while DCMs 07-12 read horizontal cells, from the top to the bottom. The thresholds are chosen to be 6 deviations from the median [93], within which the pixel is generally considered to be good. The pixels with hit rate higher than the upper threshold might have a noise problem in APD, while the ones below the lower threshold could be caused by a noisy pixel in the same APD which saturates the power from the rest. The good DCM conditions are consistent with the fact that the hit rate of vertical cells DCM 7 through 12 shows smaller deviation than the horizontal cells. For the horizontal cells, the proper signal hit rate window shifts to the lower value, as cell position goes deeper into the detector from DCM 07 to DCM 12.

- Overall hit rate: $10^{0.5}$ Hz < pixel hit rate < $10^{3.5}$ Hz.
- Signal hit rate:
  - DCMs 01-06: 13 Hz < signal hit rate < 31 Hz.
  - DCM 07: 5 Hz < signal hit rate < 45 Hz.
  - DCM 08: 4 Hz < signal hit rate < 36 Hz.
  - DCM 09: 4 Hz < signal hit rate < 30 Hz.
  - DCM 10: 2 Hz < signal hit rate < 26 Hz.
  - DCM 11: 2 Hz < signal hit rate < 23 Hz.
  - DCM 12: 1.5 Hz < signal hit rate < 20 Hz.

One challenge of applying this criteria is that since the expected number of hits is very low, there is large fluctuations in the value of the hit rate cut. The way to improve the stability of the selection is to cut on the probability of the measured rate being outside the nominal hit rate range as shown in Eq.6.1,

$$\begin{aligned} &Q(lower\ threshold \times livetime,\ measuredhits) > 0.1 \quad AND \\ &Q(upper\ threshold \times livetime,\ measuredhits) < 0.9 \end{aligned} \tag{6.1}$$

in which $Q(\mu, n) = \sum_{i=0}^{i<n} \frac{\mu^n}{n!} e^{-\mu}$, instead of applying the threshold directly. Finally, a good diblock also requires a proper asymmetry between the medians from top (vertical) and side (horizontal) DCMs (Eq.6.2),

$$0.1 < \frac{Median_{top} - Median_{side}}{Median_{top} + Median_{side}} < 0.5 \tag{6.2}$$

1.2 < number of slices per trigger per 10^4 pixels < 3.2: to make sure that a reasonable number of slices can be reconstructed per trigger per unit of the detector volume. A low reconstruction rate indicates issues with synchronization among DCMs for the case when the hits that come from the same interaction can not be matched into the same slice because of large gaps in time.

Fraction of 2D tracks over total number of tracks $< 15\%$: to ensure the low rate of reconstructed 2D tracks and a large number of 3D tracks. The 2D tracks are considered as a reconstructed track in either XZ or YZ view that failed to find a matched track in the opposite view and thus fail to form a 3D track. A high rate of 2D tracks could imply synchronization issues of a DCM that as a DCM in one view is out of synchronization, the 2D track might be truncated at the end or gapped in the middle, And interaction with this problem will have two 2D tracks in XZ and YZ views very different in Z position and fail to form a 3D track.

The cut flow table for the number of subruns and the POT is shown in Table.6.1. From the table, the good diblock cut removes the largest fraction of subruns, 7.1%. The second largest mode, which remove 3.1% of subruns, is manually removed, which are runs identified by control room shifters or by an analysis of TDU time [91]. The overall failure percentage is 11.28%. As shown in Table.6.2, among the selected subruns, 43833 of them were collected when the NuMI beam is shut down and 41281 subruns had the maximum number of consecutive diblocks less than 4. The data satisfying these conditions are not used in the first analysis. The fraction of good subruns is 54.9% which corresponds to 76.4% of total POT.

Table 6.1: The summary of the number (percentage) of subruns and the amount of POT removed by each good run selection cut for the far detector.

Failure mode	# of subruns (%)	POT (%)
Timing Stamp	291 (0.1%)	0.01×10^{20} (0.2%)
Live Time	117 (<0.1%)	$<0.01\times10^{20}$ (<0.1%)
Hit Rate	1666 (0.7%)	0.01×10^{20} (0.2%)
Diblocks	17883 (7.1%)	0.18×10^{20} (5.1%)
Slicing	1790 (0.7%)	0.04×10^{20} (1.0%)
Tracking	546 (0.2%)	$<0.01\times10^{20}$ (0.1%)
Manual	7532 (3.0%)	0.11×10^{20} (3.1%)

6.1.2 Near Detector

The good run selection is similar to the one for the far detector, with some the following differences:

Table 6.2: The summary of the number (percentage) of subruns and the amount of POT for different samples.

Description	# of subruns (%)	POT (%)
Total	251367 (100%)	3.49×10^{20} (100%)
Good	138086 (54.9%)	2.67×10^{20} (76.4%)
Good, no POT	43833 (17.4%)	0.00×10^{20} (0.00%)
Good, < 4DB	41281 (16.4%)	0.50×10^{20} (14.3%)
Bad	28167 (11.2%)	0.32×10^{20} (9.2%)

1. As the near detector is relatively small in size, the detector mask is applied to the whole detector and not per diblock;
2. NuMI spill information such as the timing peak and the number of NuMI triggers, is used in the selection;
3. As the near detector records hits mostly from NuMI beam events, the concept of rate for the near detector is understood not as the number of hits or events per livetime but per POT exposure or per spill where one spill equivalent to 2.5×10^{13} POT.

Nine cuts are designed and applied to the near detector to select ND good subruns [90]:

Timestamp of first event < timestamp of last event.

Timestamps of first and last events are later than Jan. 1st 2013: the cut is designed following the same idea explained in Sec.6.1.1.

Number of NuMI triggers > 1000: similar to the livetime requirement for the far detector, 1000 NuMI trigger ensure enough number of hits for later analysis.

217 μs $\leq$ timing peak start $\leq$ 219 μs: due to high intensity of beam events in the near detector, the data within one subrun can form clear edges of NuMI timing peak. The timing peak start and end cuts require the timing window within a correct range. The failures usually indicate problems with detector synchronization with the NuMI beam.

227 μs $\leq$ timing peak end $\leq$ 229 μs.

Fraction of empty spills over total spills $< 3\%$: a big fraction of empty spills implies some issues with the detector data acquisition system or some instability of the spill.

12 $<$ median number of signal ADC hits per 1000 spills $<$ 20 Hz: similarly to the signal hit rate test for the far detector. The purpose for this cut is to ensure that the major part of the detector is responding normally to the MIP hits.

Number of good diblocks = 4: the cut requires that all diblocks (3 full-active diblocks plus 1 muon catcher) in the near detector must be good diblocks. The criteria for a good diblock is:

- A good diblock must have 100% good DCMs.
- A good DCM must have at least 80% good FEBs.
- A good FEB must have at least 78.125% good pixels.
- A good diblock must show proper asymmetry between the median of the overall hit rate in the top and the side DCMs.
  - For full-active diblock: -0.1 < (top - side)/(top + side) < 0.1.
  - For muon catcher: -0.1 < (top - side)/(top + side) < 0.3

A good pixel is selected based on only overall hit rate information. After correcting for the difference in POT intensity, the overall hit rate is required to be $10^{0.5} Hz <$ pixel hit rate $+ 40 \times (1 - POT/spill/2.5e13) < 10^{3.5} Hz$.

3.5 $<$ number of slices per spill $<$ 5.5: the typical recorded interaction rate in the near detector usually contains 4 to 5 interactions per spill [94]. The failed subrun might have noisy APD problem which will create extra non-physics slices or synchronization issue which tend to determine less slices.

The summary table of the good run selection in ND is shown in Table.6.3 and Table.6.4. The two cuts that remove the largest number of subruns are based on spill information, the number of spills per subrun and the number of empty spills. Overall, the fraction of good subruns is 86.6%, which corresponds to 95% of the total POT.

Table 6.3: The summary of the number (percentage) of subruns and the amount of POT removed by each good run selections for the near detector.

Failure mode	# of subruns (%)	POT (%)
Timingstamp	2 (0.1%)	$<0.01\times10^{20}$ (<0.1%)
# of Triggers	205 (6.0%)	0.01×10^{20} (0.8%)
Timing Peak	132 (3.9%)	0.02×10^{20} (1.1%)
Empty Spill	240 (7.0%)	0.06×10^{20} (3.8%)
Hit Rate	175 (5.1%)	0.02×10^{20} (1.3%)
Diblocks	110 (3.2%)	0.01×10^{20} (0.8%)
Slicing	135 (4.0%)	0.02×10^{20} (1.3%)

Table 6.4: The summary of the number (percentage) of subruns and the amount of POT for different samples.

Description	# of subruns (%)	POT (%)
Total	3413 (100%)	1.65×10^{20} (100%)
Good	2956 (86.6%)	1.56×10^{20} (95.0%)
Good, no POT	4 (0.1%)	0.00×10^{20} (0.00%)
Bad	453 (13.3%)	0.08×10^{20} (5.0%)

6.2 Spill Selection

The spill selection is designed to ensure the good quality of NuMI beam as well as to improve the detector conditions for the data collected at a finer level than the good run selection.

In order to select spills with good spill quality, a few variables from the Intensity Frontier Database (IFDB) are used [96]:

spillPOT $>$ 2.0$\times10^{12}$: A stable beam performance is expected for the data used in the analysis. The typical POT for a stable spill is set to be above 2.0×10^{12} POT as shown in Fig.6.1.

-202 kA $<$ Horn Current $<$ -198 kA: Another variable to test the stability of the neutrino beam is the horn current, which has a large impact on the outgoing beam intensity. Fig.6.2 shows that for most of the spills, the horn current is measured to be between -202 kA and -198 kA.

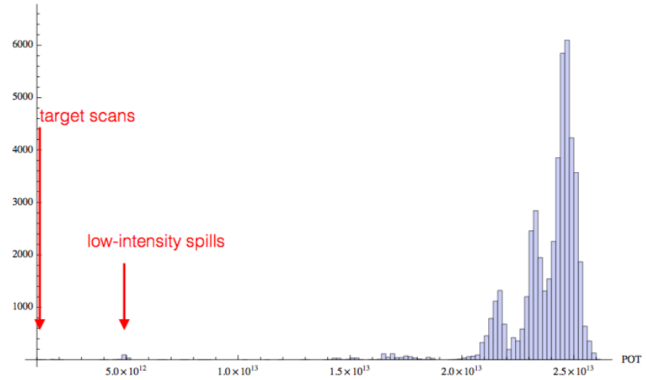


Figure 6.1: Distribution of spill POT.

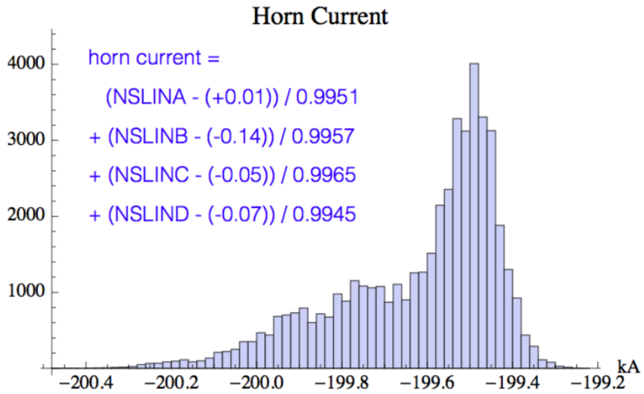
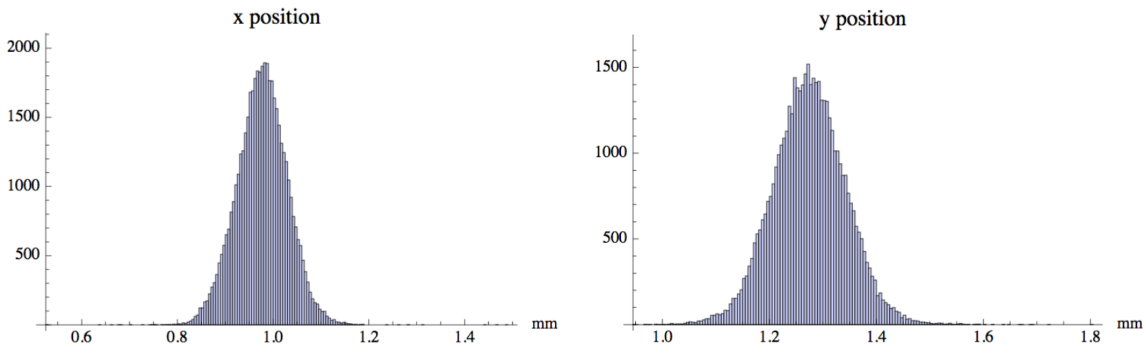


Figure 6.2: Distribution of horn current.

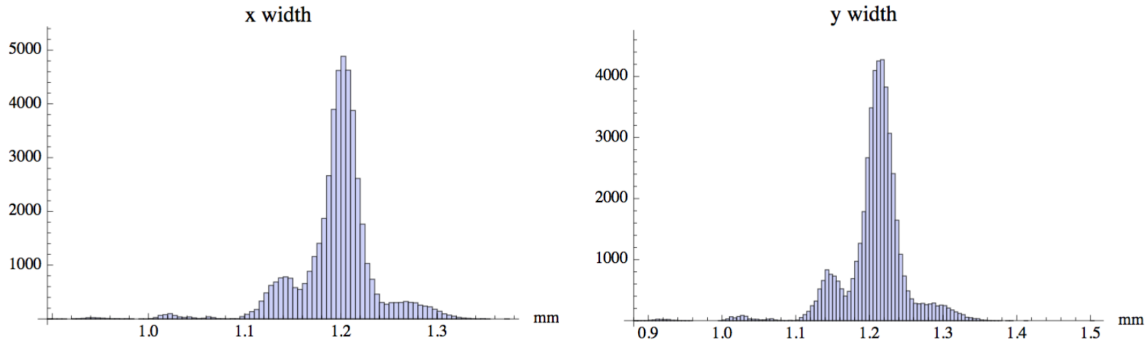
0.02 mm < Beam Position in X, Y < 2.00 mm: Input of the beam with good quality is expected to be present at the designed region on target, which is between 0.02 mm and 2.00 mm from the center of the target as shown in Fig.6.3.

0.57 mm < Beam Width in X, Y < 1.58 mm: As the beam intensity distributions in both horizontal and vertical directions are measured by the beam profile monitor, the extension of the beam is defined as σ of the gaussian function that fits to the distribution. According to Fig.6.4, most of the spills are included in the Beam Width $\in (0.57mm, 1.58mm)$ region in both X and Y directions.



(a) (b)

Figure 6.3: Distributions of the beam horizontal (left) and vertical (right) position on target.



(a) (b)

Figure 6.4: Distributions of the beam width in horizontal (left) and vertical (right) directions.

Delta T $<$ 0.5 sec: The cut is based on the time difference between the spill time recorded in NOνA files and the one recorded in the IFDB. The cut requires low time difference, which indicates small timing shift caused by beam monitoring devices. Fig.6.5 shows that most of spills present the ΔT below 50 ms, which is one-tenth of the current cut.

In addition to the cuts on spill quality, in order to select spills with good detector conditions, **no missing DCM** is required in order to avoid non-reporting DCMs. Then for the near detector, some cells are affected by the light from the detector hall. When the hall lights are

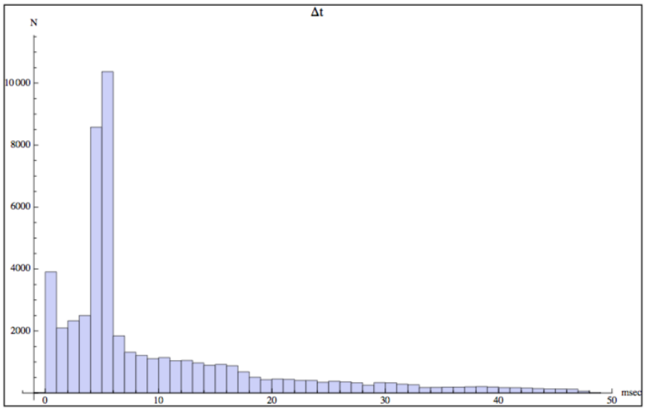


Figure 6.5: Distribution of difference between the time recorded in DAQ and the time recorded in the database.

turned on, the fraction of the out-of-spill hits over the total number of hits for the problematic cells appear to be higher than that for normal ones. The cut, **fraction of hits $>$ 0.45**, removes the spills with the problematic cells.

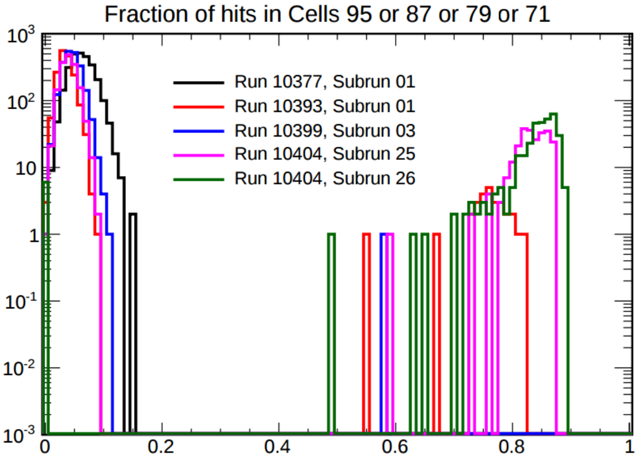
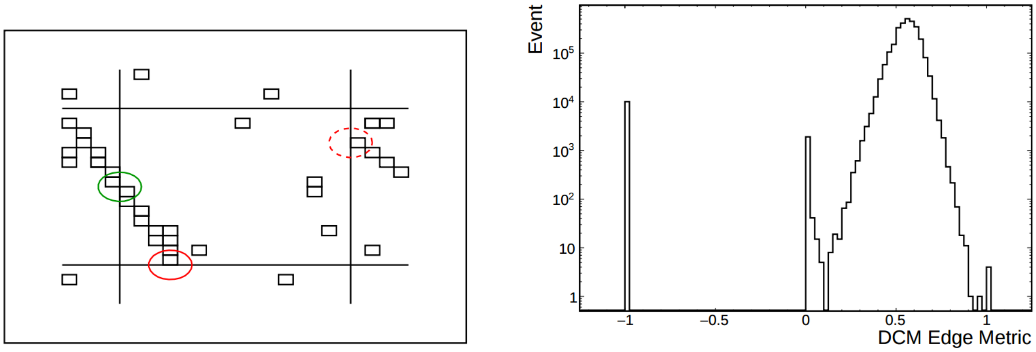


Figure 6.6: Distribution of the fraction of out-of-spill hits over total number of hits for problematic cells that have light-leakage problem.

For the far detector, the **no missing DCM** is also required. In addition, a variable called DCM edge metric is designed to improve the DCM synchronization in the data used [97]. As shown in Fig.6.7a, if a hit is found near the edge of a DCM, a match between hit and DCM edge is formed and a match opportunity is counted. If two matches are found on the neighbouring edges of a DCM, a match is counted. The DCM edge metric is defined as the ratio of the number of matches over the number of match opportunities. Fig.6.7b shows the distribution of the variable, on which the spike below 0.2 indicates the spills affected by DCM synchronization issue, in which the recorded event is truncated on the edge of the problematic DCM and only match opportunity but not match will be found. A cut, **DCM Edge Metric $>$ 0.2**, removes most of the out-of-synchronization failures.



(a) Diagram to illustrate the definition of DCM edge metric.

(b) Distribution of DCM edge metric, which presents a peak at 0.5 for normal DCMs and is below 0.2 for out-of-sync DCMs.

Figure 6.7

6.3 Far Detector Event Selection

The goal of the the far detector event selections for the ν_e appearance analysis is to distinguish ν_e CC signal interaction from all the background components [73]. The complete selections are built on following cuts: data quality cuts are applied to remove detector issues and reconstruction failures; uncontained events which are partially outside the detector and thus have poor energy reconstruction are removed by containment cuts from the analysis sample; cosmic rejection cuts are based on variables presenting good discrimination between signal and cosmic ray events; ν_e-preselections that are based on characteristic variables are then used to pre-select a ν_e enriched sample; particle identification cuts are finally applied to select signal from the background efficiently.

The study of the FD event selections is based on the official S15-05-22 cosmic triggered data to predict cosmic ray events and S15-05-22 Genie MC for beam events, including beam background and oscillated ν_e CC signal events. The simulation matches realistic configurations of the far detector that the number of diblocks in FD varies. The samples used in the selection study pass the standard good run and good spill selections described in the previous two sections. In order to match the real exposure of the first ν_e analysis, the cosmic samples are normalized to 14 diblocks and 165 sec, 14 diblock equivalent livetime of in-spill numi trigger data used in the first analysisand beam MC samples are normalized to 14 diblocks and 2.82×10^{20} POT, 14 diblock equivalent POT.

The ν_e-preselection and PID cuts are optimized preliminarily using the genie MC samples only, while for the first step in the FD selection, the cosmic ray background is assumed to be removed completely. Next, data quality, containment cuts and cosmic rejection cuts are studied at n-1 cut stage in exactly this order. In the last step, the ν_e-preselection and PID cuts are tuned again at n-1 stage. Because NOνA has two independent ν_e CC identification algorithms, LID and LEM, all the cuts were checked in both LID and LEM selected regions.

6.3.1 Data Quality Cuts

The FD data quality cuts are summarized in the first part of Table.6.5.

Table 6.5: data quality, containment and cosmic rejection cuts in LID and LEM regions. the detail of each variables will be discussed in the subsequent paragraphs.

variables	cuts
Number of hits per plane	nhit/nplane $<$ 8
Number of hits in x-view planes	ncellx $>$ 5
Number of hits in y-view planes	ncelly $>$ 5
Difference between nhit in x and y view planes	dnhit $<$ 0.4
Cosine of the angle between the two leading showers	dang $>$ -0.95
Distance of the leading shower from vertex	gap $<$ 100 cm
Fraction of nhit on showers	nhitfrac $>$ 0.7
Distance from east wall	minx0 $>$ 15 cm
Distance from west wall	maxx0 $>$ 10 cm
Distance from bottom	miny0 $>$ 10 cm
Distance from top	maxy0 $>$ 150 cm
Distance from front	minz0 $>$ 35 cm
Distance from back	maxz0 $>$ 200 cm
Fraction of event transverse momentum	ptp $<$ 0.4, when maxyall $<$ 25 cm, or ptp $<$ 0.65, when maxyall $>$ 25 cm

Number of Hits per Plane: The cut requires the number of hits per plane to be less than 8. The number of planes is counted based on the most energetic shower instead of slice in order to reduce the uncertainty due to noise hits. This cut removes FEB flash issues in data, in which case multiple contiguous hits are seen on the same plane. The FBE flash problem is often induced by high energy cosmic ray events. Fig.6.8 shows an event from cosmic trigger data presenting a pattern for the FEB flash problem. The blue contiguous boxes show the flash effect. Fig.6.9 shows the distribution of the number of hits per plane for LID and LEM selected events for signal (blue), beam background (red) and cosmic background (black) at "n-1" cut stage, at which all the cuts except for the number of hits per plane cut are applied. As cosmic background is livetime exposure normalized and beam events are POT normalized, cosmic background is much lower than the beam components.

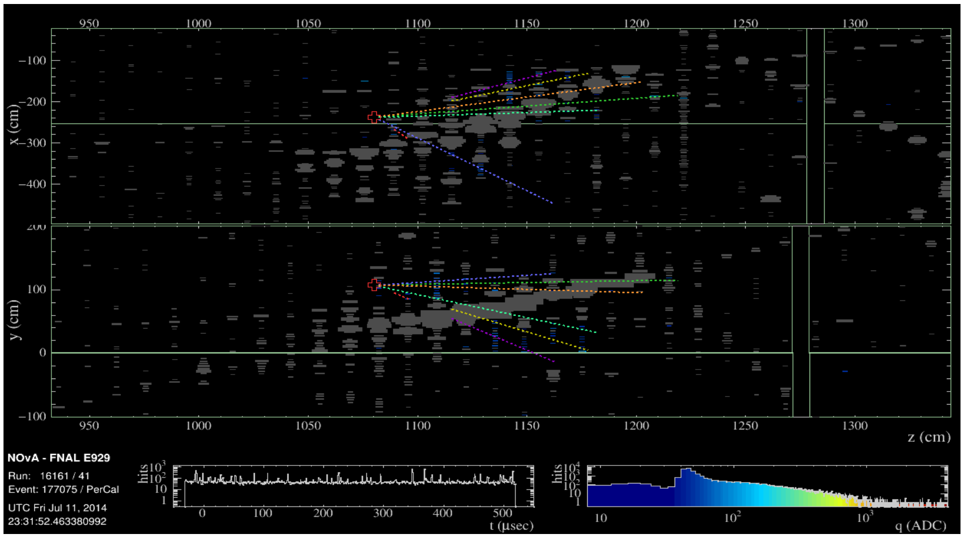


Figure 6.8: Event display of a cosmic trigger event as an example that is removed by this cut based on the number of hits per plane variable. The blue boxes represent hits, while dash lines are reconstructed prongs.

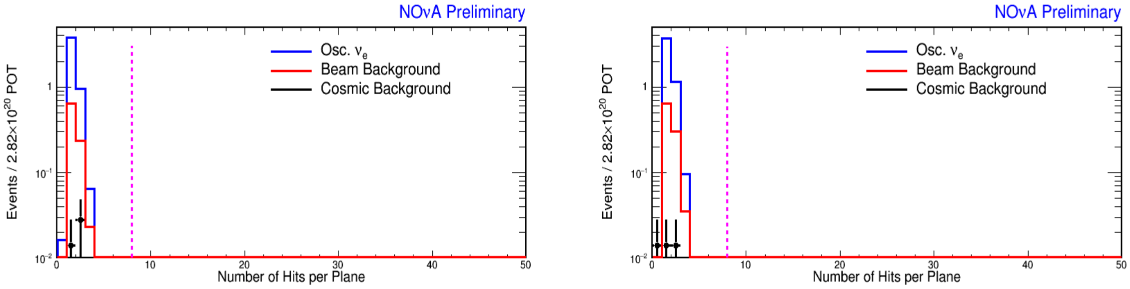


Figure 6.9: Distributions of the number of hits per plane variable at "n-1" cut stage in the LID (left) and LEM (right) selected regions for ν_e CC signal (blue), total beam background (red) and total cosmic background events (black). Magenta Lines are cut position. The cut removes events with average number of hits per plane greater than 8.

Total Number of Hits in X/Y-View Planes (nhit/nplane): This cut removes events showing little energy deposition in either view. Fig. 6.10 and Fig. 6.11 shows the distribution of total number of hits in X/Y-View planes. This cut requires that an event has at least 5 cells in both X-view and Y-view planes.

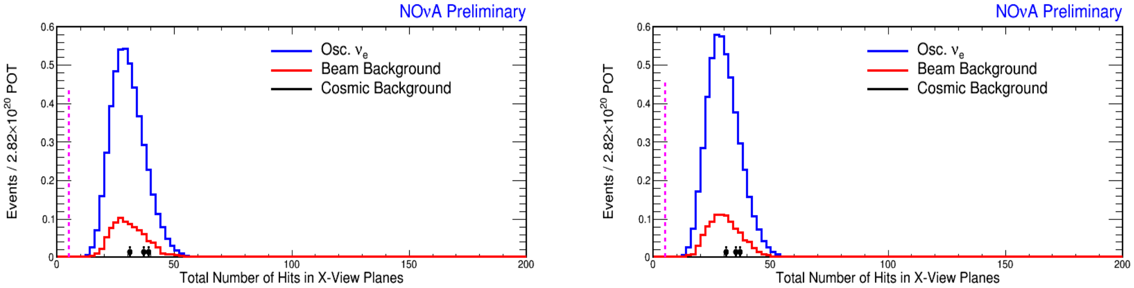


Figure 6.10: Distribution of total number of hits in X-View planes at "n-1" cut stage in the LID (left) and LEM (right) selected region.

Difference between the Number of Hits in X (ncellx) and Y (ncelly) View Planes: This cut is designed to remove the events with a large difference between the number of hits deposited in X-view planes and in Y-view planes. For a well reconstructed beam event, a high symmetry between the number of hits in X and Y views is expected. The cut also has the power of rejecting cosmic ray events, since cosmic ray events tend to enter the detector at large angle with respect to the beam direction, which ends up with more hits in X view than in Y view planes. Fig.6.12 shows an example of mis-reconstruction that has large difference between the number of hits in X and Y view planes and is re-

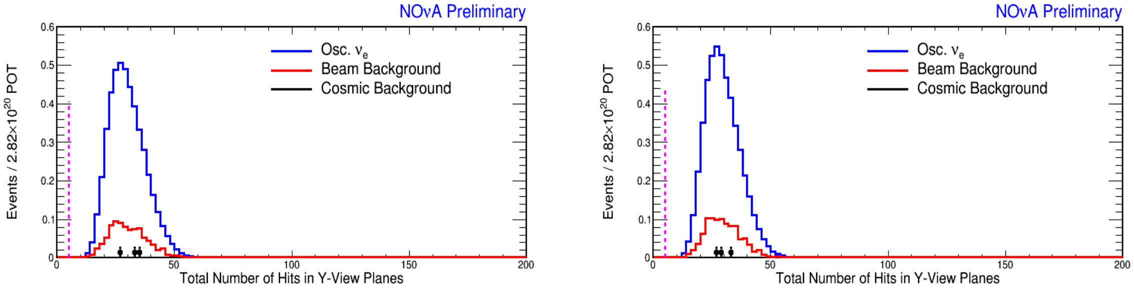


Figure 6.11: Distribution of total number of hits in Y-View planes at "n-1" cut stage in the LID (left) and LEM (right) selected regions.

jected by this cut. Fig.6.13 shows the distribution of this variable at "n-1" cut stage in LID and LEM regions. The cut is at 0.4, which preserves the majority of the signal events.

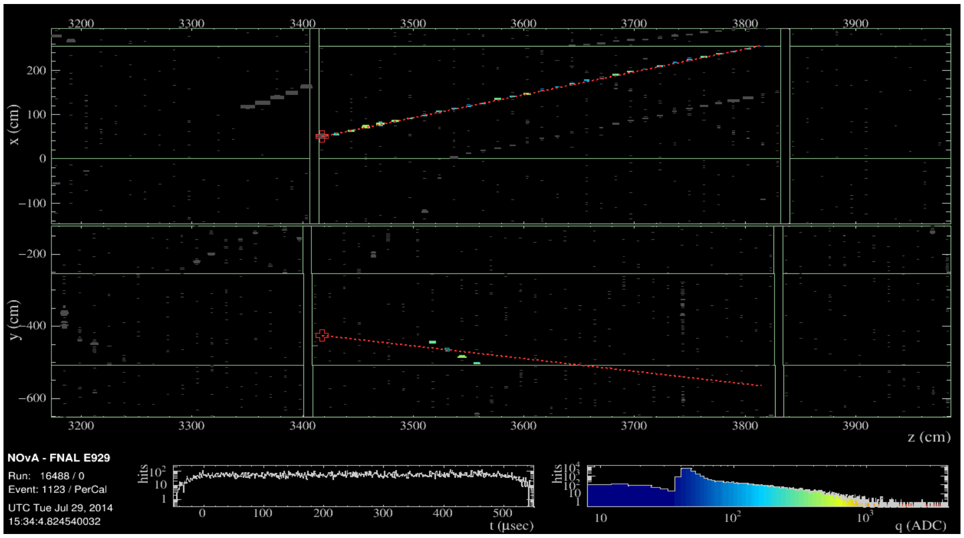


Figure 6.12: Event display of a cosmic trigger event as an example of mis-reconstruction showing large difference between the number of hits in X-view and Y-view planes.

Cosine of the Angle between the Two Leading Showers (dang): The cut removes the interactions with the two most energetic prongs that have large angle to each other. An example of such case is shown in Fig.6.14, which is probably an event with one gapped prong due to bad channel issues, non-reporting APD due to noise, and is mis-reconstructed into two prongs. Fig.6.15 shows the distribution of the variable in LID and LEM selected region. This cut selects events with the difference in angle greater than

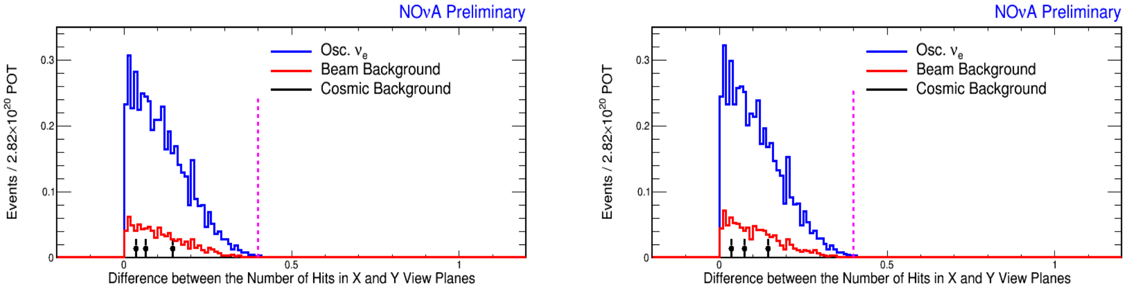


Figure 6.13: Distribution of the difference between the number of hits in X-view and Y-view planes at "n-1" stage in the LID (left) and LEM (right) selected regions.

-0.95 radians, which is ∼170 degrees. This cut is not applied for the events with only one reconstructed prong.

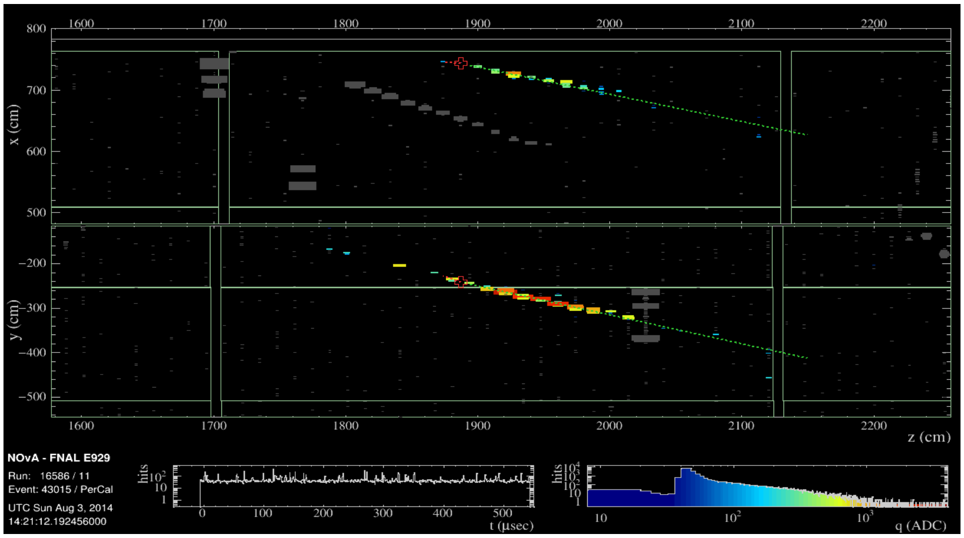


Figure 6.14: Event display of a cosmic trigger event having two reconstructed prongs, represented by green and red lines, which have a very large angle to each other.

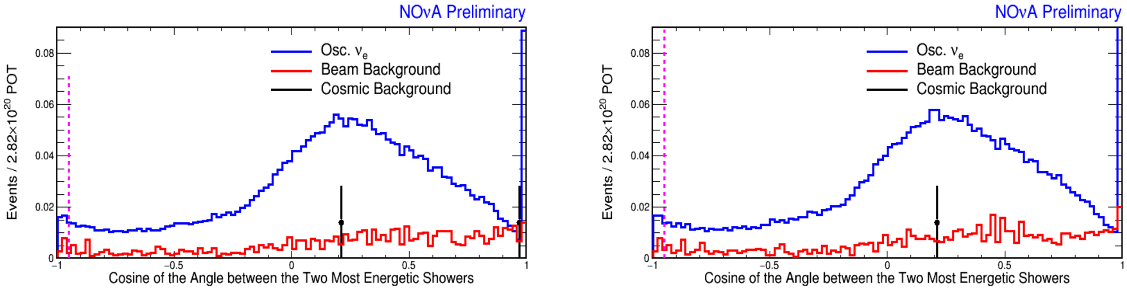


Figure 6.15: Distribution of the cosine of the angle between the first two showers in LID selected region, LID > 0.95 (left) and LEM selected region, LEM > 0.8 (right).

Distance of the Leading Shower Start from the Main Vertex (gap): This cut removes interactions where the distance of the leading shower from interacting position is greater than 100 cm, which is $\sim$ 2 times larger than the radiation length of EM showers in the NOνA detectors. Fig.6.16 shows an example of this pattern, which shows a distance of $\sim$120 cm between the vertex and prong start, in which the vertex is probably mislocated due to several very noisy cells far away from the leading shower. Fig.6.17 shows the distributions of this variable in LID and LEM selected regions. The cut at 100 cm is well above the majority of the signal (blue) events.

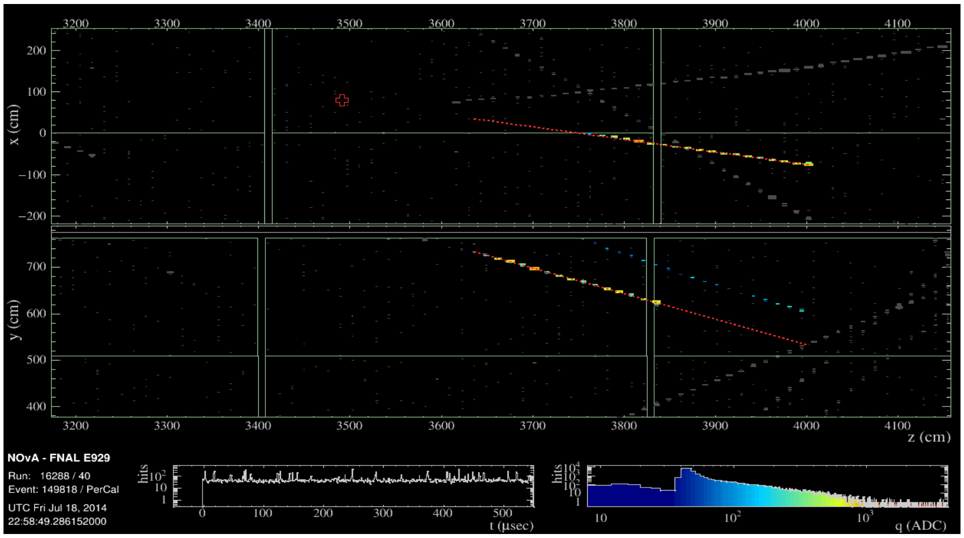


Figure 6.16: The event shows a large distance between the leading shower start and the main vertex.

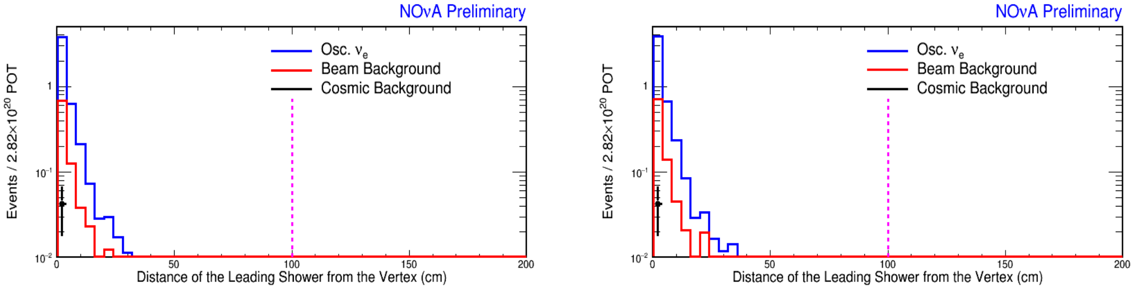


Figure 6.17: Distribution of the distance of the leading shower from the main vertex in LID (left) and LEM (right) regions.

Fraction of the Hits associated to Showers (nhitfrac): This cut selects events in which at least 70% of the hits belong to showers. The events where a small fraction of hits is reconstructed into a shower might have high contamination of noise hits. Fig.6.18 shows the distributions of this variable at “n-1” cut stage in the LID and LEM selected regions. The cut is at 0.7, which retains more than 98% of the signal events.

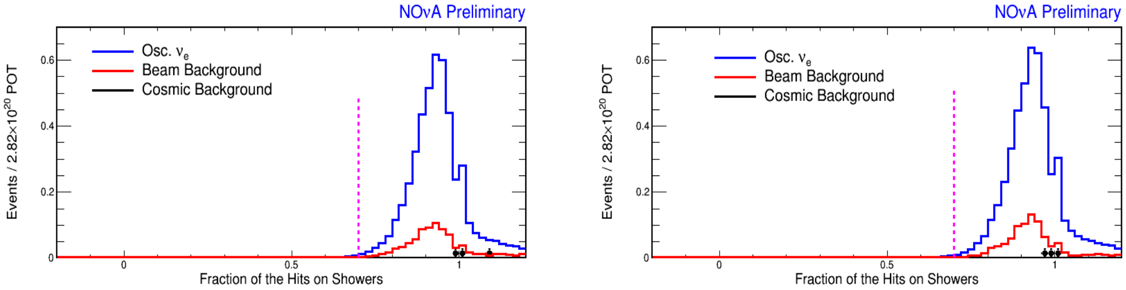


Figure 6.18: Distribution of the fraction of the hits on showers in LID (left) and LEM (right) regions.

6.3.2 Containment Cuts

The containment cuts reject events that are only partially inside the detector and thus present poor energy reconstruction, which causes uncertainty in the analysis. Moreover, for the entering events with interaction points outside the detector, they are very likely cosmic ray events. The containment cuts selects events with both start and stop points of the leading shower within the containment volume of the far detector. The cuts are applied in terms of the shorter distance of either start or stop point of the leading shower from the boundary of active area of the detector. The distance calculation is handled by LiveGeometry functions [98], which determines the detector boundaries based on the activity of DCMs. The second portion of Table.6.5 lists the containment cuts used in the first ν_e appearance analysis. Fig.6.19 shows an example of the uncontained events that will be removed by the containment cut at the top wall.

Fig. 6.20 through Fig. 6.27 are the distributions of the distance to all six walls in LID and LEM selected regions. The containment cuts at the top (Fig. 6.23) and back (Fig. 6.26) walls are set tight in order to remove the cosmic background events. Because most of the cosmic ray events are coming from the top of the detector, there is high contamination of cosmic background near the top wall. At the back wall, the detector does not have enough rocks to block large number of the cosmic ray events and hence the cosmic contamination is quite large here too. The cuts at top and back walls are studied firstly at standard n-1 cut stage (Fig. 6.23 and Fig. 6.26). The cuts are then optimized for $FOM = \frac{S}{\sqrt{B}}$ with loose PID cuts, which allows higher statistics in the cosmic sample (Fig. 6.24 and Fig. 6.27). By tuning with loose PID cuts, LID > 0.7 and LEM > 0.5 respectively, the optimal cuts were found to be 120 cm from top wall and 150 cm from the back wall. Finally, the cuts, based on the tuning results, were adjusted to 150 cm from the top wall and 200 cm from the back wall. The final cuts are set tighter than the tuned in order to avoid overtuning, because the statistics of the cosmic ray events is low near the boundary. The cuts at the other walls are determined at standard n-1 cut stage in both LID and LEM selected regions.

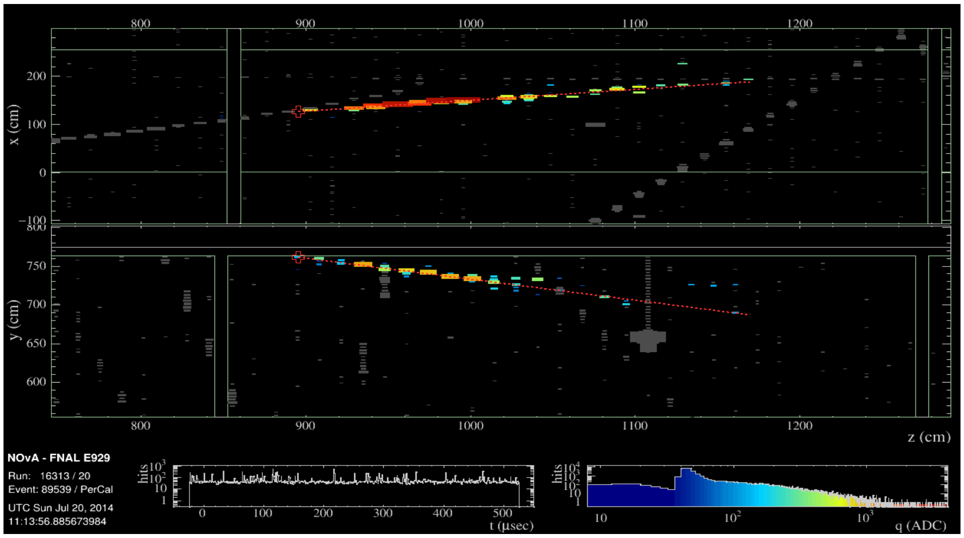


Figure 6.19: Event display of an example that is rejected by the containment cut at the top wall.

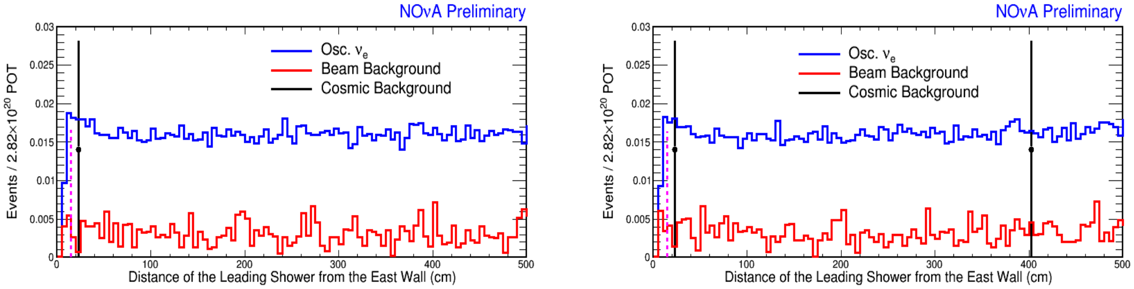


Figure 6.20: Distribution of the minimum distance of start and stop points of the leading shower from the east wall in LID (left) and LEM (right) region.

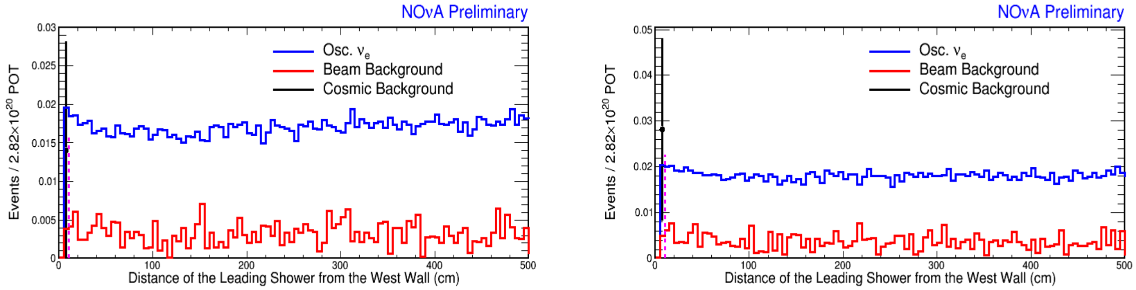


Figure 6.21: Distribution of the minimum distance of start and stop points of the leading shower from the west wall in LID (left) and LEM (right) region.

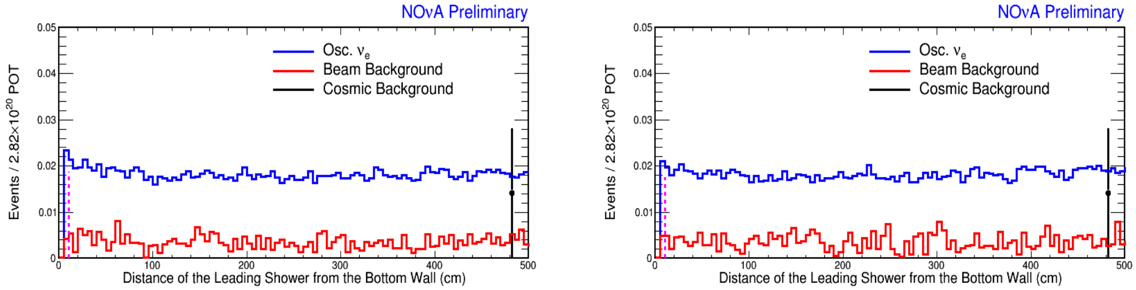


Figure 6.22: Distribution of the minimum distance of start and stop points of the leading shower from the bottom wall in LID (left) and LEM (right) region.

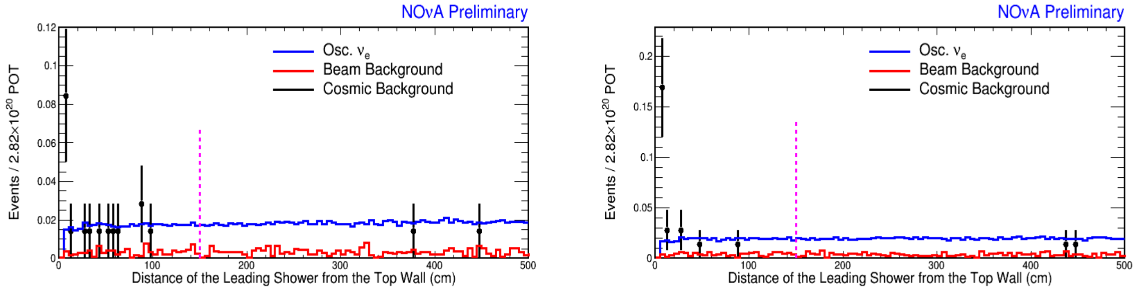
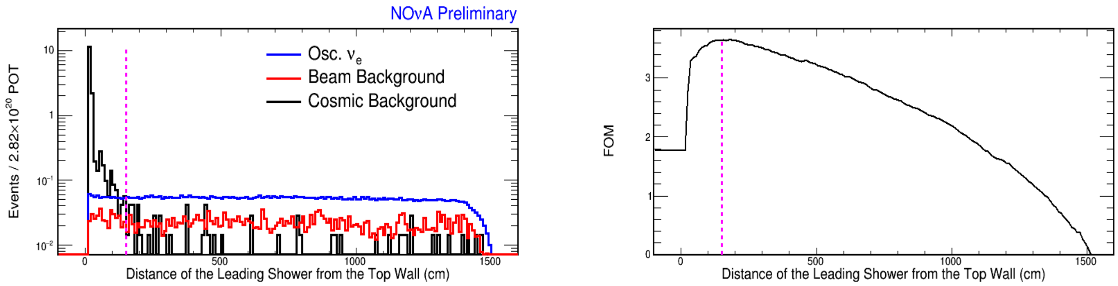


Figure 6.23: Distribution of the minimum distance of start and stop points of the leading shower from the top wall in LID (left) and LEM (right) region.



(a) Distribution of the distance of the leading shower from the top wall after loose PID cut.

(b) FOM distribution as a function of different cut positions.

Figure 6.24: The maxy cut is optimized for $S/\sqrt{B}$ with loose PID cuts, which allows higher statistics in the cosmic sample. Adjustment based on the tuning with loose PID cuts, LID $>$ 0.7 and LEM $>$ 0.5 respectively, suggests a cut on the position of the leading shower to be at 150 cm from the top wall.

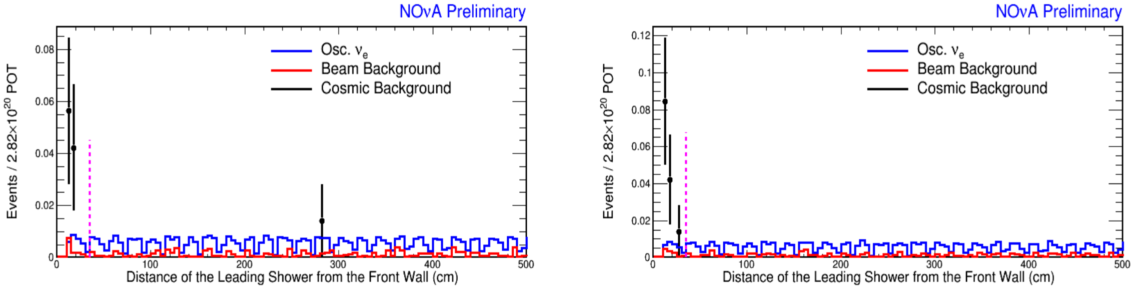


Figure 6.25: Distribution of the minimum distance of start and stop points of the leading shower from the front wall in LID (left) and LEM (right) region.

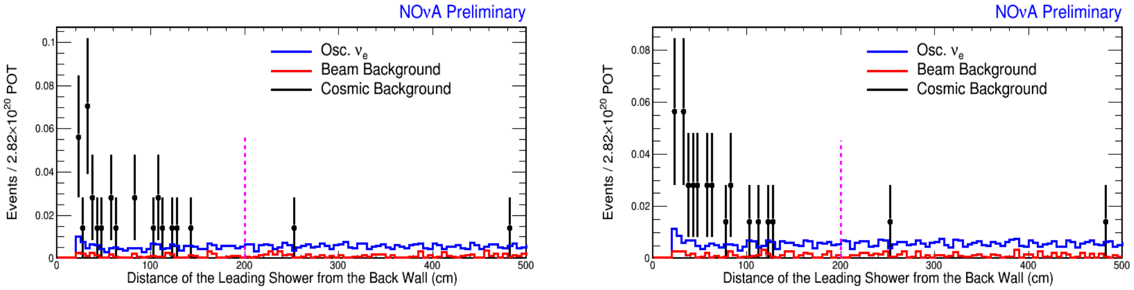
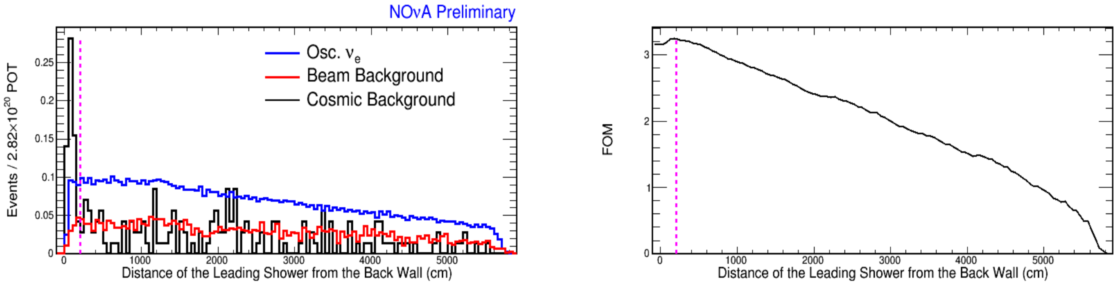


Figure 6.26: The cut is firstly studied at n-1 cut stage. To avoid overtuning, the cut is then studied after loose PID cuts, which allows more statistics in the cosmic background sample.



(a) Distribution of the distance of the leading shower from the back wall after loose PID cut.

(b) FOM distribution as a function of different cut positions.

Figure 6.27: The maxz cuts is optimized for $S/\sqrt{B}$ with loose PID cuts. The suggested cut on maxz is 200 cm.

6.3.3 Cosmic Rejection Cuts

After applying all the other cuts including containment, preselection and particle identification cuts applied, only one additional variable named, the fraction of event transverse momentum, is needed to suppress the cosmic ray background down to two orders of magnitude lower than the expected number of signal events in the far detector.

Fraction of Event Transverse Momentum The cut is based on the ratio of the event transverse momentum to the event total momentum (p_t/p) as shown in Eq.6.3:

$$p_t/p = \frac{\sum_{Prong_i}\left[\sqrt{E_i^2 - M_i^2}\cdot \sin(\theta_{i,Beam})\right]}{\sum_{Prong_i}\left(\sqrt{E_i^2 - M_i^2}\right)} \tag{6.3}$$

in which E_i, M_i are the calorimetric energy and the mass, the mass of the particle hypothesis that represents the highest likelihood value, of the examined prong and the momentum is calculated as $\sqrt{E^2 - M^2}$ for relativistic particles. The variable $\theta_{i,Beam}$ is the angle of the ith prong with respect to the beam direction. The transverse components of the ith prong is taken as $sin(\theta_{i,Beam})$ of the total. The larger p_t/p is, the less the energy depositions is concentrated along the beam direction. So the variable reflects the angle of the event with respect to z direction, which gives the capability to reject cosmic ray events.

Fig.6.28 shows the event display for a cosmic trigger event that presents a large deflection from the beam direction especially in YZ view and consequently has a high p_t/p value. The event is rejected by the cut. Fig.6.29 shows the p_t/p distribution as a function of the distance of all showers from the top wall. Since most of the cosmic ray events concentrate near the top of the detector, when an event is close to the top wall, the shortest distance of all showers from the top wall is smaller than 25 cm, p_t/p cut is tight, $p_t/p < 0.4$. when an event is far away from the wall, a loose requirement on p_t/p is placed, $p_t/p < 0.65$. (See Table.6.5) Fig.6.30 is the p_t/p distribution for the events that are within 25 cm from the top wall. The signal peaks at 0.2, while the cosmic background events peak at the high end of the distribution. The variable also has the power of separating the NC background from the signal. Because of the invisible outgoing neutrino in NC interactions, NC tends to be more asymmetric with respect to the beam direction than the charged current interaction, which leads to a higher p_t/p value.

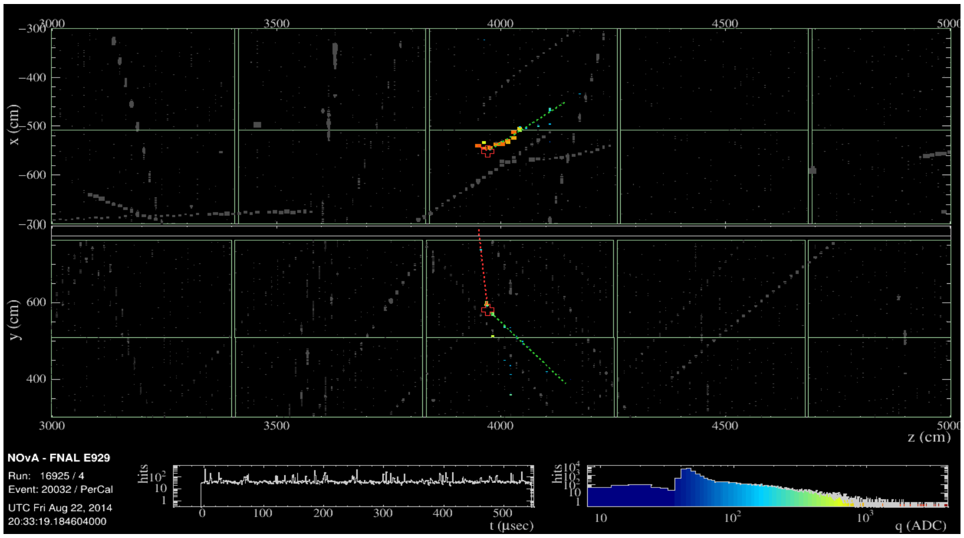


Figure 6.28: Event display of a cosmic trigger event as an example that has high ptp value. The Dash lines are showers, while the cross is the reconstructed vertex of the interaction.

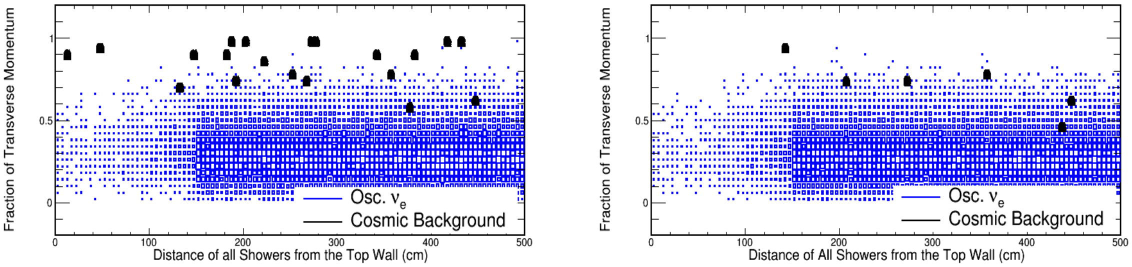


Figure 6.29: Distribution of the fraction of event transverse momentum vs the distance of all the showers from the top wall in LID (left) and LEM (right) regions. Blue box represents simulated signal events and black dots are cosmic ray background events in the studied cosmic trigger data.

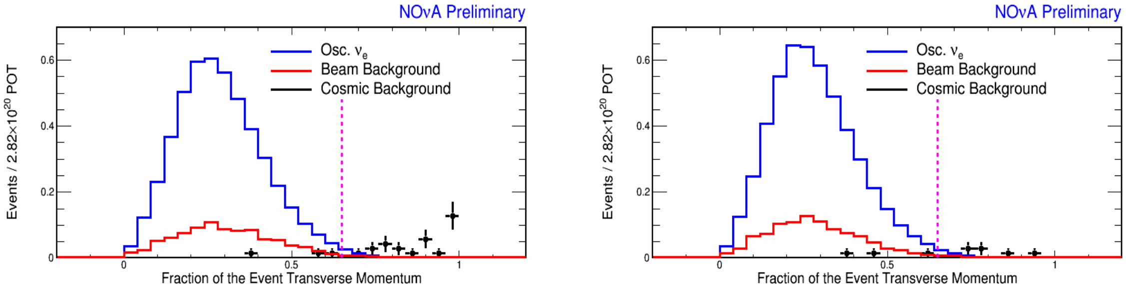


Figure 6.30: Distribution of the fraction of event transverse momentum for events that are more than 25 cm from the top in LID (left) and LEM (right).

6.3.4 ν_e-Preselection and PID Cuts

An important source of background for the ν_e appearance analysis arises from interactions of the beam components as discussed in Sec.4.1. The beam background is mainly composed of ν_μ CC, beam ν_e CC and NC interactions. The beam ν_e CC has almost the same characteristics as the oscillated ν_e CC event and represents an irreducible background for the ν_e appearance analysis. The only difference from the signal events is the energy distribution. Both ν_μ CC and NC interactions can generate additional hadronic or EM showers which causes difficulties in identifying oscillated ν_e CC interactions.

The ν_e-preselection selects a ν_e CC-like interaction enriched sample, which is to be used in PID cut study as well as for the sideband tests. The preselection cuts are based on three variables: the total number of hits ($nhit$), total calorimetric energy ($calE$) and the length of the longest shower (len) in the event. The ν_e-preselection cuts work as a complement along with the PID cut to achieve the final ν_e CC-like sample. All ν_e-preselection and PID cuts are tuned to achieve their maximal FOM ($= \frac{S}{\sqrt{B}}$) value in order to search for ν_e appearance. Tests were done that tuning to FOM ($= \frac{S}{\sqrt{S+B}}$) selects additional events for signal, however additional beam background is also selected, thus marginally improving sensitivity. For this reason the former FOM is chosen for the first analysis. After all tunings the cuts are adjusted minimally near the optimal point for rounding. Since NOνA has two independent PID algorithms, LID and LEM, different preselection cuts are needed for individual PIDs.

Table.6.6 shows the final ν_e-preselection and PID cuts in LID and LEM selected regions. The ν_e-preselection and PID cuts are preliminarily tuned assuming that there is no cosmic background and are retuned once the other cuts are determined. Fig.6.31 shows total calorimetric energy distribution for signal (blue), total beam background (red) and cosmic background (black). LID requires a tighter energy cut than LEM, because LID has more cosmic background events at the low edge than LEM. Fig.6.32 and Fig.6.33 show the distributions of the number of hits and the length of the longest shower. These two cuts are generally loose compared to the calorimetric energy cut in order to avoid large impact on the selected neutrino spectrum.

Table 6.6: ν_e-preselection and PID cuts in both LID and LEM regions. Energy cuts are different in the two PID regions.

variables	cuts in LID	cuts in LEM
Total number of hits in slice	40 < nhit < 115	40 < nhit < 115
Total calorimetric energy of slice	1.5 GeV < calE < 2.7 GeV	1.3 GeV < calE < 2.7 GeV
Length of the longest shower	140 cm < len < 500 cm	140 cm < len < 500 cm
PID	LID > 0.95	LEM > 0.8

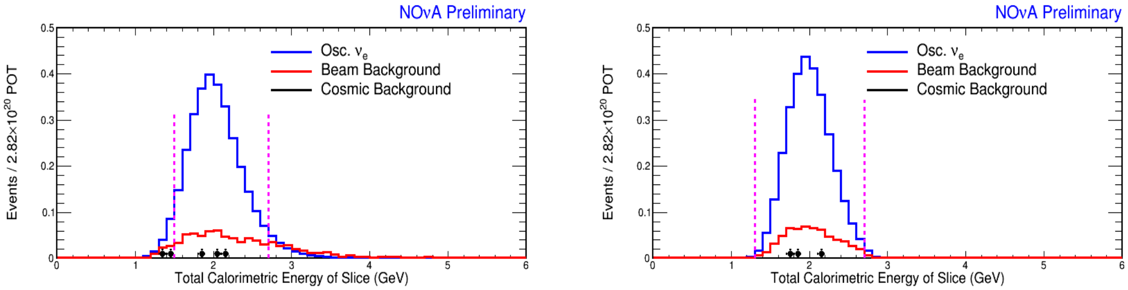


Figure 6.31: Distribution of total calorimetric energy in LID (left) and LEM (right) selected regions. The cut windows are (1.5, 2.7) GeV for LID and (1.3, 2.7) GeV for LEM.

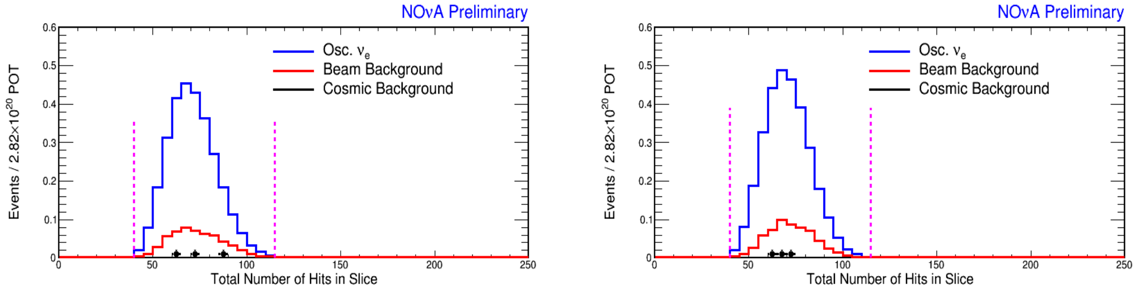


Figure 6.32: Distribution of the total number of hits in LID (left) and LEM (right) regions. The cut on number of hits is (40, 115) for LID (left) and LEM (right).

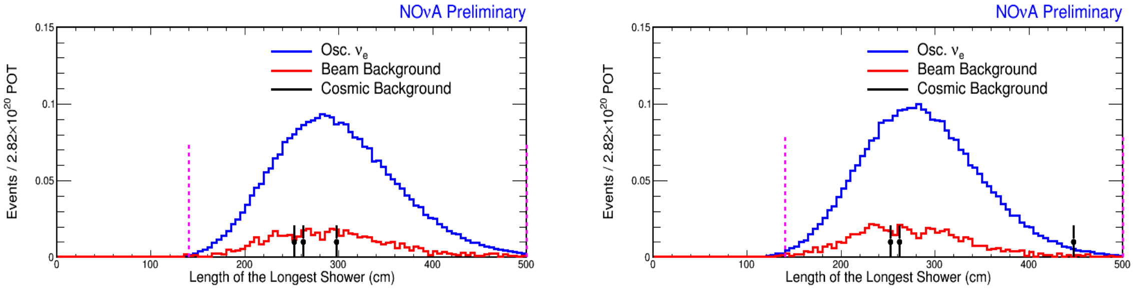
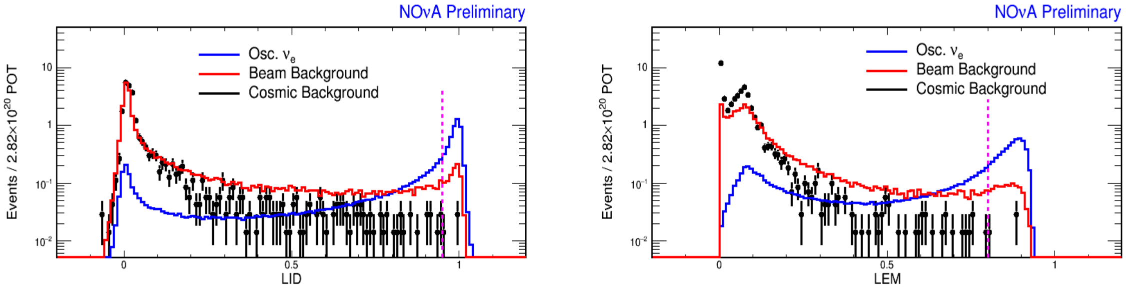


Figure 6.33: Distribution of the length of the longest shower in LID (left) and LEM (right) regions. The cut window is (140, 500) cm for LID (left) and LEM (right).

Fig. 6.34 shows the distribution of LID and LEM. Both of the two algorithms are efficient in discriminating signal from other event topologies. In the plots, signal (blue) peaks near 1, while other background components, including cosmic background which consists mostly of cosmic muons, concentrate at the low end. The PID cuts are tuned to be LID > 0.95 and LEM > 0.8 respectively.

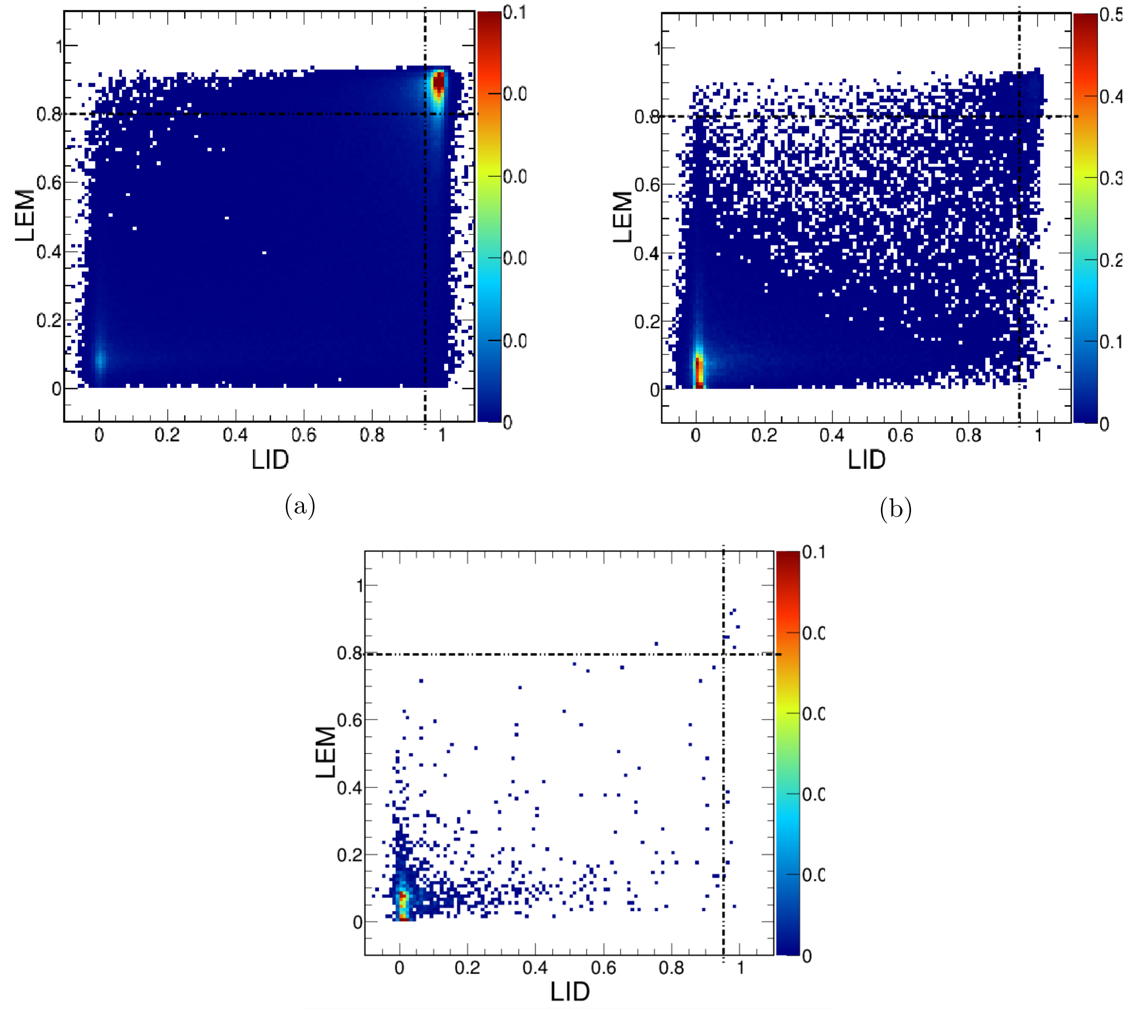


(a) LID > 0.95

(b) LEM > 0.8

Figure 6.34: Distribution of LID and LEM.

Fig. 6.35a and Fig. 6.35b show the comparison between LID and LEM after data quality, containment, cosmic rejection and loose preselection cuts for signal, beam background and cosmic background events. There is a 34% overlap in the ν_e CC signal events selected by both PIDs, 11.0% and 9.3% additionally selected by only LEM or LID respectively. For beam background, the overlap is 1.7% with 1.5% and 1.1% selected by only LEM or LID respectively. Fig. 6.35c shows that the overlap in the cosmic background events rejected by both PIDs is 99.9%.



(a)

(b)

(c)

Figure 6.35: LID vs LEM comparison for signal (a), beam background (b) and cosmic ray background (c) events after loose ν_e preselections.

Table.6.7 shows that, after all the ν_e selection cuts, including cosmic rejection, ν_e-preselection and PID cuts, the number of cosmic events in FD cosmic trigger data is reduced to a level smaller than the other beam background components. After all the FD selection cuts, the cosmic ray background is reduced to 1×10^{-8} and beam background is reduced to about 1×10^{-4}, while the efficiency of selecting signal events is about 20% with respect to the total number of events passing good run, good spill cuts. FOMs are improved to greater than 4.5 in both PID regions. Table.6.8 shows the efficiency of each cut for signal, total beam background and cosmic background from cosmic trigger. The efficiency of each cut in the data quality step is controled to be not lower than 95%, even though overall efficiency of data quality step is 70% for signal. For the ν_e-preselection, the energy cut is responsible for most of the reduction in the signal events. The energy cut is tuned to be tight in order to sufficiently remove the cosmic background events.

Table 6.7: Cut flow table for both LID and LEM.

cuts	ν_eCC signal	beam background	cosmic background Scaled (unscaled)	FOM=$\frac{S}{\sqrt{B}}$
nocut	20.22	529.65	5.75E6 (4.43E8)	0.01
data quality	14.90	100.77	3.55E5 (2.74E7)	0.03
containment	12.50	81.94	1.54E4 (1.19E6)	0.10
cosmic rejection	11.73	70.73	307.25 (23685)	0.60
ν_e-preselection (LID)	8.99	23.33	22.09 (1703)	1.33
LID	4.28	0.81	0.04 (3)	4.64
ν_e-preselection (LEM)	9.87	28.68	42.91 (3308)	1.17
LEM	4.45	0.88	0.04 (3)	4.63

Table 6.8: Cut efficiency at each cut step.

cuts	ν_eCC signal	beam background	Cosmic in Cosmic Trigger
data quality	74%	19%	6.2%
containment	84%	81%	4.3%
cosmic rejection	94%	86%	2.0%
ν_e-preselection (LID)	77%	33%	7.5%
LID	48%	3.5%	0.18%
ν_e-preselection (LEM)	84%	41%	14%
LEM	45%	3.1%	0.093%

After tuning the cuts in the independent cosmic trigger sample, the final cosmic ray background prediction is made using the cosmic ray events in the out-of-time NuMI trigger sample, which can be understood as the timing sideband of the final signal region. The prediction is made using pre shutdown and post shutdown NuMI trigger samples, which are then scaled to the sample size of the in-spill time window. The out-of-time NuMI trigger sample also works as an independent sample to check the cosmic rejection cuts which are developed using the cosmic trigger sample. Table.6.9 shows the comparison between the prediction from the two triggers. Before PID, the two samples have 1% difference. The cut efficiency at each cut stages agree between cosmic and NuMI triggers as shown in Table.6.8. The cumulative efficiencies of each cut can be found in App.C.1.

After the PID cut, two cosmic ray events are selected, the event display of which is shown in App.C.2, selected by LID and LEM. Scaling the two events by the ratio of in-spill livetime exposure to the out-of-spill livetime exposure, the cosmic ray background prediction for the in-spill sample is $0.06^{+0.06}_{-0.03}$ for LID and $0.06^{+0.06}_{-0.03}$ in LEM region, which matches the cosmic trigger data result within the statistical uncertainty. The uncertainty corresponds to the 1σ fluctuation around the selected 2 events assuming a poisson distribution for the number of selected events.

Table 6.9: Cosmic ray background for the ν_e analysis is predicted using out-of-time numi trigger data, $0.06^{+0.06}_{-0.03}$ in LID and $0.06^{+0.06}_{-0.03}$ in LEM region. The prediction from out-of-time numi trigger data matches cosmic trigger data within statistical fluctuation.

cuts	Cosmic in Cosmic Trigger Scaled (unscaled)	Cosmic in out-of-time Numi Trigger	Difference in σ	Difference in %
nocut	5.75E6 (4.43E8)	5.80E6 (1.62E8)	144.29	1.33
data quality	3.55E5 (2.74E7)	3.57E5 (1.02E7)	29.32	1.08
containment	1.54E4 (1.19E6)	1.54E4 (4.48E5)	3.55	0.62
cosmic rejection	307.25 (23685)	308.16 (5860)	0.53	0.42
ν_e-preselection (LID)	22.09 (1703)	23.24 (658)	1.19	5.71
LID	0.04 (3)	0.06 (2)	0.41	50
ν_e-preselection (LEM)	42.91 (3308)	45.63 (1241)	1.96	6.84
LEM	0.04 (3)	0.06 (2)	0.41	50

6.4 Near Detector Event Selection

There are ν_e event selection cuts for ND, similarly to the FD selection. The ND ν_e selection cuts includes:

- Data and reconstruction quality cuts are applied to ensure good quality of the selected sample:

  Data quality: The same requirement is placed for the near detector as the one for the far detector in Table.6.5. Events with more than 8 hits in one plane on average are considered to have a FEB flash problem, which are excluded by the data quality cut.

  Reconstruction quality: The cut requires that the event contains at least one reconstructed 3D shower. As some of the down-stream cuts are based on shower variables, the cut ensures that the shower object exist for the examed event.

- Fiducial and containment cuts are used to reject the rock events with vertices outside the detector and not fully contained:

  Fiducial cuts: The cuts select events with elastic arm vertices inside the fiducial volume, which is shown as the red box in Fig.6.36. The fiducial volume is about two radiation lengths from the front wall of the detector and one radiation length from east, west, bottom and top walls of the detector. An event with the vertex outside the detector is very likely a neutrino interaction with particles in the rock surrounding the detector. The distance from the fiducial volume to the back wall of the ND active detector region is 500 cm in order to allow the full extension of the selected events in the longitudinal direction so that the similar event topologies between ND and FD can be achieved and the systematic uncertainty is reduced.

  Containment cuts: The cuts require all the reconstructed showers of the tested event to be fully contained within the containment volume (green box in Fig.6.36). Demanding fully contained events improves the accuracy of the energy reconstruction of the neutrinos.

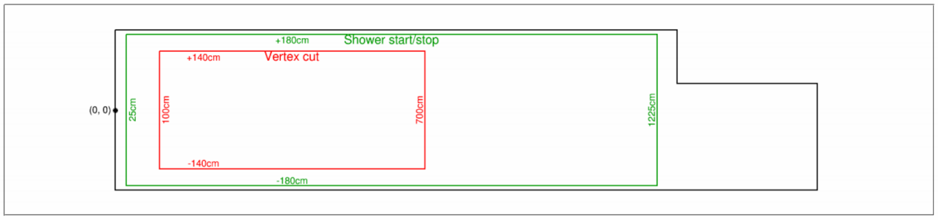


Figure 6.36: Diagram of the physical extension (black), fiducial volume (red) and containment volume (green) of the near detector.

- Preselection and PID cuts select the final set of ν_e-like events in the ND:

Front planes: The cut is equivalent to a fiducial cut in the Z direction, requiring that any hits are present in the reconstructed slice of at least 6 planes (∼40 cm) from the front wall of the detector. Because most of the rock events enter the near detector from the front wall, a tighter fiducial cut at the front wall, Front Planes > 6, is needed in order to obtain a sample free from rock.

Slice calorimetric energy: The cut requires that the slice calorimetric energy to be between 0 and 5 GeV, which is designed to be looser than the FD slice calorimetric energy cut in order to ensure that the selected ND sample covers almost all the kinematic region of the events selected in FD.

Number of slice hits: The number of slice hits should be within 20 - 200. Because the slice calorimetric energy has a strong dependence on the number of slice hits, for the same reason as the calorimetric energy cut, the ND number of slice hits cut is also chosen to be loose.

Shower length: The cut requires that the length of the longest shower is within 140 - 500 cm, the same as the FD shower length cut.

Gap: Similarly to the FD gap cut, the distance of the most energetic shower start from the reconstructed vertex should be less than 100 cm.

PID: LID and LEM cuts for ND are the same as the cuts for FD.

Table.6.11 summarizes the number of events and efficiency at different cut stages. The number of events drops significantly after the fiducial and containment cuts, since a large

Table 6.10: List of ND ν_e selections

Description	cuts
Data quality	nhit per plane < 8
Reconstruction quality	number of 3D shower > 0
Fiducial cuts	-140 cm $<$ Vertex X $<$ 140 cm
	-140 cm $<$ Vertex Y $<$ 140 cm
	100 cm $<$ Vertex Z $<$ 700 cm
Containment cuts	-180 cm $<$ Shower Stop and Start X $<$ 180 cm
	-180 cm $<$ Shower Stop and Start Y $<$ 180 cm
	25 cm $<$ Shower Stop and Start Z $<$ 1225 cm
Front planes	Shortest distance of slice from the front wall $>$ 6 planes
Slice calorimetric energy	5 GeV $<$ Slice calorimetric energy
Number of slice hits	20 $<$ Number of hits in slice $<$ 200
Shower length	140 cm $<$ Length of the longest shower $<$ 500 cm
Gap	Gap between the vertex and the leading shower $<$ 100 cm
LID	LID $>$ 0.95
LEM	LEM $>$ 0.8

number of rock events are removed. After the preselection cuts, the cut efficiency for all events is about 0.77% for MC which is consistant with the data 0.73%. The ratio of ν_e CC to ν_μ CC to NC is $\sim 1 : 30 : 20$. After PID cut, the ν_μ CC and NC components are reduced greatly. The ratio shifts to $\sim 3.5 : 1 : 2$ in LID region and $\sim 1 : 1 : 1$ in LEM region. The fraction of the ν_μ CC channel drops by the most, because the large difference in the topologies between muons and electrons. Beam ν_e CC is expected as an irreduciable component and its fraction is increased the most. After all the cuts, cosmic ray background is demonstrated to be all rejected in the near detector [102].

Table 6.11: Cut flow table of ND selections.

Cut	Total MC	Efficiency(%)	ν_μ CC	ν_e CC	NC	Data	Efficiency(%)
No cut	30049057	100.00	25742957	364671	3941430	29802297	100
Data quality	29139394	96.97	25132969	339183	3667242	29015588	97.36
Reconstruction	16338569	54.37	13577267	229803	2531499	15996841	53.68
Fiducial	1139793	3.79	758778	20924	360091	1025407	3.44
Containment	478835	1.59	253650	10236	214949	424518	1.42
Front planes	461975	1.54	243569	9815	208592	413242	1.39
Slice hits and E_{cal}	323131	1.08	177970	5437	139724	301019	1.01
Shower length	236834	0.79	145944	4859	86030	222046	0.75
Gap	230475	0.77	142987	4755	82732	217656	0.73
LID	2471	0.01	396	1292	783	2579	0.01
LEM	3225	0.01	1047	985	1193	3395	0.01

CHAPTER 7. FD BACKGROUND AND SIGNAL ESTIMATE

The ν_e appearance analysis in NOνA compares the FD prediction to the FD observation. Any excess over the background prediction is interpreted as evidence of ν_e appearance. The FD background and signal prediction is made through a method called extrapolation. We firstly identify different background interactions, ν_e CC, ν_μ CCand NC, in the ND data. Each ND decomposed channel is then separately corrected for the difference in the selection efficiency between the ND and FD and the potential difference in cross section to obtain the FD prediction. As the two NOνA detectors use identicial detection technology and read neutrinos from the same beam source, a FD beam background prediction driven by the ND data significantly reduces the systematic uncertainties of the final analysis result. On the other hand, the cosmic background in the FD is predicted using the out-of-time NuMI trigger data.

7.1 Extrapolation

One important condition that needs to be fulfilled to make the extrapolation a good approximation of the FD measurement is that the ND selected events should be good representatives of the ones selected in the FD in terms of kinematic properties. Two invariant kinematic variables, four momentum transfer, Q^2, and the mass of the system recoiling against the scattered lepton, W^2, are chosen to test the ND events representability [105]. The two variables are calculated in Eq.7.1 and Eq.7.2,

$$Q^2 = -q^2 = 2(EE' - \vec{\kappa} \cdot \vec{\kappa}) - m_l^2 - m_{l'}^2 \tag{7.1}$$

$$W^2 = (P + q)^2 = M^2 + 2(E - E') - Q^2 \tag{7.2}$$

where the basic kinematic variables are represented in Fig.7.1, in which $\vec{\kappa}$, E and m_l represent the four-momentum, energy and the mass of the incident lepton and $\vec{\kappa'}$, E' and $m_{l'}$ are the kinematic quantities for the outgoing leptons. The greater Q^2, the greater the four-momentum exchanged between the lepton and the nucleon and the smaller the mass of the virtual mediator represents. P is the four momentum of the initial nucleon with mass M. The greater W^2 is, the greater the mass of the final recoiling has. Q^2, energy exchanged, and W^2, energy of the final recoiling components which mostly consist of hadrons, are used to characterize different types of interactions, quasi-elastic, coherent, resonance and deep inelastic interactions. Fig.7.2 and Fig.7.3 show the comparison between the FD and ND 2D distributions of kinematic variables vs true neutrino energy distributions in the LID and LEM selected regions respectively. The ND distribution is represented in color, while the FD distribution is shown with boxes. The main peaks are found in the same place for both distributions, showing that the ND selected sample contains a good representative of the FD selected sample for LID. For LEM, some detector difference is observed in the peak near $W^2 = 1$, which comes from coherent single pion neutrino-nucleus scattering. The effect of the difference is then studied as part of the systematic uncertainty of GENIE parameters.

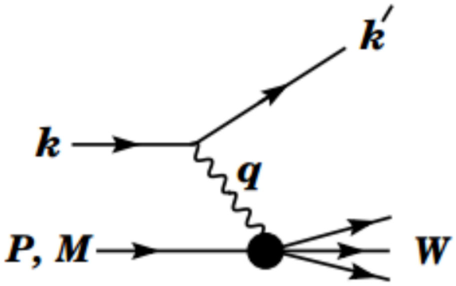
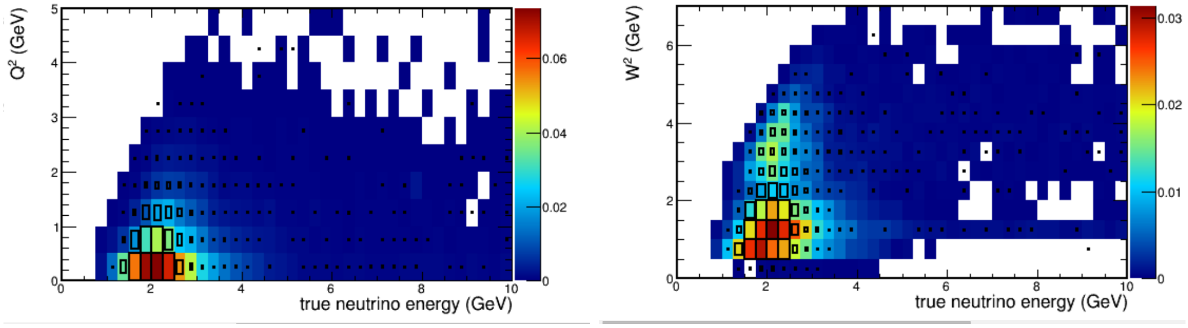


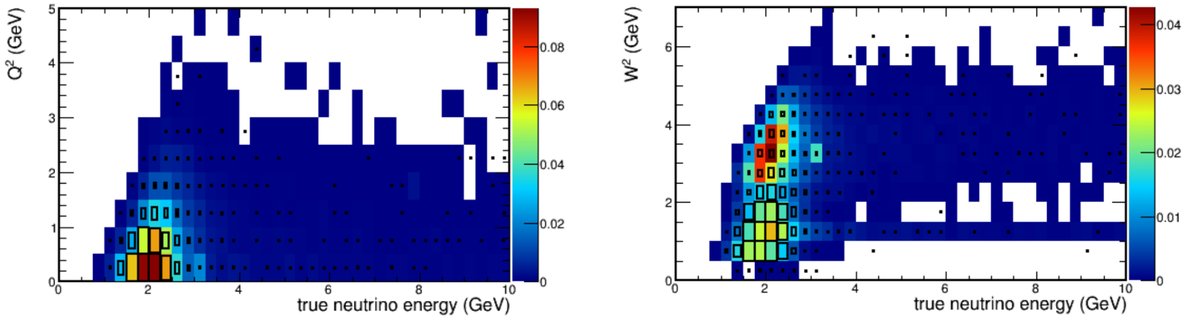
Figure 7.1: Diagram of a lepton-nucleon scattering.

For the first ν_e appearance analysis of NOνA, the background components in the near detector are proportionally decomposed based on the MC prediction. The main background components are ν_μ CC, ν_e CC and NC. The input of the decomposition process for background prediction is events selected by ν_e PID. The rate of the ND decomposed component α ($\alpha =\nu_\mu$



(a) (b)

Figure 7.2: The distribution of Q^2 (left) and W^2 (right) vs true neutrino energy in LID selected region. In color is the ND selected events and boxes are FD selected events.



(a) (b)

Figure 7.3: The distribution of Q^2 (left) and W^2 (right) vs true neutrino energy in LEM selected region. In color is the ND selected events and boxes are FD selected events.

CC, ν_e CC and NC) in reconstruction energy bin j B_j, $N^{Data}_{\alpha,S_e}(B_j)$, is calculated in:

$$N^{Data}_{\alpha,S_e}(B_j) = N^{Data}_{Tot,S_e}(B_j)\frac{N^{Data}_{\alpha,S_e}(B_j)}{N^{Data}_{Tot,S_e}(B_j)} \tag{7.3}$$

Then the decomposed component is extrapolated through Eq.7.4,

$$F^{Pred}_{\alpha\to\alpha,S_e}(B_j) = \sum_i \frac{N^{Data}_{\alpha,S_e}(B_j)F^{MC}_{\alpha,S_e}(E_i,B_j)}{N^{MC}_{\alpha,S_e}(B_j)}P_{\alpha\to\alpha}(E_i) \tag{7.4}$$

in which the event rate is scaled by the FD selection efficiency $F^{MC}_{\alpha,S_e}(E_i,B_j)$ and the oscillation probability $P_{\alpha\to\alpha}(E_i)$ and divided by the ND selection efficiency $N^{Data}_{\alpha,S_e}(B_j)$. The FD event rate at each reconstruction energy bin B_j is convoluted by $F^{MC}_{\alpha,S_e}(E_i,B^e_j)P_{\alpha\to\alpha}(E_i)$, which is a function of B_j and true energy E_i. The oscillation probability $P_{\alpha\to\alpha}(E_i)$ assumes no matter effect, $\delta_{cp} = 0$, $\Delta m^2_{32} = 2.35\times10^{-3}$, $sin^2 2\theta_{23} = 1$ and $sin^2 2\theta_{13} = 0.1$. The method above calculates the three main FD beam background channels: CC of $\nu_\mu \to \nu_\mu$, CC of $\nu_e \to \nu_e$ and NC of $\nu_\alpha \to \nu_\alpha$. The remaining eight channels: CC of $\nu_e \to \nu_\mu$, $\nu_e \to \nu_\tau$, $\nu_\mu \to \nu_\tau$, $\bar{\nu_e} \to \bar{\nu_e}$, $\bar{\nu_e} \to \bar{\nu_\mu}$, $\bar{\nu_e} \to \bar{\nu_\tau}$, $\bar{\nu_\mu} \to \bar{\nu_\mu}$ and $\bar{\nu_\mu} \to \bar{\nu_\tau}$ are predicted directly from the FD simulation, due to the fact that these rates are very small.

The FD signal channels $\nu_\mu \to \nu_e$ and $\bar{\nu_\mu} \to \bar{\nu_e}$ are extrapolated in a slightly different way. Because signal channel is expected to present as ν_μ on ND and oscillates to ν_e in FD, the input of the decomposition for the FD signal prediction is the sample passing ν_μ selection. In order to overcome the possible difference in the energy resolution between the two detectors, the ND event rate is converted into a function of true energy as:

$$N^{Pred}_{\nu_\mu,S_\mu}(E_i) = \sum_i \frac{N^{Data}_{\nu_\mu,S_\mu}(B_k) N^{MC}_{\nu_\mu,S_\mu}(E_i, B_k)}{N^{MC}_{\nu_\mu,S_\mu}(B_k)} \tag{7.5}$$

in which $N^{Data}_{\nu_\mu,S_\mu}(B_k)$ is obtained by proportionally decomposing the ND data. The extrapolation result is calculated using:

$$F^{Pred}_{\nu_\mu \to \nu_e, S_\mu}(B_j) = \sum_i \frac{N^{Pred}_{\nu_\mu,S_\mu}(E_i) F^{MC}_{\nu_e,S_e}(E_i, B_j)}{N^{MC}_{\nu_\mu,S_\mu}(E_i)} P_{\nu_\mu \to \nu_e}(E_i) \tag{7.6}$$

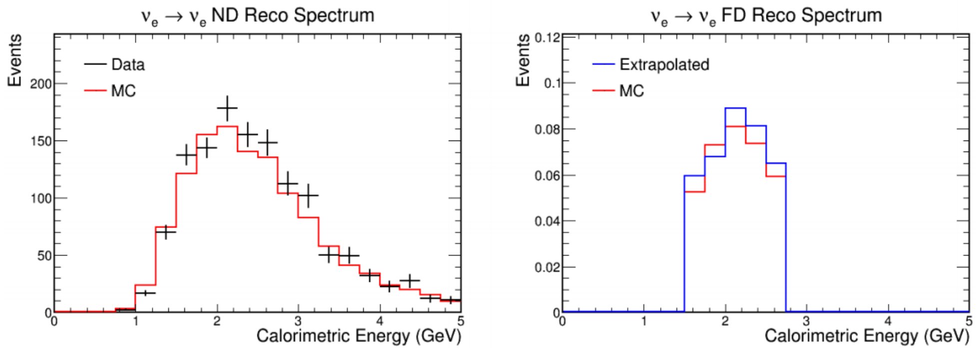
The event rate is weighted by $F^{MC}_{\nu_e,S_e}(E_i, B_j)/N^{MC}_{\nu_\mu,S_\mu}(E_i)$, to take into account the difference in the selection efficiency between ν_e in FD and ν_μ in ND and multiplied by the oscillation probability of $\nu_\mu \to \nu_e$ as a function of true energy. In the end, the FD event rate is converted back to be a function of reconstructed energy.

7.2 Background and Signal Prediction

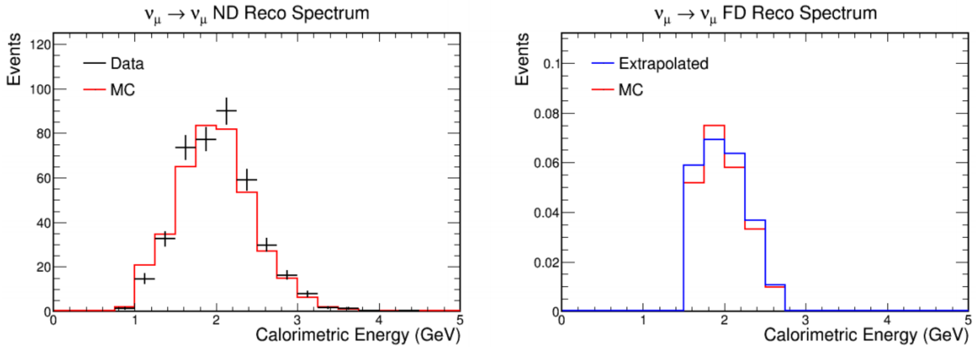
The final background prediction is made using the standard first analysis samples (Sec.5.1). The beam background is predicted using the extrapolation method described in section 7.1 and the cosmic background is predicted using the out-of-time NuMI trigger sample as presented in Sec.6.3. Table.7.1 shows the number of background events predicted by extrapolation in the first analysis data sample. The largest background component is the intrinsic beam ν_e CC, which is an irreducible background in the ν_e appearance analysis. The second largest component is the neutral current, which contains showers that are relatively easy to be misidentified as ν_e CC in the detector. The third is the ν_μ CC component, which before any selection cuts dominates the FD background. The cosmic ray background is reduced to a level equivalent to the ν_μ CC background. The total background is 0.94 in LID and 1.00 in LEM.

Table 7.1: The number of different background components in the far detector predicted by extrapolation as well as calculated using the far detector MC directly. The numbers are normalized to the POT of the first analysis data sample, $3.45 \times 10^{20} POT$.

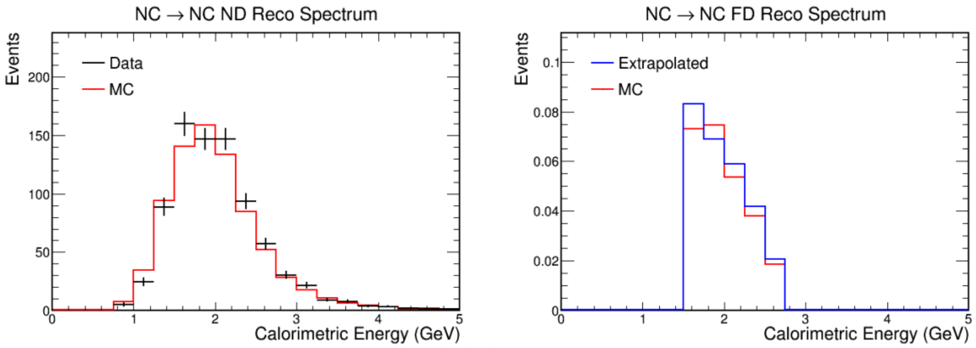
	Signal	Total Bkg	Beam ν_e CC	ν_μ CC	NC	$\nu_\tau CC$	Cosmic
LID Extrapolated	4.33	0.94	0.05	0.46	0.36	0.02	0.06
LEM Extrapolated	4.53	1.00	0.06	0.46	0.40	0.02	0.06
LID Far Detector MC	4.28	0.90	0.05	0.44	0.33	0.02	0.06
LEM Far Detector MC	4.45	0.97	0.07	0.44	0.38	0.02	0.06



(a) Distributions for FD beam ν_e CC event prediction.

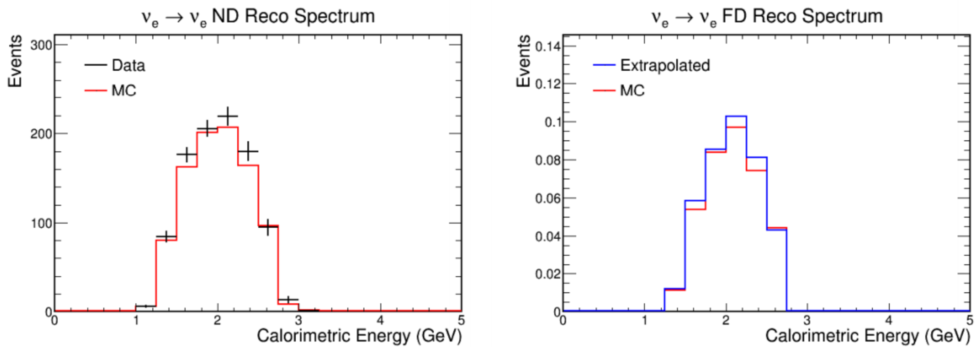


(b) Distributions for FD ν_μ CC event prediction.

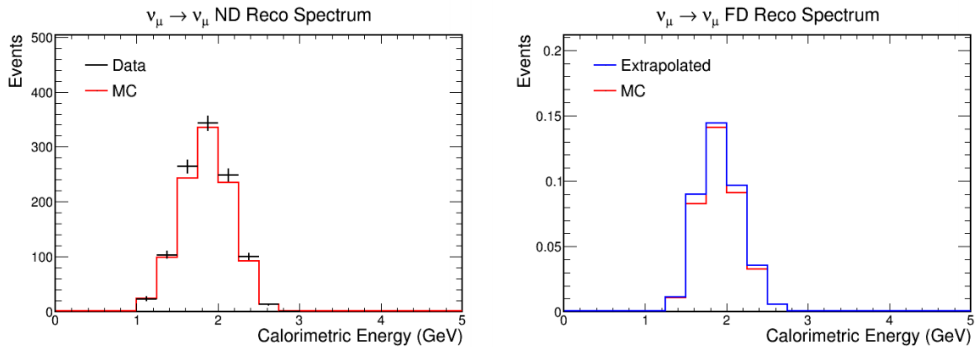


(c) Distributions for FD neutral current event prediction.

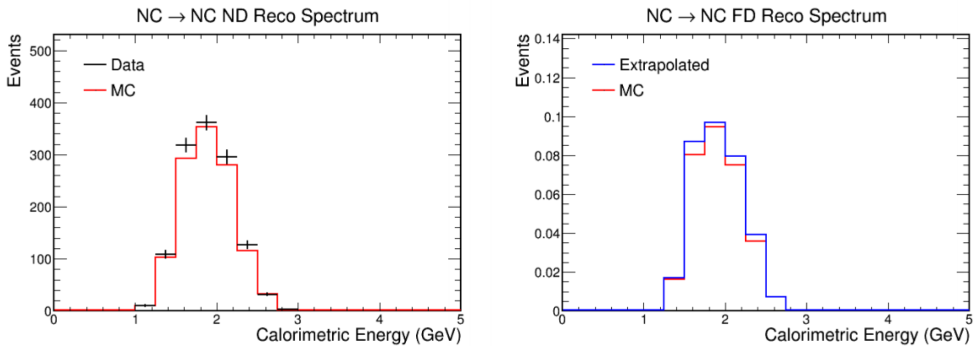
Figure 7.4: Distribution used in the beam ν_e CC (a), ν_μ CC (b) and NC (c) background extrapolation in LID region: Data vs MC comparison for NC in the near detector (left) and Extrapolation vs MC for the neutral current channel in the far detector (right).



(a) Distributions for FD beam ν_e CC event prediction.

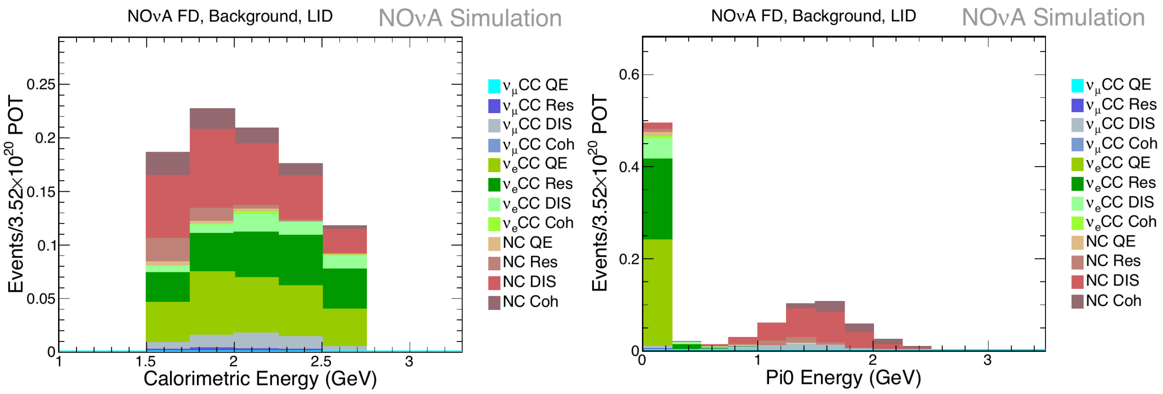


(b) Distributions for FD ν_μ CC event prediction.



(c) Distributions for FD neutral current event prediction.

Figure 7.5: Distribution used in the beam ν_e CC (a), ν_μ CC (b) and NC (c) background extrapolation in LEM region: Data vs MC comparison for NC in the near detector (left) and Extrapolation vs MC for the neutral current channel in the far detector (right).



(a) (b)

Figure 7.6: Total event energy (left) and contained π^0 energy (right) distribution for predicted FD background events categoried into different types of interactions.

Fig.7.4a through Fig.7.5c show the data vs MC comparison for different channels (beam ν_e CC, ν_μ CC and NC) in the near detector (left) and the extrapolation vs MC comparison in the FD (right) in LID and LEM selected regions. The numbers obtained by the extrapolation method which is driven by ND data is $\sim 5\%$ higher than the number suggested by FD MC, because the ND data shows a slight excess over the ND MC with equivalent POT. The composition of the FD background is also studied that, as shown in Fig.7.6, the FD beam background is dominated by beam ν_e CC and NC DIS (See Fig.7.6a) and the latter mostly contains a π^0 with energy close to 2 GeV (See Fig.7.6b).

7.3 Systematic Uncertainties

Many of the systematic uncertainties cancel due to the two functionally identical detectors. But several still remain. The main contributing systematic effects are beam flux, scintillator saturation, calibration, cross section, hadronization, containment efficiency.

There are different ways of evaluating the effect of different systematic uncertainties. Some effects, such as calibration uncertainty, scintillation saturation, require to make new modified MC sample with systematic shift applied. These uncertainties usually vary event topologies, the effect of which can be only evaluated through a complete reconstruction and particle identification process. Both nominal and shifted samples are processed through the standard analysis chain, including reconstruction and PID and standard extrapolation steps. The difference between the predicted numbers from nominal and shifted samples is recorded as the systematic uncertainty due to the effect.

The evaluation of some of the systematic uncertainty can be done by manipulating the nominal samples. For example, for the containment uncertainty, since it only changes the final selection result, the FD predictions extrapolated using different portions of the ND are calculated. The difference between the predictions is recorded as the uncertainty caused by the ND containment effect.

Table.?? lists main systematic uncertainties in the FD prediction. The total uncertainty is 8.44% and 10.59% for total background and signal in LID. Details of each individual items will be discussed in the following sections.

Table 7.2: List of systematic errors due to different effects for background and signal in both LID and LEM region. The last row corresponds to the total systematic error which is the sum in quadrature of all errors.

	LID		LEM	
	Total Bkg. (%)	Signal (%)	Total Bkg. (%)	Signal (%)
Beam	3.18	1.06	2.85	1.04
Birks	5.14	7.22	4.62	7.94
Calibration	4.44	7.58	7.90	3.68
Light Level	n/a	1.00	n/a	1.00
Neutrino Interaction	3.72	13.98	5.47	12.01
Containment	1.84	n/a	1.34	n/a
Rock contamination	0.10	n/a	0.09	n/a
Decomposition	3.90	n/a	3.50	n/a
Data & MC Stat.	3.47	0.6	2.85	0.6
Normalization	1.20	1.20	1.20	1.20
Total	10.11	17.58	12.06	14.99

7.3.1 Beam Effects

The beam systematic uncertainty comes from two type of sources: variance in the beam configuration, such as horn current, beam size, beam position, the position of different devices and the uncertainty of hadron production in the beam simulation [96], [107]. The beam effects influence both the near and far detectors. To evaluate the uncertainty caused by the beam effects, we shifted up and down each of the beam systematics and proceed through extrapolation and compare to the FD prediction made using nominal sample. Table.7.3 shows the list of the main beam related uncertainties with respect to the signal and background prediction. The overall uncertainty is dominated by item NA49, which comes from the uncertainty of hadron production simulation. NA49 is an experiment designed to measure hadron production parameters. The uncertainty due to hadron production is determined by varying hadron production parameterizations of FLUKA simulation within a physically feasible range around the best fit to the NA49 data. The uncertainty is evaluated to be 3.1% of the total background prediction in the LID region. The next largest source is the modeling of the magnetic field in the horns, which is 0.38%. The third largest source is the beam position X on target. As shown in Fig.6.3, the width of the beam position X on target is about 0.5 mm. The prediction made using the MC samples with ±0.5 mm shifts in the beam position on target shows a change by 0.36%.

Table 7.3: Percentage difference between nominal and beam systematically shifted samples of background and signal predictions in LID and LEM regions.

Diff %	LID		LEM	
	Signal	Background	Signal	Background
Horn Current	0.04	0.17	0.04	0.12
Beam Spot Size	0.05	0.08	0.05	0.13
BeamPosX	0.14	0.36	0.11	0.32
BeamPosY	0.11	0.03	0.09	0.02
H1Pos	0.23	0.33	0.21	0.40
H2Pos	0.03	0.23	0.01	0.28
TargetPos	0.02	0.14	0.02	0.04
ExpMagnField	0.01	0.38	0.02	0.35
NA49	1.01	3.10	1.01	2.77
Combined	1.06	3.18	1.04	2.85

7.3.2 Scintillator saturation

As discussed in Sec.5.2, when a large amount of energy is deposited in the scintillator, the output light yield saturates and is no longer proportional to the input energy. The Birks-Chou relation (Eq.5.1) is used to describe the effect [152]. By comparing data to the NOνA MC using the dE/dx distribution for proton candidates in the near detector, the two Birks-Chou parameters are studied to be $k_B = 0.04\text{g}^1\text{MeV}^{-1}\text{cm}^{-2}$ and $k_C = -0.0005\text{g}^2\text{MeV}^{-2}\text{cm}^{-4}$. where k_B is 4 times greater than the typical values measured by some other experiments [152] and k_C describes the secondary effect and is usually small. Smaller k_B value will lead to a larger energy loss in scintillator and thus lower total event energy and also causes missing hits which has impact on the event topology. The effect of the difference is evaluated by calculating the final background and signal prediction using the MC samples with two different parameterizations of k_B and k_C, $k_B = 0.01\ g\ MeV^{-1}\ cm^{-2}$, $k_C = 0\ g^2\ MeV^{-2}\ cm^{-4}$ (BirksB) and $k_B = 0.02\ g\ MeV^{-1}\ cm^{-2}$, $k_C = 0\ g^2\ MeV^{-2}\ cm^{-4}$ (BirksC). The result is shown in Table.7.4.

BirkC result that presents -7.22% for signal and -5.14% for background is chosen to be used in the final systematic uncertainty calculation.

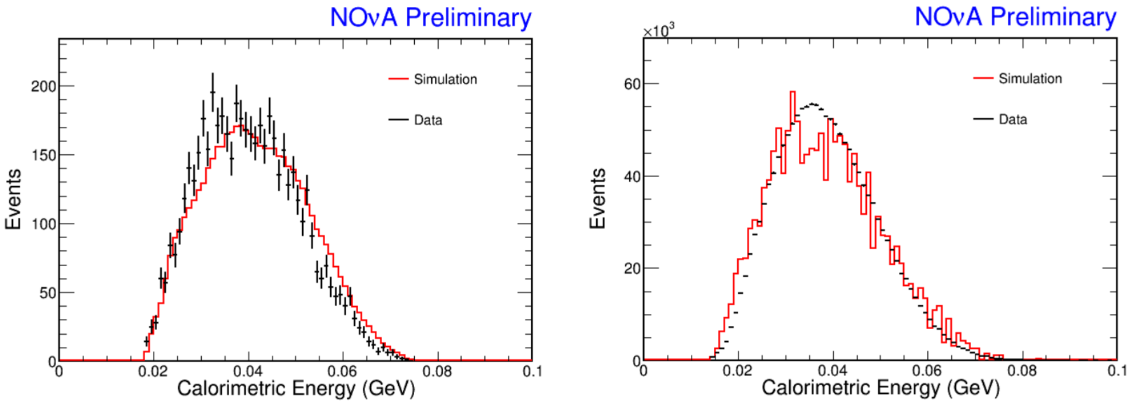
Table 7.4: Percentage difference between nominal and Birk systematically shifted samples of background and signal predictions in LID and LEM regions.

Diff (%)	Signal	Background
LID BirksB	-5.12	-8.80
LID BirksC	-7.22	-5.14
LEM BirksB	-6.28	-3.22
LEM BirksC	-7.94	-4.62

7.3.3 Calibration Systematics

The calibration systematic uncertainty comes from the imperfectness of calibration, which can take different forms: absolute calibration uncertainty and relative uncertainty involving cell to cell difference, in-cell flat shift and in-cell shift as a function of cell length. In order to study the relative miscalibration effect, shifts are applied at the calibration step and the intentionally miscalibrated samples are then processed through the standard reconstruction and particle identification chains. The absolute miscalibration effect does not require to process new samples and is studied by applying the shift to the extrapolation spectra. The miscalibration effect in the two detectors is expected to cancel largely.

Absolute miscalibration: The uncertainty of the absolute calibration for the first analysis is determined as the data vs MC difference in the reconstructed energy of Michel electrons. The Michel electron produced in a muon decay presents an upper energy limit of 53 MeV and is usually used as a standard candle to evaluate the calibration performance. Fig.7.7 shows the data/MC comparison for the calorimetric energy of Michel electrons in both ND and FD [108]. The miscalibration level is calculated as the percentage difference between the mean of the data and MC distributions, resulting in 1% for ND and 5% for FD. To take effect of the absolute calibration uncertainty into full consideration, 5% is chosen to be the shift in absolute calibration, which represents less than 1% of the total background and 6% of the signal prediction in the LID selected region.

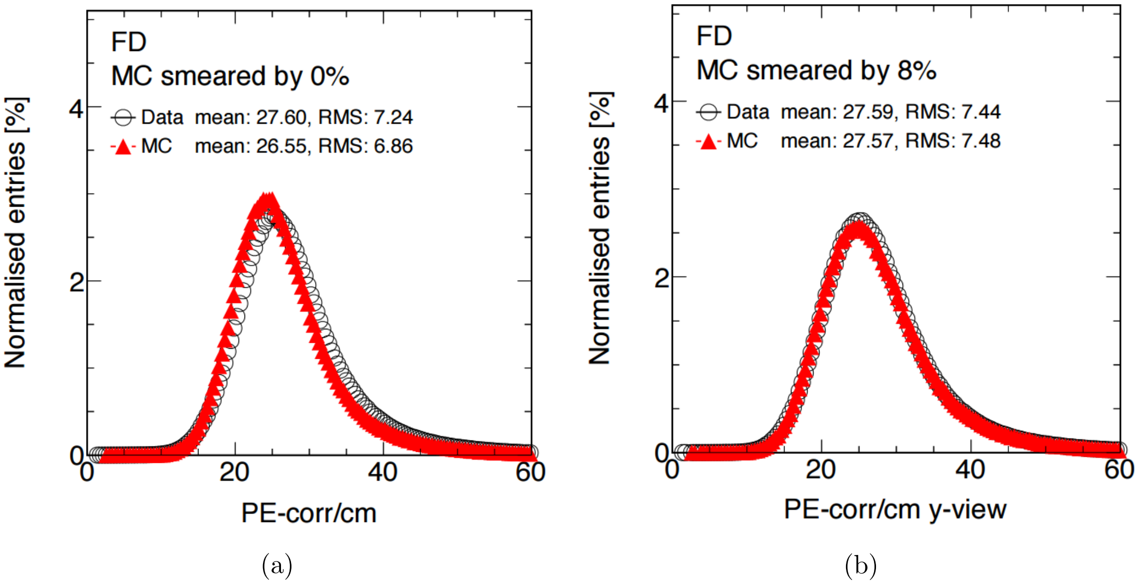


(a) (b)

Figure 7.7: The data/MC comparison in the calorimetric energy of Michel electrons in ND (left) and FD (right).

Cell-to-cell difference: The cell-to-cell difference is refered to the different performance of the attenuation calibration from cell to cell. Fig.7.8 shows the data vs MC comparison in PECorr/cm variable for the calibrated hits [109]. In order to match the MC to data, the cell response in MC needs to be smeared by 8% which is then taken as the cell-to-cell difference. The uncertainty is applied to each cell as a 8% smearing on the attenuation calibration constant. Fig.7.10 shows the comparison between the samples with nominal and shifted calibrations for the events passing pre-selection in the ND and the ones passing cosmic rejection cuts in the FD. The figures present small differences for either ND or FD [110], [111]. The resulting differences in the total background and signal prediction are less than 1% in the LID selected region.

In-cell flat shift: After the in-cell calibration, ideally the response at any position along the cell should be normalized to the response at the center of the cell. However, due to the fluctuations in data and the uncertainties of the attenuation fit, difference exists. In order to evaluate the difference between data and simulation, a polynomial function is fitted to the data/MC ratio of the attenuation constant as a function of in-cell position W (See Fig.7.9). The difference can be parameterized in two terms: flat shift which is defined as the amount of data/MC difference at the center of the cell ($W = 0$) and slope shift which is the rest of the fitted polynomial function with the flat shift subtracted and

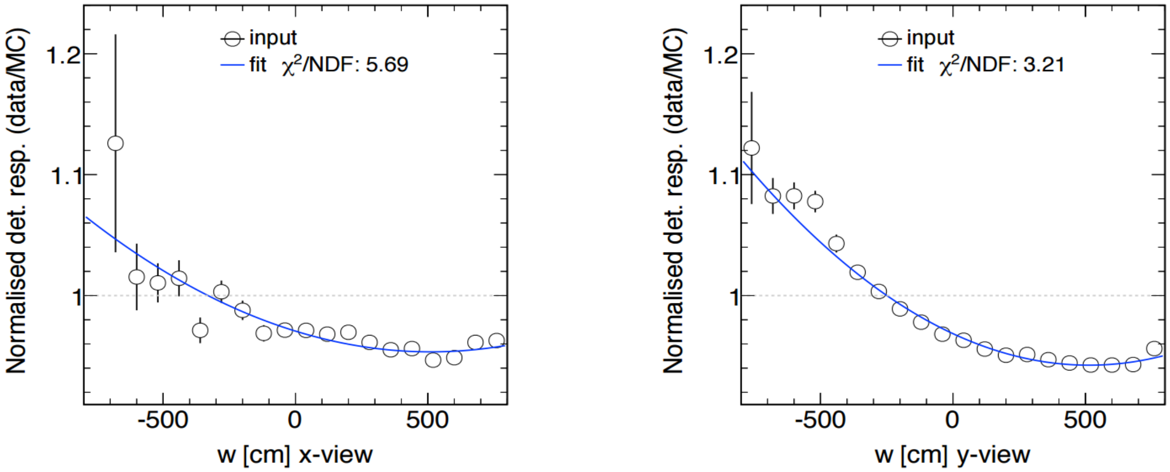


(a) (b)

Figure 7.8: The Data/MC comparison of the PECorr/cm distribution for the calibrated hits before (left) and after (right) MC smearing. In order to overlap the MC to data, MC is smeared by 8%, which is taken as the cell-to-cell difference.

contains the information of the difference in shape. In this part, only the former will be discussed and, according to the fit, the amount of the shift is 8%. Two sets of samples, one with attenuation constants shifted uniformly by 8% down and the other by 8% up, are processed through the reconstruction and PID. Fig.7.11 and Fig.7.12 show the nominal vs shifted comparison for 8% down and 8% up samples. Shifts in peaks can be clearly observed in both ND and FD for all the components. The overall uncertainty of the prediction is calculated as the average of the flat 8% down and flat 8% up samples. The effect on the final prediction is 4.22% of total background and 2.84% of signal predictions in LID region.

In-cell cell length dependent shift: The in-cell cell length dependent shift describes the difference in shape of the attenuation constant distributions between data and simulation. For the first analysis, the shift is applied as a linear function of W so that the shift in the attenuation constants is 8% at the two cell ends and is 0% the shift at $W = 0$. ”Slopeup” sample is the one with the constant shifted up at near end and down at far end by 8% so that the calibrated energy of hits at near end is overestimated and the ones at far



(a) (b)

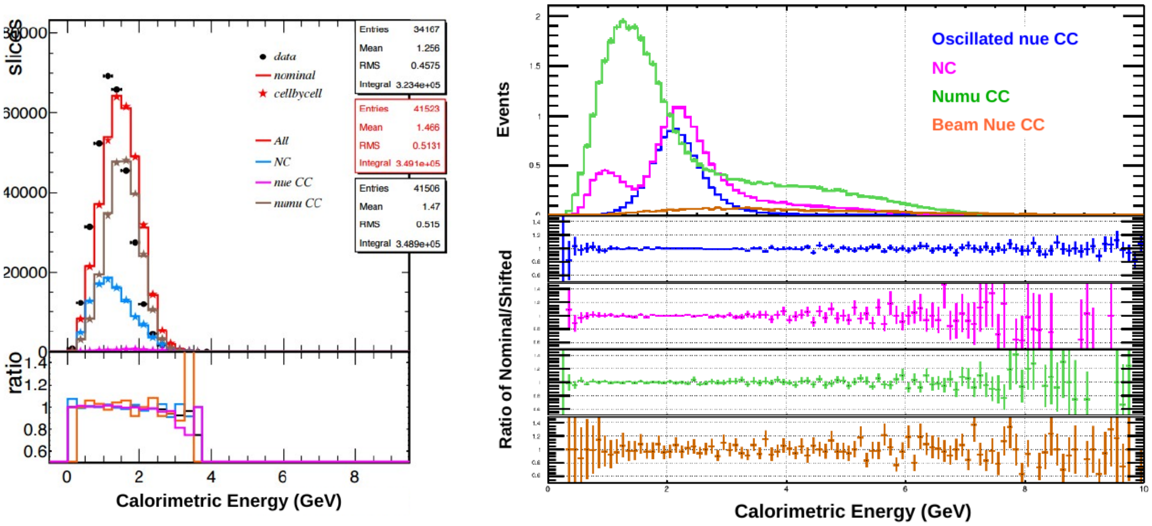
Figure 7.9: The data to MC ratio of the attenuation constant as a function of W (black) and the polynomial fit (blue) for the X-view (left) and Y-view (right) cells in FD.

end is underestimated, while ”Slopedown” sample applies an incorrect increase in the calibrated energy of hits at far end and a decrease in the ones at near end. Fig.7.13 and Fig.7.14 show the nominal to shifted comparison in the LID selected spectra for Slopeup and Slopedown samples respectively. A small difference is observed as the slope of the attenuation constant distribution changes in both ND and FD. According to the Table.7.5, the uncertainties are -3.21% for the total background and -1.51% for signal in the LID region.

Table 7.5: Calibration systematic error in terms of percentage difference w.r.t. the nominal background and signal predictions in LID and LEM regions.

	LID				
% diff	signal	total bkg.	ν_μ CC	NC	beam ν_e CC
Absolute	5.98	0.76	8.45	10.81	10.89
Relative	4.36	3.72	12.07	15.65	6.39
Random	-0.93	-0.39	1.47	1.47	1.87
Slope	1.33	2.27	2.94	10.51	3.92
Total	7.58	4.44	15.10	21.78	13.35

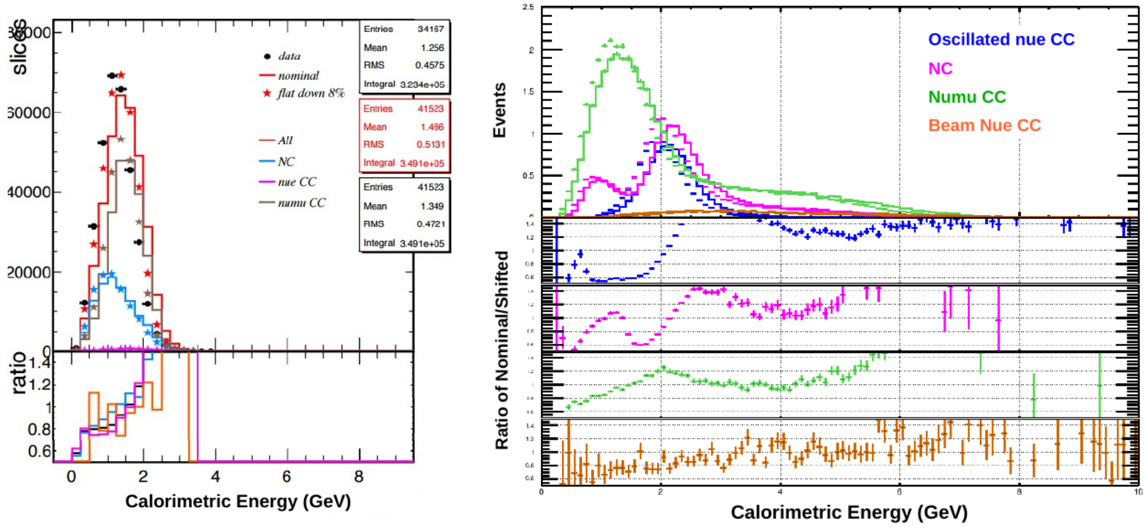
	LEM				
% diff	signal	total bkg.	ν_μ CC	NC	beam ν_e CC
Absolute	1.20	5.80	12.25	0.81	10.36
Relative	2.84	4.22	10.46	1.87	8.33
Random	-1.33	-0.80	1.09	-0.42	1.43
Slope	-1.51	-3.21	-4.35	1.06	6.45
Total	3.68	7.90	16.72	2.34	14.84



(a)

(b)

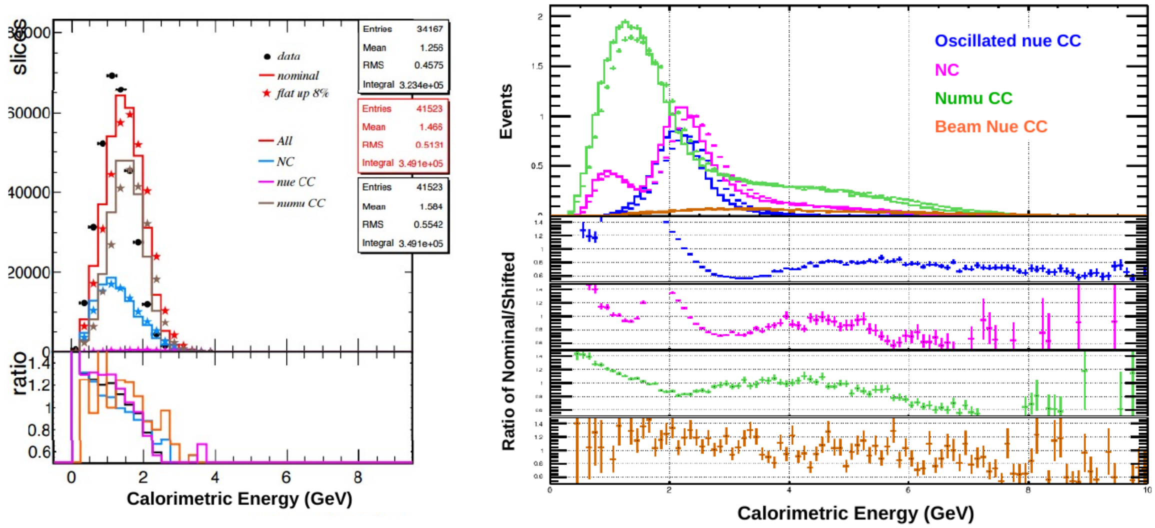
Figure 7.10: Comparison between the LID selected spectra with norminal and shifted calibration for ND (left) and FD (right).



(a)

(b)

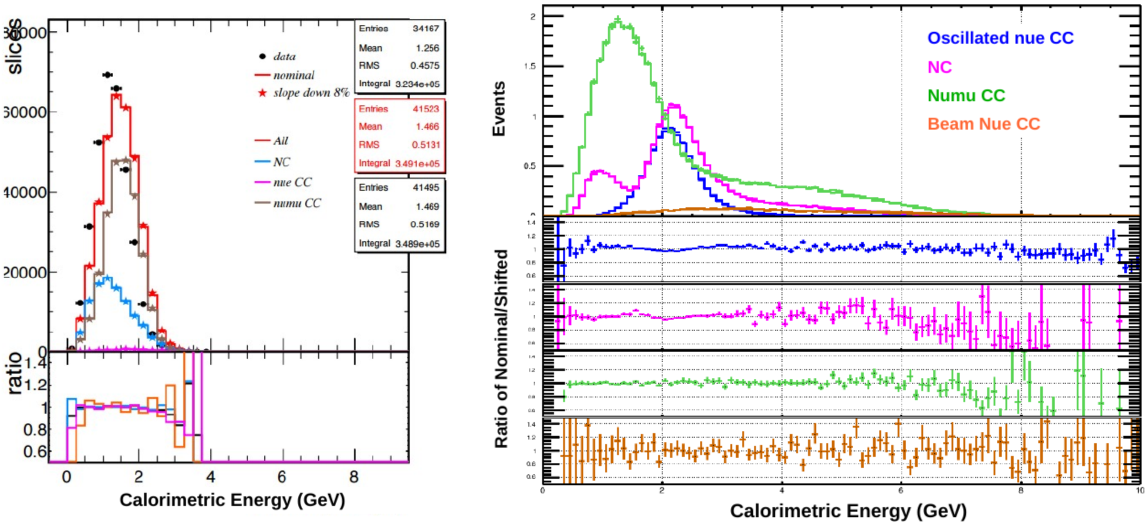
Figure 7.11: Comparison between the LID selected spectra with norminal and flatdown by 8% calibration for ND (left) and FD (right).



(a)

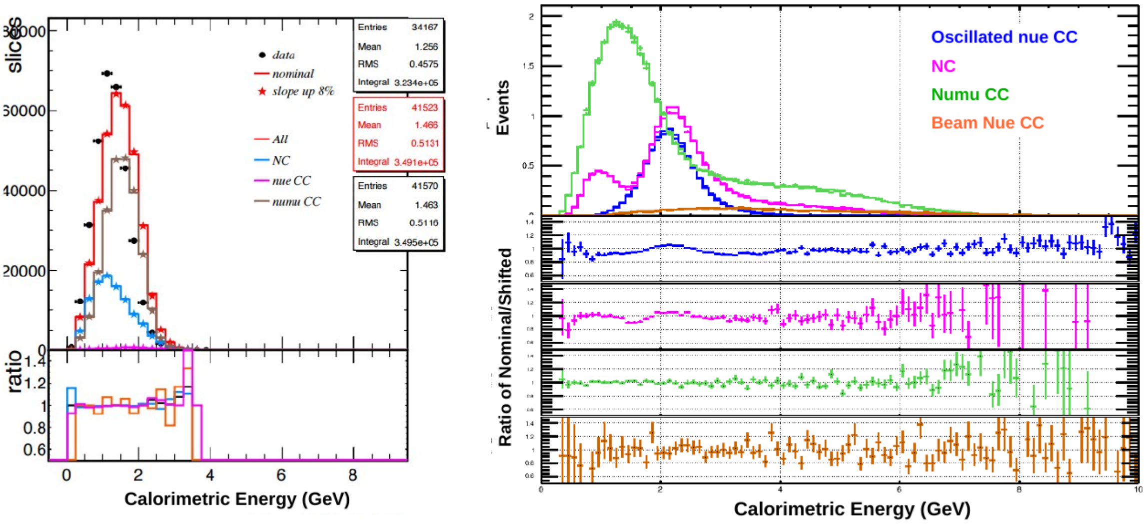
(b)

Figure 7.12: Comparison between the LID selected spectra with norminal and flatup by 8% calibration for ND (left) and FD (right).



(a) (b)

Figure 7.13: Comparison between the LID selected spectra with norminal and slopedown by 8% calibration for ND (left) and FD (right).

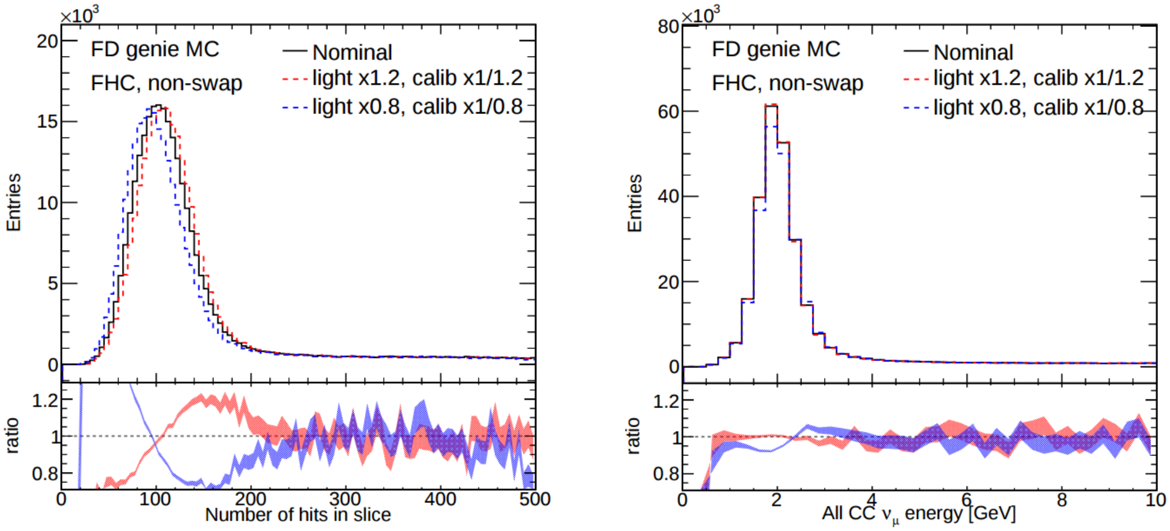


(a) (b)

Figure 7.14: Comparison between the LID selected spectra with norminal and slopeup by 8% calibration for ND (left) and FD (right).

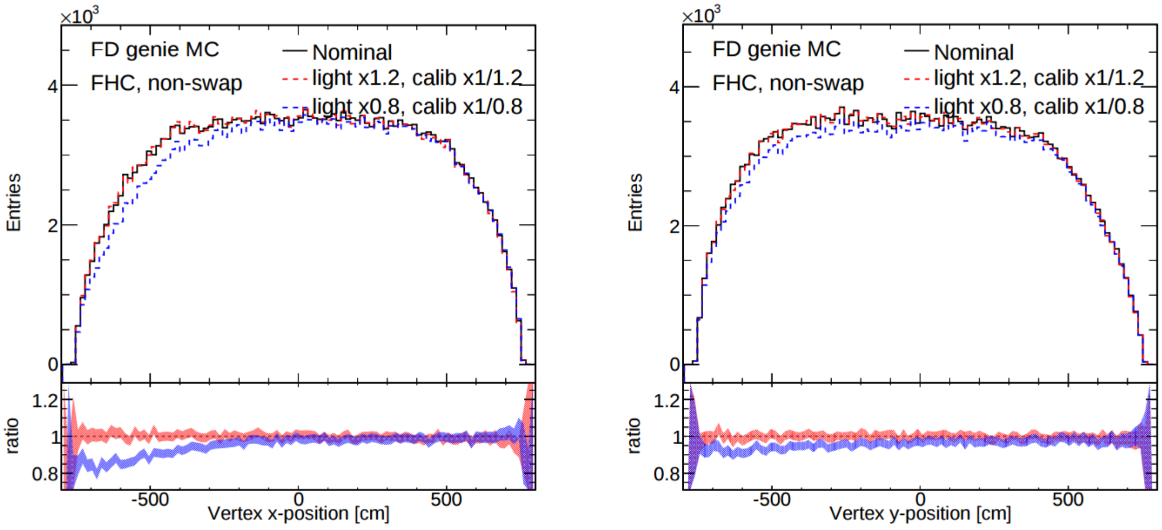
7.3.4 Light Levels

In order to reduce the impact of noise, a threshold is applied: only hits with PE higher than a threshold are recorded. The light level which determines how much photoelectrons can be produced has a big impact on the analysis result. The light-level of the NOνA far detector is known to be underestimated by 20% at the far end in the simulation [112]. The effect of the difference in light level on some basic analysis variables are studied using sample with hit PE value shifted as a function of W before the PE threshold is applied. According to Fig.7.15 and Fig.7.16, even though the increase in the light level does not change the reconstruction result by much, a decrease in the light level clearly shows a 5% decrease in the number of slice hits and pull the reconstructed vertex toward the near end [113]. The difference in light-level changes the event selection efficiency as well as the event pattern which results in different particle identification results. Because FD is much larger than ND, the effect is expected to be greater in FD than in ND and will not be cancelled efficiently in the extrapolation. The change in the final prediction due to the difference in light-level is negligible for background and is 1% for signal.



(a) (b)

Figure 7.15: Comparison between nominal (black), light-level shited up by 20% (red) and light-level shifted down by 20% (blue) for total number of slice hits (left) and the reconstructed ν_μ CCenergy (right).



(a) (b)

Figure 7.16: Comparison between nominal (black), light-level shited up by 20% (red) and light-level shifted down by 20% (blue) for the reconstructed vertex position X (left) and reconstructed vertex position Y (right).

7.3.5 Neutrino interaction

As GENIE is used as a tool by NOνA to simulate neutrino interactions with matter, the uncertainties of GENIE parameters will also cause errors in the final extrapolated prediction. The GENIE parameters are measured and updated by different experiments and a list of the uncertainties of the parameters are sugggested by the GENIE authors [153]. There are three main categories: cross-section, hadronization model and final state interaction uncertainties. The tool of the reweighting due to the different shifts in parameters is built in GENIE and the output is the reweighting factor of final number of events due to different type of systematic modifications. The reweighted spectra are used in the extrapolation to study the effect on the final prediction. Based on Table.??, the total GENIE systematic uncertainty is 13.98% for signal and 3.72% for total background in LID. Since the signal channel, $\nu_\mu \rightarrow \nu_e$, involves the simulation of neutrino interactions of two neutrino flavors in ND and FD, the ratio between the cross section of ν_e to the cross section of ν_μ has large impact on the FD signal prediction. The total uncertainty for the FD signal prediction is studied to be dominated by the change in the axial vector mass in QE events, which is used in the calculation of the axial vector component in the neutrino QECC cross section. As the parameter varies by ±25%, the final signal prediction changes by 13.91%

A data driven method is used to estimate the level of the uncertainty of the coherent π^0 cross section. Two different samples were prepared: one that is dominated by the non-coherent events and the other one after the coherent interaction selections with both coherent signal and non-coherent background. The normalization factor of the non-coherent background is extracted from the data/MC comparison in the first sample. In the second sample, the background in MC is scaled by the normalization factor obtained from the last step. Using the second sample, we computed the data/MC difference in the coherent signal. This difference is 0.59 and is chosen to be the uncertainty of coherent π^0 cross section, which is then applied to the corresponding components in the extrapolation spectra. The contribution to the final prediction is negligible for both the signal and the background in the LID region.

7.3.6 Near Detector Containment

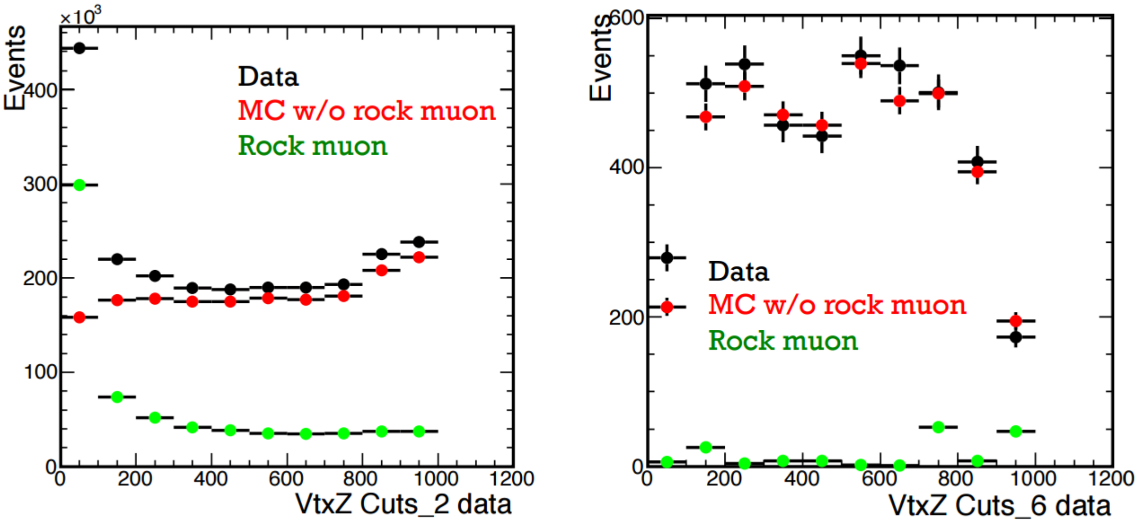
Ideally, the selected ND events should be detected uniformly in the detector. However, due to the variance in the detection efficiency of different regions of the detector, the rate of selected ND events depends on the choice of containment cut, which also presents as a potential uncertainty in the selection result. The effect exists only in the near detector. A data driven method is used to evaluate the uncertainty of the final prediction due to different containment cuts. Extrapolations using data from different parts of ND, front or rear, east and west, bottom or top and inner or outer parts, are examined (See Table.7.6). It is found that, for both PIDs, the front, west, top and inner parts of the detector present a lower rate of selected events than the standard containment region, while the rear, east, bottom and outer parts show a higher selected event rate. The maximum variance in the total background is -2.7% in LID and -2.6% in LEM and is taken as the systematic uncertainty caused by the near detector containment cut.

Table 7.6: List of Extrapolation result using data from different defined ND regions. The maximum variance is 2.7% for LID and 2.6% for LEM, which are taken as the systematic uncertainty due to the near detector containment cut.

Diff %	total bkg.	ν_μ CC	beam ν_e CC	NC	ν_τ CC
LID $100 < vZ < 400$ cm	-2.59	-1.54	-2.01	-3.66	0
LID $400 < vZ < 700$ cm	1.63	3.08	2.01	0.98	0
LID $0 < vX < 140$ cm	-2.69	-3.08	-2.75	-2.93	0
LID $-140 < vX < 0$ cm	2.11	4.62	3.11	0.49	0
LID $0 < vY < 140$ cm	-1.92	0	-0.92	-3.66	0
LID $-140 < vY < 0$ cm	0.58	1.54	0.55	0.49	0
LID $\|vX, vY\| < 99cm$ cm	-1.92	-1.54	-1.83	-2.20	0
LID $99\|vX, Vy\| < 140cm$ decomposition	1.25	3.08	2.01	0	0
LEM $100 < vZ < 400$ cm	-2.58	-1.09	-2.15	-3.65	0
LEM $400 < vZ < 700$ cm	0.86	2.17	2.15	-1.01	0
LEM $0 < vX < 140$ cm	-2.32	-1.09	-1.08	-4.26	0
LEM $-140 < vX < 0$ cm	0.77	1.09	1.25	0.20	0
LEM $0 < vY < 140$ cm	-1.46	-1.09	0	-3.45	0
LEM $-140 < vY < 0$ cm	0.43	1.09	0	-1.42	0
LEM $\|vX, vY\| < 99cm$ cm	-1.89	0	-1.61	-2.84	0
LEM $99\|vX, Vy\| < 140cm$ decomposition	0.43	1.09	1.97	-1.42	0

7.3.7 Rock Contamination

The ND rock events that come from the neutrino interaction in the rock surrounding the detector cause uncertainties in the ND analysis. Simulated rock events are mixed into the simulated ND neutrino samples and in order to increase the statistics of the rock events, each simulated rock event is reused 350 times. The question, whether the ratio of ND neutrino to ND rock events in the MC matches with realistic condition exists. Fig.7.17a shows the distributions of vertex position z for ND data, ND neutrino MC and ND rock MC after data and reconstruction quality cuts [114]. The data to MC ratio is a function of distance from the front wall, and is ∼1 near the edge of the detector but is consistantly below 1 inside the detector, which implies that the current rate of ND rock events is too high. After the LID cut (Fig.7.17b), the data vs MC agreement is better and rock event rate is lower. To estimate the effect of the ND rock event rate on the final prediction, MC samples with and without rock events are examined. The one without rock events is obtained by requiring the true vertex inside the detector. The study shows that the contribution of rock contamination to the final background prediction is far below 0.1% and negligible compared to some other effects.



(a) (b)

Figure 7.17: Distributions of Vertex position Z for data (black), neutrino MC (red) and rock MC (green) after data/reconstruction quality cuts (left) and after LID cut (right).

7.3.8 Background decomposition in ND

In the standard extrapolation, the ND data is decomposed to different channels (ν_μ CC, beam ν_e CC, NC) proportionally to the ND Monte Carlo prediction. As difference between the total number of events of data and MC exist, the method has its error in the event prediction. One way to evaluate the error is to study the extrapolation results with all the ND data/MC difference alternatively assigned to one single component. The method does not require to process new samples and the change could be made during the extrapolation. Table.7.7 shows the percentage difference in the extrapolated background components for the shifted samples. The maximum percentage difference, 3.9% of the total background prediction for LID and 3.5% for LEM, occurs, when all the ND data/MC difference is assigned to the ν_μ CC component. This is taken as the final systematic uncertainty of the decomposition method.

The prediction is also checked by some independent data-driven decomposition methods. The beam ν_e CC predicted based on the number of selected ν_μ events in ND [115] suggests about 3% more NC events in LID and 1% more in LEM compared to the standard decomposition

Table 7.7: Percentage difference in extrapolated background components for different systematically shifted samples, in which the total ND data/MC is alternatively assigned to one component.

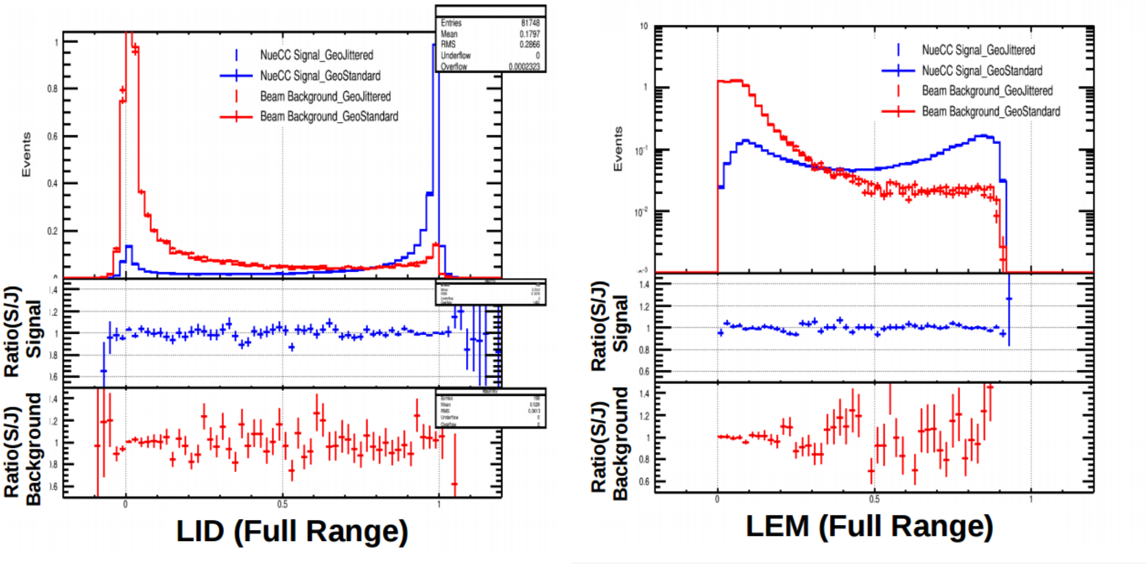
Diff %	total bkg.	ν_μ CC	beam ν_e CC	NC	ν_τ CC
LID ν_e CC decomposition	0.77	-4.62	6.41	-5.61	0
LID NC decomposition	0.86	-4.62	-6.23	11.46	0
LID ν_μ CC decomposition	-3.94	26.15	-6.23	-5.61	0
LEM ν_e CC decomposition	2.41	-4.35	10.22	-5.27	0
LEM NC decomposition	1.03	-4.35	-4.66	8.52	0
LEM ν_μ CC decomposition	-3.53	10.87	-4.66	-5.27	0

method; The SliceMEF method tags the ν_μ CC event based on the number of the selected Michel electrons in the vicinity of the event. [116] This method presents an extrapolation result that varies within 10% with respect to the standard extrapolation, which is comparable to the uncertainty prediced in Table.7.7. The MRCC method predicts the ND NC component with muon removed ν_μ CC sample [117]. The result of MRCC shows a much higher prediction of NC than the standard decomposition, which is due to the failure of the MRCC events in representing the NC event topology in the fine segmented NOνA detector.

7.3.9 Alignment

Ideally, all planes in a detector align perfectly with a baseline and the position of the energy deposition read out, reflecting the exact position of the energy deposition. However, in reality, the planes are misplaced more or less due to various effects, such as imperfection in the construction, deformation of the PVC tube caused by high inner pressure and gravity effect. The misalignment usually happens in two dimensions, displacement in X and Y direction and rotation in the XY plane, which changes the event topology and might cause extra challenges for the PID algorithms. As the FD is much larger than the ND, the misalignment is observed to be greater in the FD.

New samples with planes either shifted by 1.27 cm or tilted by 1.25 mrad, which is the level of misalignment measured in FD [118], are made to study the effect. Fig.7.18 shows that the sample with shifted geometry presents a negligible difference in the PID distributions compared to the sample with nominal geometry [119]. Table.7.8 shows that the difference between the nominal and misaligned samples in terms of the number of different FD channels is very small, ∼1% for both signal and total background.



(a)

(b)

Figure 7.18: Comparison between samples with the nominal (vertical bar) and shifted (cross) geometry in LID (left) and LEM (right) of both signal (blue) and total beam background (red) as well as the ratio of nominal to shifted distributions.

Table 7.8: Percentage difference in the extrapolated prediction of samples with misaligned geometry.

Diff %	signal	total bkg.	ν_μ CC	beam ν_e CC	NC	ν_τ CC
LID extrapolation	-0.73	-1.73	-1.54	-1.83	-1.95	0
LEM extrapolation	-0.73	-1.55	-1.09	-1.43	-1.83	0

7.3.10 Monte Carlo Statistics

Both ND and FD MC samples used in the extrapolation are limited by statistics, which will cause a systematic uncertainty in the predicted result due to the failure of including all possible topologies in MC. The effect of MC statistics is studied by examining the final prediction with the extrapolation spectra shifted by the corresponding statistics uncertainty. As the "Total" rows in Table.7.9 suggest, the difference caused by the MC statistics in the number of final FD prediction is 3.47% for total background in LID and 2.85% in LEM.

Table 7.9: Percentage difference of the extrapolation result between nominal and MC statistics shifted sample.

LID				
Channel	FD MC	ND Data	ND MC	Total (%)
beam ν_e	3.98	2.46	1.04	4.81
ν_μ	6.19	2.50	1.05	6.76
NC	2.07	2.48	1.05	3.40
Total	2.27	2.40	1.01	3.47
LEM				
Channel	FD MC	ND Data	ND MC	Total (%)
beam ν_e	3.88	1.97	0.82	4.43
ν_μ	4.91	1.88	0.79	5.31
NC	1.89	1.76	0.74	2.69
Total	2.07	1.79	0.75	2.85

7.3.11 POT Normalization

The correct prediction requires to normalize the extrapolated number to the correct POT exposure. The uncertainty of the POT exposure counting comes from several sources: a small drift of one of the beam devices results in 0.5% difference in POT counting, while the ND reconstruction efficiency represents 0.8% difference between data and MC. The mass of the detector is measured to have 0.7% uncertainty. The total normalization uncertainty is calculated as the quadratic sum of the three effects and represents 1.2% of the POT exposure counting. Since the normalization is applied as an overall scaling factor in the extrapolation, the uncertainty in the POT exposure counting is taken directly as the uncertainty of the extrapolation result.

CHAPTER 8. RESULTS

The signal region was kept blind, while the analysis was developed. Before the signal region is opened, two sidebands of the signal region, one with high energy and the other one with low PID, are examined as a final check of all the procedures. After the result inside the signal region is revealed, the properties of the observed events in the signal region are tested. In the end, the compatibility of the NOνA observation with the prediction is tested as a function of θ_{13}, mass hierarchy and δ_{cp}.

8.1 Sideband Study

8.1.1 High Energy Sideband

The energy region defined for selecting ν_e CC signal is 1.5-2.7 GeV. Data and prediction is compared in the high energy sideband, 3-10 GeV region, before the blind region is opened [127]. Because LEM has included energy information in the training and presents low score for high energy events, the check is only performed for LID not for LEM. As the slice calorimetric energy has strong dependence on the number of slice hits, the upper limit of the slice hit cut is removed in the study. The signal and beam background components are predicted by extrapolation. In order to ensure the ND selected event contains a good representative of FD events, the energy cut in ND is loosened to 0-10 GeV and the upper limit of the number of slice hits is removed as well. The cosmic background is predicted using out-of-time NuMI trigger data.

Table.8.1 shows the comparison in the number of different types of events in high energy sideband region. There are 22.68 beam events and 32.71 cosmic ray background events predicted in the whole LID preselection region, while data presents 57, which is consistent the prediction. A loose LID cut is applied, 3.06 events are predicted and 2 events are observed in data (See

Table.8.1). So data and prediction also matches well in loose PID region. In the standard PID region, the statistics of the selected events is low. Because the neutrino flux is low in the 3-10 GeV energy region as well as LID is not trained to identify electron neutrinos with high energy.

Table 8.1: The comparison between FD prediction and measurement for events passing high energy cut. The beam components are predicted by extrapolation and cosmic background is predicted using out-of-time NuMI trigger data.

	ν_e CC MC	beam Bkg. MC	Cosmic Bkg.	Total Prediction	Data
no LID cut	4.565	18.113	32.710	55.388	57
LID $>$ 0.7	2.149	0.673	0.237	3.059	2
LID $>$ 0.95	0.594	0.077	0.147	0.818	0

Among the 57 events selected in the FD data, by manually scanning based on event topology, 27 are categorized as beam candidates and 30 as cosmic ray background events. Fig.8.1 presents the comparison between cumulative FD POT exposure and the cumulation of 27 FD beam candidates in the high energy region as a function of run. A KS (Kolmogorov-Smirnov) test shows a value of 0.778 for the two distributions and prove a good compatibility between POT exposure and the number of selected beam candidates. Fig.8.2 is the comparisons in slice calorimetric energy and LID between FD data and prediction, which prove to agree well within fluctuations.

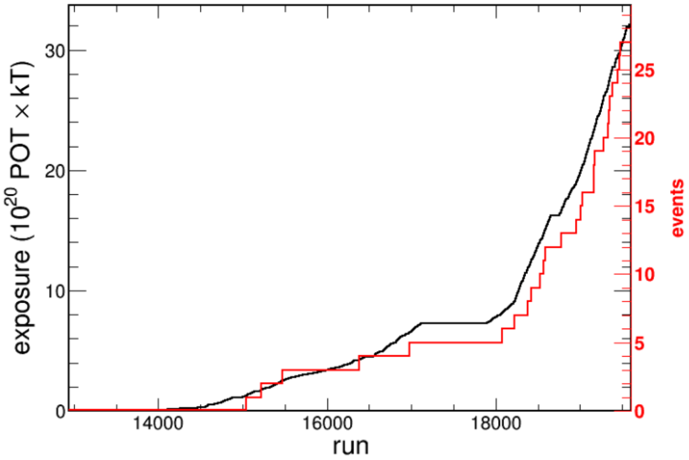
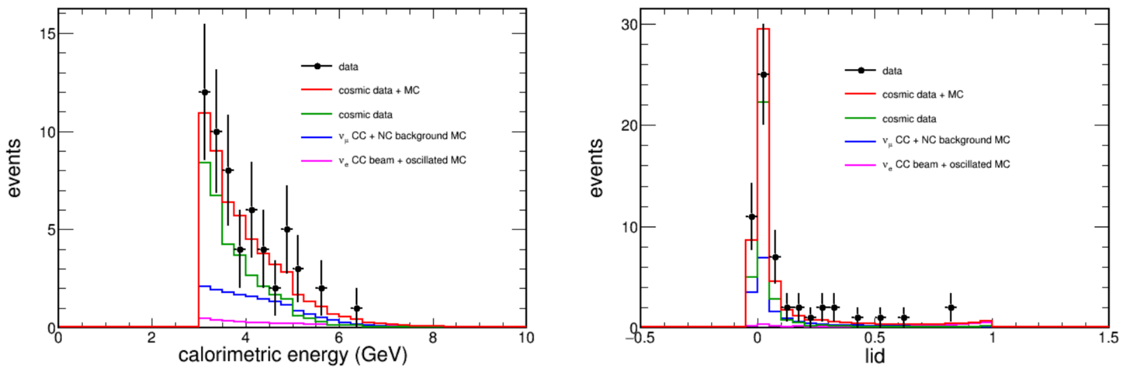


Figure 8.1: The Comparison between accumulative FD POT exposure (black) and the accumulation of 27 FD beam events (red) in the high energy region. A KS (Kolmogorov-Smirnov) test shows a value of 0.778 and proves the compatibility between the two distributions.



(a)

(b)

Figure 8.2: The comparison in energy and LID between FD data and FD prediction for events passing high energy cut, 3 GeV $<$ CalE $<$ 10 GeV.

8.1.2 Low PID Sideband

The signal regions for the two PID algorithms are optimized to LID $>$ 0.95 and LEM $>$ 0.8. The low PID regions are then defined as 0.7-0.95 for LID and 0.6-0.8 for LEM. The result in the low PID sidebands are checked before the signal region is opened [128]. In the study, only the PID cuts are changed and all other aspects of the analysis remain the same as the standard ones. Table.8.2 is the comparison in the number of events passing low PID cuts between FD prediction and data. For LID, the total prediction is 4.21, of which 2.85 as beam component which is predicted by the standard extrapolation and 0.38 from out-of-time NuMI trigger data as cosmic ray events, while 5 events are observed in the low LID sideband in the far detector. For the secondary PID, the total prediction is 2.89, while the actual observation is 2 events [135]. A good consistency between the FD prediction and measurement in the low PID regions is demonstrated.

Table 8.2: The comparison between FD prediction and measurement for events passing low PID cut. The beam components are predicted by extrapolation and cosmic background is predicted using out-of-time NuMI trigger data.

	ν_e CC MC	beam Bkg. MC	Cosmic Bkg.	Total Prediction	Data
0.7 $<$ LID $<$ 0.95	2.33	1.5	0.38	4.21	5
0.6 $<$ LEM $<$ 0.8	1.79	1.02	0.08	2.89	2

8.2 NOνA ν_e Appearance Result

After good agreement between the FD prediction and observation is proved in the two sideband, high energy and low PID regions, the signal region is opened. In LID region, 6 candidates are observed, while the secondary PID, LEM, selects 11 candidates. The observation is compared to the FD background prediction, ~ 1, in both LID and LEM selected region. A series of properties are then checked for these selected candidates to ensure the robusness of the result.

8.2.1 Properties of ν_e CC Candidates

The 6 LID candidates are also included in the 11 LEM selected events. So the ratio of the number of events selected by both LID and LEM to selected by LEM not LID to selected by LID not LEM is 6:5:0, $N_{BothLIDandLEM} : N_{LEMnotLID} : N_{LIDnotLEM} = 6 : 5 : 0$. According to the result of background prediction, the probability of the events falling in each categories is $P_{BothLIDandLEM} : P_{LEMnotLID} : P_{LIDnotLEM} = 0.2615 : 0.3154 : 0.4231$. A trinomial probability equation (Eq.8.1)

$$P = \frac{N!}{N_1!N_2!N_3!} P_1^{n_1} P_2^{n_2} P_3^{n_3} \tag{8.1}$$

is used to calculate the probability of achieving the observation of (6,5,0) given the probability of (0.2615, 0.3154, 0.4231) [164]. in which 1, 2 and 3 index different categories, BothLIDandLEM, LEMnotLID and LIDnotLEM. The calculation based on Eq.8.1 suggests a probability of 9.2% for selecting 6 candidates in both LID and LEM, 5 in only LEM and 0 in only LID region. The number of selected events is then compared to the POT exposure as a function of run (See Fig.8.3) [129]. A KS test is performed on the accumulation of the 6 LID candidates to the cumulative POT exposure and shows a KS probability of 0.789 for the two distributions, which proves that the spread of the 6 selected events is compatible with the distribution of POT exposure.

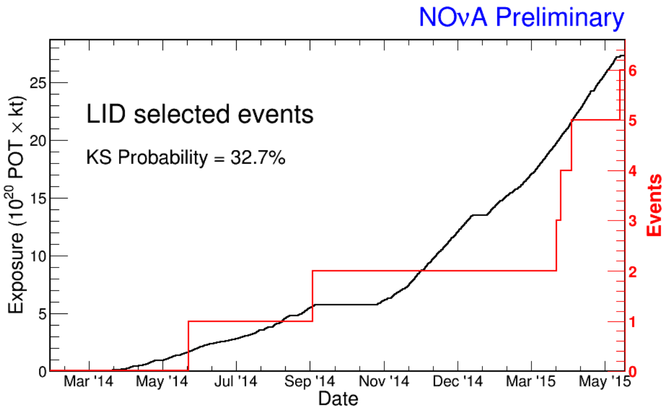


Figure 8.3: The Comparison between accumulative FD POT exposure (black) and the accumulation of 6 ν_e CCcandidates observed in the LID signal region. A KS test shows a value of 0.789 and proves the compatibility between the two distributions.

The distribution of a series of variables are checked for the candidates [131]. Fig.8.4 shows the timing distribution for all selected events compared to the in-spill window. A total of 6 ν_e CC candidate events are observed in the spill window, while 2 events present outside the in-spill window and are taken as cosmic ray background. The 6 ν_e CC candidates selected by LID concentrate near the middle of the selection window of the slice calorimetric energy variable (Fig.8.5), which is consistent with the prediction made by the extrapolation. The 6 candidates all present LID values higher than 0.98 as shown in Fig.8.6. The distribution of the variables, including the number of slice hits (Fig.8.7), the number of planes of the leading shower (Fig.8.9), the cosine of the angle between the leading shower and the beam direction (Fig.8.10), the fraction of transverse momentum (Fig.8.11) are checked for the candidates and all show good consistency with the extrapolation prediction. The plots in Fig.8.8 show the position of vertices of the candidates inside the detector. All events spread uniformly throughout the detector as expected and are well contained except for one that is close to the front wall. The distribution for LEM selected events are shown in App.E and no unexpected behavior is observed for the 11 ν_e CC candidates in LEM region.

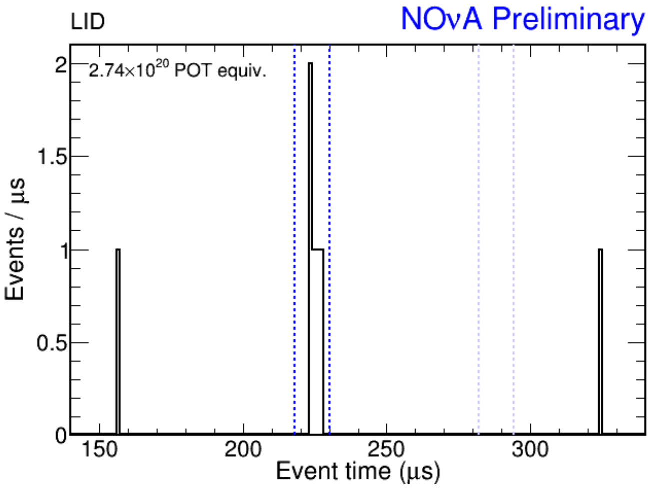


Figure 8.4: The distribution of mean time for all LID selected events (black), including the ν_e CCcandidates which are inside the in-spill window (dash light blue) as well as two cosmic ray background evnets which are outside the spill window.

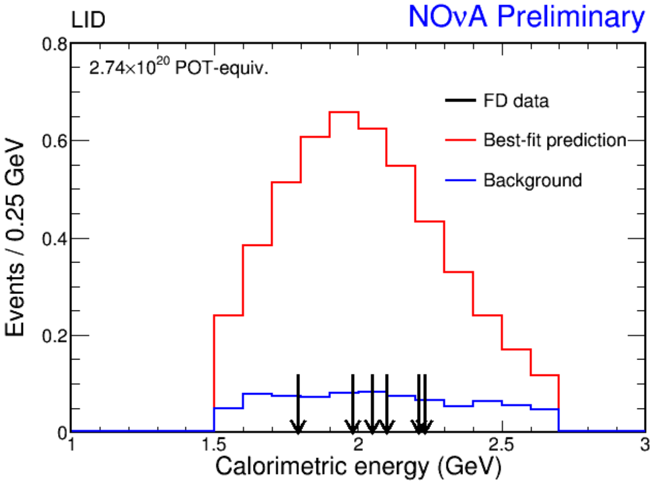


Figure 8.5: The distribution of slice calorimetric energy for the ν_e CC candidates in LID (black) overlapped with the extrapolation predicted background (blue) and predicted background + signal (red).

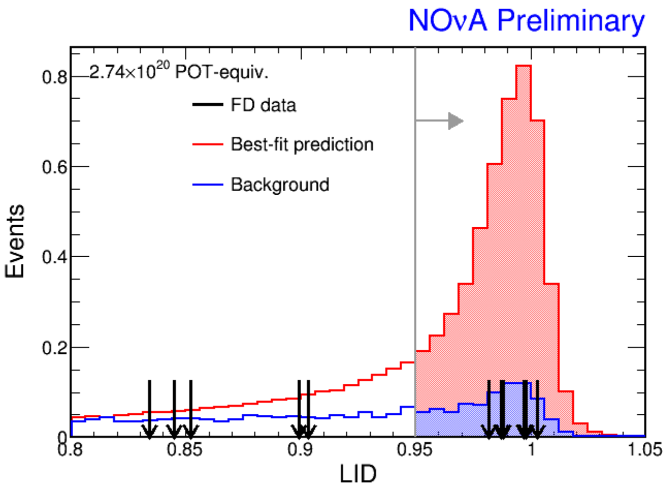


Figure 8.6: The distribution of LID for the LID selected events (black) in the signal (shaded) and low PID (line) region. The distribution for the selected events are compared to the predicted distributions from the extrapolation, the blue for just predicted background and red for the combination of predicted signal and background.

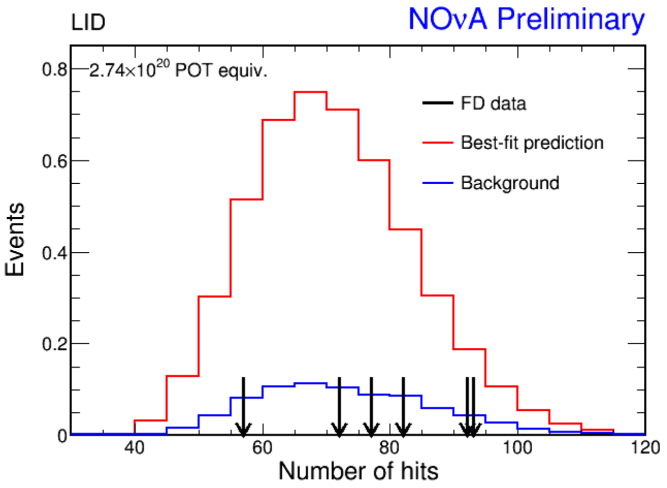
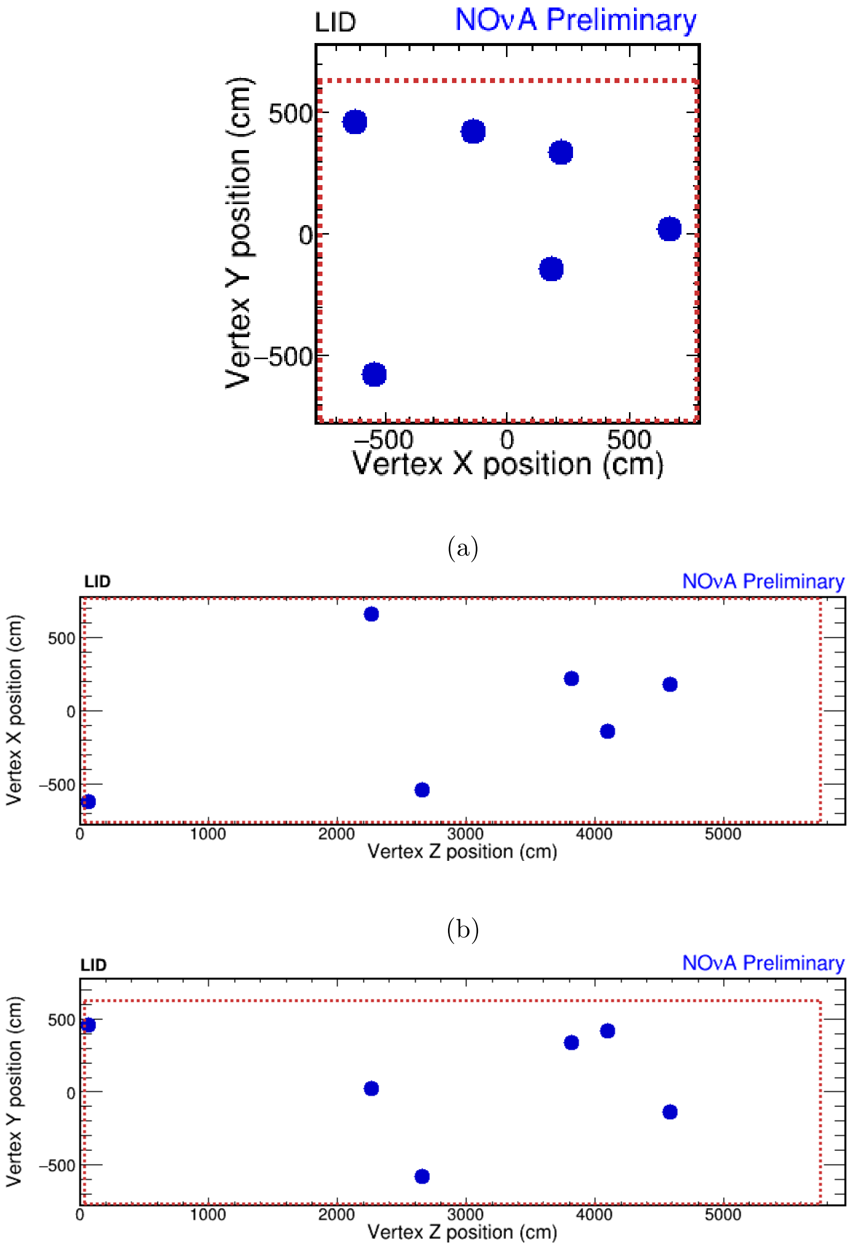


Figure 8.7: The distribution of the number of slice hits for the ν_e CC candidates in LID (black) overlapped with the extrapolation predicted background (blue) and predicted background + signal (red).



(a)

(b)

(c)

Figure 8.8: The position of the vertices of the LID selected ν_e CC candidates in XY (a), XZ (b) and YZ (c) planes. The position of the vertices is compared to the containment area (dash).

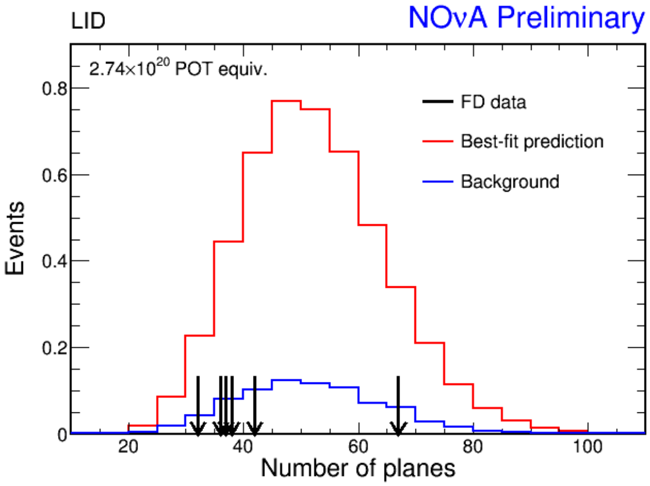


Figure 8.9: The distribution of the number of planes of the leading shoewr for the ν_e CC candidates in LID (black) overlapped with the extrapolation predicted background (blue) and predicted background + signal (red).

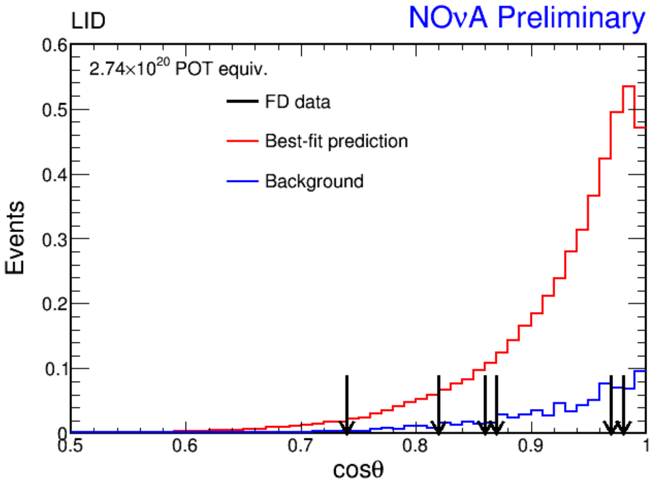


Figure 8.10: The distribution of the cosine of the angle of the leading shower w.r.t. the beam direction for the ν_e CC candidates in LID (black) overlapped with the extrapolation predicted background (blue) and predicted background + signal (red).

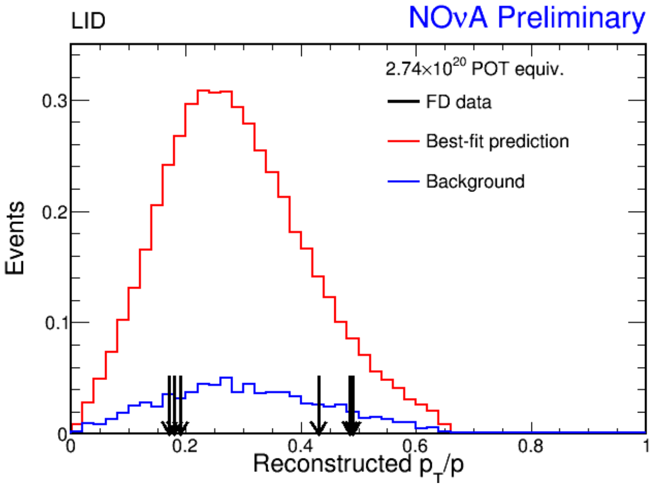
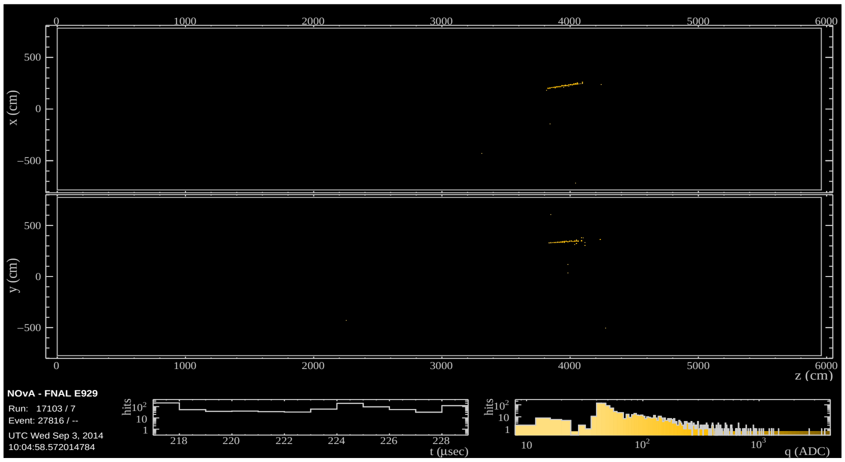
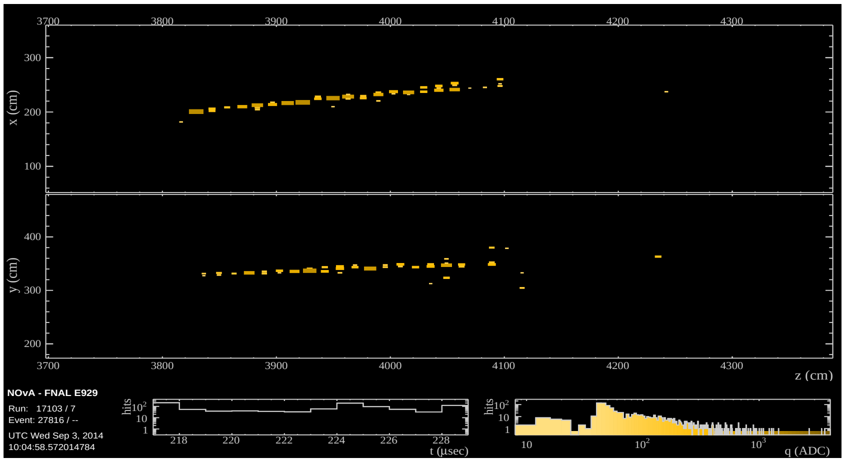


Figure 8.11: The distribution of the fraction of transverse momentum for the ν_e CC candidates in LID (black) overlapped with the extrapolation predicted background (blue) and predicted background + signal (red).

Fig.8.12 shows the full and zoomed-in event display for one of the LID selected candidates in Run 17103 [134]. The event is dominated by an energetic shower about 300 cm in length and a few cell widths in radius. The event display for the rest of the candidates can be found in App.D. Through visual scan, all the selected events present patterns, such as event position, orientation and event shape, similar to expected ν_e CC interactions. The longitudinal and transverse energy deposition rate, dE/dx, for the same event as in the event display is shown in Fig.8.13 and is compared to the simulated dE/dx distribution [132]. The dE/dx data points follow the highlighted region of the simulated distribution well. The dE/dx distributions for the rest events selected by both LID and LEM as well as by only LEM are presented in App.D and most of them show behavior of dE/dx close to the simulated distribution.

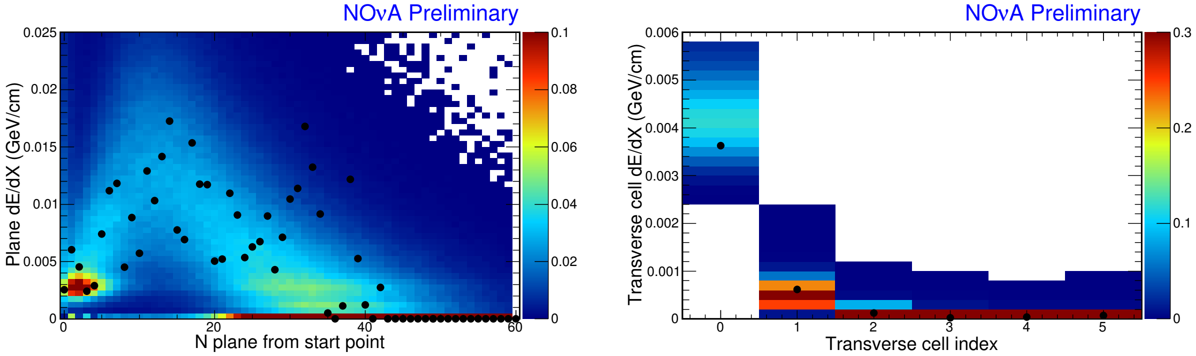


(a)



(b)

Figure 8.12: Full (top) and zoomed-in (bottom) event display of one of the ν_e CC candidates in Run17103 selected by both LID and LEM. The color of the boxes are corresponding to the amount of energy deposited in cells.

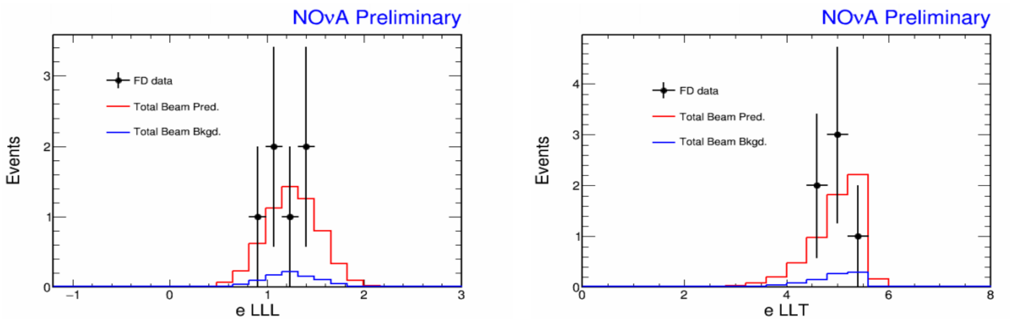


(a)

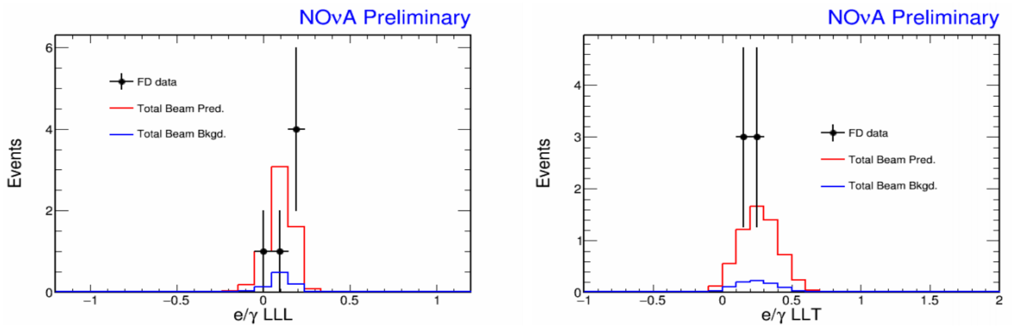
(b)

Figure 8.13: Longitudinal (left) and transverse (right) energy deposition rate, dE/dx, of one of the ν_e CC candidates in Run17103 selected by both LID and LEM. The distribution of the event (black) is compared to the simulated dE/dx (colored).

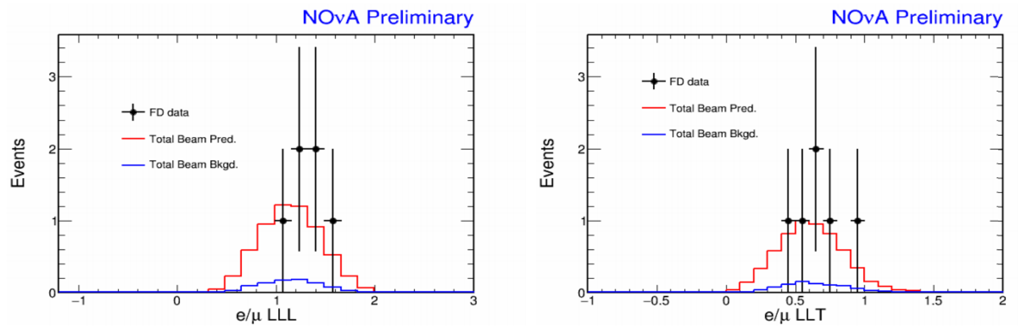
The likelihood of the leading shower of the candidate event being an electron based on the event topology in longitudinal and transverse directions are presented in Fig.8.14a and are compared to the distributions for simulated events. The likelihood of the 6 candidates in both LID and LEM falls well in the region that is predicted by the simulation. Fig.8.14b, Fig.8.14c and Fig.8.14d show the likelihood of being an electron minus the likelihood of being a photon, a muon and a neutral pion respectively. The likelihood differences for all the candidate events appear to be above 0, which implies that the leading shower of the event is more likely to be an electron than any of the other three types of particles, which originate from main background components for the ν_e appearance analysis.



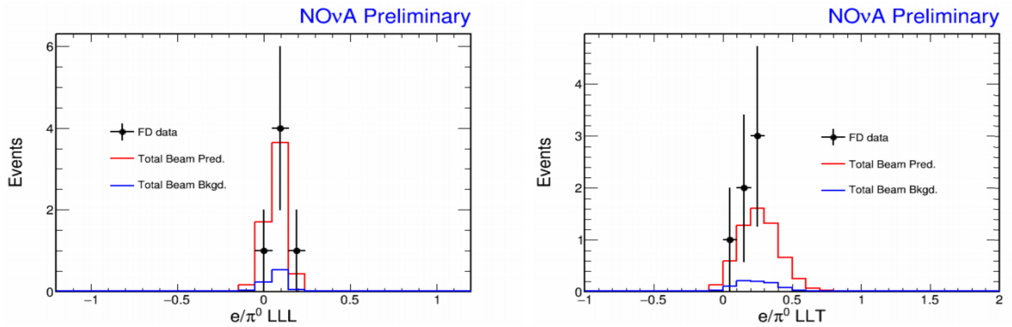
(a) Likelihood of the leading shower being an electron.



(b) Likelihood of the leading shower being an electron minus likelihood of being a photon.



(c) Likelihood of the leading shower being an electron minus likelihood of being a muon.



(d) Likelihood of the leading shower being an electron minus likelihood of being a pion.

Figure 8.14: Likelihood of the leading shower of a candidating event in data (black), simulated beam background (blue) and simulated signal + background events (red) under different particle hypotheses.

8.2.2 Rejection of “No ν_e-Oscillation” Hypothesis

Various aspects of the candidate events have been checked and the behavior of the observed events agrees with the prediction well. The hypothesis of whether the ν_e oscillation occurs is tested based on the 6 LID selected and 11 LEM selected events. Assuming the hypothesis is false, that oscillation does not happen, the FD prediction ($b_0 \pm \sigma_b$) would be equal to the background prediction, which is $0.94 \pm 0.10(syst.)$ in LID and $1.00 \pm 0.12(syst.)$ for LEM. The possible number of FD prediction events b occurs at the probability following a Gaussian distribution with mean value μ at the nominal prediction value b_0 and variance σ equal to the amount of systematic uncertainty σ_b. Then for a given number of FD prediction b, the observation should obey Poisson fluctuation. Combining the two steps gives the probability of observing x events given the prediction, $b_0 \pm \sigma_b$.

$$P(x|b_0, \sigma_b) = \int_{-\infty}^{+\infty} \left(\frac{b^x}{x!} e^{-b} \right) \times \left(\frac{1}{\sqrt{2\pi}\sigma_b} e^{-\frac{(b-b_0)^2}{2\sigma_b{}^2}} \right) \frac{db}{\sigma_b} \tag{8.2}$$

In order to determine the significance of rejecting $\theta_{13} = 0$ hypothesis, the likelihood of observing N or more than N is used in χ^2 calculation, $\chi^2(N) = -2ln[\mathcal{L}(x \geq N)] = -2ln[P(x \geq N)] = -2ln[\sum_{x \geq N} P(x)]$ and critical value for confidence level $P(x \geq N)$ is equal to $\sqrt{\chi^2(N) - \chi^2_{min}}$ The probability value is calculated numerically. For LID, given the fact that 6 ν_e CC candidates are observed, the no ν_e-oscillation hypothesis is rejected at 3.3σ. For the secondary PID, the 11 ν_e CC candidates shows a 5.4σ excess over the prediction made under no-oscillation assumption. So the no ν_e-oscillation hypothesis is rejected at more than 3σ for both main and secondary PIDs in the first analysis.

8.2.3 Confidence Interval of Oscillation Parameters

A more sophisticated analysis can be performed to study the favored or disfavored confidence interval for some oscillation parameters, including θ_{13}, δ_{cp} under different mass hierarchy cases. The nominal values of the oscillation parameters used in the analysis are listed in Table.8.3 [136], in which the value of θ_{13} is obtained from reactor neutrino oscillation experiment Daya Bay result, mixing due to θ_{23} is assumed to be fixed at the maximum.

Table 8.3: List of oscillation parameters used in the analysis.

Parameter	Value
Baseline	810 km
Matter density	2.84 gcm^{-3}
Δm_{21}^2	$(7.53 \pm 0.18) \times 10^{-5} eV^2$
$sin^2 2\theta_{21}$	0.846 ± 0.021
$sin^2 2\theta_{13}$	0.084 ± 0.005
Δm_{32}^2	$(2.40 \pm 0.10) \times 10^{-3} eV^2$
$sin^2 2\theta_{23}$	1
δ_{cp}	0

The χ^2 for a poisson variable is defined as Eq.8.3,

$$\chi^2 = -2ln\left(\frac{L(N;p)}{L(N;N)}\right) = 2(p - N + Nln\left(\frac{N}{p}\right)) \tag{8.3}$$

which with the given measurement N quantitizes the "distance" between the best fit parameter N and the tested parameter p. At the limit of large data samples, the Poisson distribution approaches a Gaussian distribution and therefore the χ^2 follow standard χ^2 distribution, the critical value of which for a specific p value can be looked up in official tables. However, the χ^2 variable used in the NOνA analysis does not obey standard χ^2 distribution, mainly because the statistics of the NOνA measurement are low and the large data sample limit is not satisfied.

Thus Feldman-Cousins (F-C) procedure is utilized to determine the relation between the p value and χ^2 value defined in Eq.8.3 [139]. The p value means that the the probability of the observation with χ^2 less than the a specific value is (1-p). For example, in order to find the χ^2 value for the significance level of 68% as a function of variables δ_{cp} and $\sin^2 2\theta_{13}$, at a specific bin with the corresponding δ_{cp} and $\sin^2 2\theta_{13}$ fixed, all the other oscillation parameters except $\sin^2 2\theta_{32}$ vary within the uncertainties and signal and background systematic uncertainties are also taken into account following a Gaussian fluctuation to obtain a fluctuated prediction. Then event count is calculated according to Poisson distribution with λ equal to the fluctuated

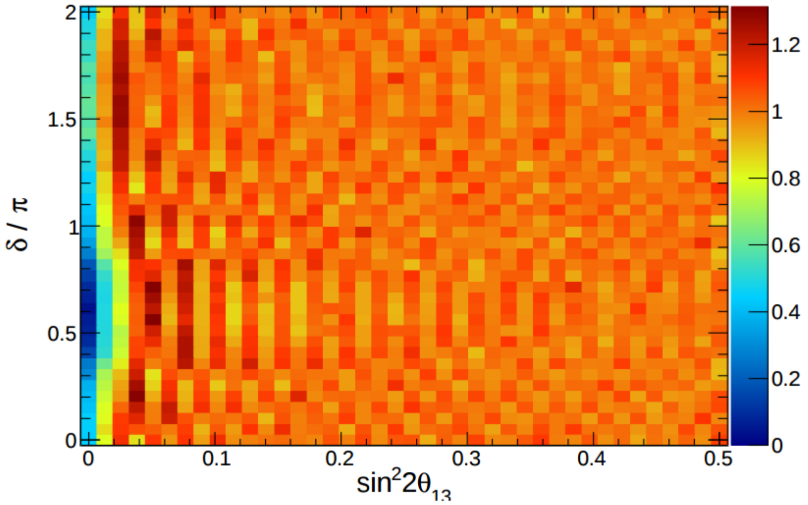


Figure 8.15: Critical value for 68% significance level as a function of δ_{cp} and $\sin^2 2\theta_{13}$ calculated using Feldman-Cousins procedure.

prediction. The fake event count along with the fluctuated prediction is filled as N and p respectively into Eq.8.3 to calculate the χ^2 value. Numerous pseudo-experiments are made and filled in this χ^2 distribution. In the end, the χ^2 value corresponding to 68% of total number of pseudo-experiments is found as the critical value at the bin. The process is iterated over all bins and get the critical value as a function of δ_{cp} and $\sin^2 2\theta_{13}$. Fig.8.15 shows the critical value for 68% confidence level in the LID region and under normal mass hierarchy assumption. According to Fig.8.15, even though the NOνA measurement is not at the large data sample limit, the critical value value for 68% appears almost at 1. The confidence interval is found as the intersected region between the $\chi^2(Measured|Pred_{Unfluctuated})$ surface and the critical value surface. Similiar procedure can be performed to get the contour for 90% confidence level. The best-fit curve contains the (δ_{cp}, $sin^2 2\theta_{13}$) pairs corresponding to minimum χ^2s, χ^2_{min}.

Fig.8.16 and Fig.8.17 present the LID and LEM results in terms of the best-fit curve as well as the confidence intervals for 68% and 90% confidence level for $\sin^2 2\theta_{13} = 0.5$ in normal mass hierarchy (NH) and inverted mass hierarchy (IH) scenarios and compare the result to the result in θ_{13} suggested by reactor neutrino experiment, Daya Bay. According to Fig.8.16, the LID result is compatible with the reactor result throughout the whole δ_{cp} region for NH and within part of the δ_{cp} region, $\delta_{cp} \in (0.8\pi,\ 2\pi)$ for IH case. The result in the secondary PID region

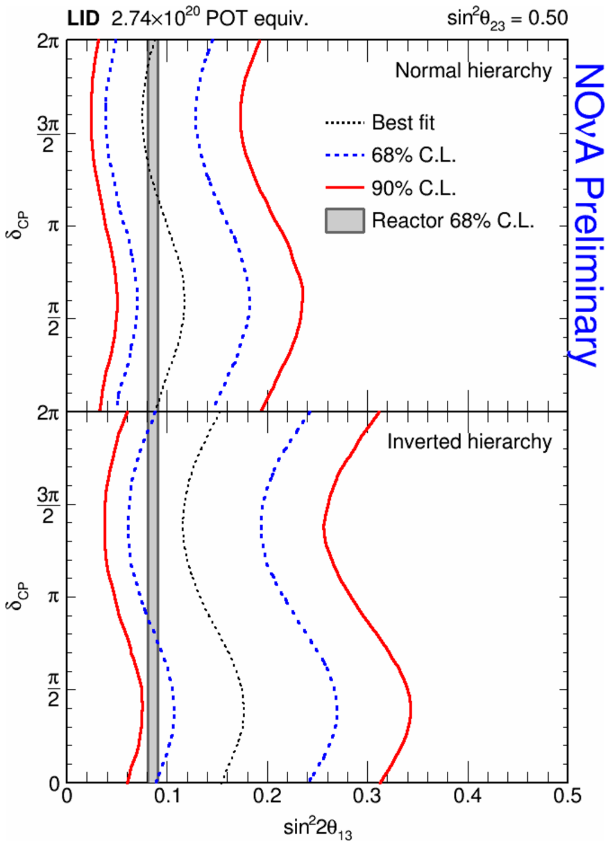


Figure 8.16: The best-fit curve (black) and the confidence Intervals for 68% (blue) and 90% (red) confidence level as a function of $\sin^2 2\theta_{13}$ and δ_{cp} for normal mass hierarchy (top) and inverted mass hierarchy (bottom) in the LID region. The result is compared to the 68% interval of $\sin^2 2\theta_{13}$ suggested by reactor experiment Daya Bay. (grey)

suggests a higher value of θ_{13} than the reactor result under the assumption of $\sin^2 2\theta_{32} = 0.5$ or a higher value of $\sin^2 2\theta_{32}$ for the θ_{13} of reactor measurement.

As seen in Fig.8.16, the significance distribution has relatively weak dependence on δ_{cp}. In order to test the compatibility with δ_{cp}, $\sin^2 2\theta_{13}$ is then constrained and varied within 0.086 ± 0.005 in the F-C procedure. At each bin of δ_{cp}, the χ^2 of the measured number of events given the prediction is calculated. The F-C procedure is then used to count the fraction of pseudo-experiments with the χ^2 value lower than the one from the real measurement. In the end the result is presented as the significance, which is converted from the fraction value counted in the last step, as a function of δ_{cp} as in Fig.8.18 and Fig.8.19. The step structure on the curves is due to the feature of the counting experiment where the χ^2 value is discrete. With the primary PID, LID, the measurement agrees with the current oscillation model with

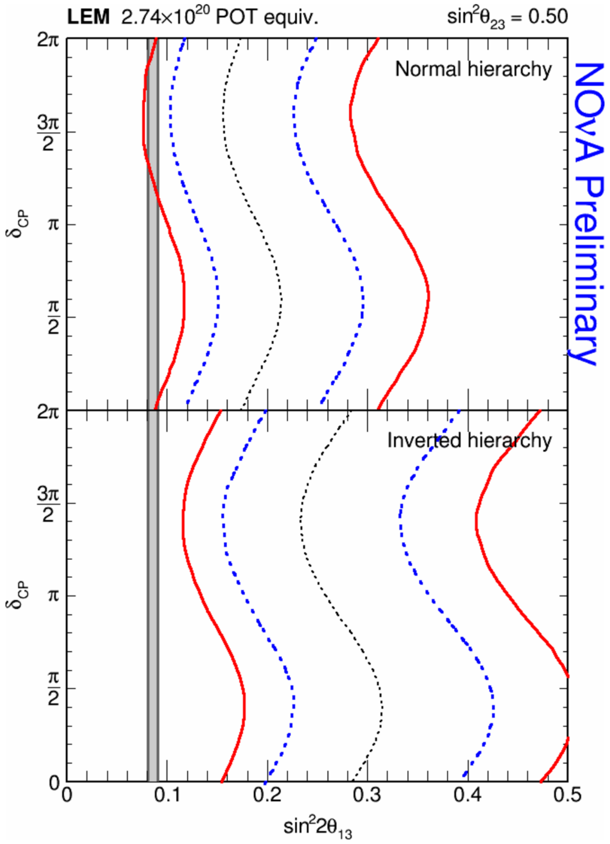


Figure 8.17: The best-fit curve (black) and the confidence Intervals for 68% (blue) and 90% (red) confidence level as a function of $\sin^2 2\theta_{13}$ and δ_{cp} for normal mass hierarchy (top) and inverted mass hierarchy (bottom) in the LEM region. The result is compared to the 68% interval of $\sin^2 2\theta_{13}$ suggested by reactor experiment Daya Bay. (grey)

$\sin^2 2\theta_{32}$ fixed at 0.5 for the NH scenario and no values in this scenario are disfavored. In the IH case the $\delta_{cp} \in (0,\ 0.65\pi)$ region is disfavored by the measurement at 90% confidence level. With the LEM selection, all values in the IH scenario with maximum θ_{32} mixing are disfavored. In the NH case the $\delta_{cp} \in (0.2\pi,\ \pi)$ region is disfavored by the LEM measurement at 90% confidence level.

A version of the significance test for the δ_{cp} measurement is also performed with the uncertainty of θ_{23} taken into account in the F-C pseudo-experiments. The result is shown in Fig.8.20. As the uncertainty of θ_{23} is included, with the primary PID, the disfavored region in IH scenario goes from $\delta_{cp} \in (0,\ 0.65\pi)$ to $\delta_{cp} \in (0.1\pi,\ 0.5\pi)$. With LEM, the disfavored region in the NH scenario goes from $\delta_{cp} \in (0.2\pi,\ \pi)$ to $\delta_{cp} \in (0.25\pi,\ 0.95\pi)$ at 90% confidence level.

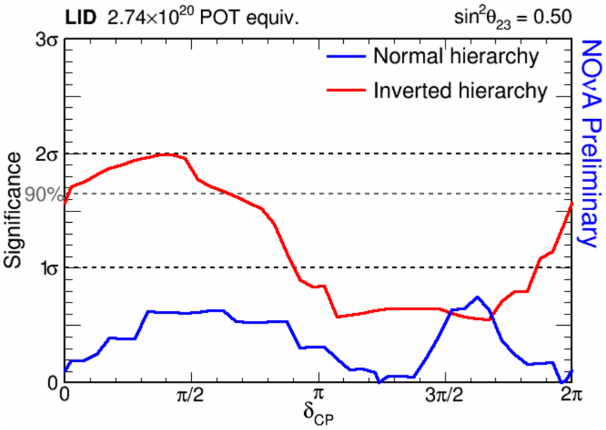


Figure 8.18: Distribution of the confidence level as a function of δ_{cp} for normal mass hierarchy (blue) and inverted mass hierarchy (red) in the LID region with $\sin^2 2\theta_{13}$ is constrained as 0.086±0.005 and θ_{23} fixed at $\pi/2$. The curve is compared to 90% confidence level.

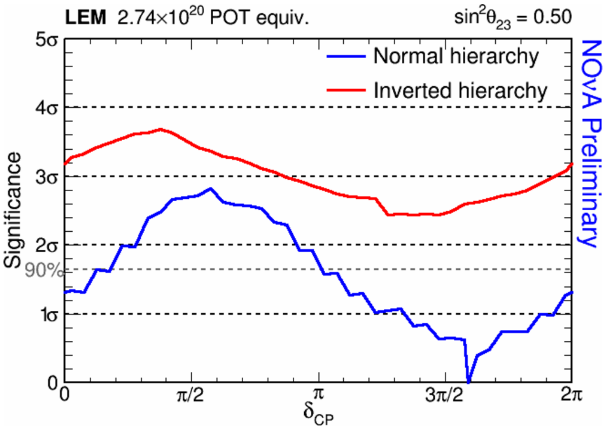


Figure 8.19: Distribution of the confidence level as a function of δ_{cp} for normal mass hierarchy (blue) and inverted mass hierarchy (red) in the LEM region with $\sin^2 2\theta_{13}$ is constrained as 0.086±0.005 and θ_{23} fixed at $\pi/2$. The curve is compared to 90% confidence level.

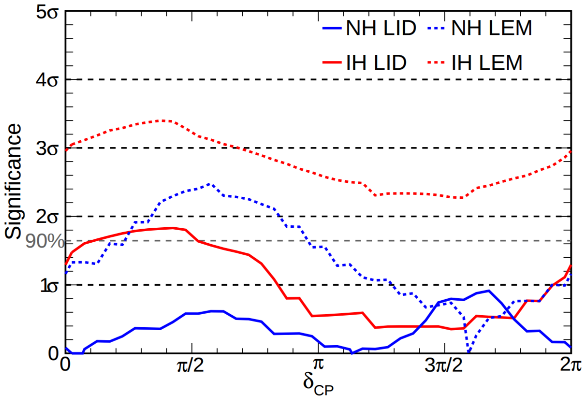
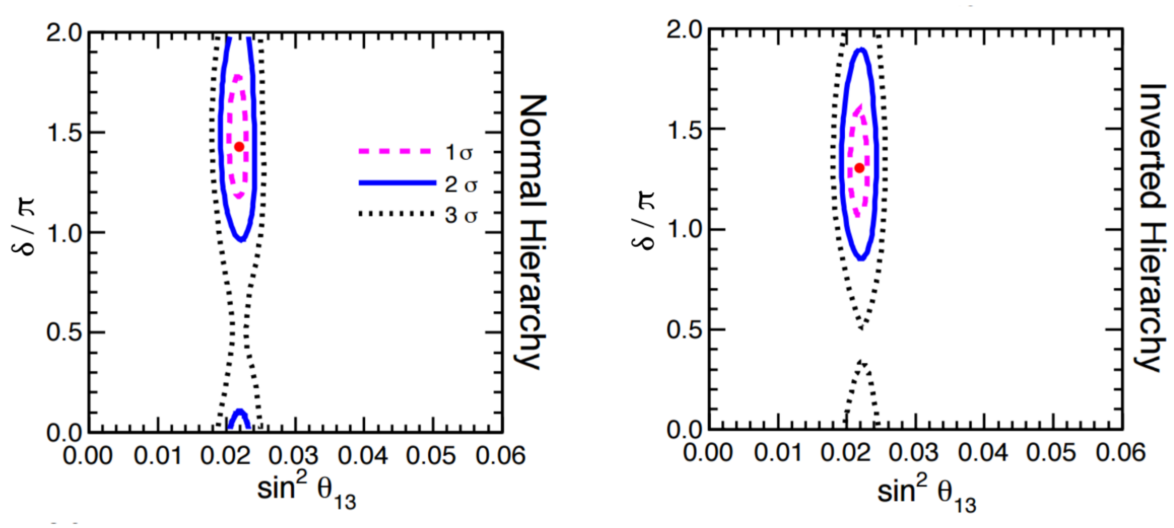


Figure 8.20: Distribution of the confidence level as a function of δ_{cp} for normal mass hierarchy (blue) and inverted mass hierarchy (red) in the LID (solid) and LEM (dash) region with $\sin^2 2\theta_{13}$ is constrained as 0.086±0.005 and the uncertainty of θ_{23} taken into account. The curve is compared to 90% confidence level.

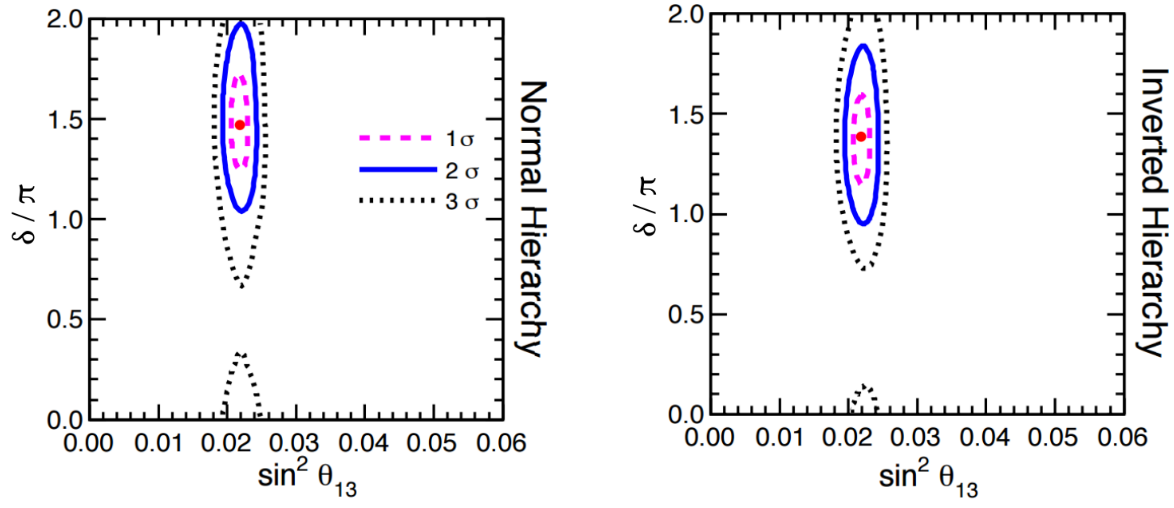
CHAPTER 9. CONCLUSION

The construction of NOνA detectors was completed in August 2014. The first ν_e appearance analysis is conducted using the data with total P.O.T. exposure of 3.52×10^{20} collected from February 2014 through May 2015. Two independent ν_e CC particle identification (PID) algorithms were developed. Both PIDs show good capability to distinguish ν_e CC from the other topologies and have 62% overlap of ν_e CC signal events. Six ν_e CC candidates are found in LID region and present 3.3σ excess over the FD prediction that is made without ν_e oscillation. The Feldman-Cousins procedure is used to study the compatibility of the measurement with the prediction. The result of θ_{13} agrees well with the measurement suggested by reactor neutrino experiments and the combination of (NH, 1.5π) is favored, which is consistent with current knowledge of mass ordering and CP violation measured in other neutrino oscillation experiments as shown in Fig.9.1 and Fig.9.2, which is the global significance contours of $sin^2\theta_{13}$ vs δ_{cp} before and after NOνA LID result is included respectively [142].



(a) (b)

Figure 9.1: Significance contour of $sin^2\theta_{13}$ and δ_{cp} for normal mass hierarchy (left) and inverted mass hierarchy (right) based on data of long baseline accelerator + Solar + short baseline reactor + atmospheric neutrino oscillation experiments before NOνA LID result is added.



(a) (b)

Figure 9.2: Global significance contour of $sin^2\theta_{13}$ and δ_{cp} for normal mass hierarchy (left) and inverted mass hierarchy (right) based on data of long baseline accelerator + Solar + short baseline reactor + atmospheric neutrino oscillation experiments after NOνA LID result is added.

9.1 Future Analysis

One important improvement in the second ν_e analysis is the increase in the neutrino statistics. NuMI beam has undergone an upgrade to a beam power of 700 kW with the technology mentioned in Sec.3.1. The 6+6 slipstacking mode is tested in July 2015 at a low power and is proved to be feasible. Due to the increase in beam power along with the full completeness of the detector, the total POT exposure for the second analysis is expected to increase by two times with respect to that of the first analysis.

As in the first analysis, the basic capability of NOνA of rejecting huge amount of cosmic ray background and detecting oscillated ν_e events is demonstrated. For the second analysis a fit for oscillation parameters based on measured spectra is planned. Therefore the selection cuts need to be modified that both PID and energy cuts could be tuned for $S/\sqrt{(S+B)}$ and not $S/\sqrt{B}$. With these looser cuts, more signal events will be accumulated with still a relatively low level of background. On the other hand, as the PID and energy cuts are loosened, in order to maintain the same level of cosmic ray background rejection, instead of applying a traditional cut flow, a multivariate technique can be used in more efficiently rejecting cosmic ray background.

The signal prediction will play a greater role than in the first analysis. In order to achieve a robust fit for oscillation parameters, a better signal prediction is needed, which means that the biggest systematic uncertainty due to the neutrino interaction uncertainty needs to be reduced. One quick way to improve it is to constrain the variation of the neutrino interaction parameters according to the cross section measurements from other neutrino experiments, like the MINERvA experiment at Fermilab.

APPENDIX A. NOνA PHOTOSENSOR

NOνA uses Avalanche Photo-diode (APD) as photosensor to transfer the scintillator light into electric signal. There are more than 10000 APDs on the far detector and about 600 on the near detector. APDs on both detectors are cooled to -15 °C in order to reduce sensor noise, and are operated at a gain of 100.

As shown in Fig.3.10, on each NOνA APD board, there are 32 $1.95 \times 1mm^2$ pixels, which correspond to 32 cells in a PVC extrusion moduler. In each pixel of APD unit, photon detection works as Fig.A.1. As a photon enters the layer of semiconductor from the top, the energy is absorbed and a pair of electron-hole is generated. Driven by the electric field, electron-hole pair is separated and the electron drifts toward the negative end (cathode). After the drift region, there is a multiplication region, in which avalanche multiplication takes place. In an avalanche multiplication process, an energetic electron ionizes and creates new electron-hole pairs, the electron of which is accelerated and continues this process. The whole multiplication process is analogy to avalanche and the number of electrons could be multiplied by two order

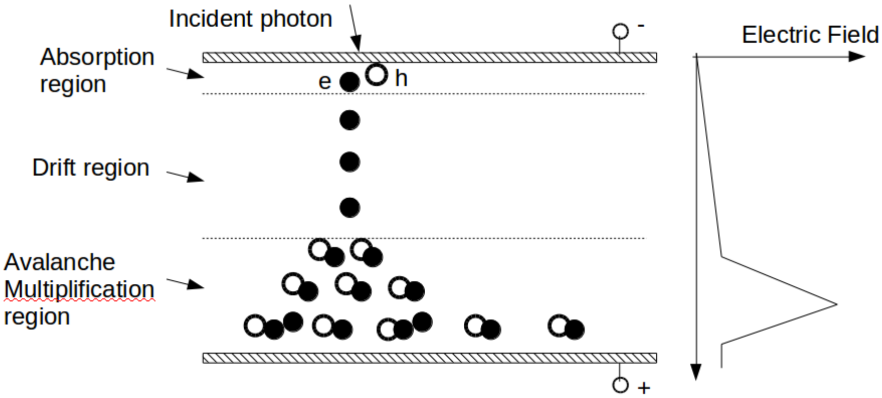


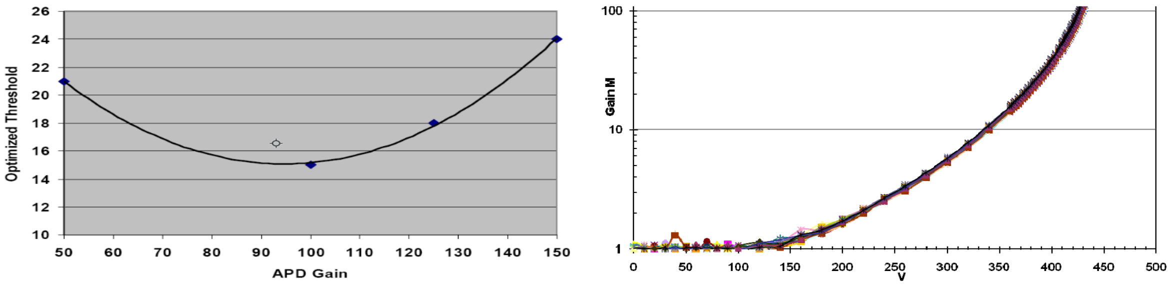
Figure A.1: Schematic diagram of a avalanche photo-diode unit, including absorption, drift and multiplication regions.

of magnitude. The curve on the right shows the strength of electric field as a function of depth. The electrons in the multiplification region are fueled by the strong electric field, which is created by high bias voltage across the APD.

Table A.1: Configurations of APD used in NOνA.

Active Area per Pixel	1.95 mm × 1.00 mm
Quantum Efficiency ($\lambda > 525nm$)	85%
Operating Voltage	375 ± 50 V
Gain at Operating Voltage	100
Operating Temperature	-15°C
Expected Signal-to-noise ratio	10:1

Table.A.1 shows some of important configurations of the APD used in NOνA [53]. As seen in 3.10, the quantum efficiency of APD is about 85% for almost the whole wavelength region of WLS fiber. The working gain of APD is optimized to be 100, which is a balance of maximizing signal and minimizing noise. The proper gain value is determined by minimizing the threshold cut that is needed to obtain 4:1 cosmic muon to noise rate ratio as shown in Fig.A.2a [59], [55]. In order to achieve a gain of 100, the operating bias voltage is about 425 V (Fig.A.2b).



(a) (b)

Figure A.2: Distributions of optimized threshold as a function of gain (left) and gain as a function of bias voltage (right). The gain is set to be 100, which is corresponding to the minimum optimized threshold value. The working voltage to achieve a gain of 100 is about 425 V.

Fig.A.3 shows the distribution of APD noise as a function of operating temperature. The APD operating temperature was finally determined to be -15°C, which is a compromisation of low APD noise and APD cooling cost. APD cooling is achieved by using thermoelectric

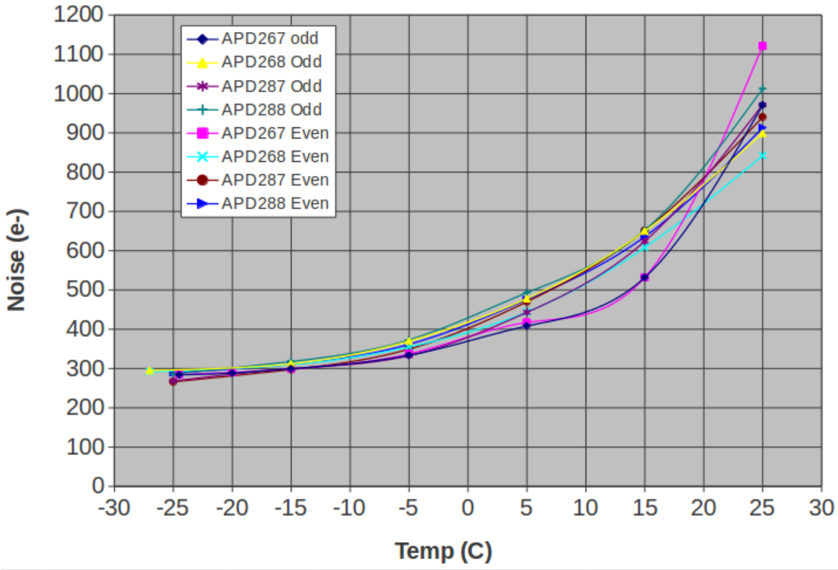


Figure A.3: Distribution of APD noise as a function of temperature at nominal working voltage.

cooler (TEC), which will be detailed later in the chapter. Combining the fluctuations in the bias voltage and temperature, the variation in the gain of APDs on NOνA detectors can be controled within 3% [53].

A.1 APD Cooling

Fig.A.4 is a diagram of the APD cooling system used in NOνAreadout unit. The heat generated by working APD is transfered to the thermoelectric cooler (TEC) and is further moved from the cold side (Al or Brass) to the hot side of TEC via thermoelectric effect, during which heat is moved along with electrons driven by electric field. The heat accumulating on the hot side of TEC is moved out from the system through water flow. The cooling process is controled by TEC controler (TECC) that The APD temperature measured by the thermistor is sent to TECC. According to the temperature, the cooling status is determined and the driven current of TEC is adjusted. The drive current is measured as the percentage of the full working current. As the TEC is in the process of cooling, the drive current is measured as 95%. When TEC isin the mode of sustaining APD at low temperature, the drive current is about 35%. The current is 0%, if cooling is turned off. Fig.A.5 shows the feature of the cooling process of a case with normal cooling function. Before the cooling starts, the temperature is close to the

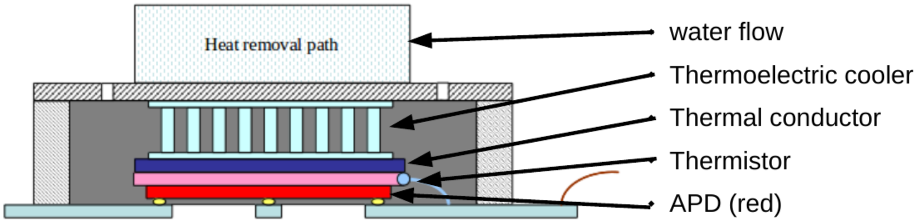
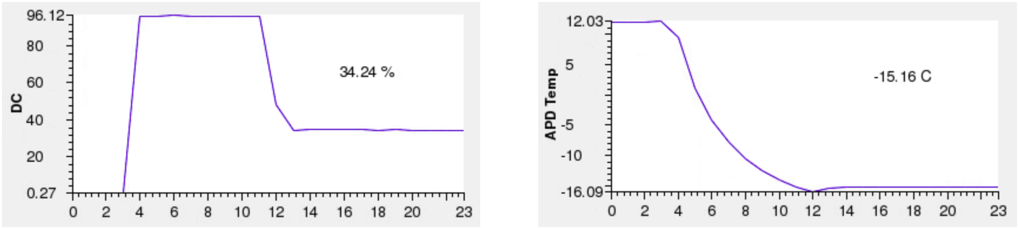


Figure A.4: Diagram of APD cooling system used in NOνA. From bottom to top: APD, conductor, thermistor, TE cooler and cooling water.



(a)

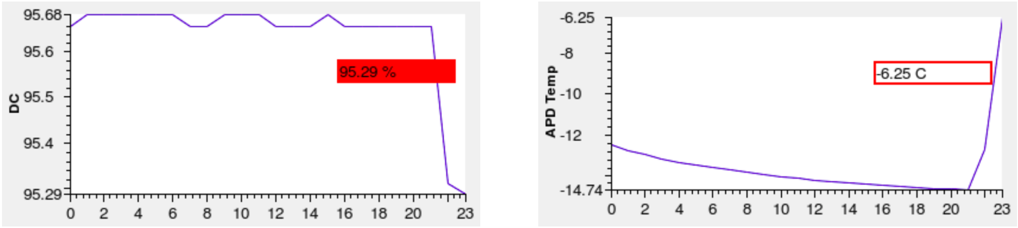
(b)

Figure A.5: Drive current vs time (left) and APD temperature vs time (right) distributions for normal cooling case.

room temperature and drive current is 0%. During the process of cooling, drive current rises up to 95% and in the mean time the APD temperature decreases. APD temperature has to reach -15 ∘C within 2 minutes after the cooling starts. In a stable cool stage, the drive current stay at around 35%.

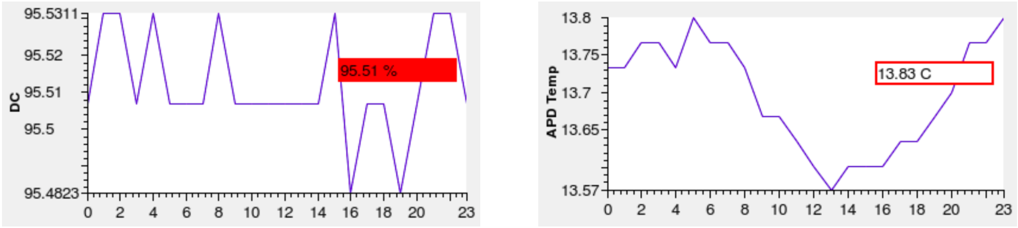
The cooling function of APD sometimes breaks down due to various reasons, which present differen features on the monitoring plots, including APD temperature, TEC drive current. A few main cooling failure mode have been found:

Slow cooling (Fig.A.6): APD temperature fails to decrease to -15 ∘C and the drive current stays at around 95% after 2 minutes. In some cases, even though the tempearture drops down to below -15 ∘C, the drive current still remains at 95%. This case will also trigger problem alarm. The problem is believed to happen in the situation that since the APD temperature is checked every 15 second, the drive current will be still at 95% if the APD temperature reaches -15 ∘C within 15 second before the 2 minute limit. The problem could be caused by bad thermal contact or bad TEC and is fixed by replacing APD in most of the cases according to Fig.A.9 [146], which is a metrics of cooling failure vs fix



(a) (b)

Figure A.6: Drive current vs time (left) and APD temperature vs time (right) distributions for failures presnting slow cooling.



(a) (b)

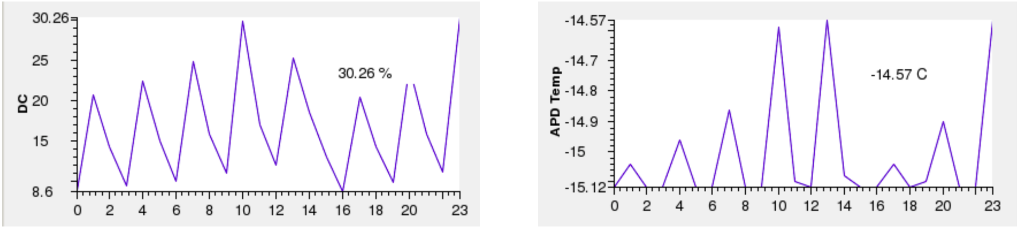
Figure A.7: Drive current vs time (left) and APD temperature vs time (right) distributions for failures presnting no cooling.

method based on the maintenance information cumulated from April 2013 to September 2014. For the slow cooling failure, The total number of problematic APDs is 132, 122 of which was fixed by replacing APD.

No cooling (Fig.A.7): After cooling starts, the drive current increases but the APD temperature does not go down. The problem is believed to be caused by bad TECC and is fixed by replacing TECC in most of the cases (Fig.A.9).

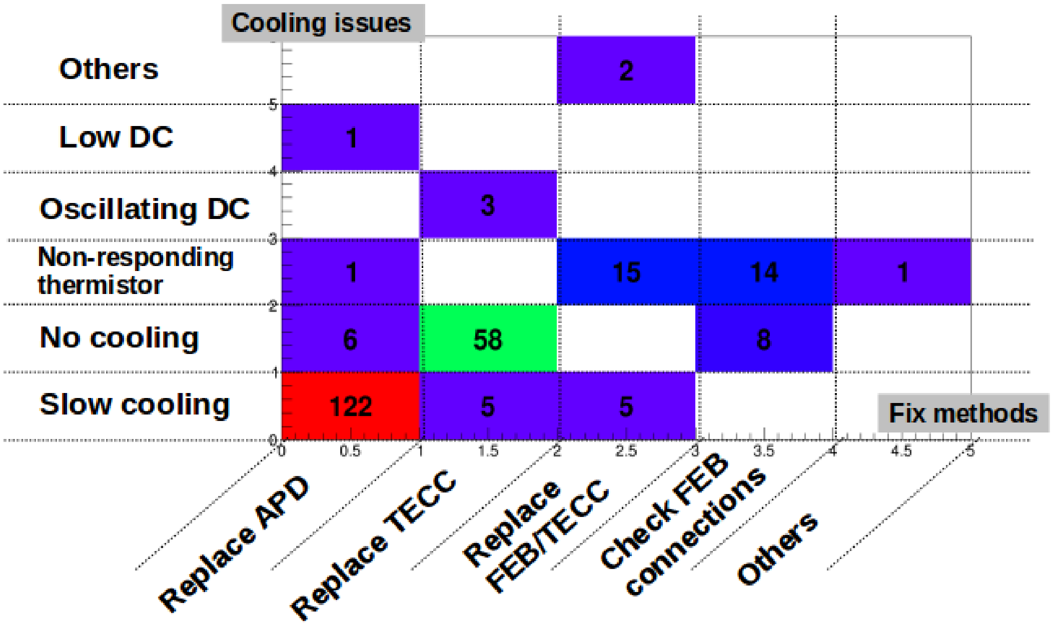
Non-responding thermistor: The failure appears as having reporting temperature value. The problem could be caused by bad ribbon cable connection between APD and FEB.

Oscillating DC (Fig.A.8): In this case, the value of TEC drive current oscillates and therefore the APD temperature varies rapidly. A bad TECC is usually responsible for the problem that the drive current fails to maintain stable.



(a) (b)

Figure A.8: Drive current vs time (left) and APD temperature vs time (right) distributions for failures presnting oscillating TEC drive current.



	Replace APD	Replace TECC	Replace FEB/TECC	Check FEB connections	Others
Others			2		
Low DC	1				
Oscillating DC		3			
Non-responding thermistor	1		15	14	1
No cooling	6	58		8	
Slow cooling	122	5	5		

Figure A.9: The problem vs fix method metrics for APD cooling failures. The table is made based on the maintenance information cumulated from April 2014 to September 2014.

A.2 Detector Noise

The detector noise is made up of different components as described in:

$$N_{Det}^2 = N_{APD}^2 + N_{FEB}^2 + N_{Env}^2 \tag{A.1}$$

APD noise: In a normal case APD noise is dominated by shot noise, which is caused by thermal fluctuation. So the APD noise is expected to be greatly reduced by running at low temperature.

FEB noise: It refers to the noise caused by the electronic parts in FEB. In the amplifier, The change in the charge cumulated on the capacitor produces current which is taken as noise.

Environment noise: The environmental noise could come from different sources. For example, two parallel wires could form a capacitor, which causes noise. The change in the current of the same two wires will generates magnetic fields that have impact on each other. Also noise could also be raised by poor grounding of different electronic parts of the detector.

The noise is evaluated in Digital Scanning Oscilloscope (DSO) procedure, which is also known as pedestal run and contains mostly noise information. The DSO data is processed through dual correlated sampling (DCS) procedured to obtain DCS data. Noise is calculated as the RMS of DCS data and is in the unit of electron. Fig.A.10 shows that the noise for cold and warm APDs is 570 e and 620 e respectively [147]. Assuming that running at low temperature removes most of the APD noise, APD noise is calculated as the difference between that for cold and warm modes, ∼240 e. The FEB noise is measured to be ∼ 430 e on average at the test-stand as shown in Fig.A.11 [146]. The overall environmental noise is then calculated to be 374 e.

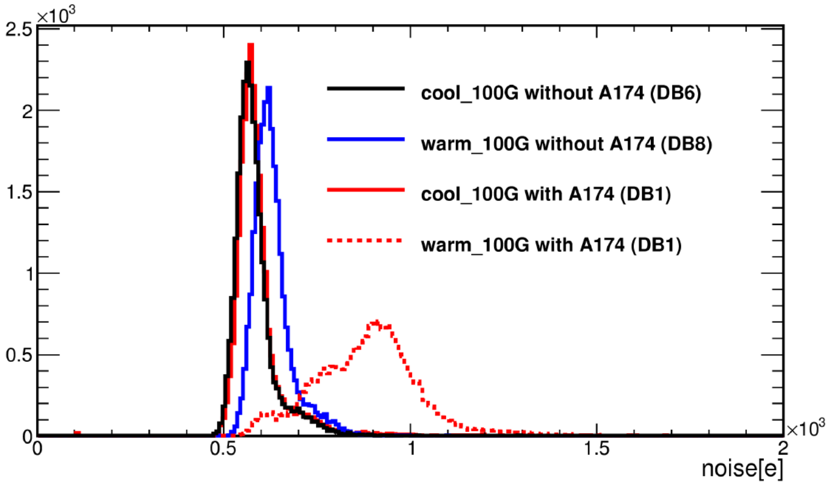


Figure A.10: The noise distribution for cold (black) and warm (blue) warm APDs.

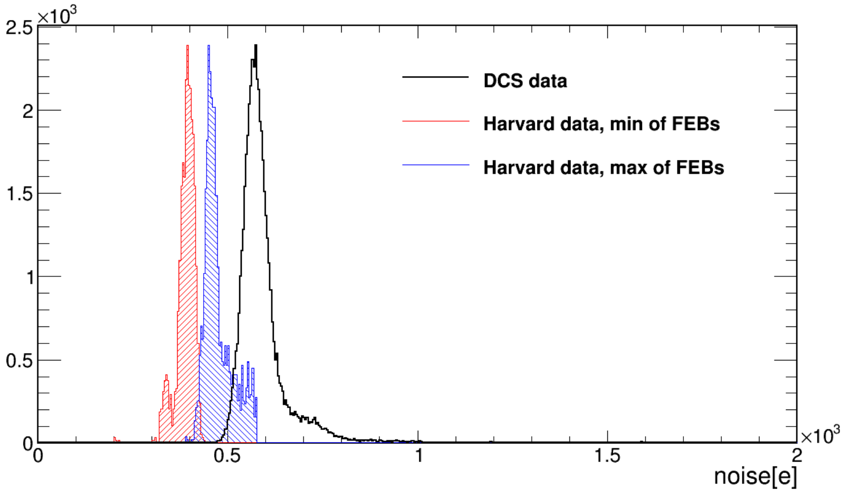
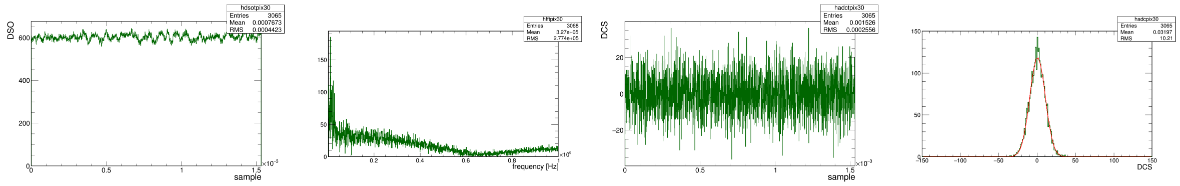


Figure A.11: The detector noise in the cold mode (black) and the minimum (red) and maximum (blue) FEB noise measured at Harvard FEB test-stand. The FEB noise is taken as 430 e on average.

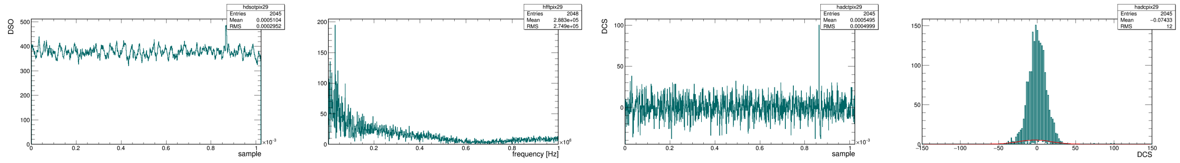
The discussion above describes the noise performance of normal channels. The DSO plots for these normal channels are shown in Fig.A.12. Fig.A.12a is the DSO trace distribution for a random normal channel. Fig.A.12b is the fast fourier transform (FFT) of DSO trace. Fig.A.12c is the DCS trace and Fig.A.12d is the Y projection of DCS trace. The DCS trace and its Y projection show that for normal channels, the noise performs as a Gaussian fluctuation with center at 0 and RMS of around 11.



(a) (b) (c) (d)

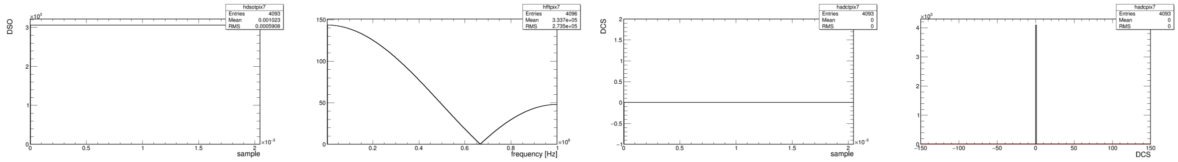
Figure A.12: DSO plots, DSO trace (a), fast fourier transformation of DSO (b), DCS trace (c) and DCS Y projection (d) for normal channels. The DCS follows Gaussian fluctuation with RMS of 11 ADC.

For the problematic channels which usually show very high or low hit rate, the DSO distributions present more details of them and are ploted in Fig.A.13 through Fig.A.25. The properties, such as baseline, noise spikes and periodicity, as well as the fix rate for different types of noise modes are listed in Table.A.2. The reason for different modes could be the bad connection of a wire, an unexpected mechanical pressure on APD board, environmental viberation or some intermmmitent issues. Although the underlying causes still need more investigation, an empirical table of fix rate in terms of noisy APD mode vs fix method helps increase the success rate of repairing noisy APD issues.



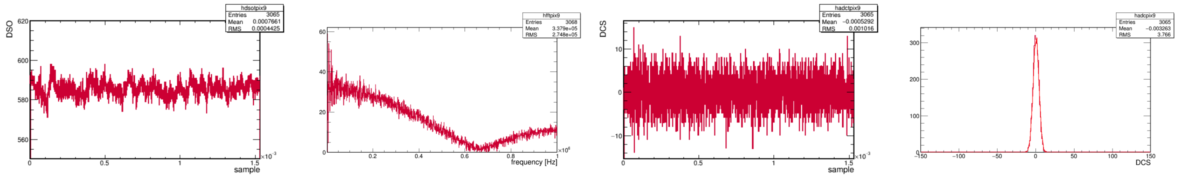
(a) (b) (c) (d)

Figure A.13: DSO plots, DSO trace (a), fast fourier transformation of DSO (b), DCS trace (c) and DCS Y projection (d) for Type One problematic channels that everything appears fine except that only even DCS values present according to DCS Y projection distribution.



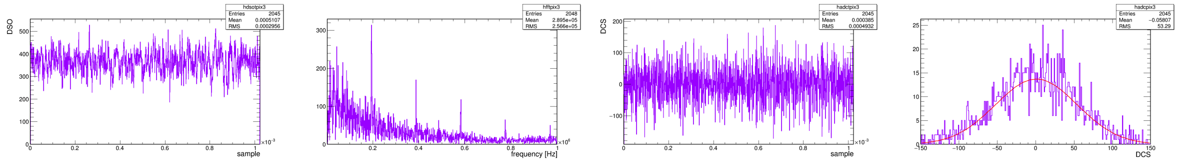
(a) (b) (c) (d)

Figure A.14: DSO plots, DSO trace (a), fast fourier transformation of DSO (b), DCS trace (c) and DCS Y projection (d) for Type Two problematic channels that the DCS baseline is 0 in RMS.



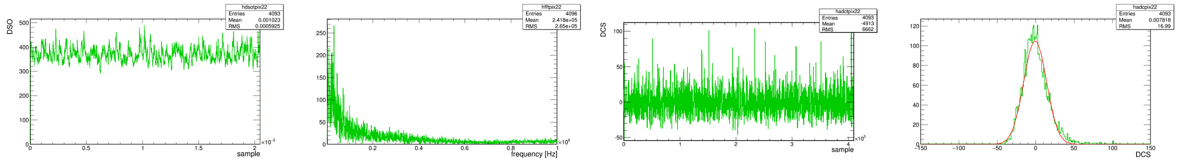
(a) (b) (c) (d)

Figure A.15: DSO plots, DSO trace (a), fast fourier transformation of DSO (b), DCS trace (c) and DCS Y projection (d) for Type Three problematic channels that the DCS baseline is less than 5 in RMS.



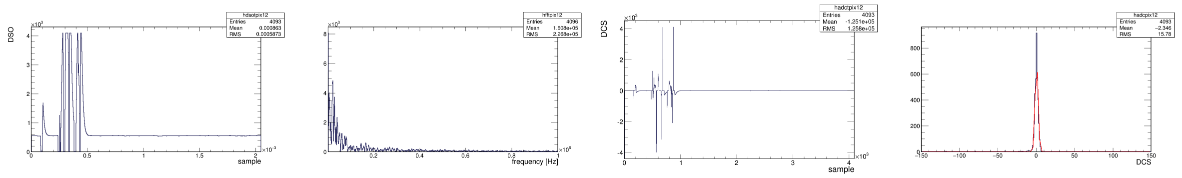
(a) (b) (c) (d)

Figure A.16: DSO plots, DSO trace (a), fast fourier transformation of DSO (b), DCS trace (c) and DCS Y projection (d) for Type Four problematic channels that the DCS baseline is greater than 50 in RMS.



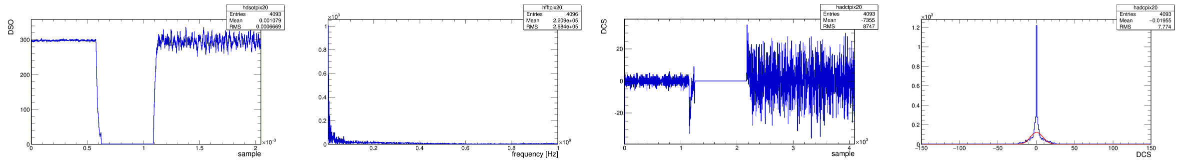
(a) (b) (c) (d)

Figure A.17: DSO plots, DSO trace (a), fast fourier transformation of DSO (b), DCS trace (c) and DCS Y projection (d) for Type Five problematic channels that the DCS baseline is normal and multiple slow DSO noise spikes (>50 ADC) exist.



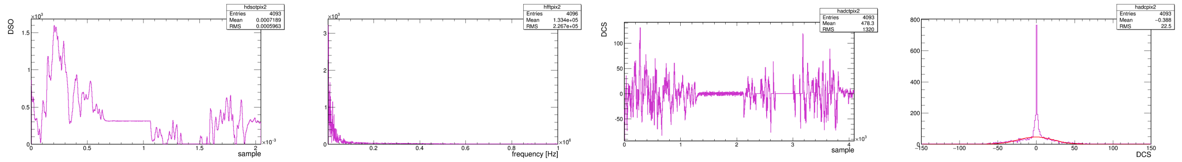
(a) (b) (c) (d)

Figure A.18: DSO plots, DSO trace (a), fast fourier transformation of DSO (b), DCS trace (c) and DCS Y projection (d) for Type Six problematic channels that the DCS baseline is less than 5 in RMS and slow noise spikes (>1000 ADC) exist.



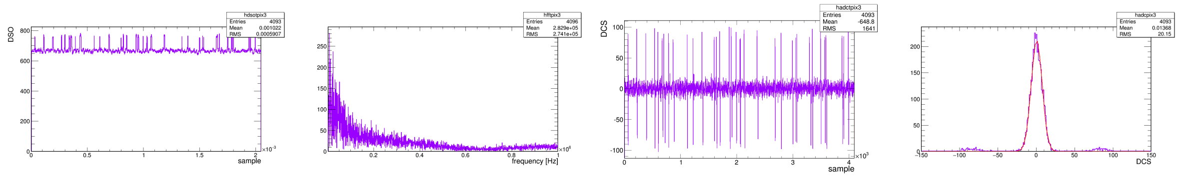
(a) (b) (c) (d)

Figure A.19: DSO plots, DSO trace (a), fast fourier transformation of DSO (b), DCS trace (c) and DCS Y projection (d) for Type Seven problematic channels that the DCS baseline varies, 0 or less than 5 ADC or $\sim$ 10 ADC.



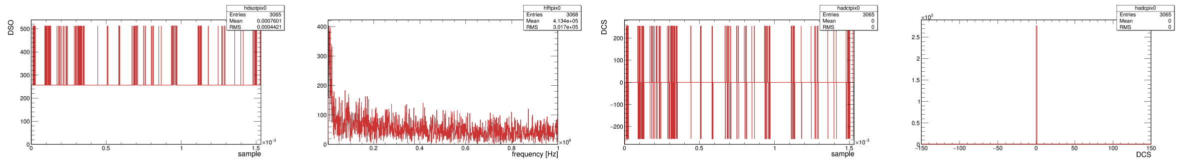
(a) (b) (c) (d)

Figure A.20: DSO plots, DSO trace (a), fast fourier transformation of DSO (b), DCS trace (c) and DCS Y projection (d) for Type Eight problematic channels that the DCS baseline varies, 0 or less than 5 ADC or $\sim$ 10 ADC or even $>$ 50 ADC.



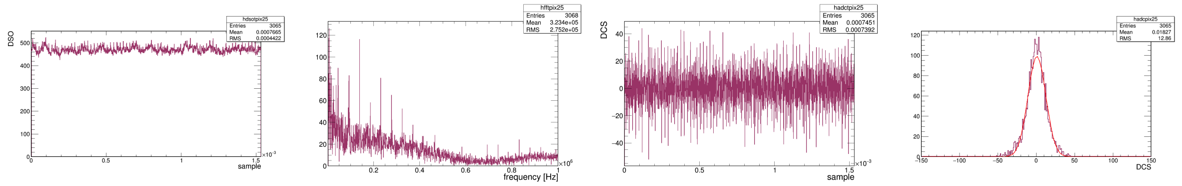
(a) (b) (c) (d)

Figure A.21: DSO plots, DSO trace (a), fast fourier transformation of DSO (b), DCS trace (c) and DCS Y projection (d) for Type Nine problematic channels that the DCS baseline is normal but rapid DSO spikes with amplitude 50 exist.



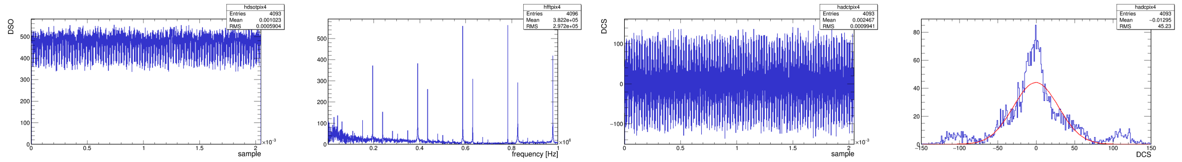
(a) (b) (c) (d)

Figure A.22: DSO plots, DSO trace (a), fast fourier transformation of DSO (b), DCS trace (c) and DCS Y projection (d) for Type Ten problematic channels that the DCS baseline is 0 in RMS and rapid DSO spikes with amplitude > 50 exist.



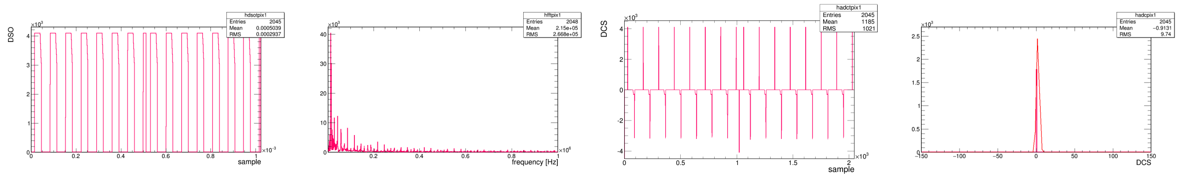
(a) (b) (c) (d)

Figure A.23: DSO plots, DSO trace (a), fast fourier transformation of DSO (b), DCS trace (c) and DCS Y projection (d) for Type Eleven problematic channels that the DCS baseline is normal in RMS and periodic rapid DSO spikes with frequency of 500 kHz exist.



(a) (b) (c) (d)

Figure A.24: DSO plots, DSO trace (a), fast fourier transformation of DSO (b), DCS trace (c) and DCS Y projection (d) for Type Twelve problematic channels that the DCS baseline is normal in RMS and periodic rapid DSO spikes with two frequencies of 50 kHz and 200 kHz exist.



(a) (b) (c) (d)

Figure A.25: DSO plots, DSO trace (a), fast fourier transformation of DSO (b), DCS trace (c) and DCS Y projection (d) for Type Thirteen problematic channels that the DCS baseline is 0 in RMS and DSO baseline is 0 and periodic wide DSO spikes with frequencies of $\sim$ 150 kHz exist.

Table A.2: Properties and fix rate for different types of noise modes.

	Properties			Fix Rate		
	Baseline (DCS trace)	Spikes (DSO trace)	Periodicity (FFT of DSO)	Replace APD	Replace FEB	Tighten Connection
Normal	∼10	N	N	0/0	0/0	0/0
Type 1	∼10 but only even DCS	N	N	1/1	0/0	0/0
Type 2	0	N	N	1/1	0/1	0/0
Type 3	<5	N	N	1/1	0/0	0/0
Type 4	>50	N	N	3/3	0/0	0/0
Type 5	∼10	Y (slow, >50)	N	48/54	2/3	5/5
Type 6	<5	Y (slow, >1000)	N	8/10	0/0	0/0
Type 7	0 or <5 or ∼10	N	N	0/3	0/0	0/0
Type 8	0 or <5 or ∼10 or > 50	N	N	1/3	0/1	0/2
Type 9	∼10	Y (rapid, >50)	N	0/1	0/0	0/0
Type 10	0	Y (rapid, >50)	N	0/0	1/1	0/0
Type 11	∼10	Y (rapid)	Y (500 kHz)	22/25	0/1	5/12
Type 12	∼10	Y (rapid)	Y (500 kHz and 2000 kHz)	6/6	0/0	0/0
Type 13	0 in both DSO and DCS	Y (wide, >1000)	Y (150 kHz)	0/0	2/3	0/1

The cooling and noisy APD fix metrics are integrated with other work done in the operation group and helps to reduce the far detector APD issue rate, including all cooling and noisy APD issues, from 3.5% in November 2013 [144] to 0.4% in April 2015 [145] at the early stage of far detector commissioning.

APPENDIX B. DATA-DRIVEN THRESHOLD CORRECTION

Threshold and shadow effects vary the shape of PE/path distribution as a function of W and Y and thus shift the mean of (PE/path) value which is later used in the attenuation fit. So correction for these two effects is needed before attenuation calibration. In the first analysis a MC driven method was used to take the two effects account. A data driven threshold and shadow correction method is being studied in StopperThreshold module in Calibration package and is supposed to reduce the uncertainties caused by any imperfectness of the related simulation. The correction is applied before attenuation calibration, which is supposed to correct for the attenuation effect in fibers.

B.1 Data-Driven Shadow Correction

As a charged particle deposits energy during the propergation in the detector, the momentum changes and therefore the spectrum of the energy deposition rate ($PE/path$) shifts, which is known as shadow effect [148]. Since cosmic trigger sample is used for calibration and contains mostly “vertical” cosmic ray events, the shadow effect is a function of Y Because the attenuation calibration is calculated based on the average $PE/path$ in the cell, in order to ensure that the mean deposition rate has the same representivity of spectra at any point of the detector, a shadow correction is needed to shift the to a standard, which in this case is the $PE/path$ of hits from minimium ionizing particles (MIP). In the simulation, the correction factor is calculated as

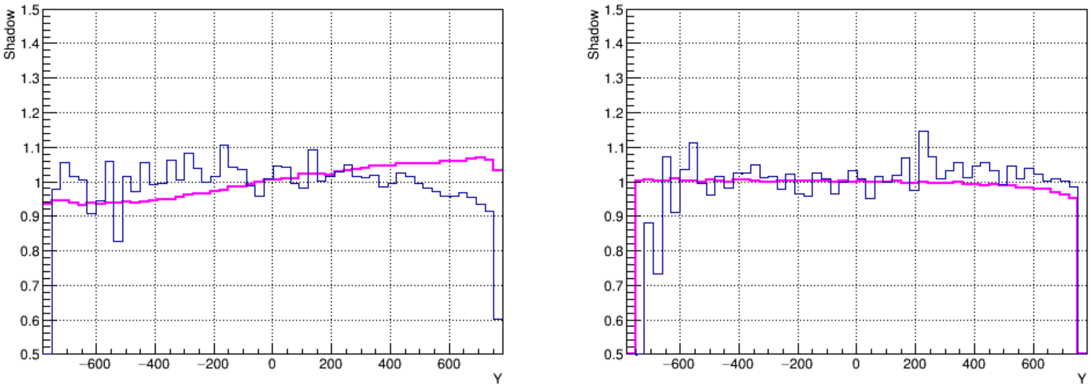
$$Corr_{Shadow,MC}(Y) = Mean\left(\frac{trueGeV}{MIP_{Reco}}\right) = Mean\left(\frac{trueGeV}{Path_{reco} * 1.78}\right) \tag{B.1}$$

in which $trueGeV$ is the true energy of FLS hit which is only accessible from MC and MIP_{Exp} is the predicted MIP PE, $Path[cm] * 1.78$ MeV, given the path length. The MC driven shadow correction used in the first analysis was applied as $PE_{Corr} = PE/Corr_{Shadow}(Y)$ [149].

In the data-driven method, the response of MIP hits is used as standard candle, the position of which does not shift in the spectrum of energy deposition rate. The MIP hits are selected as being within 100 - 200 cm from the end of stopping muons. The shadow effect is corrected by normalizing the mean response of all hits to the mean response of MIP hits (Eq.B.2).

$$Corr_{Shadow,Data}(Y) = \frac{Mean\left((PE/Path)_{(MIP+NonMIP)\&long}\right)}{Mean\left((PE/Path)_{MIP\&long}\right)} \tag{B.2}$$

The ratio of the two variables is expected to cancel the attenuation effect and as long hit, $Path > 6$ cm for X-view cells and $Path > 4.5$ cm for Y-view cells, are required, the influence of threshold effect is supposed to be reduced. The correction is designed to be applied as $(PE/Path)_corr = (PE/Path)_uncorr/Corr_{Shadow,Data}$. Fig.B.1 shows the MC driven shadow correction compared to the factor calculated by the data driven method [150]. So far, because hits fulfilling both MIP and long are used in the denominator, the distribution suffers low statistics.



(a)

(b)

Figure B.1: The shadow correction factors for X (left) and Y (right) cells calculated using MC (pink) and data-driven method (blue).

B.2 Data-Driven Threshold Correction

The number of photoelectrons (PE) generated in APD follow Poisson distribution. As threshold cut is applied, which is usually at 45 PE, truncates the lower part of the Poisson

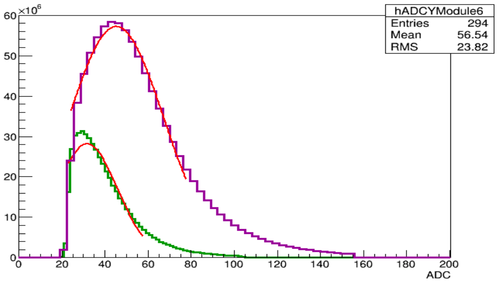


Figure B.2: The distribution of PE for MIP hits. A cut off at around ADC = 20, which is equivalent to 40 PE, can be seen in each of the curves which represents PE distribution at different depth of cells.

distribution, it shifts the mean of PE higher than it should be. The target of the threshold correction is to remove the shifting effect due to threshold cut and to scale the mean of truncated Poisson distribution back to the mean of complete Poisson distribution, which is not affected by the threshold cut and is believed to be proportional to the actual energy deposition. Due to the attenuation effect in the fiber, the shift effect due to the threshold cut is mainly a function of W. In the simulation, the correction factor is calculated as:

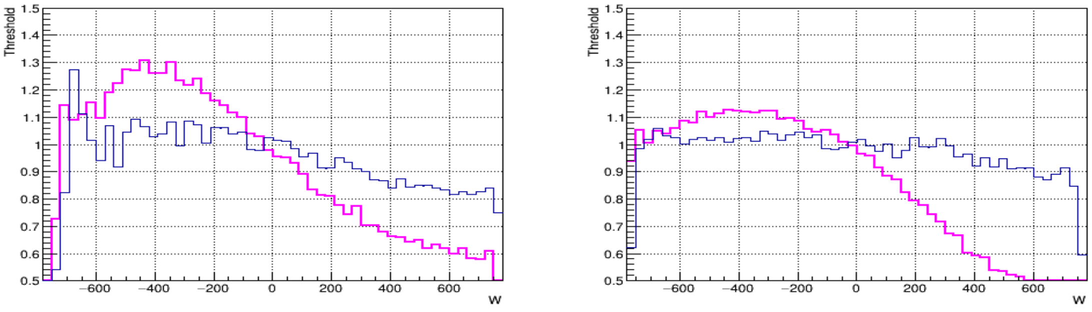
$$Corr_{Threshold,MC}(W) = Mean(PE_{Reco}/\lambda_{Poisson}) \tag{B.3}$$

in which PE is the reconstructed PE value and the denominator is the λ of the Poisson distribution which is based on the true MC information. The correction is applied as $PE_{Corr} = PE/Corr_{Threshold}(Y)$ [149].

In the data-driven method, the correction factor is designed to be:

$$Corr_{Threshold,Data}(W) = \frac{Mean\left((PE/Path)_{MIP\&AllLength}\right)}{Mean\left((PE/Path)_{MIP\&long}\right)} \tag{B.4}$$

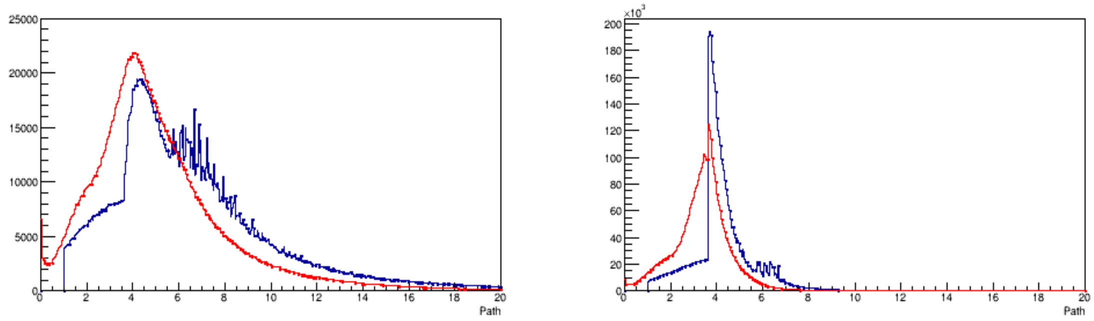
The calculation uses only MIP hits so that the threshold factor is not influenced by the shadow effect. Due to the attenuation effect, the threshold effect is greater at the far end than at the near end of a cell. Because the difference caused by the attenuation effect is greater than the difference between the spectra of all hits and MIP hits. The correction factor calculated using MIP hits is believed to have a good representative of all hits in terms of correcting for threshold effect. The denominator uses hits fulfilling long hits which is the same criteria as in data driven shadow correction to reduce the threshold effect in the denominator. Fig.B.3



(a)

(b)

Figure B.3: The threshold correction factors for X (left) and Y (right) cells calculated using MC (pink) and data-driven method (blue).



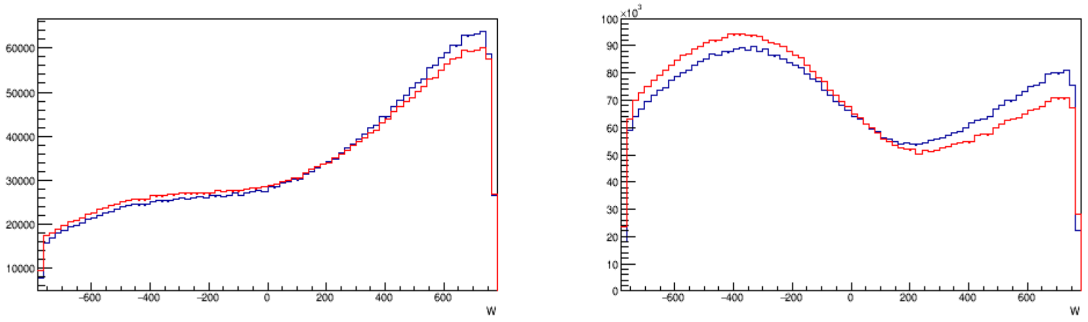
(a)

(b)

Figure B.4: The threshold correction factors for X (left) and Y (right) cells calculated using MC (red) and data-driven method (blue).

shows the threshold correction factor calculated using MC true information compared to the factor calculated by the data driven method. Even with the fluctuations due to low statstics in the sample, correction factors calculated by data and MC driven method show the same trend that the correction is greater at the far end than the near end. However the MC threshold correction still shows larger slope than the data driven method.

So far the data driven method is still in the development in StopperThreshold module in the Calibration package using MC samples that several variables, such as Path (See Fig.B.4) and W (See Fig.B.5), are checked between reconstructed MC and true MC. Unexpected plateau is seen at 6 cm in both X and Y view cells in the reconstructed MC. The plan to resolve the issue is to first check the distribution with all hits included and then the event display for events that contain hits in that 6 cm plateau. As observed in the previous section, the



(a)

(b)

Figure B.5: The threshold correction factors for X (left) and Y (right) cells calculated using MC (red) and data-driven method (blue).

data driven correction result is experiencing low statistic issue. One way that was discussed to improve the statistics is to enlarge the MIP window from 100-200 cm to maybe 100-400 cm to include more hits. Other than these issues above, some potential questions remain: definition of "long" hits; ratio of mean (data driven) or mean of ratio (mc driven).

APPENDIX C. COSMIC RAY BACKGROUND ESTIMATE

C.1 Cumulative Cut Efficiency Table

Table C.1: Cumulative cut efficiency at each cut flow step.

cuts	ν_eCC signal	beam background	Cosmic in Cosmic Trigger	Cosmic in out-of-time Numi Trigger
no cut	100%	100%	100%	100%
data quality	74%	19%	6.2%	6.2%
containment	62%	15%	0.27%	0.27%
cosmic rejection	58%	13%	0.0053%	0.0053%
ν_e-preselection (LID)	44%	4.4%	0.00038%	0.0004%
LID	21%	0.5%	0.00000070%	0.0000010%
ν_e-preselection (LEM)	49%	5.4%	0.00075%	0.00079%
LEM	22%	0.17%	0.00000070%	0.0000010%

C.2 Out-Of-Time Numi Trigger Events Passing ν_e Selections

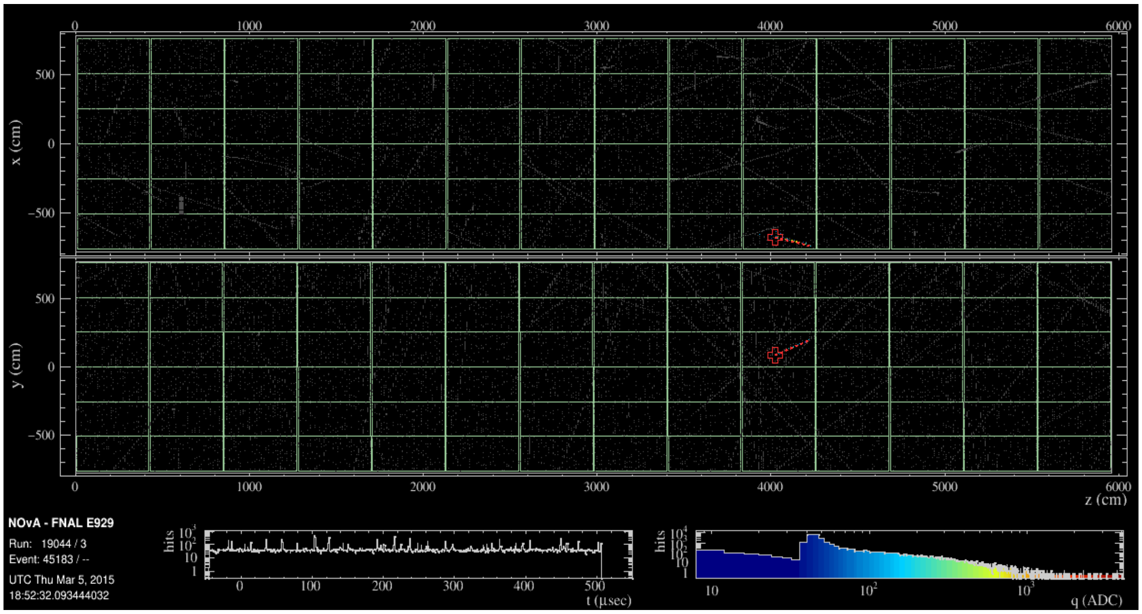


Figure C.1: This event is selected by both LID and LEM. LID is 0.98. LEM is 0.92. This event is 23 cm from the east wall.

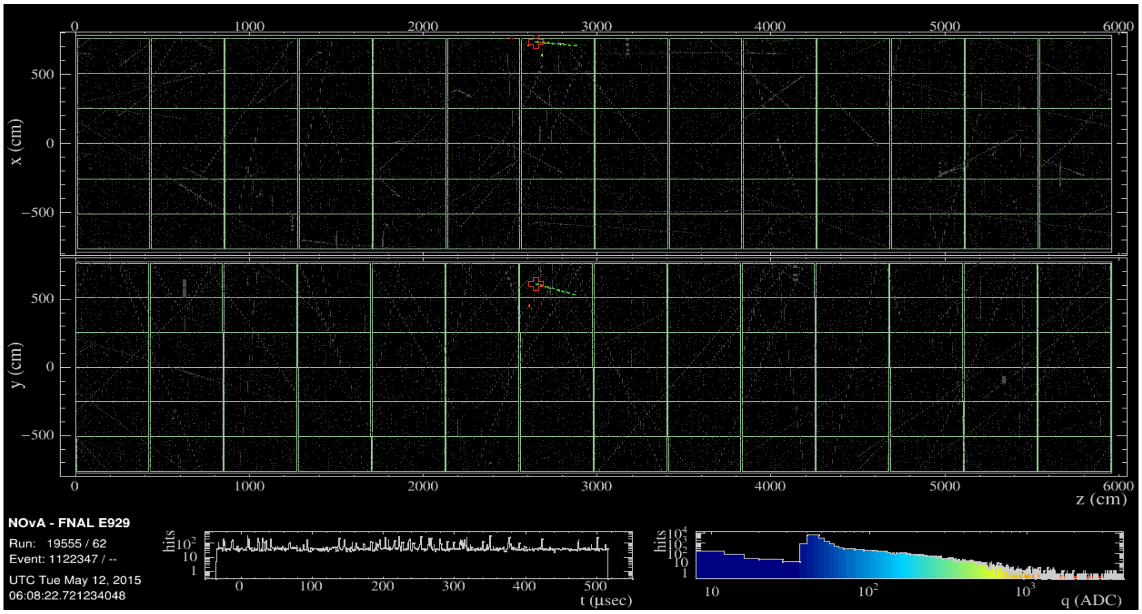
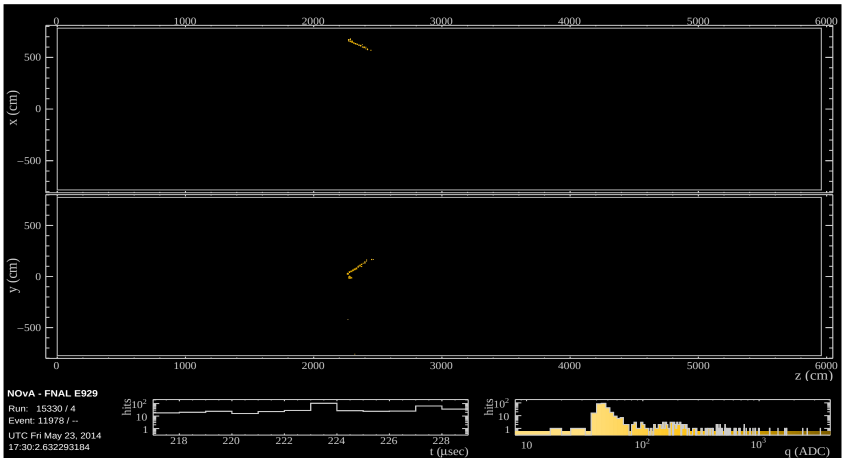


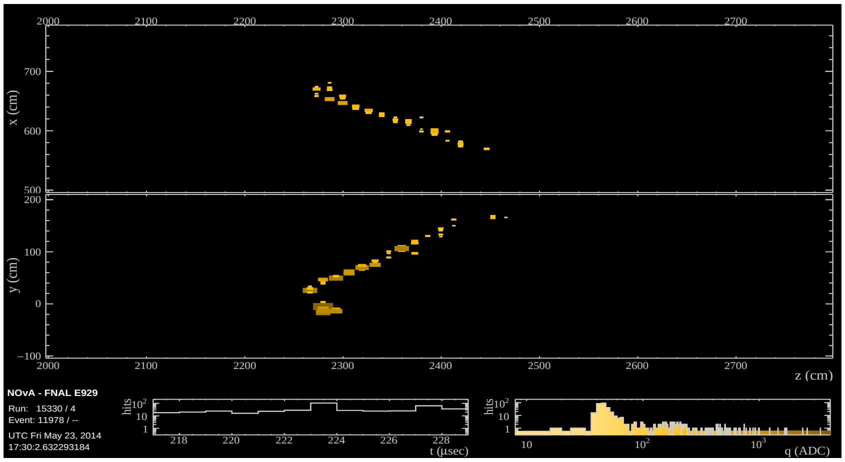
Figure C.2: This event is selected by both LID and LEM. LID is 0.96. LEM is 0.82. This event is 26 cm from the west wall.

APPENDIX D. EVENT DISPLAY AND dE/dx DISTRIBUTIONS OF ν_e CC CANDIDATES

D.1 Events Selected by Both LID and LEM

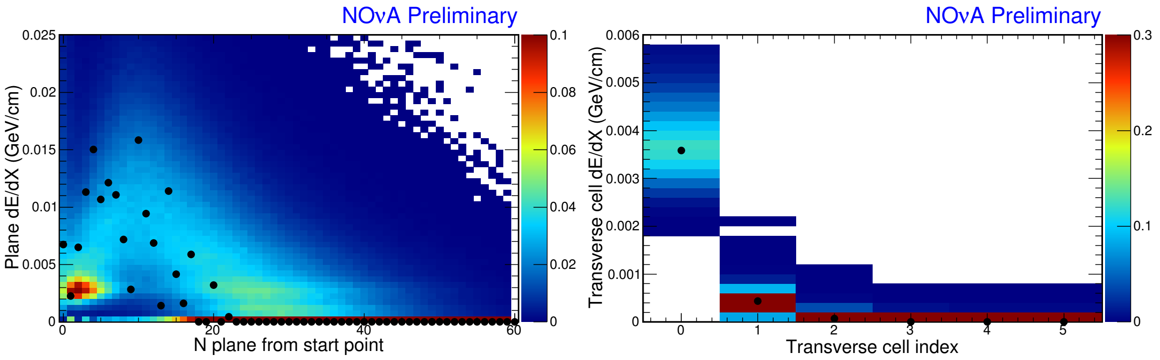


(a)



(b)

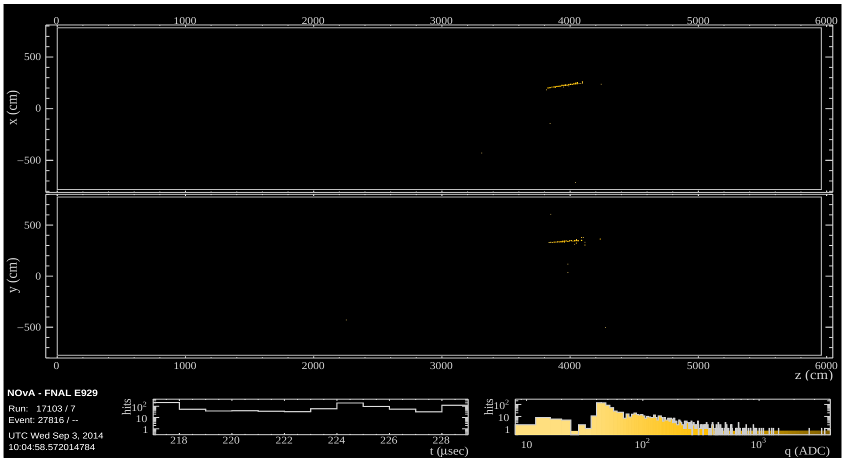
Figure D.1: Full (top) and zoomed-in (bottom) event display of one of the ν_e CC candidates in Run15330 selected by both LID and LEM.



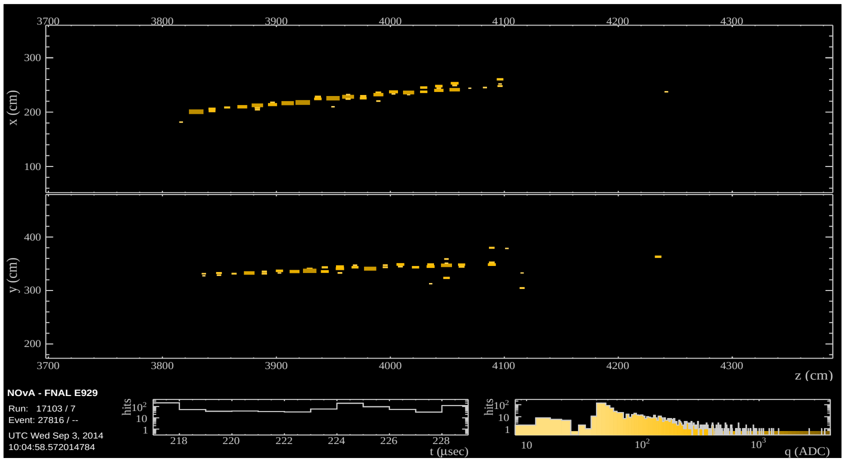
(a) (b)

Figure D.2: Longitudinal (left) and transverse (right) energy deposition rate of one of the ν_e CC candidates in Run15330 selected by both LID and LEM.

D.2 Events Selected by Only LEM

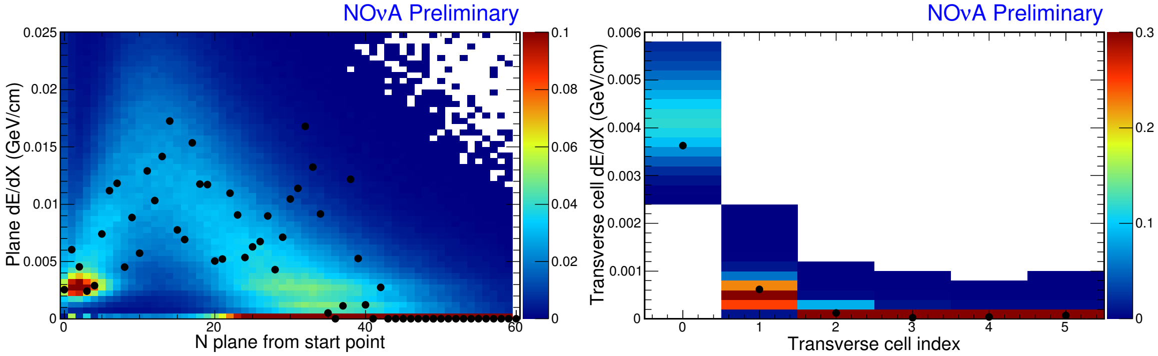


(a)



(b)

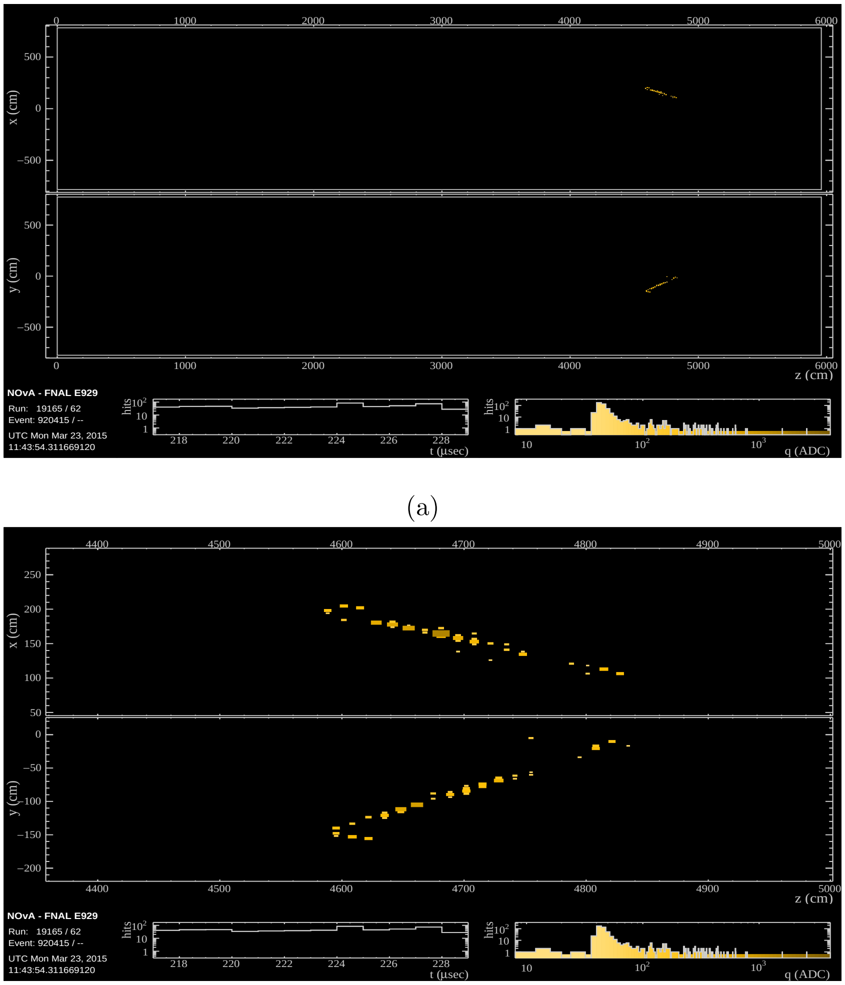
Figure D.3: Full (top) and zoomed-in (bottom) event display of one of the ν_e CC candidates in Run17103 selected by both LID and LEM.



(a)

(b)

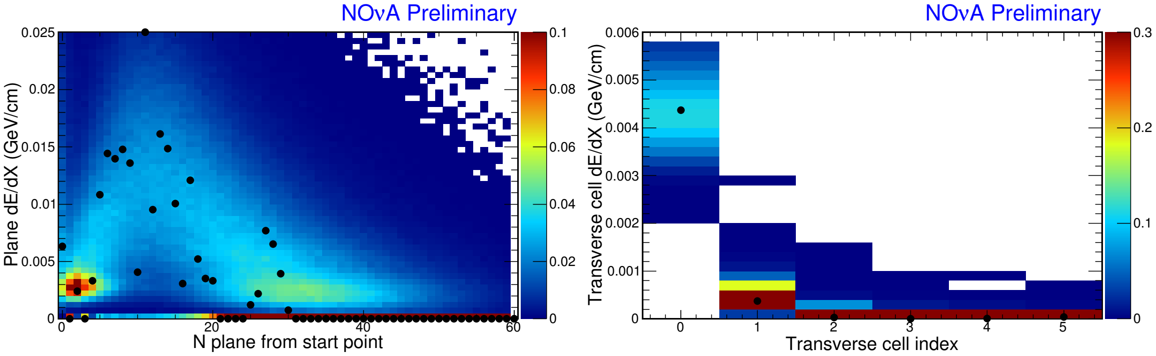
Figure D.4: Longitudinal (left) and transverse (right) energy deposition rate of one of the ν_e CC candidates in Run17103 selected by both LID and LEM.



(a)

(b)

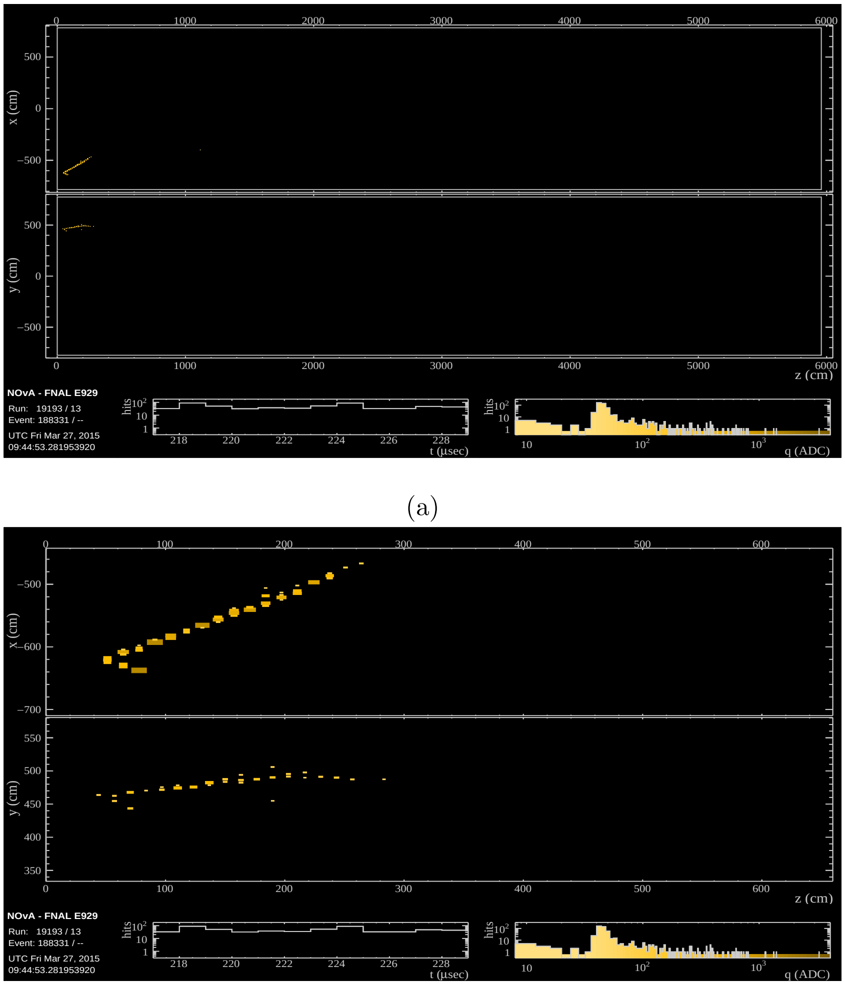
Figure D.5: Full (top) and zoomed-in (bottom) event display of one of the ν_e CC candidates in Run19165 selected by both LID and LEM.



(a)

(b)

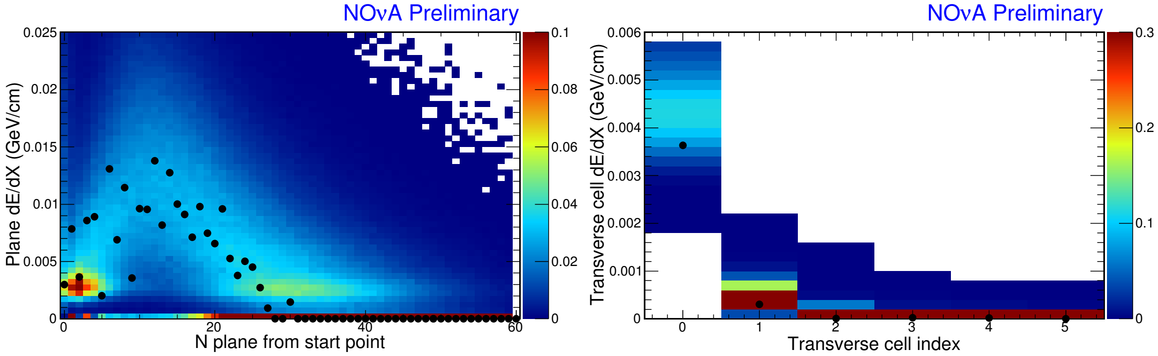
Figure D.6: Longitudinal (left) and transverse (right) energy deposition rate of one of the ν_e CC candidates in Run19165 selected by both LID and LEM.



(a)

(b)

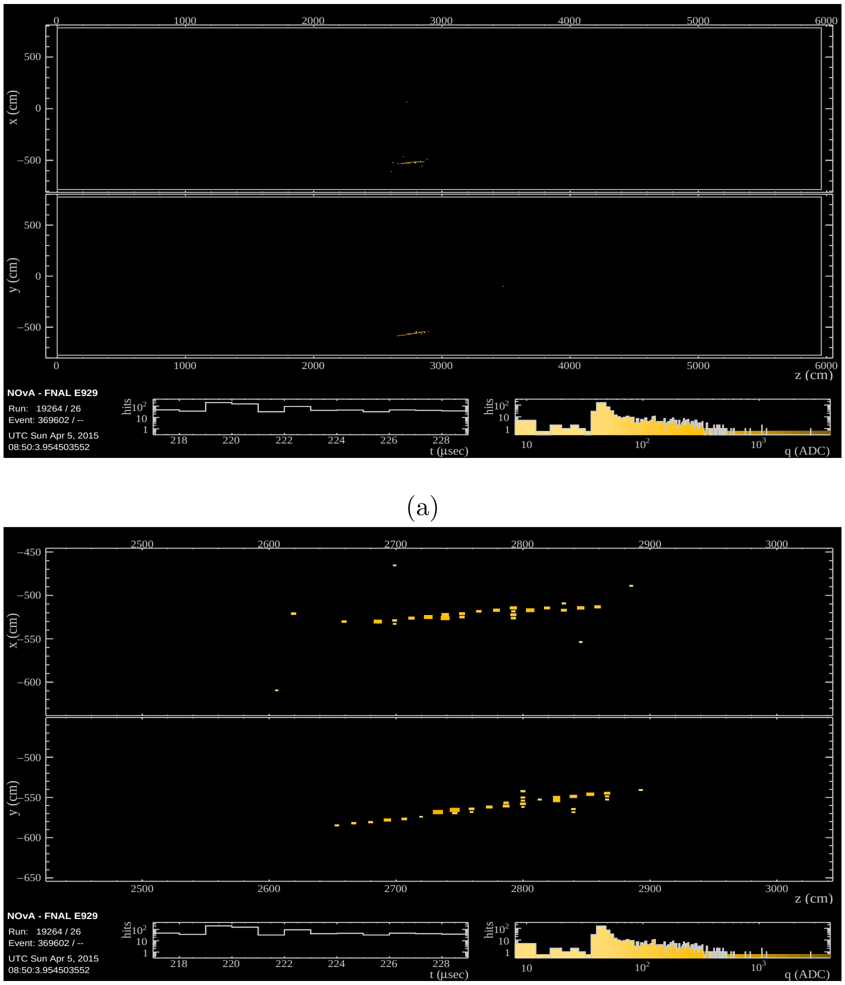
Figure D.7: Full (top) and zoomed-in (bottom) event display of one of the ν_e CC candidates in Run19193 selected by both LID and LEM.



(a)

(b)

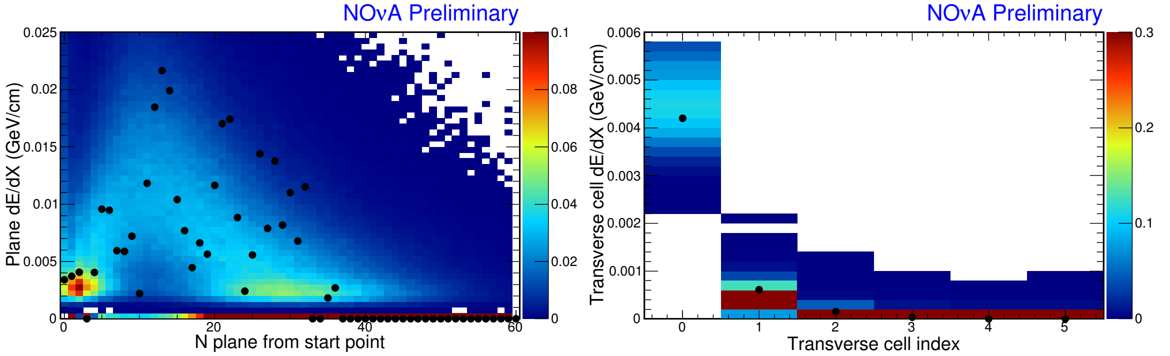
Figure D.8: Longitudinal (left) and transverse (right) energy deposition rate of one of the ν_e CC candidates in Run19193 selected by both LID and LEM.



(a)

(b)

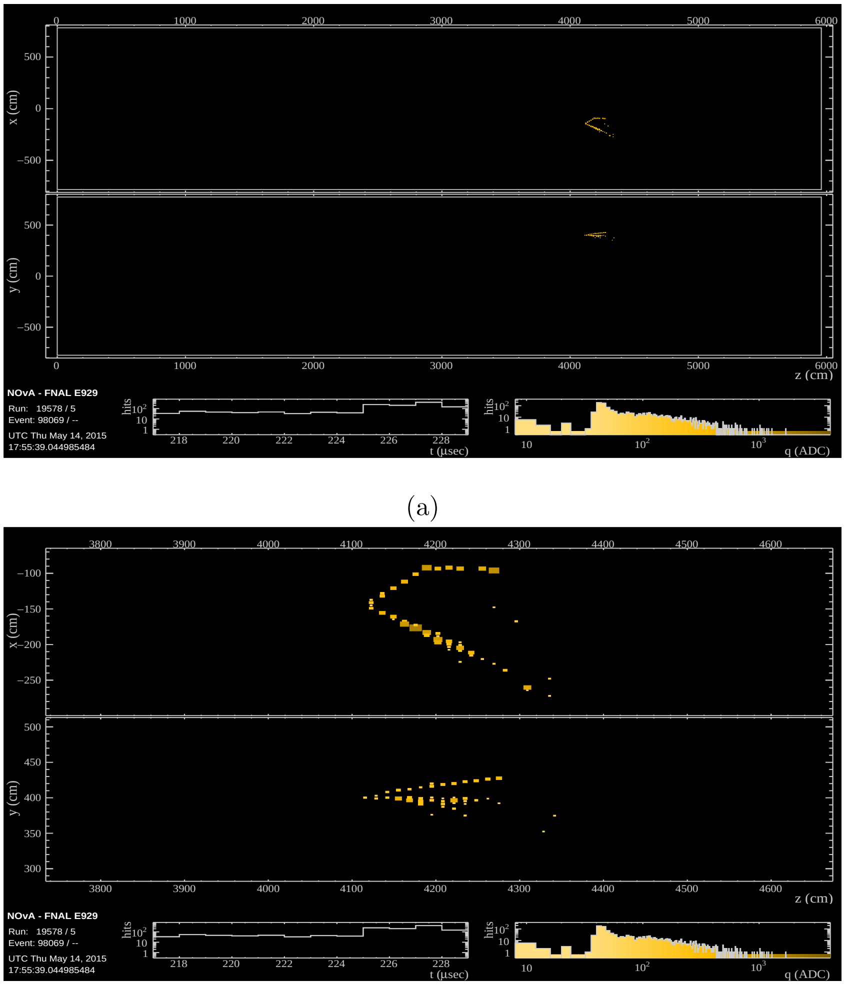
Figure D.9: Full (top) and zoomed-in (bottom) event display of one of the ν_e CC candidates in Run19264 selected by both LID and LEM.



(a)

(b)

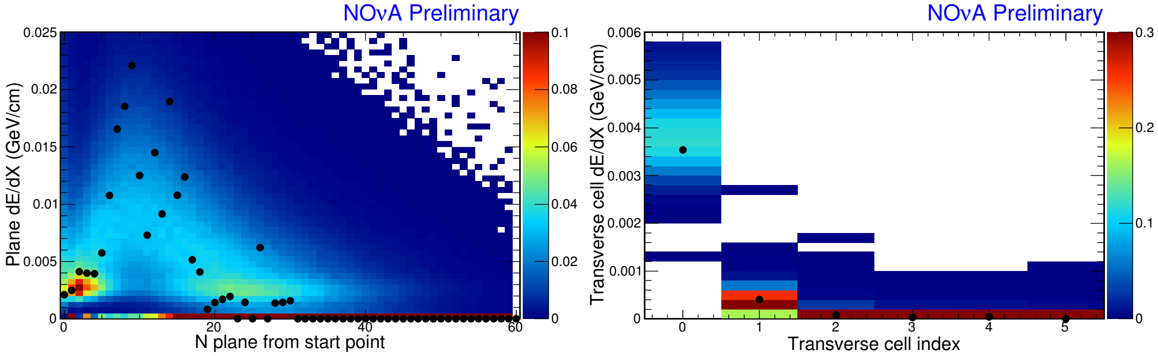
Figure D.10: Longitudinal (left) and transverse (right) energy deposition rate of one of the ν_e CC candidates in Run19264 selected by both LID and LEM.



(a)

(b)

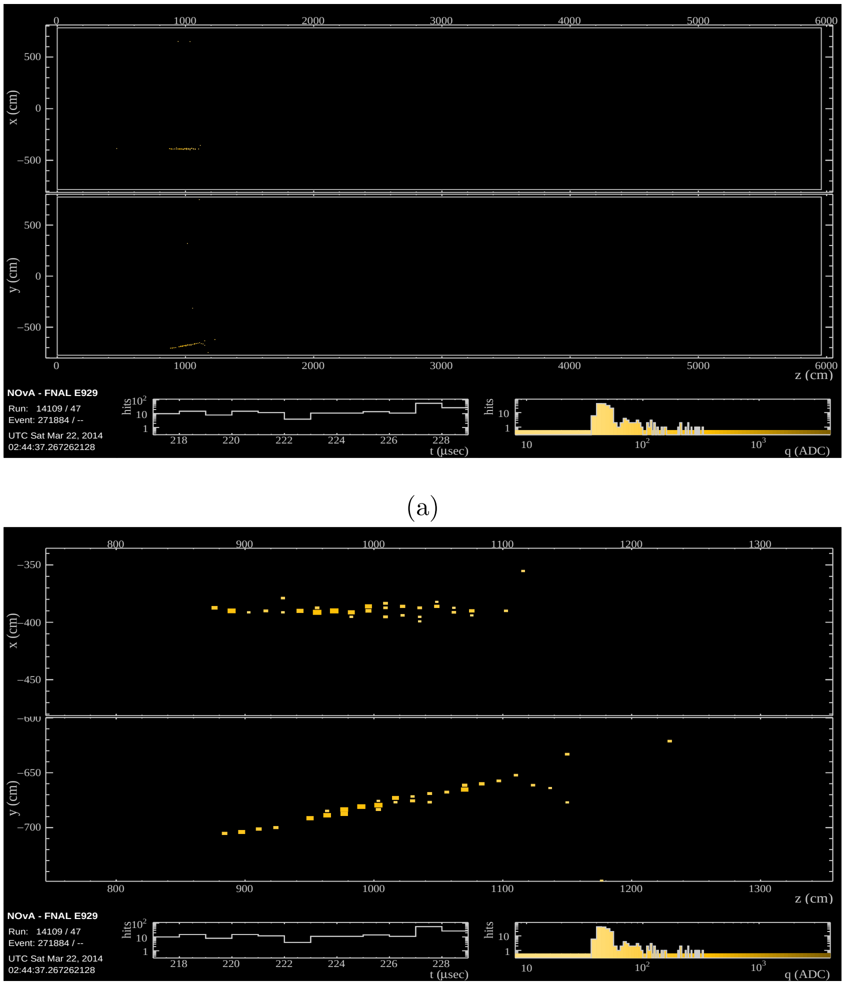
Figure D.11: Full (top) and zoomed-in (bottom) event display of one of the ν_e CC candidates in Run19578 selected by both LID and LEM.



(a)

(b)

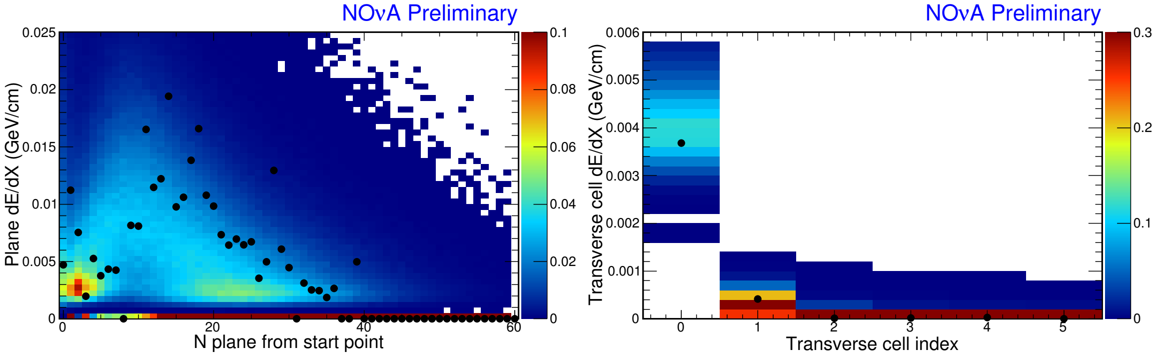
Figure D.12: Longitudinal (left) and transverse (right) energy deposition rate of one of the ν_e CC candidates in Run19578 selected by both LID and LEM.



(a)

(b)

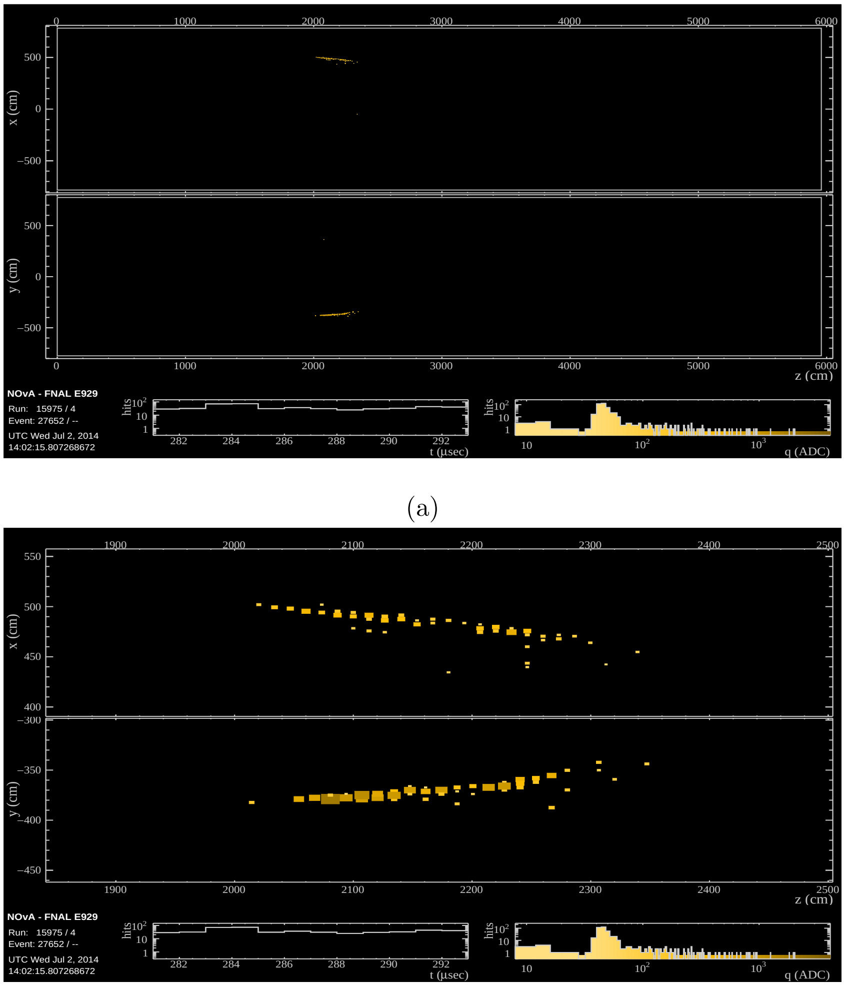
Figure D.13: Full (top) and zoomed-in (bottom) event display of one of the ν_e CC candidates in Run14109 selected by only LEM.



(a)

(b)

Figure D.14: Longitudinal (left) and transverse (right) energy deposition rate of one of the ν_e CC candidates in Run14109 selected by only LEM.

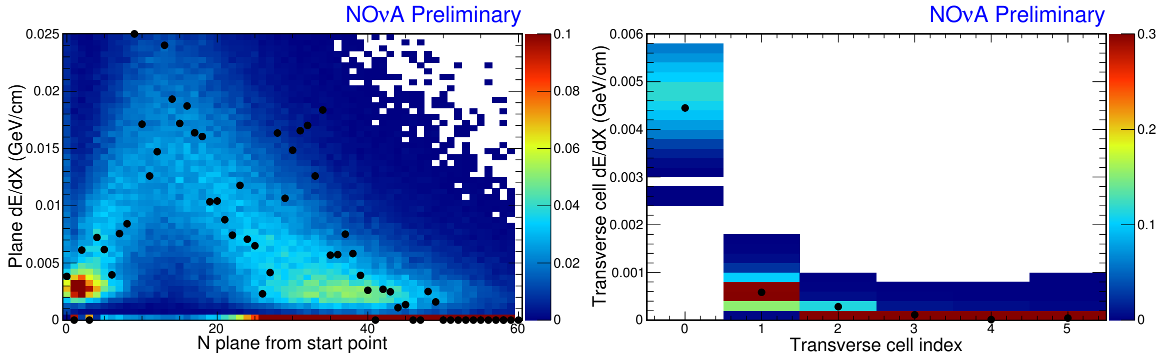


(a)

(b)

Figure D.15: Full (top) and zoomed-in (bottom) event display of one of the ν_e CC candidates in Run15975 selected by only LEM.

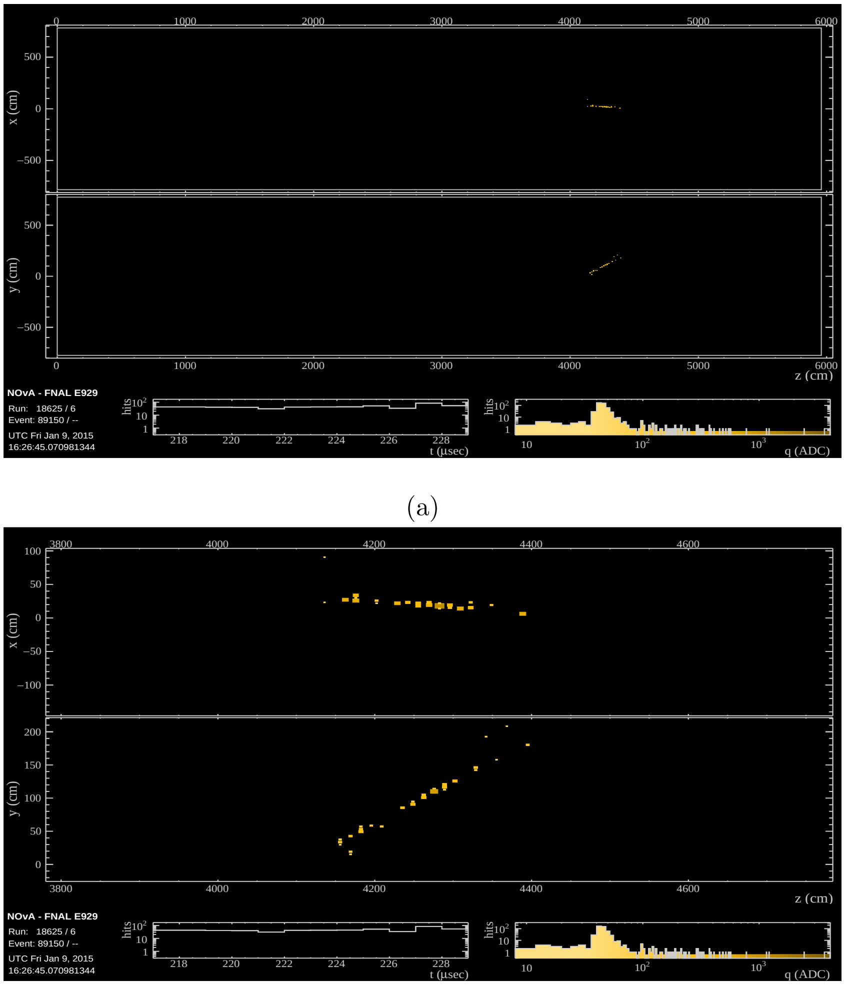
D.3 Likelihood Difference Distributions for LID selected candidates



(a)

(b)

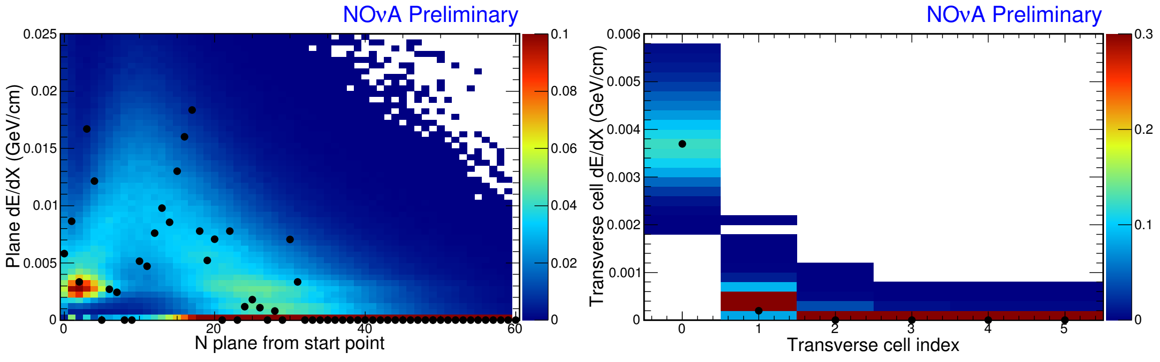
Figure D.16: Longitudinal (left) and transverse (right) energy deposition rate of one of the ν_e CC candidates in Run15975 selected by only LEM.



(a)

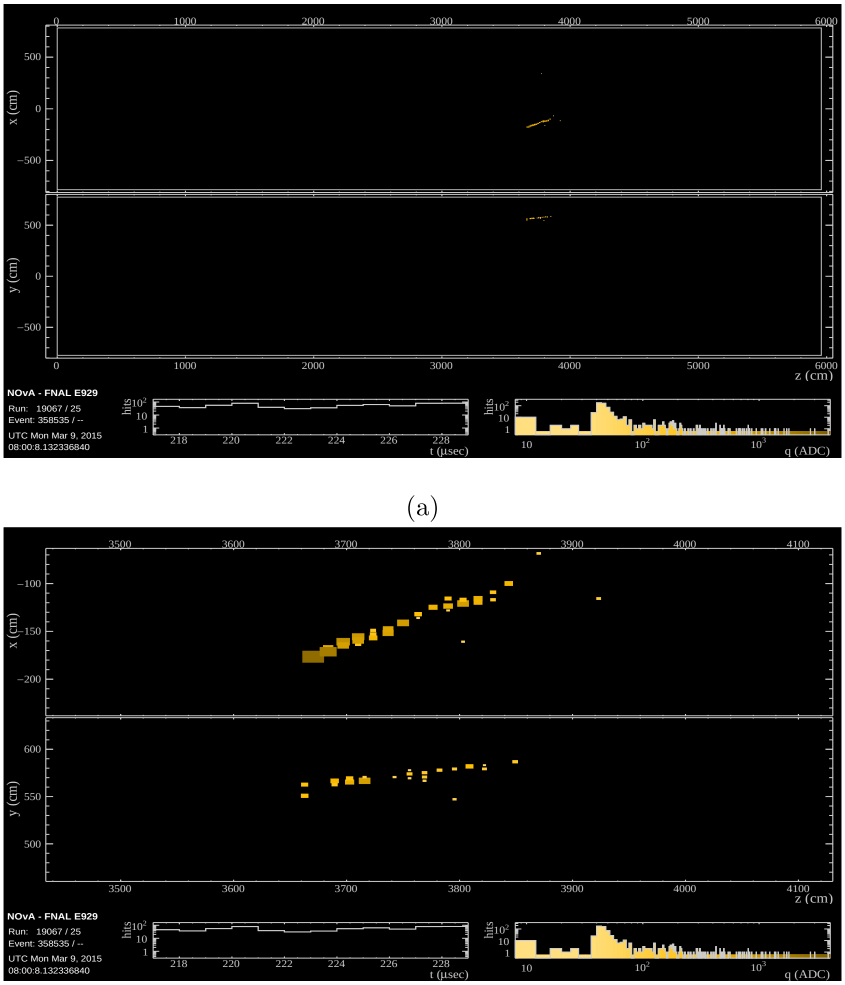
(b)

Figure D.17: Full (top) and zoomed-in (bottom) event display of one of the ν_e CC candidates in Run18625 selected by only LEM.



(a) (b)

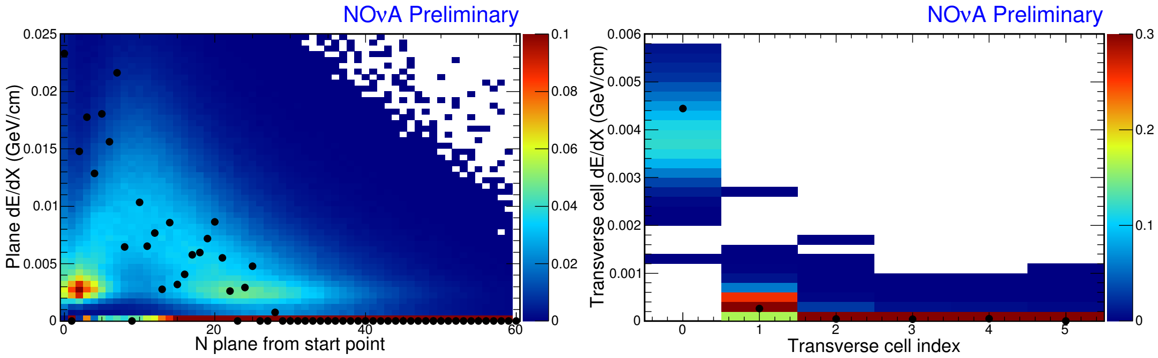
Figure D.18: Longitudinal (left) and transverse (right) energy deposition rate of one of the ν_e CC candidates in Run18625 selected by only LEM.



(a)

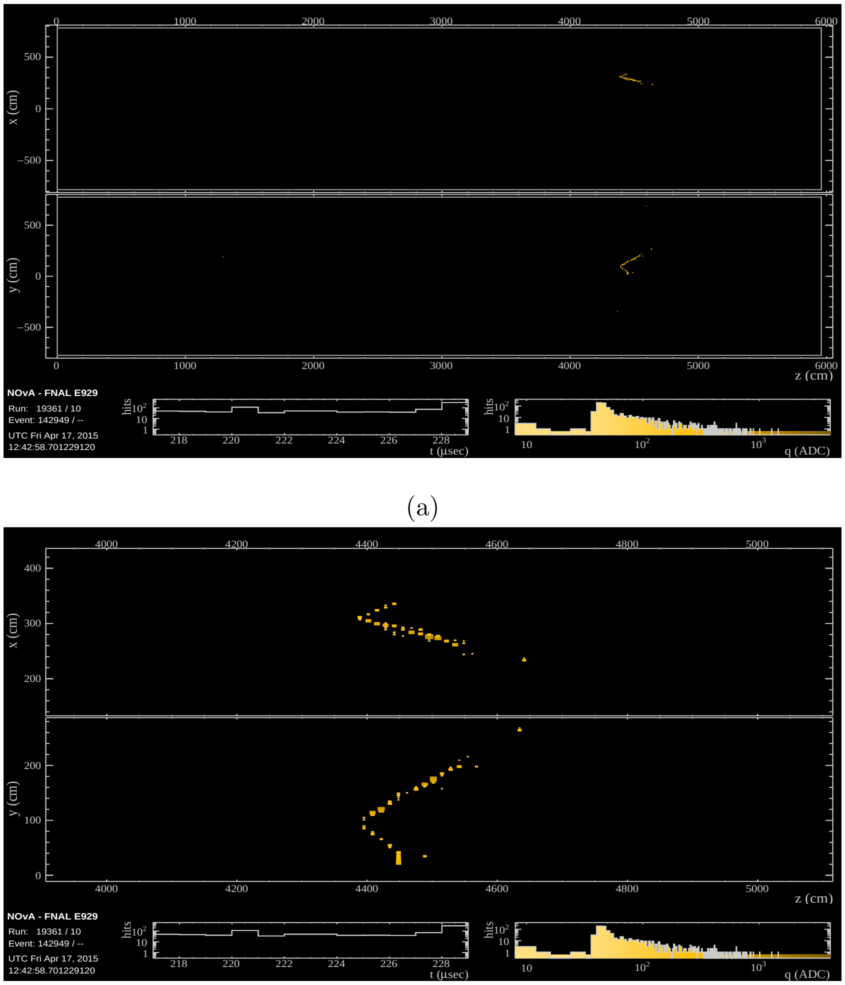
(b)

Figure D.19: Full (top) and zoomed-in (bottom) event display of one of the ν_e CC candidates in Run19067 selected by only LEM.



(a) (b)

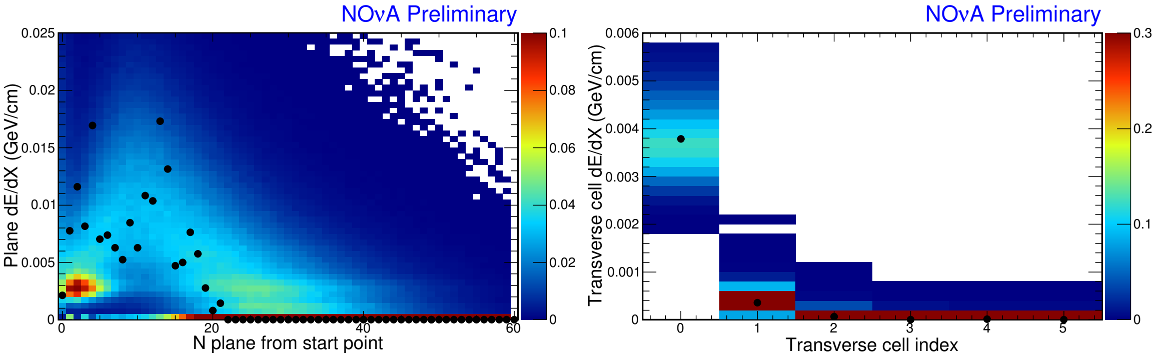
Figure D.20: Longitudinal (left) and transverse (right) energy deposition rate of one of the ν_e CC candidates in Run19067 selected by only LEM.



(a)

(b)

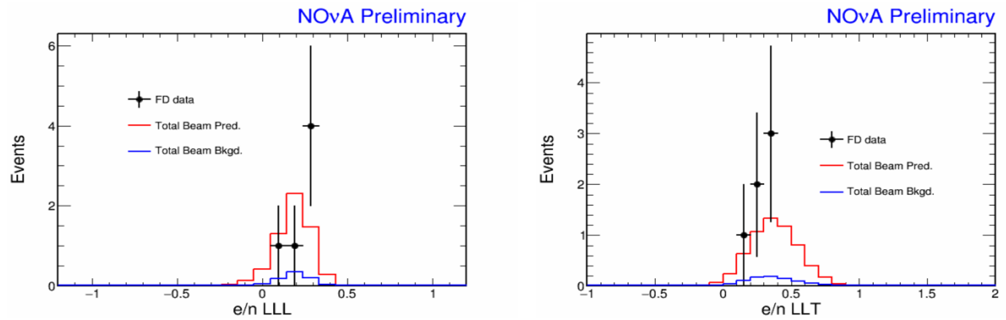
Figure D.21: Full (top) and zoomed-in (bottom) event display of one of the ν_e CC candidates in Run19361 selected by only LEM.



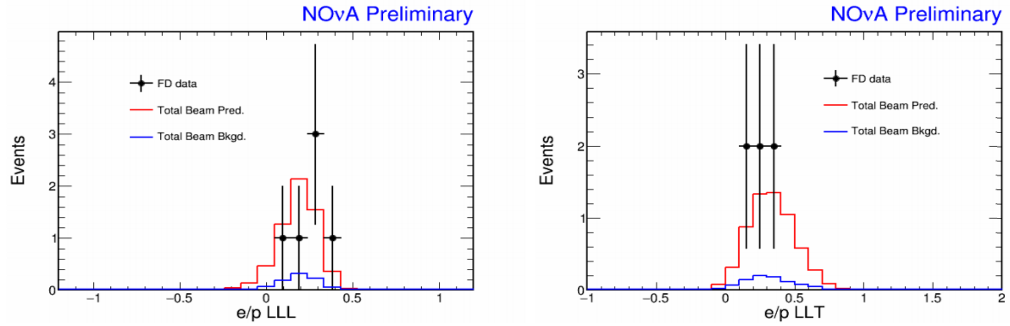
(a)

(b)

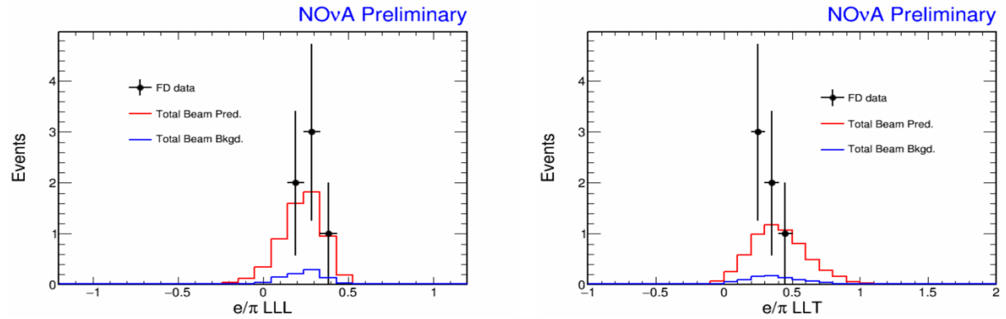
Figure D.22: Longitudinal (left) and transverse (right) energy deposition rate of one of the ν_e CC candidates in Run19361 selected by only LEM.



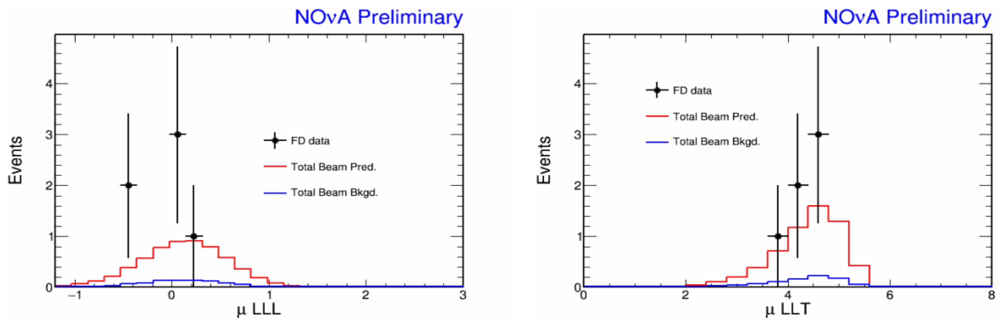
(a) Likelihood of the leading shower being an electron minus likelihood of being a neutron.



(b) Likelihood of the leading shower being an electron minus likelihood of being a proton.



(c) Likelihood of the leading shower being an electron minus likelihood of being a charged pion.



(d) Likelihood of the leading shower being a muon.

Figure D.23: Likelihood of the leading shower of a candidating event in data (black), simulated beam background (blue) and simulated signal + background events (red) under different particle hypotheses.

APPENDIX E. DISTRIBUTIONS FOR LEM SELECTED ν_e CC CANDIDATES

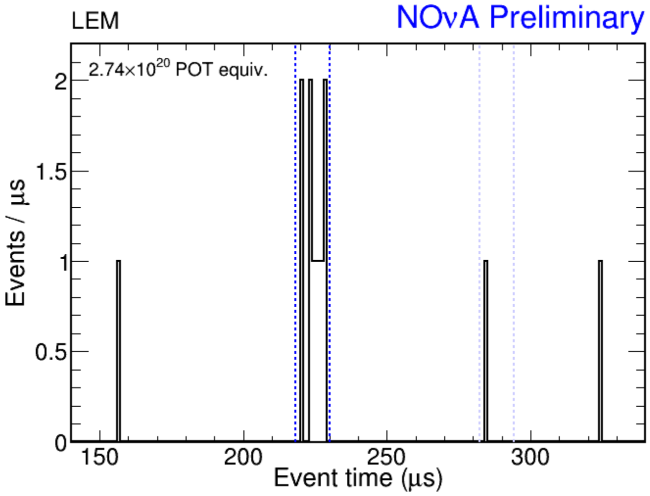


Figure E.1: The distribution of mean time for all LEM selected events (black), including the ν_e CCcandidates which are inside the in-spill window (dash light blue) as well as two cosmic ray background evnets which are outside the spill window.

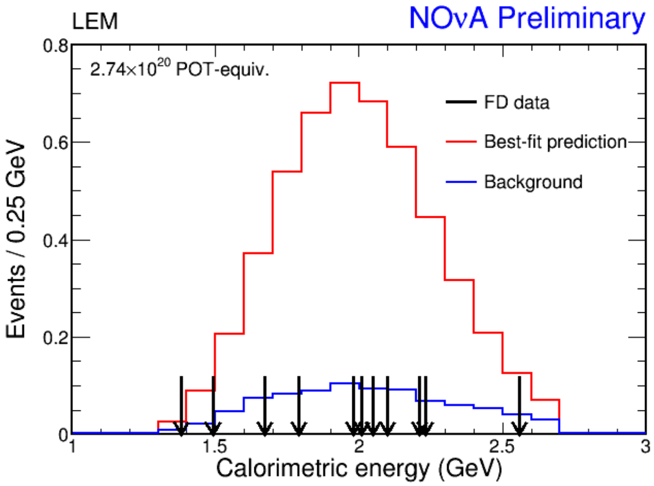


Figure E.2: The distribution of slice calorimetric energy for the ν_e CC candidates in LEM (black) overlapped with the extrapolation predicted background (blue) and predicted background + signal (red).

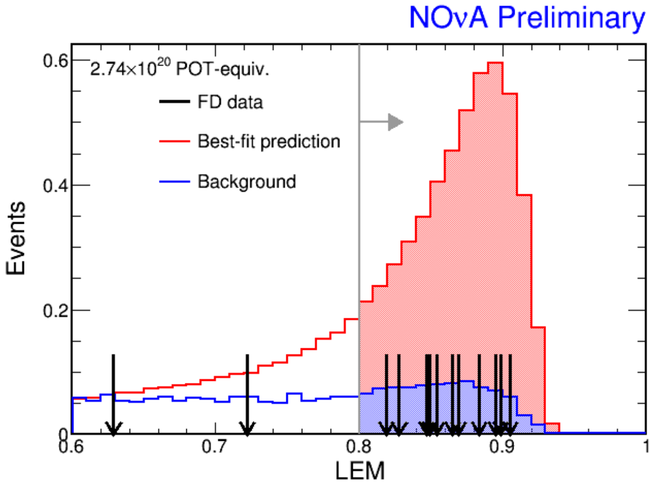


Figure E.3: The distribution of LID for the LEM selected events (black) in the signal (shaded) and low PID (line) region. The distribution for the selected events are compared to the predicted distributions from the extrapolation, the blue for just predicted background and red for the combination of predicted signal and background.

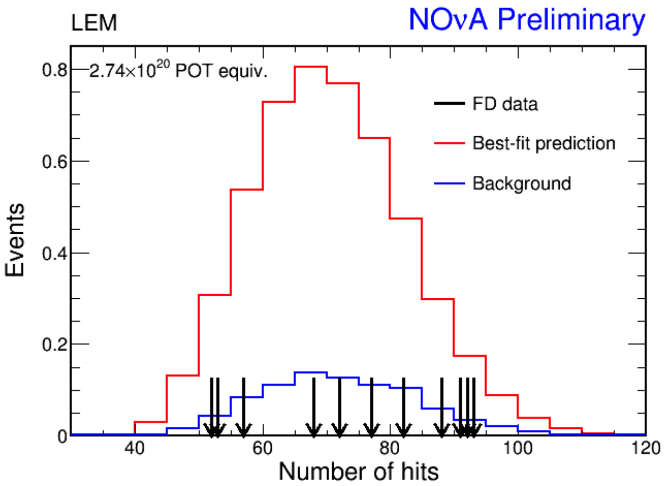


Figure E.4: The distribution of the number of slice hits for the ν_e CC candidates in LEM (black) overlapped with the extrapolation predicted background (blue) and predicted background + signal (red).

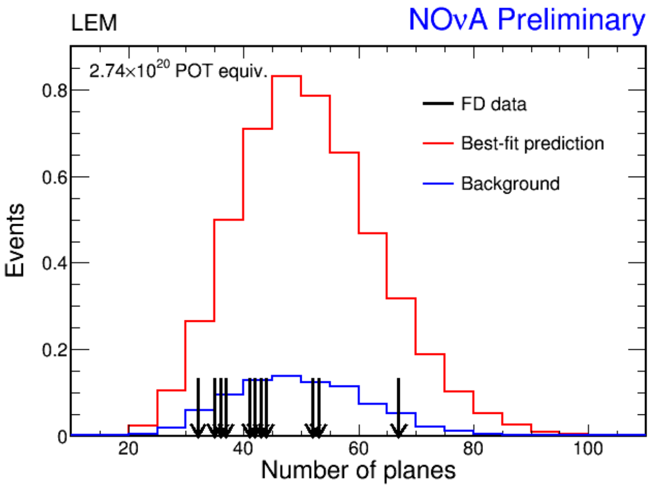


Figure E.5: The distribution of the number of planes of the leading shoewr for the ν_e CC candidates in LEM (black) overlapped with the extrapolation predicted background (blue) and predicted background + signal (red).

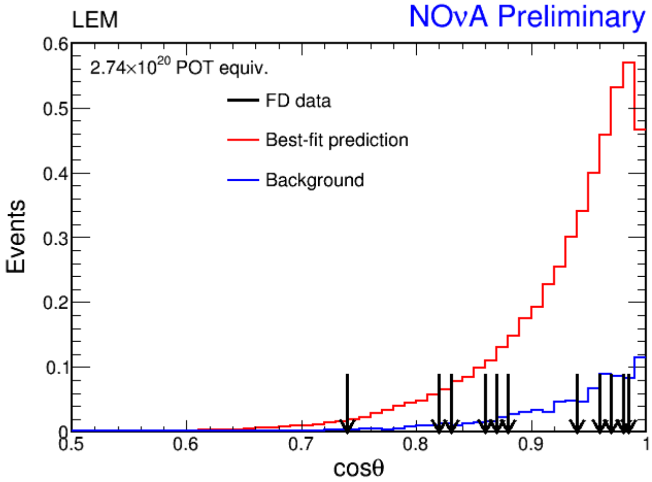


Figure E.6: The distribution of the cosine of the angle of the leading shower w.r.t. the beam direction for the ν_e CC candidates in LEM (black) overlapped with the extrapolation predicted background (blue) and predicted background + signal (red).

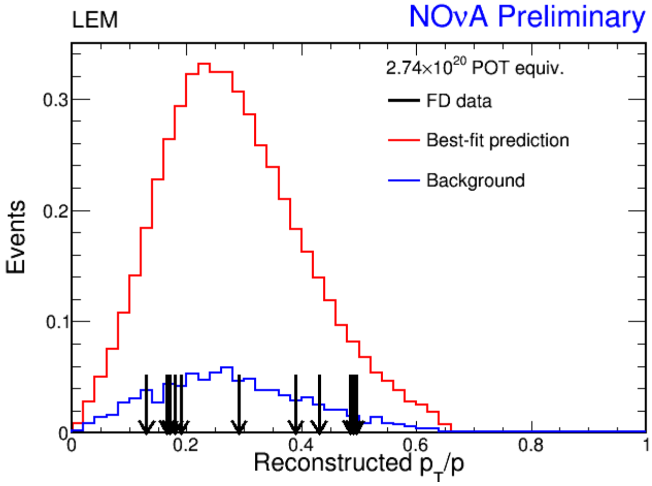
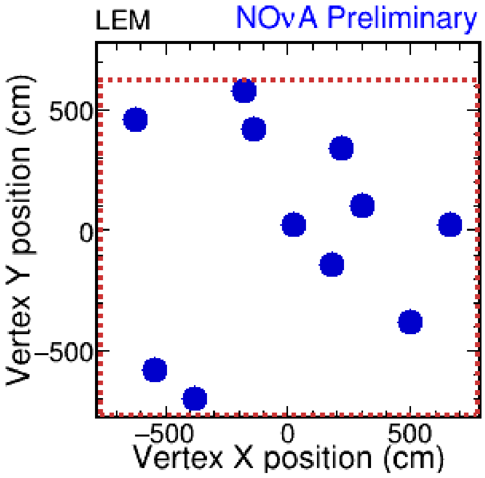
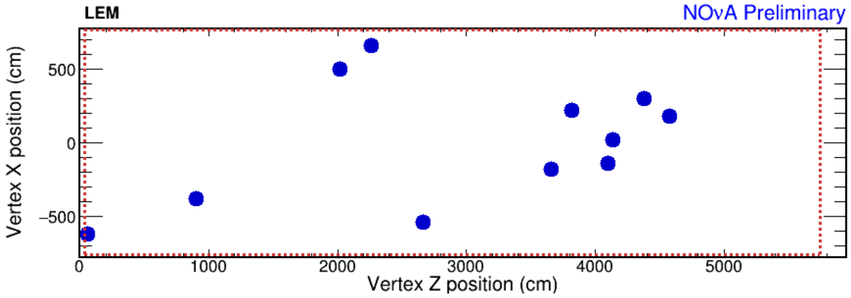


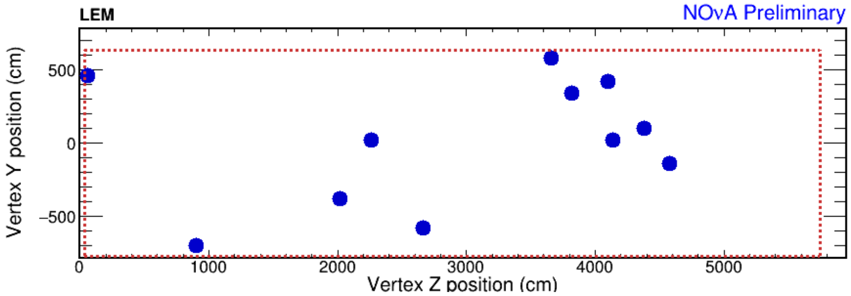
Figure E.7: The distribution of the fraction of transverse momentum for the ν_e CC candidates in LEM (black) overlapped with the extrapolation predicted background (blue) and predicted background + signal (red).



(a)



(b)



(c)

Figure E.8: The position of the vertices of the LEM selected ν_e CC candidates in XY (a), XZ (b) and YZ (c) planes. The position of the vertices is compared to the containment area (dash).

BIBLIOGRAPHY

[1] J. Chadwick. Distribution in intensity in the magnetic spectrum of the -rays of radium. Deutch. Phys. Gessel., 16.8:883891, 1914.

[2] Wolfgang Pauli. Letter to the physical society of Tubingen, 1930.

[3] J. Chadwick. Possible Existence of a Neutron. Nature, 129:312, 1932.

[4] E. Fermi. Tentativo di una teoria dei raggi . Ricerca Scientifica, 1933.

[5] H. Bethe and R. Peierls, Nature 133, 532, 1934.

[6] F. Reines and C. L. Cowan, Phys. Rev. 92, 830 (1953); C. L. Cowan et al., Science 124, 103, 1956.

[7] M. Goldhaber, L. Grodzins and A. W. Sunyar, Phys. Rev. 109, 1015, 1958.

[8] R. P. Feynman and M. Gell-Mann, Phys. Rev. 109, 193, 1958.

[9] S. L. Glashow, Nucl. Phys. 22, 579, 1961.

[10] A. Salam, in Elementary Particle Theory, N. Svartholm, ed.(Stockholm: Almquist and Wiksell, 1968).

[11] S. Weinberg, Phys. Rev. Lett. 19, 1264, 1967.

[12] G. Arnison et al. [UA1 Collaboration], Phys. Lett. B 122, 103, 1983.

[13] F. J. Hasert et al. [Gargamelle Neutrino Collaboration], Phys. Lett. B 46, 138, 1973.

[14] P. Bagnaia et al. [UA2 Collaboration], Phys. Lett. B 129, 130, 1983.

[15] G. S. Abrams et al., Phys. Rev. Lett. 63, 2173, 1989.

[16] D. Decamp et al. [ALEPH Collaboration], Phys. Lett. B 231, 519, 1989.

[17] G. Danby, J. M. Gaillard, K. Goulianos, L. M. Lederman, N. B. Mistry, M. Schwartz and J. Steinberger, Phys. Rev. Lett. 9, 36, 1962.

[18] M. L. Perl et al., Phys. Rev. Lett. 35, 1489, 1975.

[19] K. Kodama et al. [DONUT Collaboration], Phys. Lett. B 504, 218, 2001 [arXiv:hep-ex/0012035].

[20] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. Phys.Rev.Lett., 81:15621567, 1998.

[21] Kitagaki T et al. Phys. Rev. D 28 436, 1983.

[22] Horstkotte J et al. Phys. Rev. D 25 2743, 1982.

[23] K. Nakamura et al. (Particle Data Group), JP G 37, 075021, 2010.

[24] P.A.R. Ade et al. Planck 2015 results. XIII. Cosmological parameters, 2015.

[25] Agostini, M. et al. Physical Review Letters 111 12, 2013.

[26] Maki, Z; Nakagawa, M.; Sakata, S. Progress of Theoretical Physics 28, 870, 1962.

[27] S. P. Mikheyev and A. Yu. Smirnov, Yad. Fiz. 42, 1441 (1985); Nuovo Cim. 9 C, 17 (1986).

[28] V. D. Barger, K. Whisnant, S. Pakvasa and R. J. Phillips, Phys. Rev. D 22, 2718 (1980)

[29] L. Wolfenstein. Neutrino Oscillations in Matter. Phys.Rev., D17:23692374, 1978.

[30] B. T. Cleveland; et al. Astrophysical Journal 496: 505526, 1998.

[31] Gamow G Phys. Rev. 53 595, 1938.

[32] Bethe H A and Critchfield C L, Phys. Rev. 54 248, 862, 1938.

[33] Bellerive, A. "Review of solar neutrino experiments", Int. J. Mod. Phys. A19: 11671179, 2004.

[34] P. Anselmann et al. [GALLEX Collaboration], Phys. Lett. B 285, 376, 1992.

[35] D. N. Abdurashitov et al., Phys. Lett. B 328, 234, 1994.

[36] M. Altmann et al. [GNO COLLABORATION Collaboration], Phys. Lett. B 616, 174, 2005.

[37] Ahmad, QR; et al. Physical Review Letters 87 (7): 071301, 2001.

[38] B. Aharmim et al, [SNO Collab.],Phys. Rev. C72, 055502, 2005.

[39] Nick Jelley, Arthur B. McDonald, and R.G. Hamish Robertson, Annu. Rev. Nucl. Part. Sci. 59:431-465, 2009.

[40] S. Abe et al. Precision Measurement of Neutrino Oscillation Parameters with KamLAND. Phys.Rev.Lett., 100:221803, 2008.

[41] T. J. Haines et al., Phys. Rev. Lett. 57, 1986; D. Casper et al., Phys. Rev. Lett. 66, 2561, 1991.

[42] Allison, W.W.M. et al. [Soudan2 Collab.] Phys.Lett. B391, 1997.

[43] K. S. Hirata et al. [KAMIOKANDE-II Collaboration], Phys. Lett. B 205, 416, 1988.

[44] P. Adamson et al. [MINOS Collaboration] Phys. Rev. Lett. 110, 251801, 2013.

[45] M. Apollonio et al. Phys.Lett., B466:415430, 1999.

[46] F. P. An et al. [Daya Bay Collaboration] Phys. Rev. Lett. 108 171803, 2012.

[47] K. Abe et al. [T2K Collaboration] Phys. Rev. Lett. 112, 061802, 2014.

[48] K. Abe et al. [T2K Collaboration] Phys. Rev. D 91, 072010, 2015.

[49] S. M. Bilenky and Carlo Giunti. Mod. Phys. Lett., A27, 1230015, 2012, 1203.5250. filbreak

[50] F. Capozzi, G.L. Fogli, E. Lisi, A. Marrone, D. Montanino, et al. Phys.Rev. D89(9):093018, 2014.

[51] P. Adamson, K. Anderson et al. "The NuMI Neutrino Beam", arXiv:1507.06690v2 [physics.acc-ph] 29 Jul, 2015.

[52] NuMI Technical Design Handbook, 2012.

[53] NOvA colleagues, NOvA technical design report, Oct, 2010.

[54] R. Schroeter. NOvA internal docdb 13097-v4.

[55] NOvA internal docdb 4371.

[56] L. Vinton et al. "First Analysis Calibration Technotes". NOvA internal docdb 13579.

[57] A. Radovic et al. "Attenuation Calibration Bless Plots Package". NOvA internal docdb 13898.

[58] L. Vinton et al. "Calorimetric Energy Scale Calibration". NOvA internal docdb 13607.

[59] L. M. Mualem "Electronics and DAQ production - plenary talk for DOE CD-1 Review". NOνA internal docdb 715.

[60] L. M. Mualem "DAQ Threshold". NOνA internal docdb 83.

[61] R. Fruhwirth, M. Regler "Data Analysis Techniques for High-Energy Physics", Second Edition.

[62] Fabjan, C.W. Calorimetry in High Energy Physics. In: Concepts and Techniques in High Energy Physics III, p. 281, Plenum Press, New York, 1985.

[63] Holder, M. et al. Performance of a Magnetized Total Absorption Calorimeter between 15 GeV and 140 GeV. Nuclear Instruments and Methods 151, 69, 1978.

[64] Raphael Schroeter. Wrong sign neutrino contamination and beam nue content for FHC/RHC beam configurations. NOνA internal docdb 13097.

[65] M. Baird, "Slicing Module Comparison Tech Note". NOνA internal docdb 9195.

[66] M. Baird, "Introduction to Slice4D". NOνA internal docdb 9127.

[67] M.Baird, "Tech Note for the Multi-Hough Transform". NOνA internal docdb 8241.

[68] M. D. Messier, "Vertex Reconstruction Based on Elastic Arms". NOνA internal docdb 7530.

[69] E. Niner, "Vertex Clustering with Possibilistic Fuzzy-K Means Algorithm". NOνA internal docdb 7648.

[70] E. Niner, NOνA internal docdb 8773.

[71] E. Niner, "Likelihood Based nue Identifier (LID)". NOνA internal docdb 13590.

[72] J. Bian, "EID tech note". NOνA internal docdb 9923.

[73] T. Xin, "Note of Selection Cuts for Nue First Analysis". NOνA internal docdb 13592.

[74] C. J. Backhouse, "The Library Event Matching nue PID – technote". NOνA internal docdb 13588.

[75] E. Niner, "Position paper: POT accounting". NOνA internal docdb 13556.

[76] B. Zamorano, "Technical note: FD timing peak monitoring". NOνA internal docdb 13441.

[77] M. Baird, "NearDet Timing Peak Blessing Packag". NOνA internal docdb 11373.

[78] A. Aurisona, "The NOvA Simulation Chain". NOνA internal docdb 13129.

[79] Bohlen T, Cerutti F, Chin M, Fasso A, Ferrari A, Ortega P, Mairani A, Sala P, Smirnov G and Vlachoudis V 2014 Nucl. Data Sheets 120 211 214.

[80] Ferrari A, Sala P R, Fasso A and Ranft J 2005 FLUKA: A multi-particle transport code (Program version 2005) Tech. Rep. CERN-2005-010 CERN.

[81] Campanella M, Ferrari A, Sala P and Vanini S 1999 First Calorimeter Simulation with the FLUGG Prototype Tech. Rep. CERN-ATL-SOFT-99-004 CERN.

[82] Andreopoulos C et al. 2010 Nucl. Instrum. Meth. A 614 87104.

[83] Hagmann C, Lange D and Wright D 2007 Nuclear Science Symp. Conf. Rec. (Honolulu, HI) vol 2 (IEEE) pp 11431146.

[84] Agostinelli S et al. 2003 Nucl. Instrum. Meth. A 250303.

[85] Allison J et al. 2006 IEEE Trans. Nucl. Sci. 53(1) 270278.

[86] A. Aurisona, "The NOvA Detector Simulation". NOνA internal docdb 13577.

[87] C. J. Backhouse, "Near Detector data/MC comparisons for the nue analysis". NOνA internal docdb 13587.

[88] T. Xin, "data/mc plots for cosmic rejection variables". NOνA internal docdb 11335.

[89] T. Xin, "data/mc plots for EID variables". NOνA internal docdb 11337.

[90] J. Coelho, "Tech note: Good data selection". NOνA internal docdb 13546.

[91] M. Messier, "Timing system notes". NOνA internal docdb 13354.

[92] J. Paley, "BadChannels Technical Note". NOνA internal docdb 12771.

[93] J. Coelho, "FD good runs modifications". NOνA internal docdb 12479.

[94] J. Coelho, "ND Good Data Selection Improvement". NOνA internal docdb 13213.

[95] K. Sachdev, X. Bu, "Spill Level Data Quality- Technical Note". NOνA internal docdb 12437.

[96] L. Goodenough, "Technical Note on the NOvA Beam Monitoring for 2015 Summer Analysis". NOνA internal docdb 13572.

[97] S. Lein, "DCM Edge Metric". NOνA internal docdb 13527.

[98] K. Bays, "LiveGeometry tech note". NOνA internal docdb 11470.

[99] K. Bays, "Cosmic rejection technical note". NOνA internal docdb 11205.

[100] NOvA colleagues. (2007, October). NOvA technical design report, ch.4.

[101] J.J. Beatty, J. Matthews. (2011, August). Particle Data Group Review, ch.26, Cosmic Rays.

[102] G. Davies, "ND cosmics". NOνA internal docdb 13259.

[103] H. Duyang, "Cosmic and Rock Muon Induced EM Showers Technote". NOνA internal docdb 13594.

[104] K. Sachdev, "Muon-Removed Charged Current: A Technical Note". NOνA internal docdb 9729.

[105] E. Niner, "Position paper: nue extrapolation". NOνA internal docdb 13585.

[106] G. Davies, "Position paper: nue systematics". NOνA internal docdb 13597.

[107] R. Schroeter, "A Technote Describing the Derivation and Size of NuMI Flux Uncertainties Used in the First NOvA Analyses". NOνA internal docdb 13584.

[108] D. Pershey, "Michel E Calib Check". NOνA internal docdb 13519.

[109] M. Tamsett, "Energy Systematic". NOνA internal docdb 13331.

[110] J. Liu, "ND Calibration Systematics shift". NOνA internal docdb 13208.

[111] T. Xin, "FD calibration systematics study". NOνA internal docdb 13054.

[112] J. Hartnell, "NOvA Light/ADC Levels and Thresholds". NOνA internal docdb 13505.

[113] J. Hartnell, "Plots/Discussion from light level/threshold studies". NOνA internal docdb 13589.

[114] J. Liu, "Rock muon in ND". NOνA internal docdb 13563.

[115] I. M. Anghel, "Beam Nue Prediction in ND". NOνA internal docdb 13634.

[116] D. Pershey, "Michel E Tagging". NOνA internal docdb 13207.

[117] K. Sachdev, "MRCC tech note". NOνA internal docdb 9729.

[118] X. Tian, "Geometry update". NOνA internal docdb 11805.

[119] T. Xin, "Alignment systematics for nue analysis". NOνA internal docdb 12674.

[120] L. Suter, "NuMu ND selection tech note". NOνA internal docdb 13212.

[121] K. Bays, "Hadronic E discussion". NOνA internal docdb 13635.

[122] S. Lein, "FA Numu Energy Retune". NOνA internal docdb 13327.

[123] E. Niner, "Position paper: nue high energy sideband". NOνA internal docdb 13557.

[124] E. Niner, "Nue low pid sideband prediction". NOνA internal docdb 13649.

[125] C. J. Backhouse, "The steps of the calibration chain". NOνA internal docdb 8221.

[126] K. Sachdev, "Note on nue energy reconstruction". NOνA internal docdb 13633.

[127] E. Niner, "Position paper: nue high energy sideband". NOνA internal docdb 13557.

[128] E. Niner, "Nue low pid sideband prediction". NOνA internal docdb 13649.

[129] E. Niner, "blessing: nue ks plots". NOνA internal docdb 13807.

[130] G. Feldman, "How Probable is the Observed Lid Only/LEM Only/Both Distribution?". NOνA internal docdb 13720.

[131] G. Davies, "Blessing plots: Nue FD Data/MC". NOνA internal docdb 13874.

[132] J. Bian, "blessing package: Far Detector Event dE/dx and Electron Probability Surface". NOνA internal docdb 13801.

[133] G. Davies, "First Analysis Far Detector Data vs MC comparison". NOνA internal docdb 13732.

[134] F. Psihas, "First Ana Event Displays". NOνA internal docdb 13846.

[135] C. J. Backhouse, "Results of the nue box opening". NOνA internal docdb 13715.

[136] A. Radovic, "Oscillation Parameters for First NOvA Analyses". NOνA internal docdb 13640.

[137] E. Catano-Mur, "Nue - Significance cross-checks". NOνA internal docdb 13730.

[138] S. Magill, "Longitudinal Development of Electromagnetic Showers in the NOvA Detector". NOνA internal docdb 6689.

[139] C. J. Backhouse, "Contours, slices, and Feldman-Cousins corrections for the first nue analysis – technote". NOνA internal docdb 13897.

[140] J. Bian, "First Results of ν_e Appearance Analysis and Electron Neutrino Identification at NOνA". arXiv:1510.05708 [hep-ex].

[141] C. J. Backhouse, "Blessed plots for nue First Ana contours/slices". NOνA internal docdb 13871.

[142] A. Marrone, "Three-neutrino mixing: status and prospects", TAUP 2015.

[143] I. Kourbanis, "Longer Term Plans and Risks for 700 KW". NOνA internal docdb 14217.

[144] R. J. Tesarek, "Project Status and Far Detector Commissioning". NOνA internal docdb 10544.

[145] J. Zelesak, "Operations Group for Collaboration meeting". NOνA internal docdb 13226.

[146] T. Xin, "FarDet APD operation". NOνA internal docdb 11047.

[147] T. Xin, "FEB/APD noise technical note". NOνA internal docdb 10168.

[148] C. J. Backhouse, "FD self-shielding calibration effects". NOνA internal docdb 8756.

[149] A. Radovic, "Attenuation, Threshold, and Absolute Calibration Overview". NOνA internal docdb 12201.

[150] T. Xin, "data-driven threshold correction". NOνA internal docdb 12483-v7.

[151] T. Xin, "data-driven threshold correction". NOνA internal docdb 12483-v14.

[152] C.N. Chou, "The Nature of the Saturation Effect of Fluorescent Scintillators". Phys. Rev. 87.904, 1952.

[153] The GENIE collaboration, "The GENIE Neutrino Monte Carlo Generator: Physics and User Manual". arXiv:1510.05494 [hep-ph].

[154] P. Vahle, "Nue PID Selection". NOνA internal docdb 13702.

[155] A. Gando et al. [KamLAND Collaboration]. "Reactor on-off antineutrino measurement with KamLAND". Phys. Rev. D 88, 033001 Published 2 August 2013

[156] Y. Ashie et al., [Super-Kamiokande Collab.], Phys. Rev. D71, 112005, 2005.

[157] P. Adamson et al. [MINOS Collab.] Phys. Rev. Lett. 112, 191801, 2014.

[158] F.P. An et al, [Daya Bay Collab.] Phys. Rev. D 90, 071101, 2014.

[159] F.P. An et al, [Daya Bay Collab.] Nucl. Instr. Meth A 685, 78-97, 2012.

[160] S. R. Phan Budd. "PVC Dimensions.pdf". NOνA internal docdb 13757.

[161] NOνA Nearline website: http://nusoft.fnal.gov/nova/datacheck/nearline/nearline.html.

[162] NOνA POT Monitor website: http://nusoft.fnal.gov/nova/datacheck/nearline/POTmonitoring.html.

[163] R. Dharmapalan. NOνA internal docdb 13823.

[164] G. Feldman. NOνA internal docdb 13720.