

DISSERTATION

A MEASUREMENT OF MUON NEUTRINO CHARGED-CURRENT INTERACTIONS  
WITH A CHARGED PION IN THE FINAL STATE USING THE NO $\nu$ A NEAR  
DETECTOR.

Submitted by

Paul Nelson Rojas

Department of Physics

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2023

Doctoral Committee:

Advisor: Norm Buchanan

Siu Au Lee

John Harton

Piotr Kokoszka

Copyright by Paul Nelson Rojas 2023

All Rights Reserved

## ABSTRACT

### A MEASUREMENT OF MUON NEUTRINO CHARGED-CURRENT INTERACTIONS WITH A CHARGED PION IN THE FINAL STATE USING THE NO $\nu$ A NEAR DETECTOR.

The NO $\nu$ A experiment is a long-baseline neutrino experiment hosted by Fermilab. The intense NuMI neutrino beam, combined with the NO $\nu$ A Near Detector, provides the opportunity to study neutrino interactions at an unprecedented level. The goal of this analysis is to measure the rate of muon-neutrino charged-current interactions in the NO $\nu$ A near detector resulting in the production of one muon and at least one charged pion. This thesis will present the result of the double differential cross section measurement of this process in muon kinematics of energy and angle. Excesses in the extracted signal (greater than 25%), relative to the simulation, were found at large scattering angles. These excesses were greater than the estimated uncertainties ( $\sim 15\%$ ).

## ACKNOWLEDGEMENTS

I would like to acknowledge the people that assisted me throughout my time at CSU and throughout my graduate school career. First, I'd like to thank my family, including my parents and siblings. Their encouragement and support over the years was very appreciated. Also to my extended family, grandparents, abuelos, aunts, uncles, tio and tia, cousins, and primos: thank you all as well.

I'd also like to thank my advisor, Norm Buchanan, his guidance and patience over the years was very helpful throughout the analysis process. You also taught me about navigating academia, for which I am grateful. Additionally, thanks to all of my CSU colleagues over the years, Matt Judah, Tyler Boone, Connor Johnson, Derek Doyle, Shih-kai Lin, and Steven Calvez. You all were great help with either working with me or assisting me whenever I needed it. I'd also like to thank all of my NOVA colleagues as well, and in particular: Leo Aliaga, Linda Cremonesi, and Matt Judah, their feedback as cross section conveners was very helpful, always giving good feedback.

Finally, I'd like to thank some specific people in my life that inspired me to pursue my PhD. My middle name namesake and grandfather, Nelson Rudolph Nunnally. I remember when I was a child he would always gift me books about space and dinosaurs, fostering my growing interest in science at a young age. I'd also like to thank my uncle, Emmanuel Anemogiannis. As a child, I remember asking you science questions whenever I visited, which you always entertained. While in college and graduate school, you were the only person in my family I could talk to that could relate to my experiences and research interests.

## TABLE OF CONTENTS

ABSTRACT .....	ii
ACKNOWLEDGEMENTS .....	iii
Chapter 1. Introduction .....	1
1.1. Physics Motivation .....	1
1.2. Additional Related NO $\nu$ A Effort .....	3
Chapter 2. A Brief Neutrino Physics Primer .....	5
2.1. The Standard Model .....	5
2.2. The History of the Neutrino .....	7
2.3. Neutrino Oscillations .....	8
2.4. Neutrino-Nucleon Interactions and Pion Production .....	11
2.5. Previous Measurements of Inclusive Pion Production Cross Sections .....	14
Chapter 3. The NO $\nu$ A Experiment .....	18
3.1. Fermilab Accelerator Complex .....	18
3.2. NuMI Beamline .....	19
3.3. NO $\nu$ A Far Detector .....	21
3.4. NO $\nu$ A Near Detector .....	21
3.5. The NO $\nu$ A Detector Unit Cell .....	24
3.6. Readout Electronics .....	27
Chapter 4. NO $\nu$ A Simulation and Reconstruction, and Analysis Software .....	30

4.1.	Data.....	30
4.2.	Simulation.....	30
4.3.	Simulation Datasets.....	32
4.4.	Reconstruction.....	35
4.5.	Analysis Variables.....	38
Chapter 5. Analysis.....		43
5.1.	Measurement Strategy.....	43
5.2.	Event Preselection.....	52
5.3.	Signal Event Selection.....	54
5.4.	Pion Identification.....	57
5.5.	Measurement Binning.....	61
5.6.	Background Estimation.....	63
5.7.	Signal Estimation Using a Data-Driven Template Fitting Procedure.....	65
5.8.	Unfolding.....	70
5.9.	Efficiency Correction.....	74
5.10.	Flux Calculation.....	76
5.11.	Determination of Target Nucleon Number.....	78
5.12.	Uncertainty Calculation.....	79
5.13.	Mock Data Test.....	81
Chapter 6. Results.....		87
6.1.	Fit Results.....	87
6.2.	Measured Cross Section.....	87
6.3.	Uncertainty on the Measurement.....	92

Chapter 7. Conclusions and Future Work .....	95
7.1. Conclusions .....	95
7.2. Future Work .....	95
Bibliography .....	97
Appendix A. Additional Figures .....	101
Appendix B. List of Abbreviations .....	110

## CHAPTER 1

# INTRODUCTION

This document will describe the work done in the measuring muon neutrino induced interactions that produce a muon and at least one pion in the NO $\nu$ A near detector.

### 1.1. PHYSICS MOTIVATION

Understanding pion cross sections in neutrino experiments is vital to understanding the interaction measurements produced in these experiments. There have been many measurements of inclusive cross sections and charged-current quasi-elastic cross sections. Making measurements of semi-inclusive pion production events and improving pion modelling can reduce uncertainties in the NO $\nu$ A oscillation measurements. Figure 1.1 shows a list of the largest uncertainty contributions to NO $\nu$ A oscillation parameter measurements. The uncertainty from neutrino cross sections are currently one of the largest contributions to these measurements. Tools and techniques developed during this measurement can be used to assist in future pion analyses, measuring event rates of different processes such as deep inelastic scattering (DIS), resonant interactions, and coherent charged pion production, further improving our modelling of these events. Figure 1.2 shows the expected and measured cross section for various neutrino charged-current channels as well as their sum. The total cross section, shown in black, has been measured by many experiments at various energies. The quasi-elastic curve, shown in red, has also been measured by many experiments at various energies. However, the resonant curve, shown in blue, and the deep inelastic scattering curve, shown in green, are more difficult to measure and more measurements are necessary to understand these processes. Identification of pions is a challenging endeavor, and as new

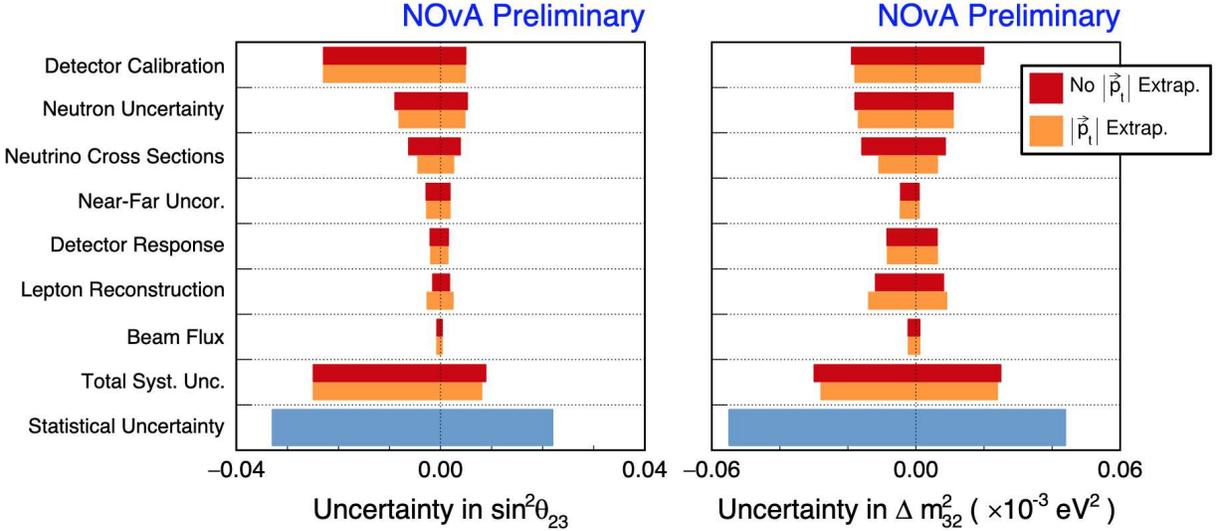


FIGURE 1.1. The individual uncertainties in the measurement of oscillation parameters for NO $\nu$ A [1]. The uncertainty from cross sections is a noticeable contribution to the total uncertainty.

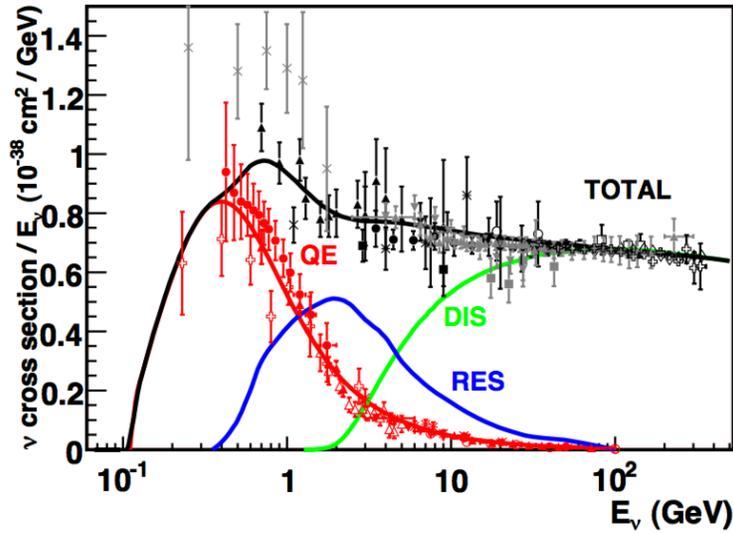


FIGURE 1.2. Plot showing the the total  $\nu_\mu$  CC cross section as a function of neutrino energy as well as the estimated cross section curves for quasi-elastic, deep inelastic scattering, and resonance[2]. Note that resonant and DIS curves do not have measurements associated with them yet.

techniques are developed for identification, such as Convolutional Neural Nets, its important to cross check with previous, well established techniques. Any identification technique developed for this analysis can be used for future analyses as a cross-check or as a tool usable for those analyses.

## 1.2. ADDITIONAL RELATED $\text{NO}\nu\text{A}$ EFFORT

In addition to the pion measurement, I contributed to other  $\text{NO}\nu\text{A}$  work. I took part in production campaigns where simulated events were generated for use by the collaboration. I generated Monte Carlo simulated data and processed them for analyzers. This also includes the production of additional datasets such as systematically shifted samples. I also assisted in producing a small dataset of simulation for analyzers to validate their code and the computing infrastructure before the implementation of software updates.

I also significantly overhauled the simulation production code to make it more user friendly so that any member of the  $\text{NO}\nu\text{A}$  Collaboration could create their own bespoke simulation with ease. This included writing documentation describing how to create simulation using the production software .

The  $\text{NO}\nu\text{A}$  experiment received a grant from Amazon to test Amazon Web Services (AWS)<sup>1</sup> for production capabilities. Working with the FNAL Computing division, two datasets, one a simulated sample and the other real data, were transferred to their servers. These datasets required reprocessing to use newer PIDs and be processed into analysis files in a single job. This was an ideal use for the AWS grant because this would yield minimal data transfer charges. The processing chain was successfully run on AWS and the output files were transferred back to the FNAL computers. While further usage of AWS services was not pursued, our test was successful.

Finally, part of my duty as a  $\text{NO}\nu\text{A}$  collaboration member was to monitor the detector during operation shifts with other  $\text{NO}\nu\text{A}$  collaboration members. Initially, these shifts were

---

<sup>1</sup>A list of all acronyms used in this thesis can be found in the appendices.

performed at FNAL using their Remote Operations Center (ROC) which collected information from the detectors and beamline into one location. I helped design, setup, and operate a local computer system, CSU-ROC, which allowed remote shift-taking at CSU, reducing the need for travel to FNAL. This facility is also used by other other groups at CSU to monitor their experiments at FNAL.

## CHAPTER 2

# A BRIEF NEUTRINO PHYSICS PRIMER

### 2.1. THE STANDARD MODEL

The Standard Model (SM) describes the elementary particles that make up the matter of the universe and the forces through which they interact[3]. It was developed by a combination of theoretical frameworks and empirical data.

2.1.1. MATTER PARTICLES IN THE STANDARD MODEL. The particles described in the SM come in two types: matter particles and force mediators. Matter particles are fermions, with spin- $\frac{1}{2}$ , and can split into two groups, quarks and leptons, with each group containing three families. Quarks are typically bound in either three quark states, called baryons, or quark-antiquark pairs, called mesons. Figure 2.1 shows the SM and describes all known subatomic particles. The particles shown in purple are quarks, the particles shown in green are leptons, the particles shown in red are the force mediators, and shown in yellow is the Higgs Boson. In the first family of quarks is the up and down quarks, with fractional electric charges  $+\frac{2}{3}$  and  $-\frac{1}{3}$  respectively. These quarks make up most ordinary matter, with a baryon quark configuration uud yielding a proton, and udd yielding a neutron, which make up the nuclei of elemental matter. Charm and strange quarks make up the second family of quarks with same electrical charges as the up and down but with larger masses. The third family contains the top and bottom quarks, still with the same electric charges as the previous families but once again with larger masses. Each family also has anti-matter quarks. In addition to electric charge, all quarks have a color charge as well, which is the fundamental charge of Quantum Chromodynamics (QCD).

Leptons also have three families, each containing a pair: a "massive" lepton and a light neutrino partner. The first family contains the electron neutrino and the familiar electron. Unlike the quarks, the leptons have integer electric charge. The electron has electric charge  $-1$ , while the electron neutrino is electrically neutral and interacts solely through the weak force. The second family of leptons contains the muon and the muon neutrino. They have similar properties to the first family particles, but the muon has a larger mass relative to the electron. The third family contains the tau and tau neutrino. Again, the spin and charge for these remain the same as in previous families, and the mass of the tau is larger than both the muon and electron.

2.1.2. FORCES IN THE STANDARD MODEL. Three forces are described by the Standard Model: the electromagnetic force, the strong nuclear force, and the weak nuclear force. Each of these forces are mediated by a vector boson of spin 1. The electromagnetic force is mediated by the photon particle, denoted as  $\gamma$ . Any particle with an electric charge, denoted by a  $+$  or  $-$ , can interact through the electromagnetic force. This is the same electromagnetic force experienced at the macro scale, such as two magnets repelling, or two electrically charged objects following Coulomb's Law. The top left pane of Figure 2.2 shows an example diagram of electron scattering interaction mediated by the photon. The strong force is mediated by the gluon particle, which unlike the photon, carries the charge of QCD. This leads to a more complicated theory than Quantum Electrodynamics (QED). Any particle with a color charge may interact through the strong force. The charge in the strong force is called color and can be one of three possible values: red, green, and blue, or their corresponding anti-charge: anti-red, anti-green, and anti-blue. This color charge is not related the color seen in the visible spectrum, and the assigned names are arbitrary.

Figure 2.2 shows an example of quark scattering mediated by the strong force in the top right pane. The third force is the weak force, this is the force that mediates the decay of particles. Both hadrons and leptons can interact through the weak force. This force mediated by the  $W^\pm$  and  $Z^0$  bosons. A unique property of the weak force is the ability to change flavor, or type, of particles in interactions involving the  $W^\pm$  boson. For example, a muon neutrino that interacts with a quark and exchanges a  $W^\pm$  boson, will transform in a muon, or in the quark sector, a down quark can decay, through the weak force, into an up quark, such as in neutron decay. The bottom pane of Figure 2.2 shows an example of muon decay mediated by the weak force. The muon is converted into a muon neutrino and the  $W^-$  produces an electron-anti-electron pair. The last boson is the Higgs boson, a massive scalar boson, which is electrically neutral, has spin 0, no color charge, and couples to mass.

## 2.2. THE HISTORY OF THE NEUTRINO

A massless and chargeless particle was first hypothesized by Wolfgang Pauli to explain the missing energy and momentum from  $\beta$ -decay[4],  $n \rightarrow e^- + p + \bar{\nu}_e$ , which is the decay of a neutron in an atomic nucleus. As a two body decay, scientists could measure the energy of two resultant particles: the proton and the electron. However, the sum of the energy of those two particles is less than the initial energy of the neutron, seemingly violating energy conservation with “missing” energy. Pauli suggested that perhaps there was a third particle that carried off the remaining energy, which was chargeless and massless. He further suggested that these properties would make the hypothetical particle difficult to detect. Fermi dubbed this mystery particle the “neutrino”, meaning “little neutral one”[5]. The neutrino was eventually detected in 1956 by Clyde Cowan and Frederick Reines[6], through an inverse  $\beta$ -decay process:  $\bar{\nu}_e + p \rightarrow n + e^+$ , the capture of an anti-neutrino by an atomic

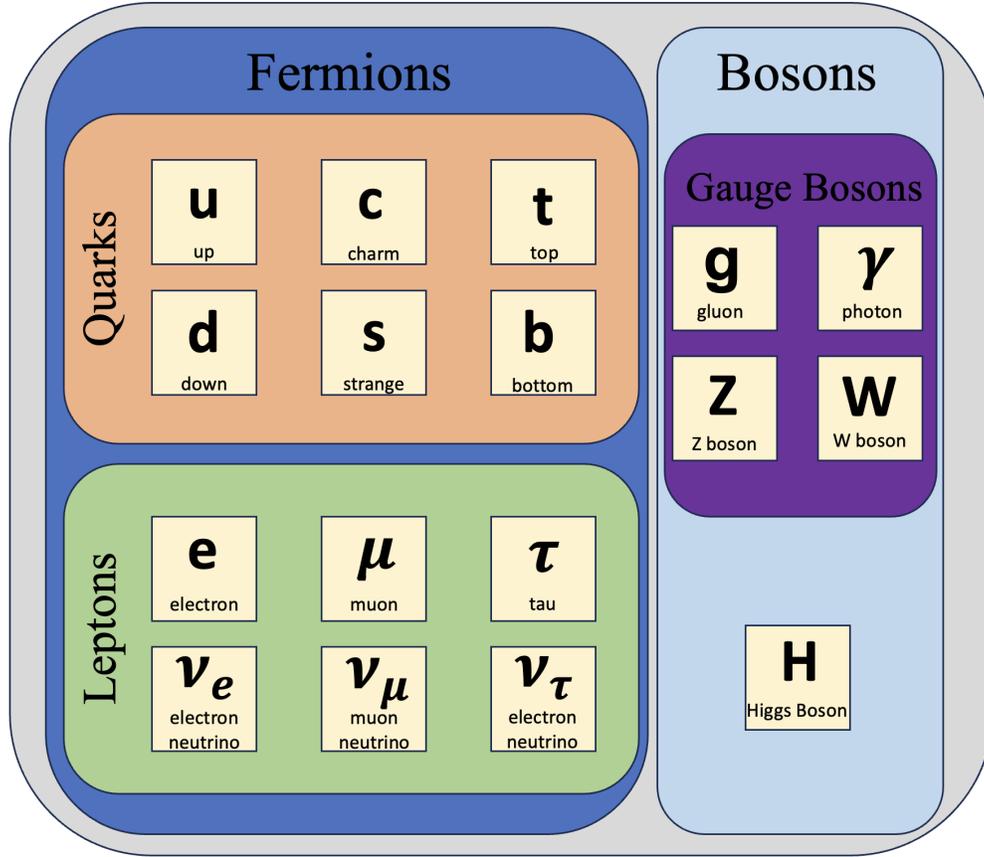


FIGURE 2.1. Standard Model of Particle Physics. This diagram describes all currently known subatomic particles and places them in groups: lepton, quarks, and bosons.

proton. Since the discovery of the neutrino, there have been many experiments to detect and understand the three types of neutrinos. Neutrinos are currently believed to be the only elementary particle to interact through the weak nuclear force alone.

### 2.3. NEUTRINO OSCILLATIONS

The Homestake experiment by Raymond Davis in the 1960's resulted in the finding that the number of neutrinos from the sun was less than predicted based on the expected flux from the internal nuclear fusion of the sun[7]. This unexpected deficit became known as the “solar neutrino problem”. In 1968 Bruno Pontecorvo put forward his theory that neutrinos could potentially oscillate between flavor states if they had mass and corresponding

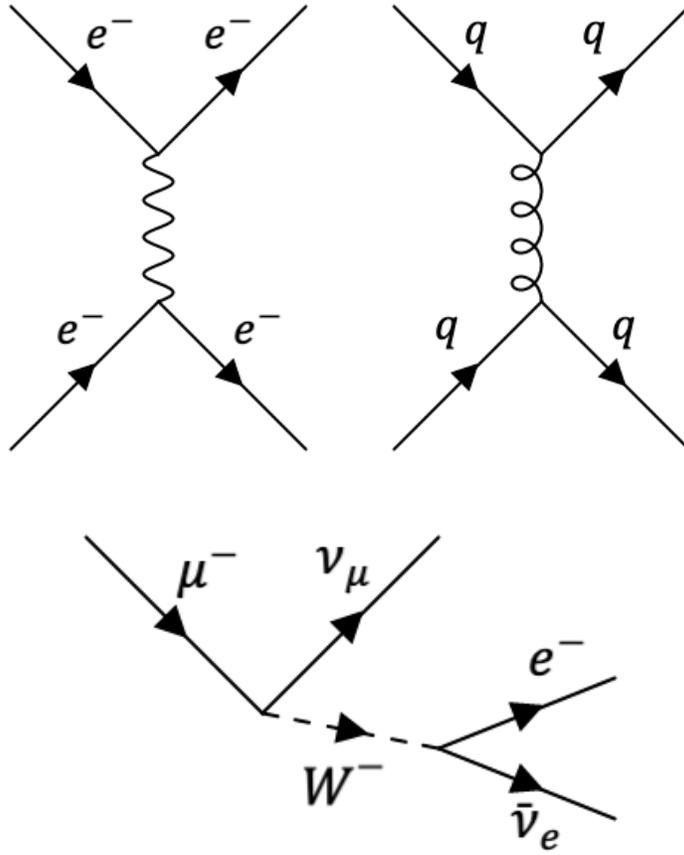


FIGURE 2.2. Simple diagrams of the three forces in the standard model. The top-left shows two electrons scattering mediated by a photon. The top-right shows two quarks scattering such as might occur in a proton-proton collider, mediated by a gluon. The bottom shows the decay of a muon into a muon neutrino as well as an electron and anti-electron, this being mediated by the  $W^-$  boson.

mass states. Since this was before the discovery of the third neutral lepton,  $\nu_\tau$ , the theory only incorporated a two flavor basis using a 2D rotational matrix. This was similar to the previously observed  $K^0 \rightarrow \bar{K}^0$  oscillation in the quark sector. After the discovery of the third lepton family, Pontecorvo's theory had to be reworked as 3-dimensional rotation

matrix, called the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix:

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} c_{13}c_{12} & c_{13}s_{12} & s_{13}e^{-i\delta_{cp}} \\ -c_{23}s_{12} - s_{13}s_{23}c_{12}e^{i\delta_{cp}} & c_{23}c_{12} - s_{13}s_{23}s_{12}e^{i\delta_{cp}} & c_{13}s_{23} \\ s_{23}s_{12} - s_{13}c_{23}c_{12}e^{i\delta_{cp}} & -s_{23}c_{12} - s_{13}c_{23}s_{12}e^{i\delta_{cp}} & c_{13}c_{23} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}$$

Note that  $c$  and  $s$  represent cosine and sine respectively, while the indices represent which angle, e.g.  $c_{13}$  means  $\cos(\theta_{13})$ . The parameter  $\delta_{cp}$  is a phase angle used to describe charge-parity violation in the lepton sector. This parameter is important because it may help to explain the matter-anti-matter asymmetry in the universe. A value of 0 means that there is no charge-parity violation and if a non-zero value is eventually measured for this parameter it would indicate charge-parity (CP) violation. The  $\delta_{cp}$  is important because a non-zero result may help to explain the matter-anti-matter asymmetry in the universe. The value of this parameter is currently unknown, but is being studied. T2K[8] and NO $\nu$ A [9] have provided initial studies. There is a similar parameter in the hadron sector describing charge-parity violation in quark mixing, which was found to be non-zero, but too small to explain the matter-anti-matter imbalance. Using this matrix allows the prediction of the probability of flavor change from one family to the other. For example the probability of a muon staying a muon would be:

$$P(\mu \rightarrow \mu) \approx 1 - \sin^2(2\theta_{32})\sin^2\left(\frac{\Delta m_{32}^2 L}{4E}\right)$$

The last product in this equation comes from treating the propagation of the neutrino as a plane wave and using the relativistic approximation  $p \gg m$ . This equation depends on the mass splitting,  $\Delta m_{32}^2$  as well as the distance that the neutrino travelled,  $L$ , and the energy of the neutrino,  $E$ . From this equation, it becomes apparent that for an experiment to detect

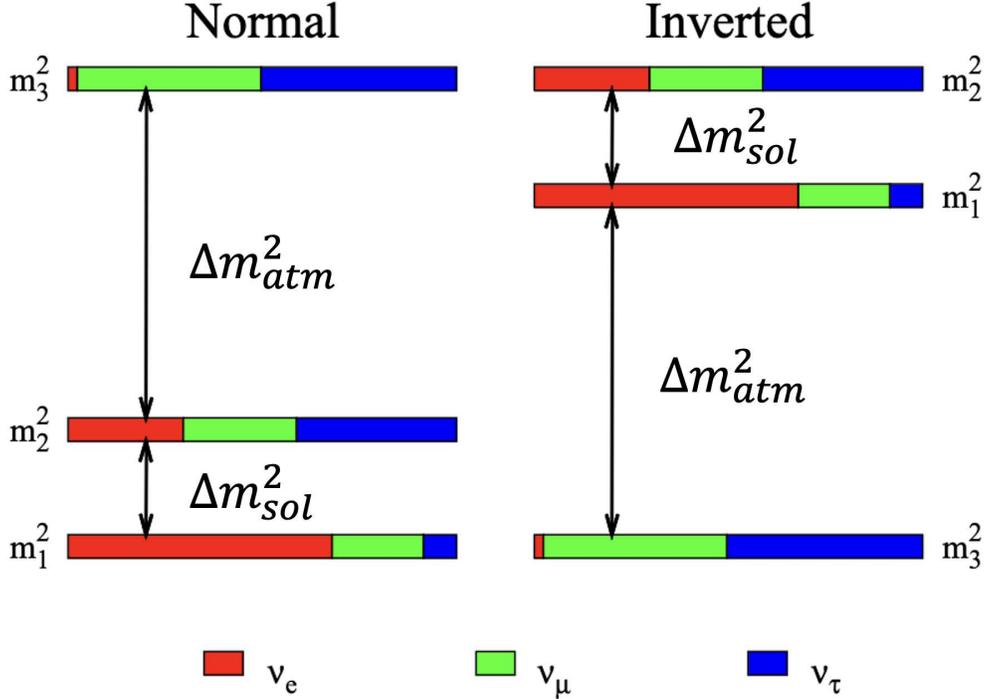


FIGURE 2.3. Diagram illustrating the two different possibilities of the mass ordering, with the normal hierarchy on the left and inverted hierarchy on the right.[10]

this oscillation, it would need to use an appropriate  $L/E$  to yield maximal probability. Figure 2.3 shows the two possible neutrino mass hierarchies. While the three mass differences and  $\Delta m_{32}^2$  have been measured to some significance, as shown in Figure 2.4, the hierarchy of the splitting has not yet been determined, though current data suggests normal hierarchy is preferred. Long-baseline oscillation experiments use two detectors, a near detector close to the neutrino beam source used to measure interactions and study the beam, and a far detector that is used to measure the oscillated neutrinos.

#### 2.4. NEUTRINO-NUCLEON INTERACTIONS AND PION PRODUCTION

Due to the neutrino being electrically neutral, it is not directly detectable. Specifically, what is detected in neutrino experiments is the charged particles resulting from the neutrino interactions. The neutrino will interact either with a nucleon inside an atomic nucleus, the

		Normal Ordering (best fit)		Inverted Ordering ( $\Delta\chi^2 = 2.3$ )	
		bfp $\pm 1\sigma$	$3\sigma$ range	bfp $\pm 1\sigma$	$3\sigma$ range
without SK atmospheric data	$\sin^2 \theta_{12}$	$0.303^{+0.012}_{-0.011}$	$0.270 \rightarrow 0.341$	$0.303^{+0.012}_{-0.011}$	$0.270 \rightarrow 0.341$
	$\theta_{12}/^\circ$	$33.41^{+0.75}_{-0.72}$	$31.31 \rightarrow 35.74$	$33.41^{+0.75}_{-0.72}$	$31.31 \rightarrow 35.74$
	$\sin^2 \theta_{23}$	$0.572^{+0.018}_{-0.023}$	$0.406 \rightarrow 0.620$	$0.578^{+0.016}_{-0.021}$	$0.412 \rightarrow 0.623$
	$\theta_{23}/^\circ$	$49.1^{+1.0}_{-1.3}$	$39.6 \rightarrow 51.9$	$49.5^{+0.9}_{-1.2}$	$39.9 \rightarrow 52.1$
	$\sin^2 \theta_{13}$	$0.02203^{+0.00056}_{-0.00059}$	$0.02029 \rightarrow 0.02391$	$0.02219^{+0.00060}_{-0.00057}$	$0.02047 \rightarrow 0.02396$
	$\theta_{13}/^\circ$	$8.54^{+0.11}_{-0.12}$	$8.19 \rightarrow 8.89$	$8.57^{+0.12}_{-0.11}$	$8.23 \rightarrow 8.90$
	$\delta_{\text{CP}}/^\circ$	$197^{+42}_{-25}$	$108 \rightarrow 404$	$286^{+27}_{-32}$	$192 \rightarrow 360$
	$\frac{\Delta m_{21}^2}{10^{-5} \text{ eV}^2}$	$7.41^{+0.21}_{-0.20}$	$6.82 \rightarrow 8.03$	$7.41^{+0.21}_{-0.20}$	$6.82 \rightarrow 8.03$
	$\frac{\Delta m_{3\ell}^2}{10^{-3} \text{ eV}^2}$	$+2.511^{+0.028}_{-0.027}$	$+2.428 \rightarrow +2.597$	$-2.498^{+0.032}_{-0.025}$	$-2.581 \rightarrow -2.408$
	with SK atmospheric data	$\sin^2 \theta_{12}$	$0.303^{+0.012}_{-0.012}$	$0.270 \rightarrow 0.341$	$0.303^{+0.012}_{-0.011}$
$\theta_{12}/^\circ$		$33.41^{+0.75}_{-0.72}$	$31.31 \rightarrow 35.74$	$33.41^{+0.75}_{-0.72}$	$31.31 \rightarrow 35.74$
$\sin^2 \theta_{23}$		$0.451^{+0.019}_{-0.016}$	$0.408 \rightarrow 0.603$	$0.569^{+0.016}_{-0.021}$	$0.412 \rightarrow 0.613$
$\theta_{23}/^\circ$		$42.2^{+1.1}_{-0.9}$	$39.7 \rightarrow 51.0$	$49.0^{+1.0}_{-1.2}$	$39.9 \rightarrow 51.5$
$\sin^2 \theta_{13}$		$0.02225^{+0.00056}_{-0.00059}$	$0.02052 \rightarrow 0.02398$	$0.02223^{+0.00058}_{-0.00058}$	$0.02048 \rightarrow 0.02416$
$\theta_{13}/^\circ$		$8.58^{+0.11}_{-0.11}$	$8.23 \rightarrow 8.91$	$8.57^{+0.11}_{-0.11}$	$8.23 \rightarrow 8.94$
$\delta_{\text{CP}}/^\circ$		$232^{+36}_{-26}$	$144 \rightarrow 350$	$276^{+22}_{-29}$	$194 \rightarrow 344$
$\frac{\Delta m_{21}^2}{10^{-5} \text{ eV}^2}$		$7.41^{+0.21}_{-0.20}$	$6.82 \rightarrow 8.03$	$7.41^{+0.21}_{-0.20}$	$6.82 \rightarrow 8.03$
$\frac{\Delta m_{3\ell}^2}{10^{-3} \text{ eV}^2}$		$+2.507^{+0.026}_{-0.027}$	$+2.427 \rightarrow +2.590$	$-2.486^{+0.025}_{-0.028}$	$-2.570 \rightarrow -2.406$

FIGURE 2.4. Best fit neutrino oscillation parameters for both normal and inverted ordering.[11]

entire nucleus, or the atomic electron. The resulting particles from the interaction can have electric charge, which allows them to be visible within detectors.

During an interaction, the neutrino will exchange a boson, which mediates the interaction, with the target particle. Two different bosons can mediate the interaction: the electrically neutral Z-boson ( $Z^0$ ) or the charged W-boson ( $W^\pm$ ). When a  $Z^0$  is exchanged, it is defined to be a neutral current (NC) interaction, due to no electrical charge being exchanged between the hadron and lepton. These types of interactions result with a neutrino in the final state.

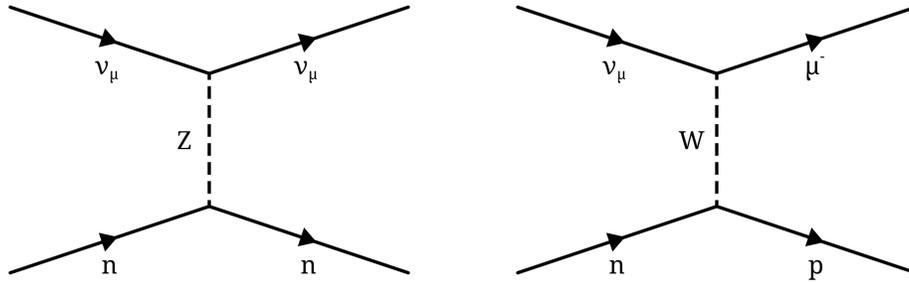


FIGURE 2.5. A simple neutral current elastic interaction on the left and an example of charged-current quasi-elastic interactions on the right.[12]

These events do not produce an observable lepton, and must be inferred by any resultant hadrons. Conversely, when a  $W^\pm$  is mediating the interaction, it is defined to be a charged-current (CC) interaction, because the charged boson exchanges an electric charge to the final state lepton.

In CC neutrino events, there are three main neutrino-nuclear interaction types that can create charged pions: resonant, coherent, and deep inelastic scattering. Resonant pion production (Res), is facilitated through a  $\Delta$  resonance.. The interaction can be described as:  $\nu_\mu + N \rightarrow \mu^- + \Delta \rightarrow \mu^- + N + \pi^{0/\pm}$ , where the nucleon ( $N$ ) can be ejected from the nucleus. An example diagram of a muon neutrino interacting on a neutron is shown in the left pane of Figure 2.6. Coherent production is when the neutrino interacts on the nucleus as a whole instead of specific nucleon, and when the nucleus de-excites a pion is emitted. An example diagram of the coherent process is shown in the right pane of Figure 2.6. Deep inelastic scattering (DIS) is when the neutrino interacts with a nucleon deep within the nucleus, this can cause nuclear breakup with many nucleons and hadrons being ejected. Though pion production is not guaranteed through DIS, it often occurs. This is not an exhaustive list of all pion production processes, but instead the most common types of interactions to occur in the NO $\nu$ A detectors. The relative probabilities of these processes was shown previously

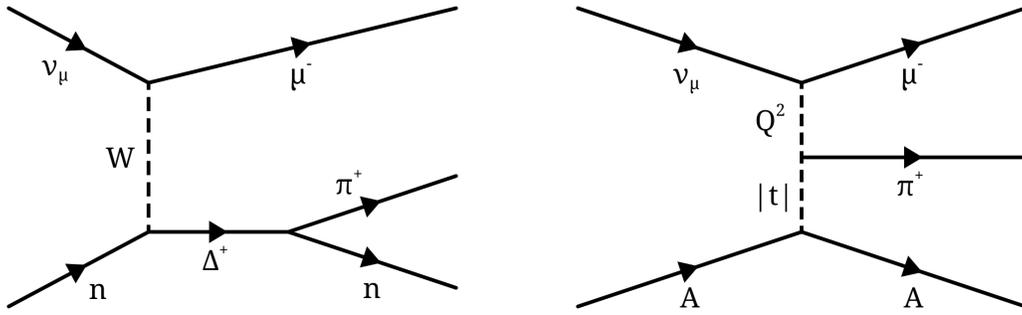


FIGURE 2.6. Example diagrams of pion production processes with  $\Delta^+$  resonance on the left and coherent on the right.[12]

in Figure 1.2. There are multiple models for describing the production of pion in neutrino interactions. For this measurement, the simulation used the Rein-Sehgal model[13], which treats the nucleon as a bound state of 3 quarks.

## 2.5. PREVIOUS MEASUREMENTS OF INCLUSIVE PION PRODUCTION CROSS SECTIONS

Many measurements of pion production in neutrino events have been made; however, only the most recent and relevant measurements will be discussed. These recent measurements have been made by: MINERvA[14], MiniBooNE[15], and T2K[16] collaborations. The MINERvA collaboration measured a cross section of multi-pion production on hydrocarbon as a function of pion energy and pion angle with respect to the beam, with their results shown in Figure 2.7. The MiniBooNE collaboration measured a multi-pion production cross section on mineral oil, measured as a function of both energy and angle for both the muon and the pion. Their results are shown in Figure 2.8. The T2K collaboration performed a single-pion cross section measurement on water, measured as a function of both muon and pion kinematics. Their results can be seen in Figure 2.9.

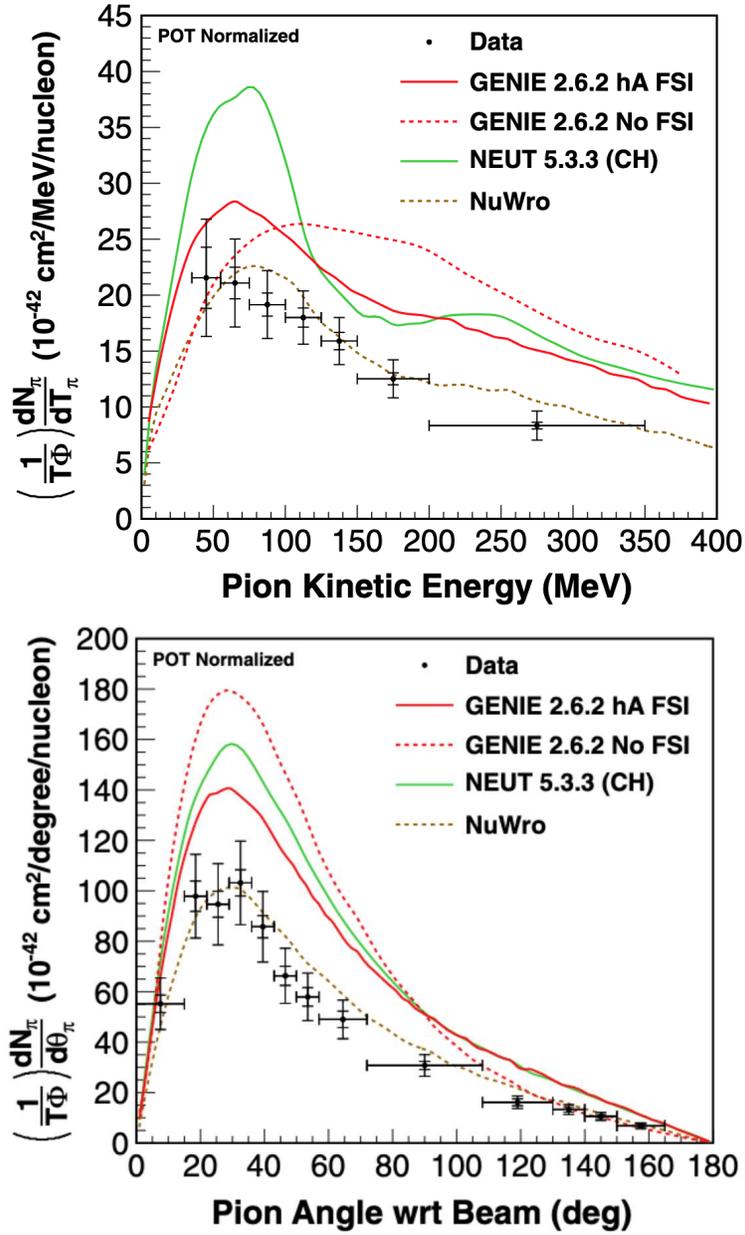


FIGURE 2.7. The measured differential cross section with respect to pion kinetic energy on top and the measured differential cross section with respect to pion angle on bottom from MINERvA.[14]

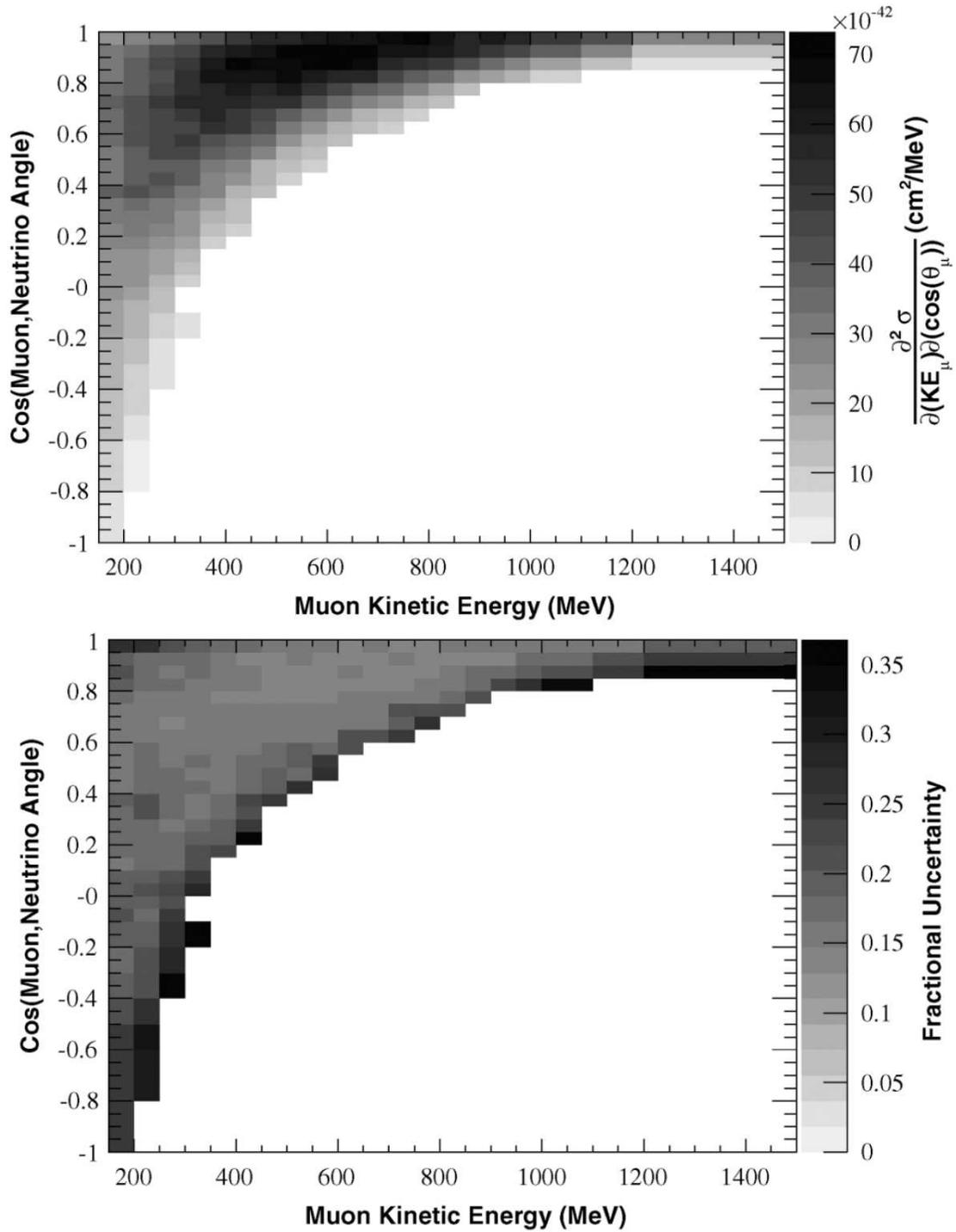


FIGURE 2.8. The measured double differential cross section on top and the estimated fractional uncertainty on the measurement on bottom from MiniBooNE.[15]

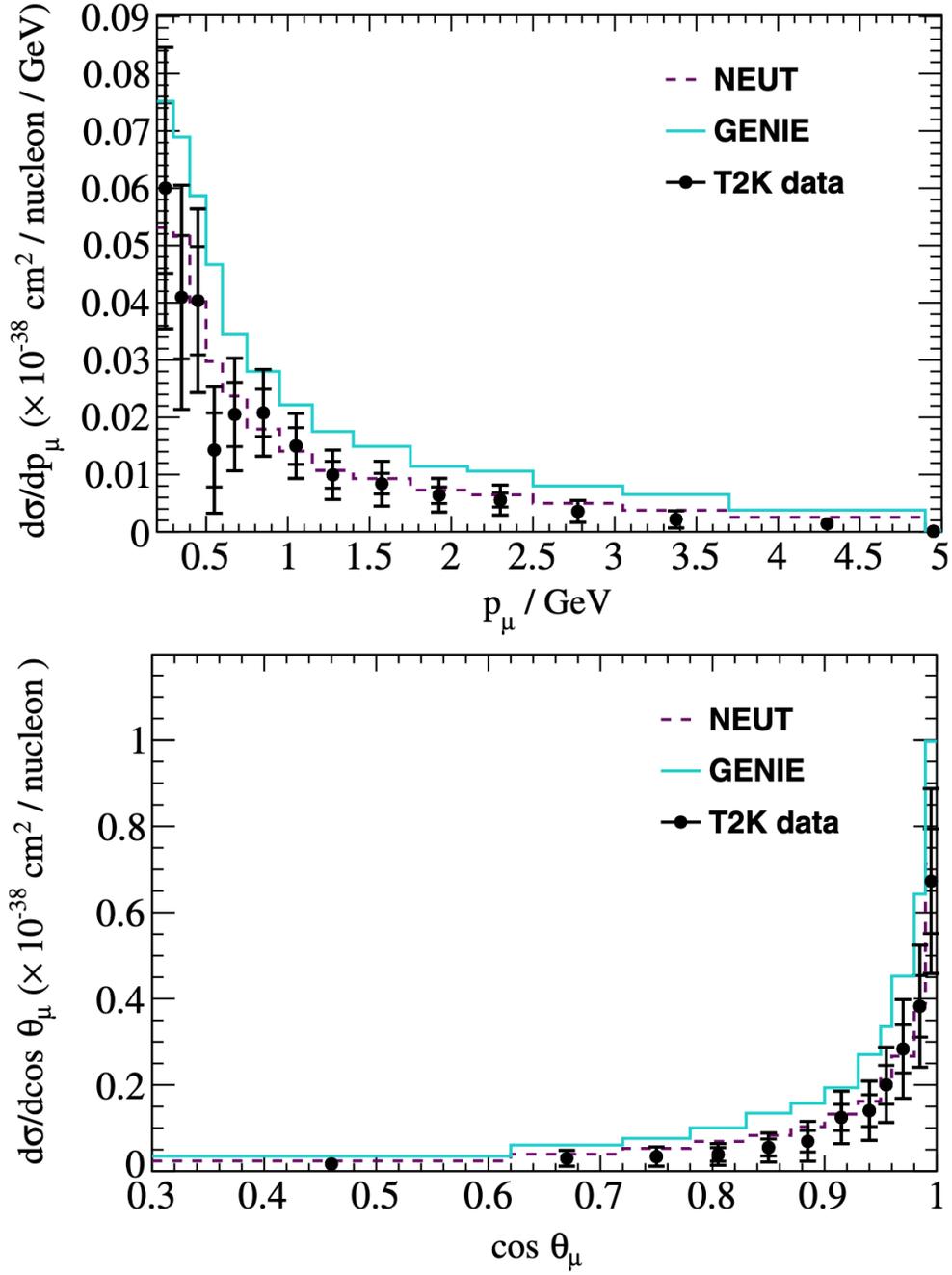


FIGURE 2.9. The measured differential cross section in muon momentum on top and the differential cross section in  $\cos\theta_\mu$  on the bottom from T2K.[16]

## CHAPTER 3

# THE NO $\nu$ A EXPERIMENT

The NuMI Off-Axis  $\nu_e$  Appearance (NO $\nu$ A) experiment is a long-baseline neutrino oscillation experiment hosted by Fermilab in Batavia, Illinois. The primary goal of NO $\nu$ A is to measure the neutrino oscillation parameters from the PMNS matrix, such as  $\theta_{23}$ ,  $\Delta m_{32}^2$ , and attempt to measure  $\delta_{cp}$  and determine the neutrino mass hierarchy, whether it is normal or inverted. To achieve these goals, NO $\nu$ A utilizes a neutrino beam generated at Fermilab, which is aimed through the earth, and oscillated neutrinos are measured with its detector. NO $\nu$ A consists of two detectors, a near detector located on-site at Fermilab, 800 m from the beam source, which is used to measure the unoscillated beam, and a far detector in Ash River, Minnesota, 810 km from the beam source, which is used to measure the oscillated beam. NO $\nu$ A started operation in 2014 and is expected to continue operating until 2028.

### 3.1. FERMILAB ACCELERATOR COMPLEX

The Fermilab Accelerator Complex, illustrated in Figure 3.1, creates all of the particle beams used by experiments hosted at Fermilab. The complex has four accelerators: the Linac, the Booster, the Recycler Ring, and the Main Injector, each one feeding the next to generate high-energy protons. The beam originates at the ion source, which is ionized hydrogen (electrons and protons). The protons travel through the Linear Accelerator (Linac) which accelerates them up to 0.4 GeV. Following the Linac, the ions move to the Booster ring, further accelerating the ions to 8 GeV. Before entering the next accelerator, the hydrogen ions impinge on carbon foil, with only protons passing through. These protons enter the

## Fermilab Accelerator Complex

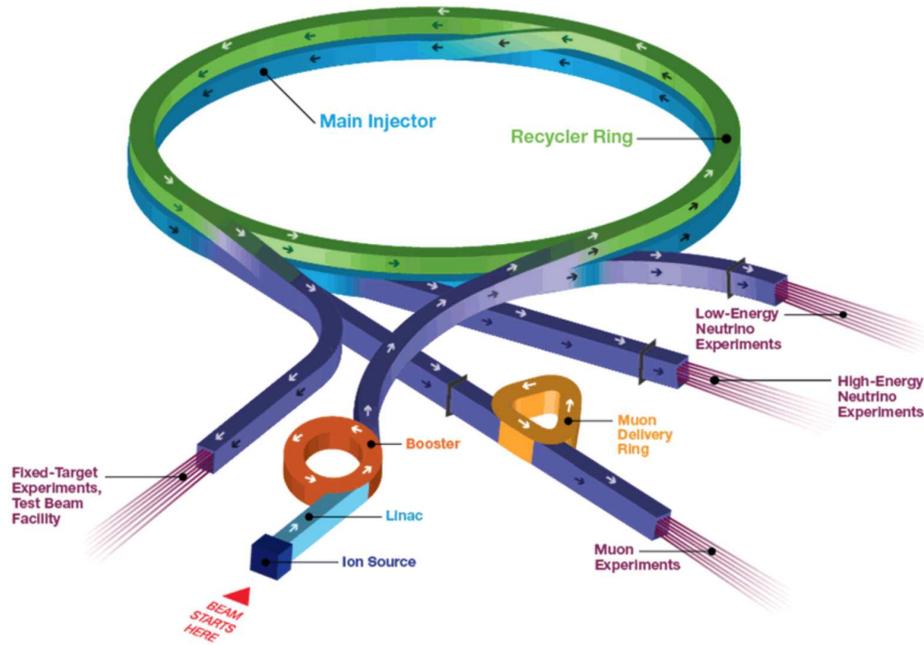


FIGURE 3.1. A diagram of the Fermilab Accelerator Complex.[17]

Recycler Ring, which bunches the protons together into groups. The bunched protons are then fed into the Main Injector and accelerated to 120 GeV.

### 3.2. NUMI BEAMLINE

Neutrinos at the Main Injector (NuMI) is the beamline used by  $\text{NO}\nu\text{A}$ , as well as other FermiLab based experiments. A schematic of the NuMI beamline is shown in Figure 3.2. Accelerated protons from the Main Injector are smashed into a carbon target generating mesons, primarily pions and kaons. The number of protons that interacting with the target is called the “protons on target” (POT) and allows for data comparisons between experiments. These mesons are directed by a electromagnetic focusing horn, which collimates the beam and directs wrong sign particles out of the beam. The direction of electric current in the horns can be reversed, allowing for opposite signed mesons to be selected. After the horn is a decay

pipe where the mesons decay, primarily via the  $\pi^- \rightarrow \mu^- + \nu_\mu$  or  $K^- \rightarrow \mu^- + \nu_\mu$  decay channel, producing the muon neutrinos used by the experiment, along with background neutrinos. This configuration is called Forward Horn Current (FHC), referring to the direction of the electric current in the horns. The background neutrinos produced by the beam in the FHC configuration are electron neutrinos, and both muon and electron anti-neutrinos. The electron neutrinos produced by the beam are indistinguishable from the oscillated electron neutrinos and interactions produced by anti-neutrinos can be difficult to distinguish from those produced by normal neutrinos as the NO $\nu$ A near detector is not magnetized. The other beam configuration is Reverse Horn Current (RHC), where the current in the horns runs in the opposite direction relative to the current direction in the FHC configuration. In the RHC configuration the most likely decay channels in the decay pipe are  $\pi^+ \rightarrow \mu^+ + \bar{\nu}_\mu$  and  $K^+ \rightarrow \mu^+ + \bar{\nu}_\mu$ . The energy of the produced neutrino is determined by the following kinematic equation, relating energy with the scattering angle:

$$(1) \quad E_\nu = \frac{0.43E_\pi}{1 + \gamma^2\theta^2},$$

where  $E_\nu$  is the energy of the produced neutrino,  $E_\pi$  is the energy of the parent pion,  $\theta$  is the angle between the decaying pion direction and the resulting neutrino direction, and  $\gamma$  is the Lorentz factor, which is given by  $\gamma = E_\pi/m_\pi$ , where  $m_\pi$  is the mass of the pion. The simulated neutrino energy spectra for NO $\nu$ A is shown in Figure 3.3. Downstream from the decay pipe is a beam absorber which is used to reduce muon contamination in the beam. This is followed by rock layered between the Muon Monitors, which are used to monitor the muons in the beam. The protons are arranged into bunches with twelve bunches being delivered every 500 ms. The produced neutrinos have energies ranging from 1 to 3 GeV while

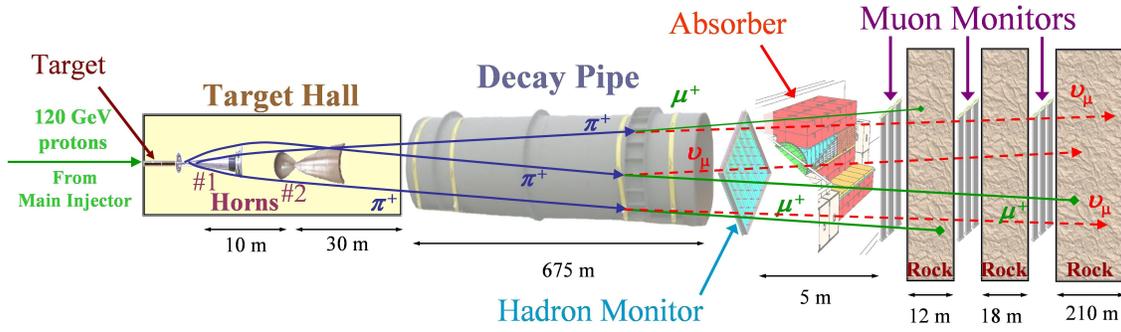


FIGURE 3.2. A schematic of the Beamline past the Main Injector.[18]

a small fraction exist outside of these range. There is some background contamination in the  $\text{NO}\nu\text{A}$  detectors from the Booster beam, which is used in the short baseline experiments at Fermilab. The cumulative POT as a function of time, as of 2020, can be seen in Fig 3.4.

### 3.3. $\text{NO}\nu\text{A}$ FAR DETECTOR

The  $\text{NO}\nu\text{A}$  Far Detector (FD), shown in Figure 3.5, is located in Ash River, Minnesota, 810 km from the NuMI beam source. The FD consists of alternating planes of extruded Poly-Vinyl-Chloride (PVC) pipes,  $4\text{ cm} \times 6\text{ cm} \times 16\text{ m}$  in size, containing liquid scintillator. The far detector is located  $14.6\text{ mrad}$  off axis relative to the center of the NuMI beam to narrow the energy spectrum of the neutrino centered at 2 GeV and to reduce high energy backgrounds to the  $\nu_e$  oscillation signal. This, along with the distance from the beam source, provide an  $L/E$ , previously described in Chapter 2.3, that maximizes the  $\nu_e$  appearance probability in the far detector. The FD is above ground, which means that background contamination from cosmic muons in the detector is quite high: 150kHz.

### 3.4. $\text{NO}\nu\text{A}$ NEAR DETECTOR

The  $\text{NO}\nu\text{A}$  Near Detector (ND), shown in Figure 3.6, is located 800 m downstream of the beam target in the Near Detector Hall at Fermilab. The ND is also located  $14.6\text{ mrad}$

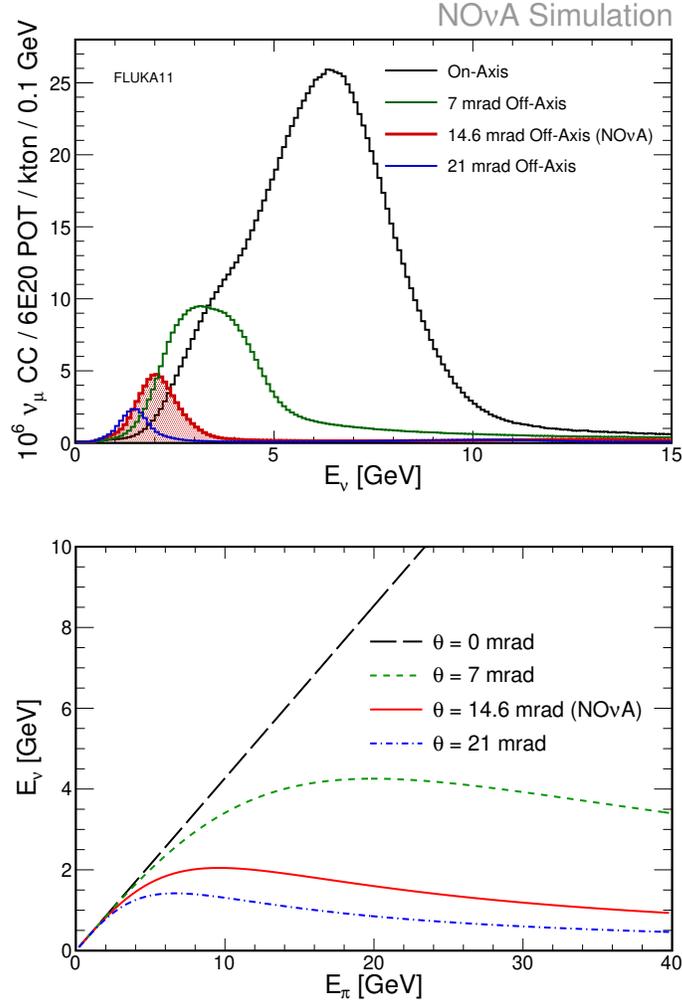


FIGURE 3.3. TOP: The simulated neutrino energy spectra at various angles from the beam direction.[19] BOTTOM: The simulated neutrino energy as a function parent pion energy, with the various possible off-axis angles. Each angle refers to a possible site for the Far Detector during the design phase.[19]

off-axis from the NuMI beam axis. The ND has cross-sectional area of  $4.2 \text{ m} \times 4.2 \text{ m}$  and is 14.3 m in length, while weighing about 300 tons. It is functionally identical to the FD; however, located downstream of the active scintillator region is a non-active steel region called the Muon Catcher, which is used to contain long tracks. The ND is composed of 80 planes of alternating horizontal and vertical PVC cells filled with scintillating liquid and a wavelength shifting fiber. The purpose of the near detector is to allow the observation and characterization of the unoscillated off-axis NuMI beam and  $\nu - N$  interactions. Though

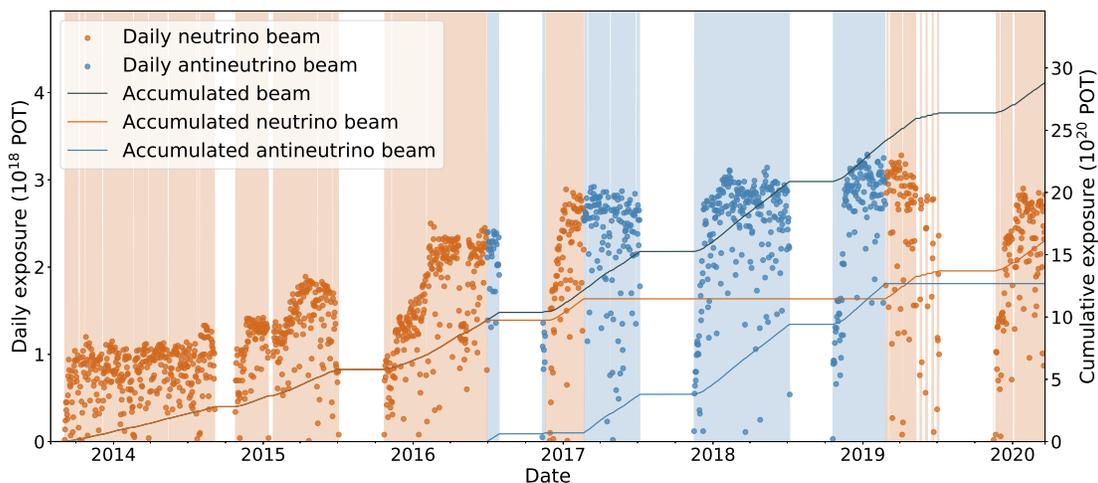


FIGURE 3.4. The daily and cumulative exposure, in POT.[19]



FIGURE 3.5. The fully constructed Far Detector in the Ash River facility. Note the person on the third floor catwalk for a sense of scale.[19]

the ND is considerably smaller than the FD, it is subjected to high neutrino fluxes due to its close position to the NuMI beam source, which enables cross section measurements. The



FIGURE 3.6. A picture of the front face of the Near Detector inside the Near Detector Hall at Fermilab. The beam direction is into the visible detector face.[19]

ND is in Near Detector Hall at Fermilab, which is underground, therefore the contamination from cosmic rays is relatively low compared to the FD.

### 3.5. THE $\text{NO}\nu\text{A}$ DETECTOR UNIT CELL

As stated above, both detectors utilize the same components and technologies in their construction. The smallest active component is called the unit cell. The unit cell is an extruded PVC plastic pipe measuring 4 cm by 6 cm. The inside of the PVC pipe is coated in titanium-dioxide, which reflects light back towards the interior of the cell. Contained in the cell is a scintillating liquid and a wavelength shifting fiber used to collect the scintillated light. Both ends of the fiber terminate at an avalanche photodiode (APD) at the end of the cell. Signals from the APDs are read out and converted to digital values.

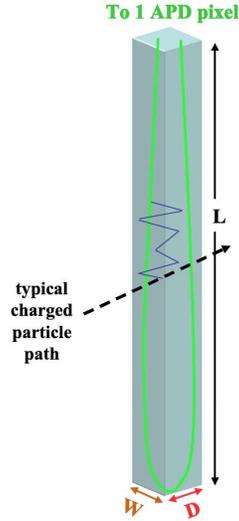


FIGURE 3.7. Diagram of the NOVA detector unit cell, filled with liquid scintillator, and a looped wave-length shifting fiber.[17]

3.5.1. LIQUID SCINTILLATOR. The liquid filling the unit cell is primarily mineral oil, acting as a solvent, mixed with pseudocumene, which facilitates the scintillation. Figure 3.8 shows all of the components comprising the scintillating liquid. Light produced by the scintillation of pseudocumene will be peaked around 360 nm - 390 nm. Additional additives, PPO [2,5-diphenyloxazole] and bis-MSB [1,4-di(methylstyryl)benzene], are used for their wavelength shifting properties to shift the light to a peak of 400 nm - 450 nm, allowing collection by the wavelength shifting fiber.

3.5.2. WAVELENGTH SHIFTING FIBER. Each unit cell contains a looped piece of wavelength shifting fiber that captures the scintillated light from the liquid and shifts the wavelength of the light to a longer wavelength. The longer wavelength, 520 nm - 550 nm, is the final step of wavelength-shifting process to occur before it exits the fiber at the end.

3.5.3. MULTI PIXEL PHOTON COUNTER. At the end of each PVC extrusion is the photodetector, which is a multi-pixel photon counter (MPPC) made up of a  $4 \times 8$  APD array. Each pixel of the MPPC is an APD that converts absorbed light into electrical

<b>component</b>	<b>purpose</b>	<b>mass fraction</b>	<b>volume (gal)</b>	<b>tot mass (kg)</b>
<b>mineral oil</b>	solvent	95.8%	3,082,145	9,917,109
<b>pseudocumene</b>	scintillant	4.1%	128,439	425,908
<b>PPO</b>	waveshifter #1	0.091%		9,373
<b>bis-MSB</b>	waveshifter #2	0.0013%		131
<b>Stadis-425</b>	antistatic agent	0.0003%		46.6
<b>tocopherol (Vit.E)</b>	antioxidant	0.0010%		104
<b>Total</b>		<b>100.0%</b>	<b>3,210,584</b>	<b>10,352,551</b>

FIGURE 3.8. Table of the composition of the liquid used in the NO $\nu$ A experiment.[17]



FIGURE 3.9. Photographs of the fiber ends in the connector and the pixel array of the MPPC. The MPPC fits into the connector lining up each pair of fiber ends from a cell to one of the 32 pixels.[19]

current. Figure 3.9 shows a photograph of the 32 fiber ends that connect to the MPPC on the left, and the 32 pixel array on the MPPC on the right. These APDs have a quantum efficiency of 85% for light between 520 nm - 550 nm. To reduce thermal noise, the APDs use thermo-electric coolers to keep their temperature at -15 C°.

3.5.4. ARRANGEMENT OF THE UNIT CELLS. The cells are arranged in planes of one unit cell thickness, with planes of vertical extrusions used to measure horizontal position

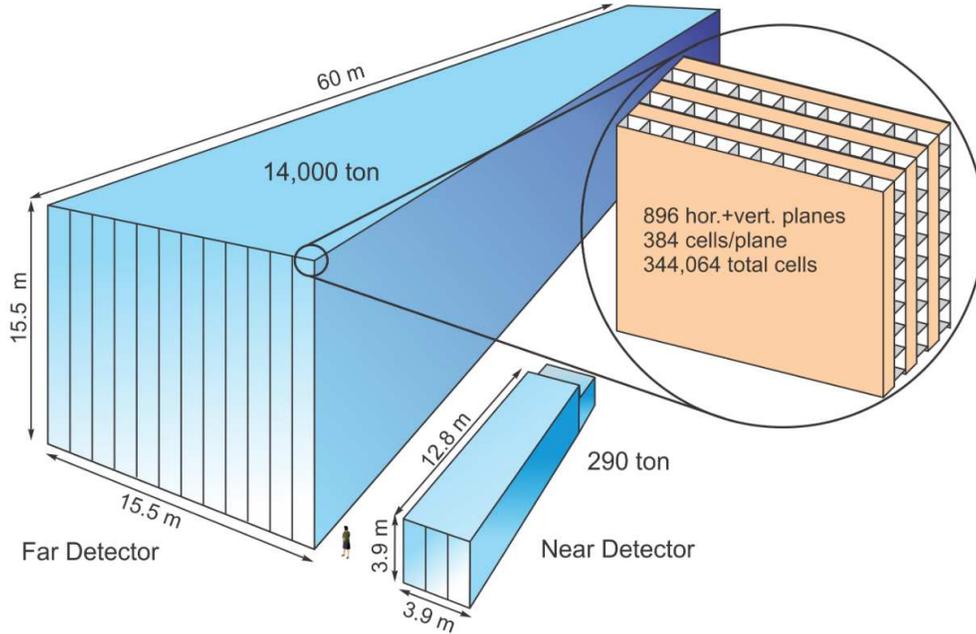


FIGURE 3.10. A diagram of the NO $\nu$ A Far Detector and the NO $\nu$ A Near Detector showing the arrangement of planes in the detectors.[17]

and planes of horizontal extrusions used to measure vertical positions. A diagram of these alternating planes is shown in Figure 3.10. These planes, called the extrusion module, are layered one after the other, in the direction of the beam to allow full three dimensional reconstruction of events. A diagram of an extrusion module is shown in Figure 3.11.

### 3.6. READOUT ELECTRONICS

There are two MPPCs per extrusion module and each MPPC pair connects to a Front End Board (FEB), which converts the current through the APDs into a digital signal. The FEBs are connected together in a gigabit ethernet network, and transferred to the data acquisition system (DAQ) which processes all of the data taken by the detector. The FEBs stream the data from the detectors, however, only data correlated to the beam spill is saved to disk. All signals within  $30 \mu\text{s}$  of the  $11 \mu\text{s}$  spill window are recorded. Additional triggers are used to record data outside of the beam spill, such as the cosmic ray trigger and exotic data-driven

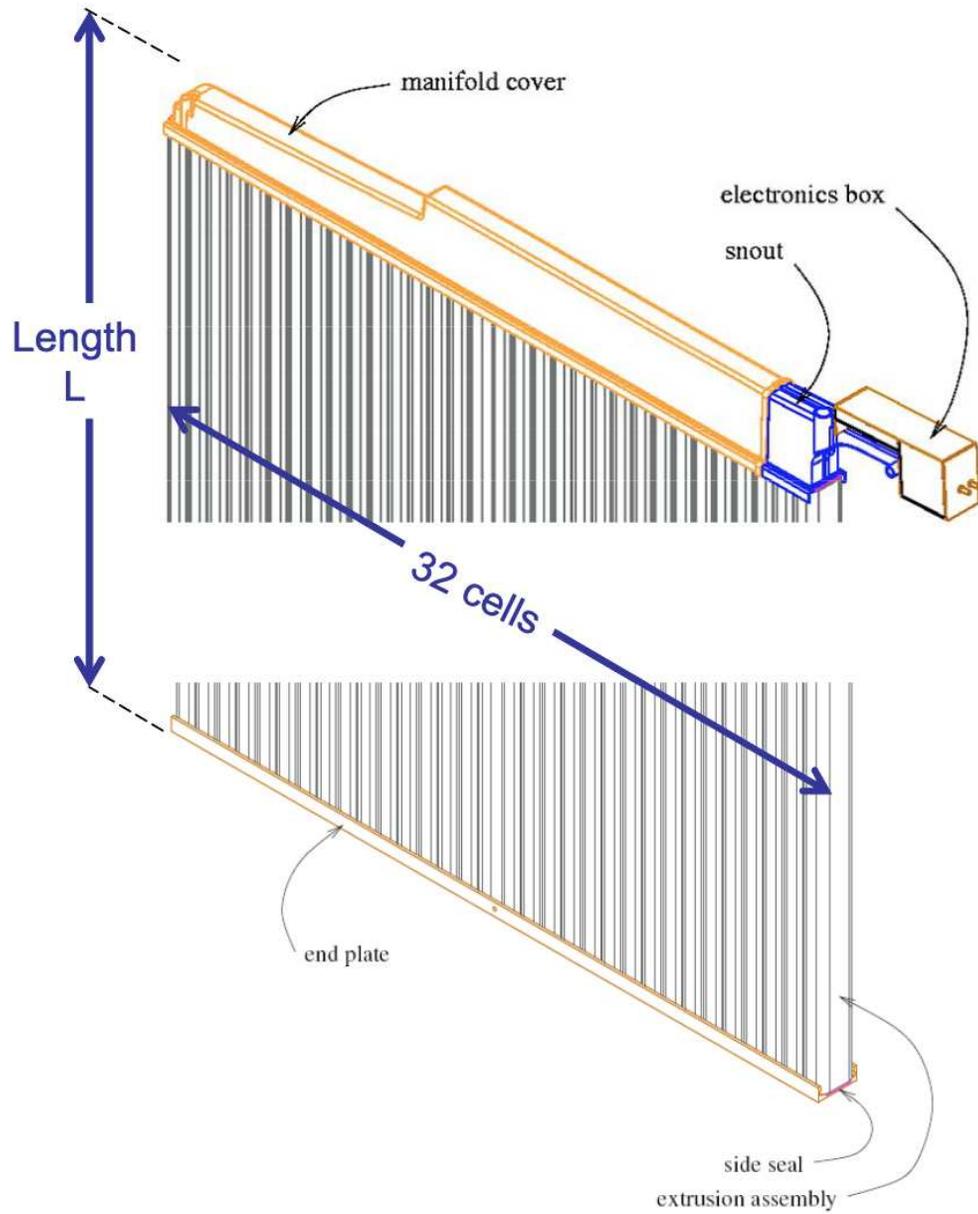


FIGURE 3.11. Shown here is a diagram of an extrusion module.[17]

triggers, such as the magnetic monopole trigger. Figure 3.12 illustrates the individual steps in the DAQ.

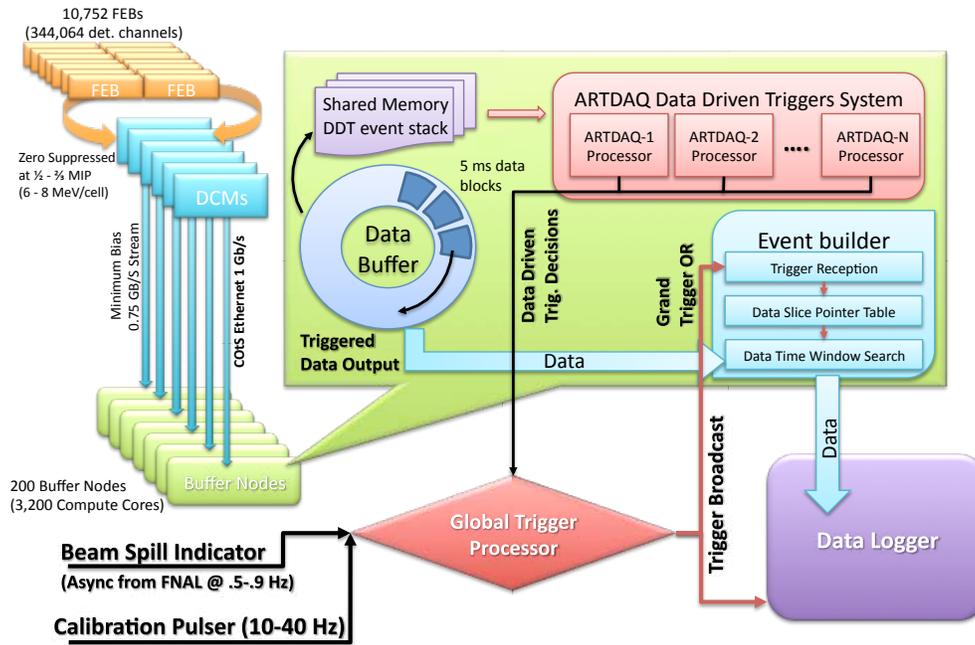


FIGURE 3.12. This schematic shows the individual steps in data acquisition.[17]

## CHAPTER 4

# NO $\nu$ A SIMULATION AND RECONSTRUCTION, AND ANALYSIS SOFTWARE

A robust software suite that includes simulation, event reconstruction, and data analysis tools has been developed for NO $\nu$ A analyses. The NuMI beam and detector electronics, discussed in Chapter 3, are simulated to generate neutrino events that allow the testing and development of analysis code prior to the use of real data. Event reconstruction is performed offline and processed after the data have been taken or after the simulation has been generated. Reconstruction is an important aspect as it processes the electronic information from the NO $\nu$ A detectors into physical parameters allowing measurements by analyzers.

### 4.1. DATA

This analysis used FHC data, described in Chapter 3, from the Near Detector, processed with the fourth version of the NO $\nu$ A software, containing  $8.09 \times 10^{20}$  POT.

### 4.2. SIMULATION

NO $\nu$ A analyzers employs the use of simulation to build the tools for each analysis. Every aspect of the NO $\nu$ A experiment is simulated. The emulate the neutrino beam, hadrons interacting on a carbon target are simulated, replicating how the real beam generates neutrinos. The neutrinos generated from the beam simulation are then passed to the simulated detector and neutrino interactions are simulated in the detector using GENIE (Generates Events for Neutrino Interaction Experiments) version 2[20].

4.2.1. NEUTRINO BEAM SIMULATION USING FLUKA AND G4NuMI. To simulate the NO $\nu$ A neutrino beam, two packages are used: FLUKA[21] and G4NuMI[22]. FLUKA simulates the hadronic interactions to replicate the NuMI beam, whereas G4NuMI is used to simulate beamline geometry and the propagation of particles through the beamline. The kinematics and characteristics of the neutrino, as well as those of their parent hadrons, are saved and recorded for next stage of the simulation, the propagation of the final-state particles through the NO $\nu$ A detectors.

4.2.2. NEUTRINO INTERACTION SIMULATION USING GENIE. With the simulated neutrinos produced, the next step is to generate the neutrino interactions themselves. GENIE is a neutrino event generator that simulates neutrino-nucleus interactions using neutrino cross section models constrained through measurements from other experiments and theory. The neutrino information from the flux simulation is convolved with the cross sections from GENIE to generate all the different neutrino interaction types producing the resultant particles from each interaction. Interactions outside of the detector are also simulated to create out-of-detector events such as neutrino interactions that occur outside of the detector but which the products enter into the detector. These kinds of interactions are important to include in the simulation as they are seen in the data. These extra-detector events are mixed with events that occurred in the detector in a process called “overlaying”. To save on computational power and time, events that occur inside the detector are simulated separately from events that occur outside the detector. Most of the particles that originate outside of the detector yet pass through it are muons; minimum ionizing particles creating long straight tracks. Most of these events look very similar to the detector so only a smaller sample of events are needed as many of them can be reused for the same effect.

4.2.3. PARTICLE PROPAGATION AND DETECTOR SIMULATION. Geant4 is also used to propagate the daughter particles from the simulated neutrino interactions produced by GENIE. Additionally, the detector is modelled and simulated, allowing simulation of the particles and their interactions with detector components.

The detector electronics response is also modeled, allowing simulation of the signals collected in the detector. The final output of the simulation is the APD response, defined in Chapter 3, which is then forwarded to the full chain reconstruction to allow analysis of the simulation for analyzers.

4.2.4. SIMULATION TUNING. While the simulation uses best prediction models, the results alone won't perfectly match the data, therefore the simulation is tuned to match data. The flux is tuned by matching the simulation to the hadronic activity measured in the beamline after the proton collisions. The model parameters within GENIE are also tuned so that the simulated hadronic energy distribution matches what is seen in the Near Detector data from  $\nu_\mu$ CC events as seen in Figure 4.1. This is done by modifying the individual model parameters for CCQE and MEC processes.

### 4.3. SIMULATION DATASETS

Multiple Monte Carlo (MC) simulation datasets were created for various purposes, such as nominal simulation, which is the best prediction for the data. This simulated dataset allows for testing and optimization of analyses and measurements prior to examining the real data. Additionally, datasets were created to assist in uncertainty calculations for measurements.

4.3.1. NOMINAL MC. Approximately four times the amount of MC has been generated than data captured. This is done to reduce the statistical uncertainty in the simulated

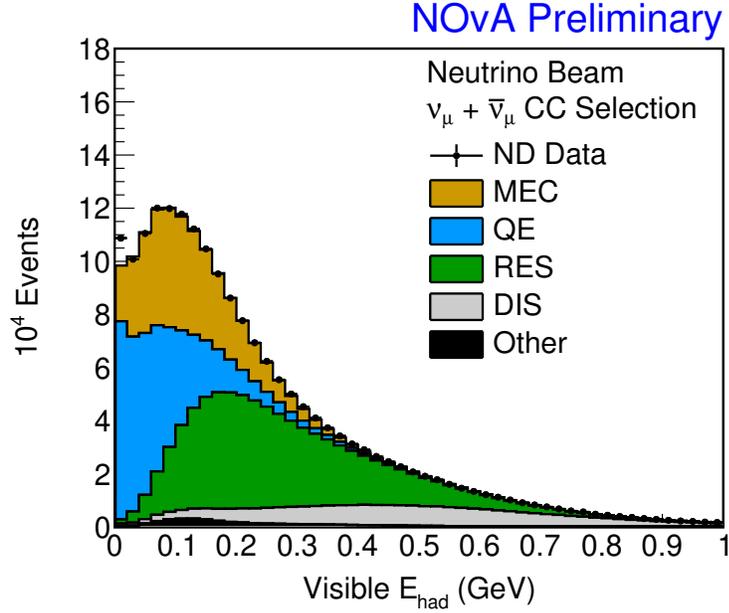


FIGURE 4.1. The reconstructed hadronic energy comparison between the data and the simulation after it has been tuned to best match the data.[19]

samples. The total MC generated is equivalent to  $3 \times 10^{21}$  POT. For this analysis, the nominal MC has been split into three statistically independent samples: a particle identification training dataset that is 1/6 of the total MC dataset, a fake data set using 1/6 of the total MC, and the remaining 2/3 was used for simulation studies, testing, and optimizing the analysis selection.

4.3.2. VARIATION OF SIMULATED SAMPLES. To estimate the model uncertainties in an analysis, special tools and datasets have been created that allow the variation of underlying model parameters.

4.3.2.1. *Detector Electronics Model-Variied Samples.* Seven MC samples have been generated resulting from systematic shifts of specific detector simulation model parameters by  $1\sigma$  of deviation.

- Light Response Up/Down

In these samples the light response of the scintillator is increased/decreased, which therefore change the light calibration constant.

- Cherenkov

In this sample, the absorption efficiency of Cherenkov light is changed.

- Positive/Negative Calibration Offset

In these samples, the absolute hadronic energy scale in the detector is varied by  $\pm 5\%$ .

- Calibration Shape

In this sample, the shape of the electronic response to scintillated light is changed.

- Neutron

In this sample, the neutron response in the detector is varied. Initially it was assumed that neutron capture in the detector would produce low energy, undetectable gammas. However, it was found that gammas of energy 1 MeV - 8.6 MeV were possible depending on the nucleus.

4.3.2.2. *Multiverse Samples.* An additional two samples can be generated by NO $\nu$ A analysers using the nominal MC already produced. To generate these samples, a “multiverse” approach is used. An individual universe has each model parameter varied from the nominal value by a random fraction of the standard deviation for each parameter. Distributions produced by the analyzer are then reweighted, corresponding to the varied parameters. Performing this process multiple times produces multiple universes, called the multiverse.

- GENIE

Individual model parameters in GENIE are varied to affect the various FSI and nuclear interaction models. Both shape and scale parameters can be varied.

- Flux

Using the Package to Predict the Flux (PPFX), the flux model parameters used by NO $\nu$ A are varied, which change the shape and scale of the neutrino flux from the NuMI beam.

#### 4.4. RECONSTRUCTION

The recorded hits during each spill are passed through the Slicer algorithm[23], which groups the hits together by space and time coincidence and removes noise hits. The goal of this algorithm is to separate hits from varying sources from one another, whether they are from cosmic rays passing through the detector or beam-neutrino events. This allows the removal of cosmic ray events and leaves the analyzer with the reconstructed neutrino interactions that occurred during the beam spill. An example of the slicing process is shown in Figures 4.2, 4.3, and 4.4 using event displays of an interaction in the Far Detector. Event displays illustrate activity that occurs in the detectors by showing two 2D images of reconstructed objects from different views. The top image in the event display is the XZ view of the detector as a top-down perspective. The bottom image in each event display is the YZ view from a side of the detector perspective. In Figure 4.2 the calibrated hits are shown prior to any reconstruction. There are many hits in this figure; this is primarily due to noise coming from the readout electronics. Additionally, many cosmic rays are visible, with the neutrino event seemingly hidden. An example of this same event after it has passed the slicer algorithm is shown in Figure 4.3. In this display, the noise hits have been suppressed and

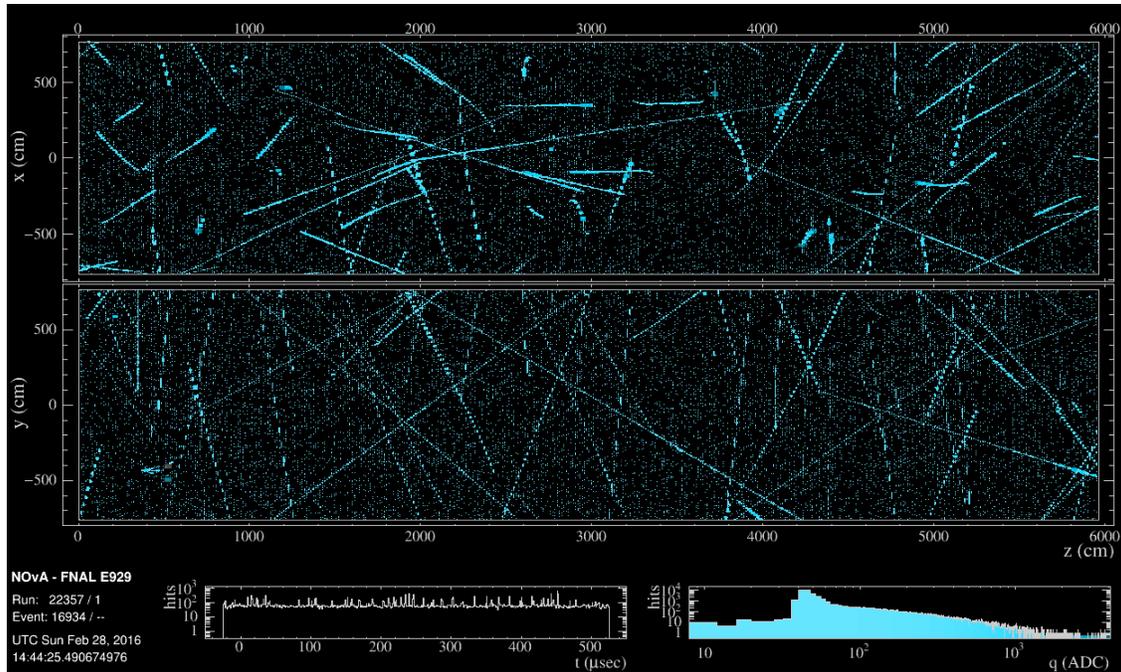


FIGURE 4.2. Event display showing an example neutrino event in the Far Detector prior to any reconstruction. These hits span the entire readout time of  $550 \mu\text{s}$ . [19]

each object in the spill has been grouped together into their slice. Each color represents a different slice, with the same color used in the XZ and YZ views for the same slice. Figure 4.4 shows this same event once again, but now the cosmic rays have been removed. The event display has been zoomed in with the neutrino event now clearly visible. Once the hits have been separated, they are passed through the Kalman Filtering algorithm [24] which creates a reconstructed object, called a track, by grouping hits in each slice following a linear like path. This algorithm is preferred to reconstruct particles that follow long, straight paths in the detector, such as muons, high energy protons, and pions. For non-track like particles, such as photons, electrons, and low energy protons, other algorithms are preferred and used such as the fuzzy k-means algorithm [25], which is used to cluster contiguous hits. With the cosmic rays removed, and the neutrino events fully reconstructed, the analysis variables are determined using the information from the reconstructed objects.

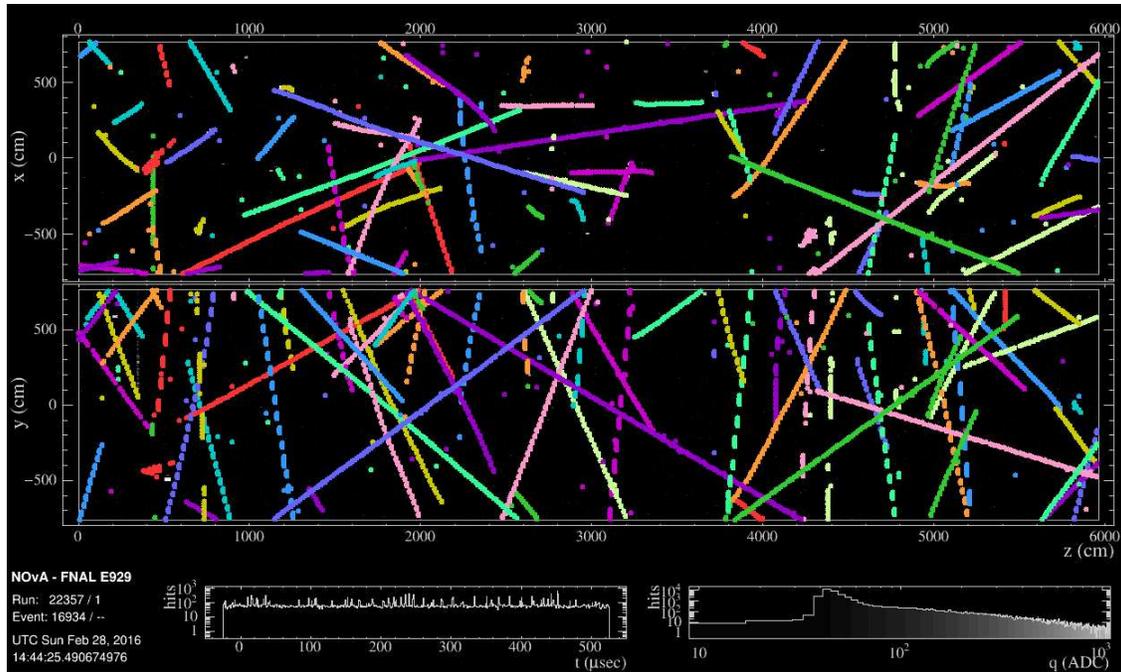


FIGURE 4.3. Event display showing the individual reconstructed slices for each object in Far Detector. These slices exist within the entire readout time of  $550 \mu\text{s}$ . [19]

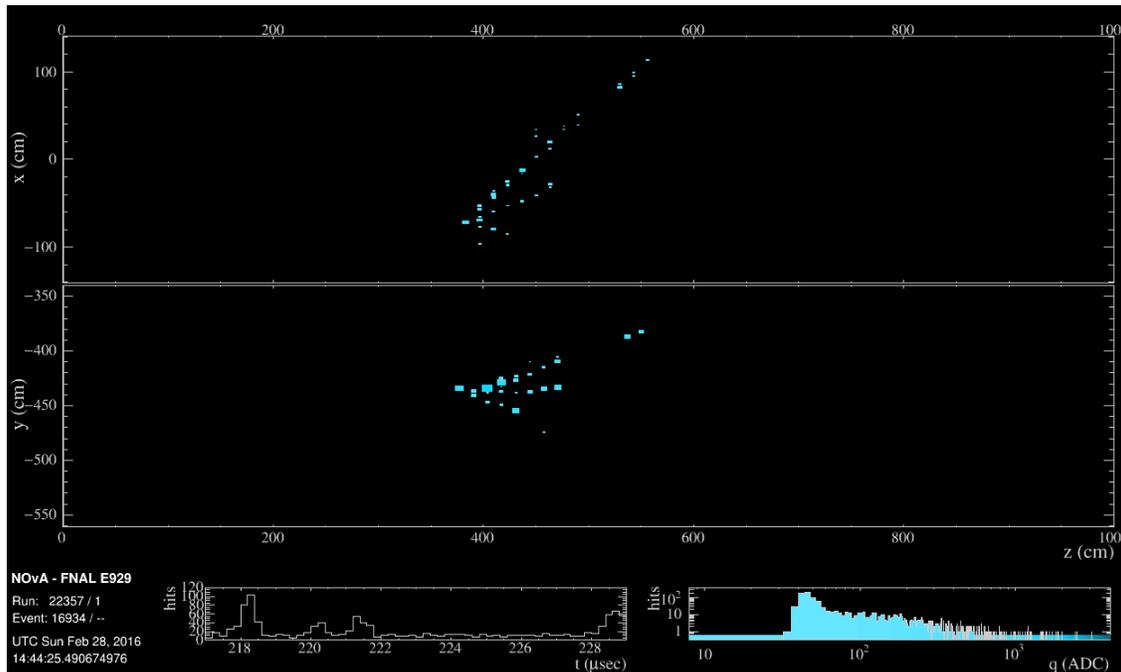


FIGURE 4.4. Zoomed in event display showing the reconstructed neutrino event slice in the Far Detector after noise suppression and cosmic ray removal. Note this is zoomed in by time as well, only looking at a specific  $12 \mu\text{s}$  window. [19]

## 4.5. ANALYSIS VARIABLES

Once particle trajectories have been reconstructed, the calculation and measurement of physical quantities and particle identifiers can be applied. These quantities are stored in files to be used by analyzers.

4.5.1. MUON IDENTIFICATION. To identify muons, a multivariate approach was used. A particle classifier was created by using a Gradient Boosted Decision Tree (BDT)[26]. These quantities were chosen due shape differences in their distributions, as seen in Figure 4.5, between muons, in blue, and the rest of the commonly produced particles in the detector: gammas, pions, and protons, in red. These four inputs are:

(1)  $\frac{dE}{dx}$  Log-Likelihood

This is the difference between the  $\frac{dE}{dx}$  log-likelihood with a muon assumption and the  $\frac{dE}{dx}$  log-likelihood with a pion assumption. The distribution of this quantity is shown in the top left of Figure 4.5.

(2) Scattering Log-Likelihood

This value is the scattering log-likelihood with a muon assumption subtracted by the scattering log-likelihood with a pion assumption. The scattering is calculated as  $\frac{\theta^2}{d}$  where  $\theta$  is defined as the scattering angle and  $d$  is the distance from the last scatter. The distribution of this quantity is shown in the top right of Figure 4.5.

(3) Average  $\frac{dE}{dx}$  in the Last 10 cm of the Track

This variable is the average energy deposited at the last 10 cm of the track divided by 10 cm. The distribution of this quantity is shown in the bottom left of Figure 4.5. This is an important quantity because it is relatively effective in identifying the Bragg peak in protons and the Michel electron at the end of muon tracks.

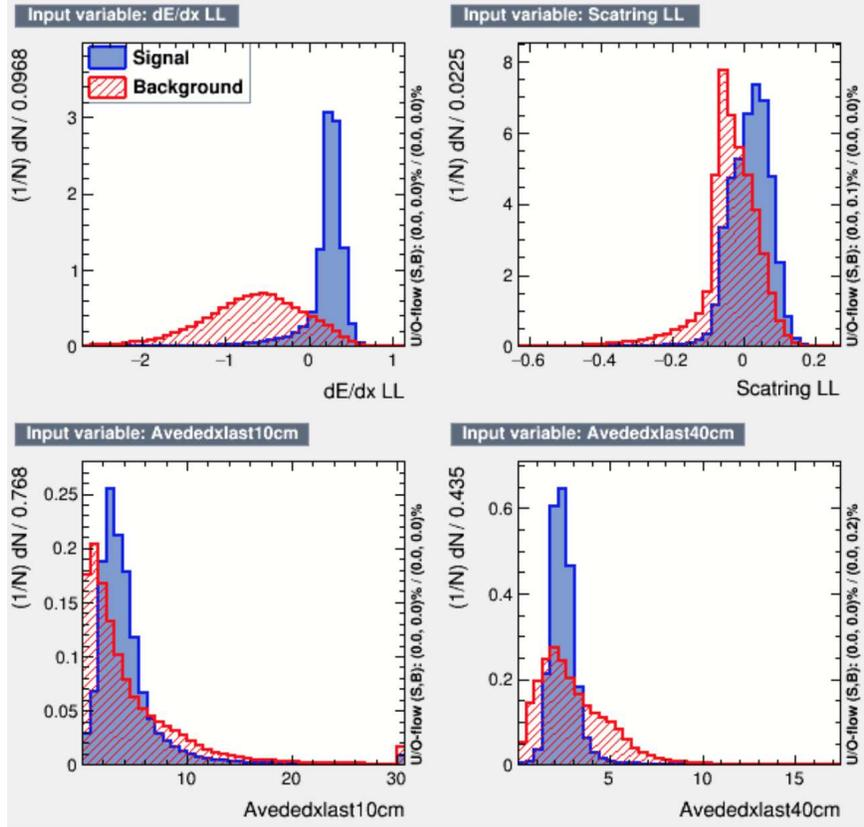


FIGURE 4.5. These are the distributions in the training dataset of the four inputs that were used to train the BDT. The blue distributions are measured from true muons, where the red distributions are measured from any object that was not a true muon.[27]

(4) Average  $\frac{dE}{dx}$  in the Last 40 cm of the Track

This variable is the average energy deposited at the last 40 cm of the track divided by 40 cm. The distribution of this quantity is shown in the bottom right of Figure 4.5. This is an important quantity because it is relatively effective in identifying the Bragg peak in protons and the Michel electron at the end of muon tracks.

These inputs were used in the training of the BDT classifier, of which the output is called the MuonID. The final distribution is shown in Figure 4.6, which give great separation between signal (true muons) and background (all other particles). To remove the large

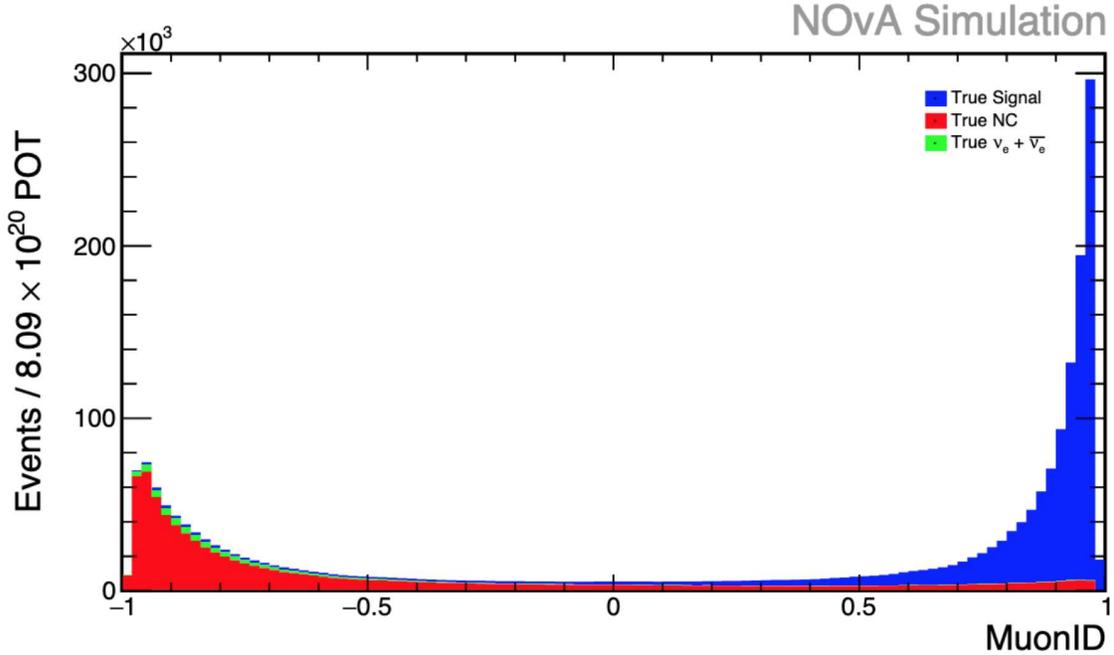


FIGURE 4.6. This is the distribution of the MuonID classifier. The blue curve represents events that truly do have a muon, the red curve is from  $\nu_\mu$ -NC events, and the green are from  $\nu_e$  and  $\bar{\nu}_e$  events.[27]

background peak in the low MuonID region, a minimum value for the MuonID was implemented. This minimum value was optimized by minimizing the total uncertainty, described in Chapter 5.1.2, on the  $\nu_\mu$ -CC inclusive cross section. A minimum uncertainty was found at a MuonID value of 0.24 as shown in Figure 4.7.

4.5.2. VERTEX RECONSTRUCTION. It is important to understand where the neutrino interacted in the detector. This point is called the vertex. For  $\nu_\mu$ CC analyses, the true vertex would be the point where the muon was created. By reconstructing and identifying the muon, the start of the muon trajectory can be declared as the vertex.

4.5.3. MUON KINEMATIC RECONSTRUCTION. The two most useful kinematic quantities for theory and modelling purposes are the energy of the particle and the angle of the particle

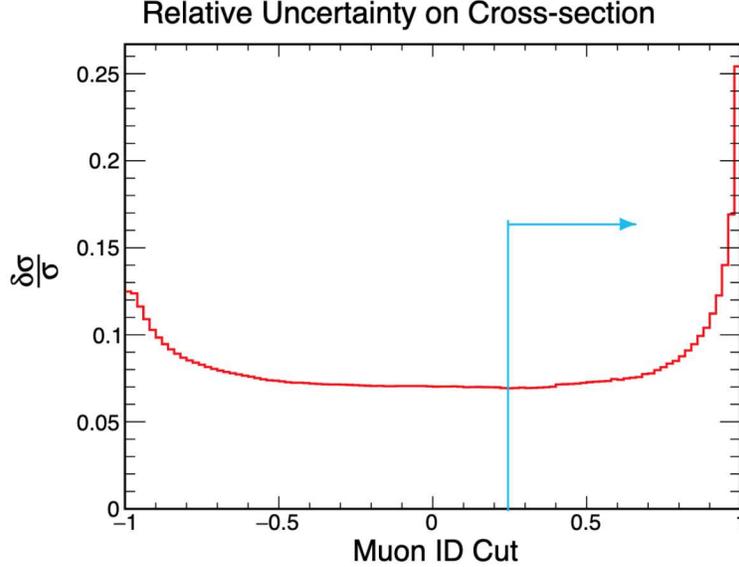


FIGURE 4.7. This plot shows the total relative uncertainty on the  $\nu_\mu$ -CC inclusive cross section as a function of a cut on the MuonID.[27]

relative to the neutrino direction. The muon scattering angle and the muon kinetic energy were used as observables for this measurement.

4.5.3.1. *Muon Angle.* Angle reconstruction is possible as once the particle has been reconstructed and identified. The position and direction information is then known. Calculations can be performed to measure the angle relative to the known neutrino-beam direction. The angle between the particle and the neutrino can be defined as  $\theta = \cos^{-1}(\hat{n}_\nu \cdot \hat{n}_i)$  where  $\hat{n}_\nu$  is the unit vector of the neutrino beam direction, and  $\hat{n}_i$  is the unit vector of the reconstructed particle track direction.

4.5.3.2. *Muon Energy.* Muons are reconstructed as relatively straight tracks while traversing through the detector as minimum ionizing particles. As particles travel through matter, they lose energy. Particles losing energy at the minimum energy loss rate are called minimum ionizing particles. The distribution of true muon kinetic energy as a function of length is a linear relationship. For muons, this relationship is highly correlated with there being little off-axis activity in an event, as seen in Figure 4.8. By fitting a polynomial to this

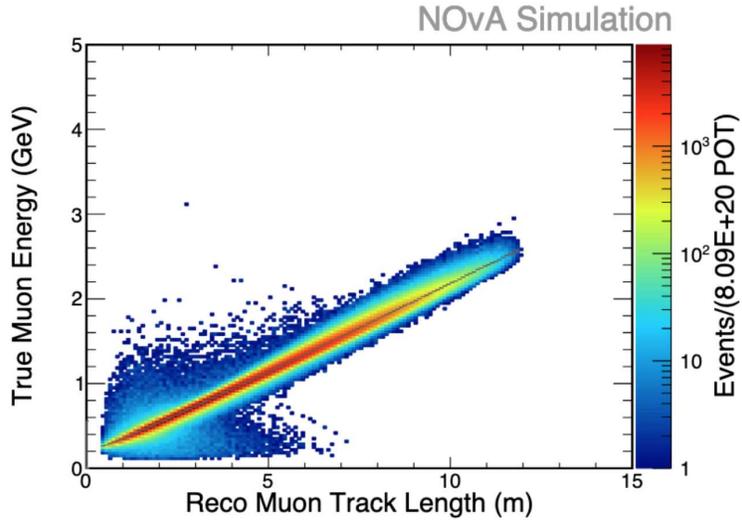


FIGURE 4.8. True muon energy vs the reconstructed track length for the muon. The majority of events have a linear relationship[27]

distribution, the muon energy can be reconstructed as a function of the track length of the selected muon.

## CHAPTER 5

### ANALYSIS

#### 5.1. MEASUREMENT STRATEGY

The goal of this analysis is to measure the double differential cross section of the semi-inclusive pion production process that yields one observable muon and at least one observable charged pion as a function of the kinematic variables of the produced muons. To achieve this measurement, both the muon and the pion were identified utilizing classifiers. The MuonID classifier has already been discussed previously in Chapter 4 and the PionID classifier will be discussed later in this chapter in Section 5.4. Additionally, methods for selecting and discriminating signal from background were determined and optimized. Next, a template fitting procedure was performed to extract a data-driven estimation of the number of signal events. Using the result from the fitting procedure, the signal events were unfolded to unsmear detector effects in the measured distributions. Finally, the unsmear distributions were used to calculate the final measured cross section parameters, a double-differential cross section, presented in muon kinematics: the kinetic energy ( $T_\mu$ ) and the angle ( $\cos\theta_\mu$ ). As this measurement builds off of a more inclusive measurement, the inclusive  $\nu_\mu$  CC cross section measurement, previous selections, tools, and binning were used in this analysis. This measurement was performed in a blinded fashion, meaning that no analysis was performed on data in the signal region until the measurement was well-understood using simulated samples.

A cross section is a physical quantity that measures the probability of a specific interaction occurring. If one imagines a large beam of particles that is aimed towards a smaller fixed target made of matter, only some of the particles from the beam will interact, resulting in a

scatter. The cross section,  $\sigma$ , for this scenario can be defined as:

$$(2) \quad \sigma = \frac{N^{Scattered}}{\frac{N^{Incident}}{A} N^{Target}}$$

Where  $N$  is the number of particles, and  $A$  is the cross sectional area of the target. From this equation it is evident that the cross section will have units of area. The benefit of measuring the cross section, instead of a percentage probability, is that this quantity is constant for a given interaction type and therefore does not depend on parameters from individual experiments, allowing this number to be easily compared between different experiments. For neutrino physics experiments, a similar cross section definition is used,

$$(3) \quad \sigma = \frac{N^{Selected} - N^{Background}}{\epsilon N^{Target} \phi}$$

where  $N^{Selected}$  is the number of selected events,  $N^{Background}$  is the number of selected background events. The difference of these two quantities will yield the number of selected signal events, where signal refers to the specific interaction type being measured.  $N^{Target}$  is the number of target nucleons,  $\phi$  is the total integrated flux of the neutrino beam, and  $\epsilon$  is the true signal selection efficiency which is defined as

$$(4) \quad \epsilon = \frac{N^{Selected \ True \ Signal}}{N^{Total \ True \ Signal}}$$

If the signal can be effectively estimated, the equation simplifies to

$$(5) \quad \sigma = \frac{N^{Signal}}{\epsilon N^{Target} \phi}$$

The cross section can also be measured as a function of kinematic variables from the particles resulting from the interactions. This analysis will measure the differential cross section as a function of two kinematic variables of the resulting lepton. A double differential cross section is defined as:

$$(6) \quad \left( \frac{d^2\sigma}{d\cos\theta_\mu dT_\mu} \right)_i = \frac{\sum_j U_{ij} N_{sig}^{sel}(\cos\theta_\mu, T_\mu)_j}{\epsilon(\cos\theta_\mu, T_\mu)_i (\Delta\cos\theta_\mu)_i (\Delta T_\mu)_i N_{target} \phi}$$

In the numerator,  $N_{sig}^{sel}$  is the selected signal, binned in  $\cos\theta_\mu$  and  $T_\mu$ ,  $U_{ij}$  is an element of the unfolding matrix where  $j$  refers to measured reconstructed quantities and  $i$  refers to the true quantities. This is due to the fact that detectors will measure different values for quantities than what the true values are for those quantities. In the denominator is  $\epsilon$ , the true signal selection efficiency binned in  $\cos\theta_\mu$  and  $T_\mu$ ,  $\Delta\cos\theta_\mu$  and  $\Delta T_\mu$  are the muon kinematic bin widths,  $N_{target}$  is the number of target nucleons, and  $\phi$  is the integrated neutrino flux.

5.1.1. SIGNAL DEFINITION. The signal for the measurement made in this thesis is defined as an event with one muon and at least one charged pion with a kinetic energy of 250 MeV or more, any additional hadrons that are produced in the interaction are allowed. The charged pion may be either negatively or positively charged and there is no simple way to disentangle their signs. These signal event vertices must occur in the fiducial volume of the near detector, described in Section 5.2.2. Additionally, the event must be contained within the detector. A cartoon showing the topology of an event can be seen in Figure 5.1. An event display of a simulated signal event is shown in Figure 5.2. The top of the figure is an XZ view of the simulated signal event in the near detector, where the bottom of the figure shows the YZ view of the same event.

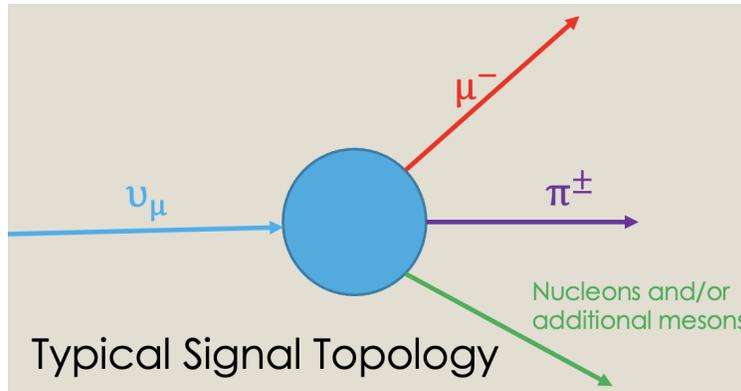


FIGURE 5.1. Cartoon illustrating the topology of the signal.

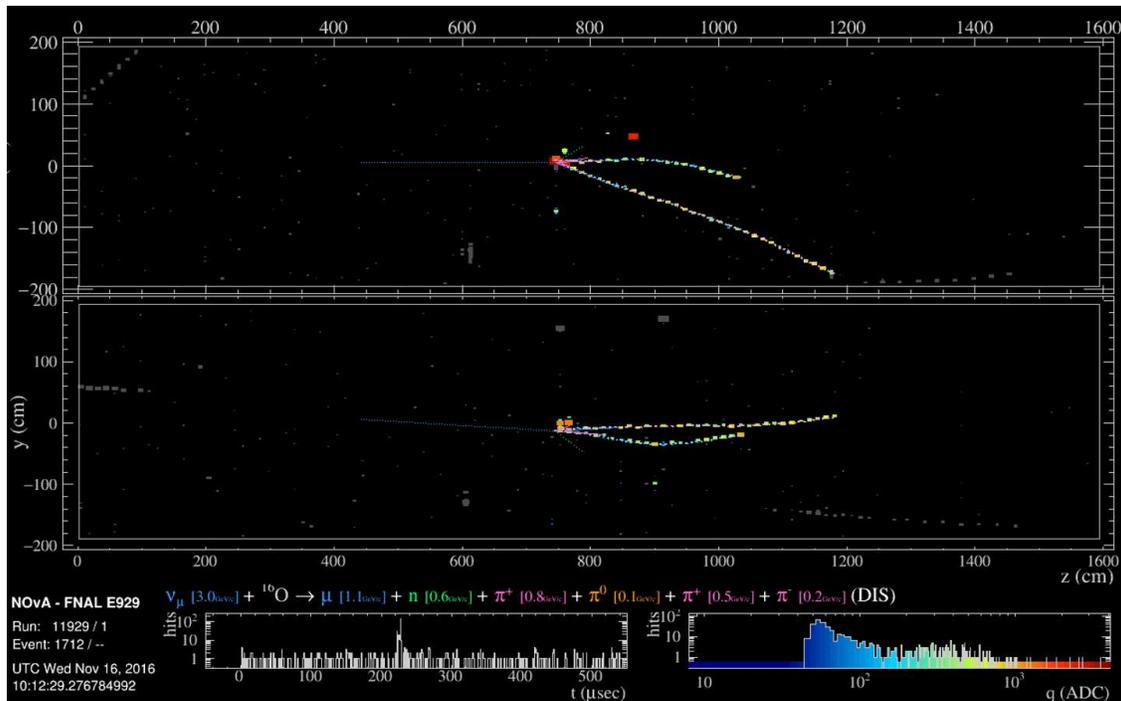


FIGURE 5.2. Event display showing a simulated pion event consistent with the signal definition. In this event, the shorter trajectory in both views is due to the produced muon from the neutrino interaction, while the longer trajectory is due to the produced pion.

5.1.2. SELECTION OPTIMIZATION. When creating a procedure to select signal events from a larger dataset, the total uncertainty on the final measurement must be considered as uncertainty is sensitive to the sample selection. For this cross section measurement, a figure

of merit is used, derived from Eq. 3[28],

$$(7) \quad \frac{\delta\sigma}{\sigma} = \sqrt{\frac{(\delta N_{sel}^{stat})^2 + (\delta N_{bkg}^{stat})^2 + (\delta N_{bkg}^{syst})^2}{(N_{sel} - N_{bkg})^2} + \left(\frac{\delta\epsilon}{\epsilon}\right)^2}$$

This equation is the fractional uncertainty on the cross section and does not include the individual uncertainties from the flux and number of targets as those uncertainties are independent of the selection and applied post-fit. By treating the statistical uncertainties as Poisson-distributed, the figure of merit can be further simplified to,

$$(8) \quad \frac{\delta\sigma}{\sigma} = \sqrt{\frac{N_{sel} + N_{bkg} + (\delta N_{bkg}^{syst})^2}{(N_{sel} - N_{bkg})^2} + \left(\frac{\delta\epsilon}{\epsilon}\right)^2}$$

The systematic uncertainties referenced in this formula are uncertainties from the models used to produce the simulation and will be discussed in detail later in this chapter.

5.1.3. PION KINEMATIC RECONSTRUCTION. Prior to choosing to present the cross section in bins of muon energy and angle, the pion angle and energy were reconstructed first to determine the feasibility of using these parameters in the measurement.

5.1.3.1. *Pion Angle*. Reconstructing the pion angle was challenging because pions produced in the NO $\nu$ A near detector are near the  $\Delta$ -baryon resonance peak and therefore pions have a high chance of reinteracting in the detector, causing scattering. This can affect the angle reconstruction in events, as a pion that scatters can produce multiple tracks, or if the scattering occurs near the vertex, it may not be possible to reconstruct the initial trajectory. An example of the latter occurring in a simulated event is illustrated in Figure 5.3 as a zoomed in event display. The dashed pink line is the reconstructed pion track, with the start of the track being located at the left most end of the pink line. The other track in this event

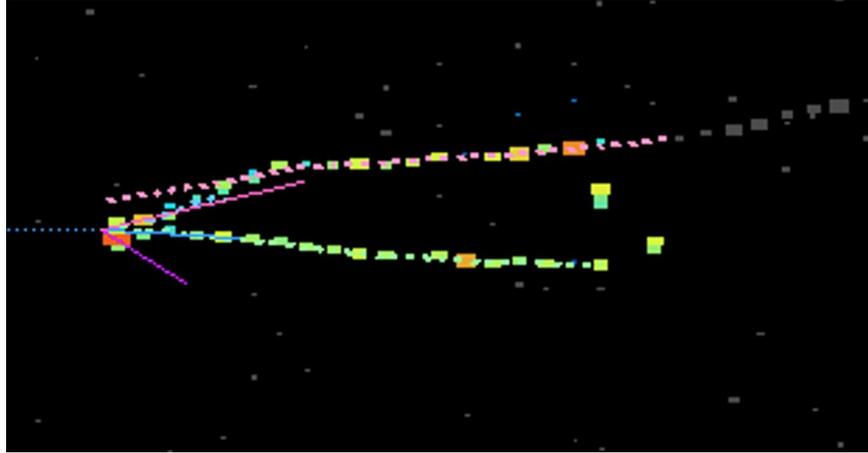


FIGURE 5.3. This is a zoomed in event display that shows an example signal interaction where the pion was misreconstructed due to scattering that occurred near the start of the track.

is due to the muon. The incoming neutrino is visible as the blue dotted line incoming from the left side of the image. Another challenge with reconstructing the pion direction came from events with high amounts of activity. Figure 5.4 shows a simulated CC DIS signal event with many produced particles, the trajectories from many of the particles overlap, leading to reconstruction errors. To be able to properly measure the pion direction in events such as this a more sophisticated technique would need to be employed. For the reasons listed above, it was decided to present the results in muon kinematics rather than pion kinematics.

5.1.3.2. *Pion Energy.* Much like the muon, a linear relationship exists for the pion, however the linearity is not as strong due to scattering and decay of the pion as it traverses the detector as seen in Figure 5.5. A third order polynomial fit was applied to this plot to obtain a best fit of the energy to the length of the reconstructed pion track without adding superfluous fit parameters. Plotting the difference between the reconstructed pion energy from this method and the true energy as a function of the true energy, as shown in Figure 5.6, shows

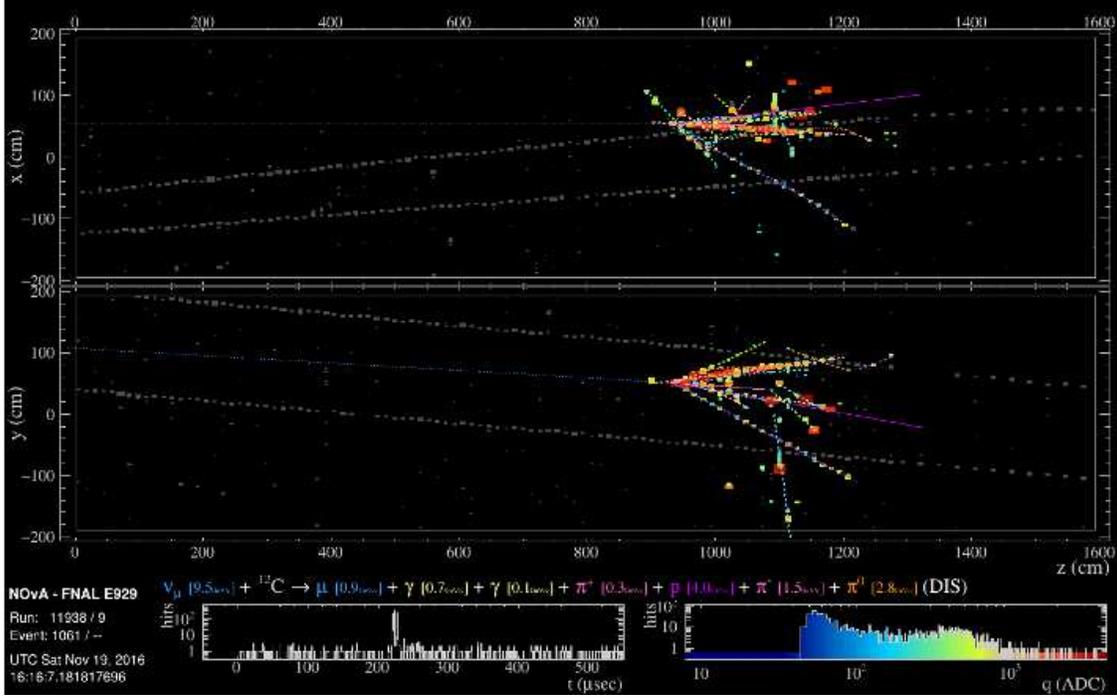


FIGURE 5.4. Event display of a simulated high energy DIS event with significant final state activity.

a large bias<sup>1</sup> throughout the entire energy range. The red curve is the mean value for the bias in each energy bin. There is a collection of events with only a small amount of bias, these are the same events that lie along the diagonal in Figure 5.5. A correction to this bias was attempted, by subtracting the bias mean, this is shown in Figure 5.7. Unfortunately, this method introduced a bias in the events where the relationship between length and energy were quite linear. To be able to properly estimate pion energy, a simple polynomial fit on a track will not work and other techniques will be needed.

5.1.4. MINIMUM PION ENERGY. Pions can be reconstructed in the NO $\nu$ A Near Detector over a wide energy range, but when the energy of the pion is too low, it may not be detectable due to either not creating enough hits to form a Kalman track, or it may not deposit enough energy to form any reconstructable object at all. Therefore, a minimum energy

<sup>1</sup>Bias in this context is defined as the difference between the reconstructed and true values.

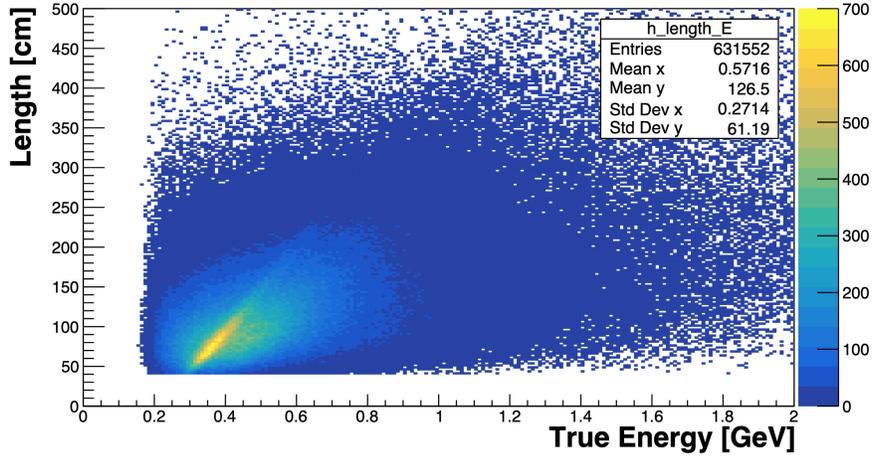


FIGURE 5.5. This plot shows the reconstructed pion track length vs the true energy of the pion in simulation.

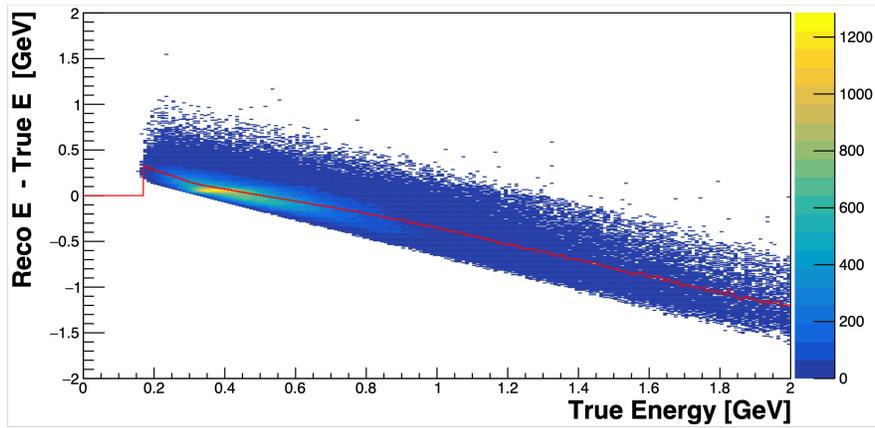


FIGURE 5.6. Bias between the truth and reconstruction as a function of the true energy of the pion. The red line is the mean value for the bias as a function of the true pion energy.

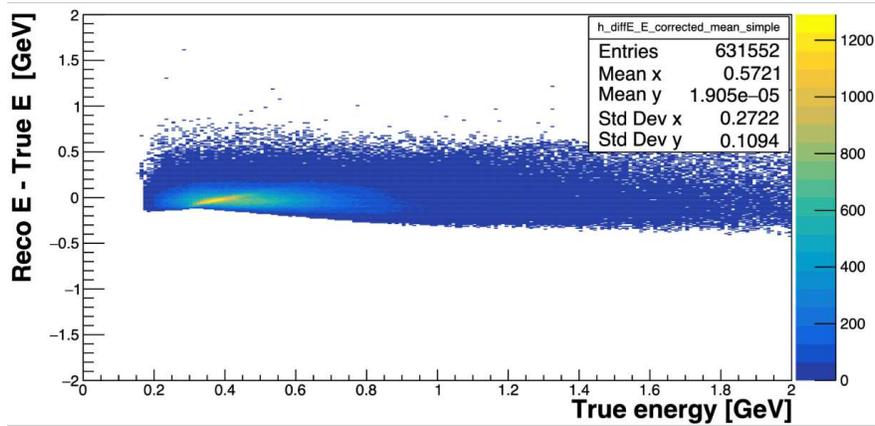


FIGURE 5.7. Bias between the truth and reconstruction as a function of the true energy of the pion after the bias correction is applied.

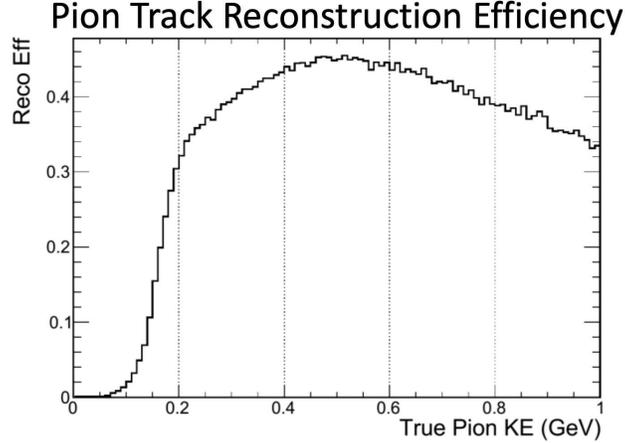


FIGURE 5.8. Pion reconstruction efficiency as a function of the true kinetic energy of the simulated pion.

selection must be applied to pions in the signal definition to ensure that selected events have well reconstructed pions and don't otherwise contribute large uncertainties in the final measurement. The minimum energy cutoff was determined by examining the reconstruction efficiency of pions as a function of their true energy. The reconstruction efficiency is defined as:

$$(9) \quad \epsilon_{\pi} = \frac{N_{\pi}^{Reco}}{N_{\pi}^{True}}$$

Where  $N_{\pi}^{True}$  is the true number of pions created in the simulation and  $N_{\pi}^{Reco}$  is the number of reconstructed pion tracks, which have been matched to a pion generated by the simulation. Figure 5.8 shows the pion reconstruction efficiency, from Equation 9, as a function of the true energy of the pion using simulated data. Pions become visible at 150 MeV, below which pions cannot be observed by the NO $\nu$ A detectors. There is a steep slope from 100 MeV to 220 MeV, after which the reconstruction efficiency decreases less rapidly. The efficiency decreases at higher kinetic energies due to reinteractions in the detector. Pions produced in the NO $\nu$ A detectors have energies near the  $\Delta$ -baryon resonance peak and therefore the

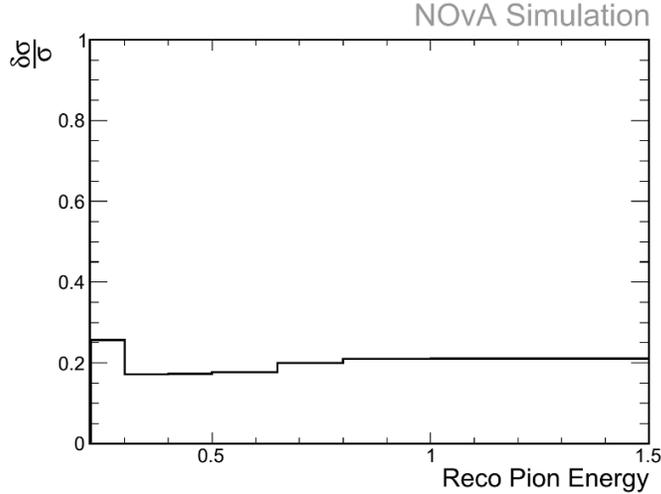


FIGURE 5.9. Estimated fractional uncertainty on the cross section as a function of the reconstructed kinetic energy of the simulated pion.

produced pions are very likely to reinteract within the detector, which can reduce the ability to accurately identify pions and result in poorly reconstructed energies. Figure 5.9 shows the estimated fractional uncertainty on the final cross section measurement in each bin of pion energy. The uncertainty on cross section below 220 MeV is relatively high, compared to higher energies. However, comparing to the reconstruction efficiency from Figure 5.8, this value exists in the region where the efficiency is rapidly changing, therefore 250 MeV was chosen as the minimum reconstructed value for the energy of the pion.

## 5.2. EVENT PRESELECTION

Event preselection is the initial selection performed on the dataset designed to create a collection of events of good quality without discriminating between signal and background events. The following sections describe the quantities used in the preselection.

5.2.1. DATA QUALITY. The data quality preselection ensures the data taking occurred during good running conditions, as defined by the beam and the near detector, and that the

reconstruction quality of each slice meets a minimum standard. The reconstruction quality requirement ensures the selection of events with these qualities:

- **Number of Tracks:** The slice must contain at least one 3D Kalman track.
- **Number of Hits:** The slice must contain at least 20 hits. This primarily removes reconstructed noise, though some NC events are removed as well.
- **Number of Contiguous Planes:** The slice must contain at least 4 contiguous planes with hits to remove events that have a high likelihood of reconstruction errors.

5.2.2. FIDUCIAL. The fiducial volume is defined as the region of the detector considered to be well understood. For this analysis, the start of the selected muon track is treated as the reconstructed vertex for the event, where the vertex is defined to be the position where the neutrino interaction occurred. To ensure that the event originated from within the detector and did not leak in from outside, the vertex is required to be between -130 cm to 140 cm in the X and Y dimensions and in the Z dimension its required to be within 100 cm to 1000 cm away from the front face of the detector with (0,0,0) being defined as the center of the front face of the detector. The asymmetry in the X and Y dimension is due to the beam direction not perfectly aligning with the Z dimension. Figure 5.10 shows a simulated signal event with the fiducial volume, shown overlaid in green. The illustrated event passed the fiducial preselection.

5.2.3. EVENT CONTAINMENT. Containment requires that the energy from the event doesn't escape the detector. This is an important requirement because it is challenging to reconstruct the energy of the selected track without requiring all of its energy be deposited into the detector. The event must be at least 20 cm from the detector edge in the X and Y

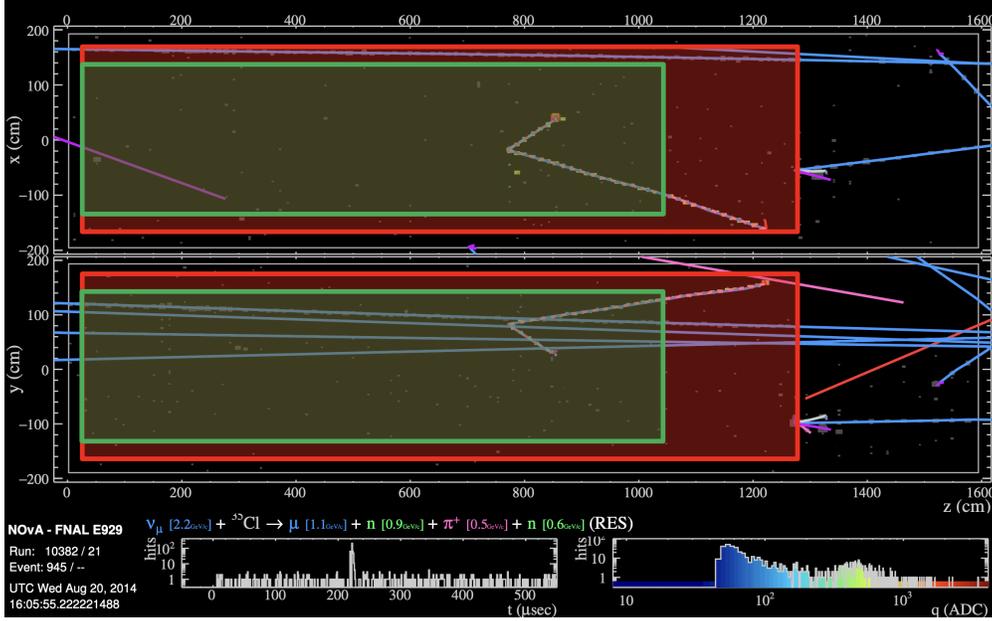


FIGURE 5.10. This event display shows a simulated signal event with the fiducial and containment volumes overlaid in green and red respectively.

dimensions, and in the Z dimension must be between 20 cm to 1525 cm from the front face of the detector. Finally, events that have non-muon tracks in the muon catcher are removed. Figure 5.10 shows a simulated signal event overlaid with the containment volume, shown in red. The illustrated event passed the containment preselection.

### 5.3. SIGNAL EVENT SELECTION

Following preselection, a more detailed selection is performed to select a higher purity signal event sample. The following selections were performed.

5.3.1. MUONID. In order to identify a muon in an event, MuonID was used. It was required that each event have at least one track with a minimum MuonID value of 0.24. Events passing this requirement were considered candidate events containing a reconstructed muon.

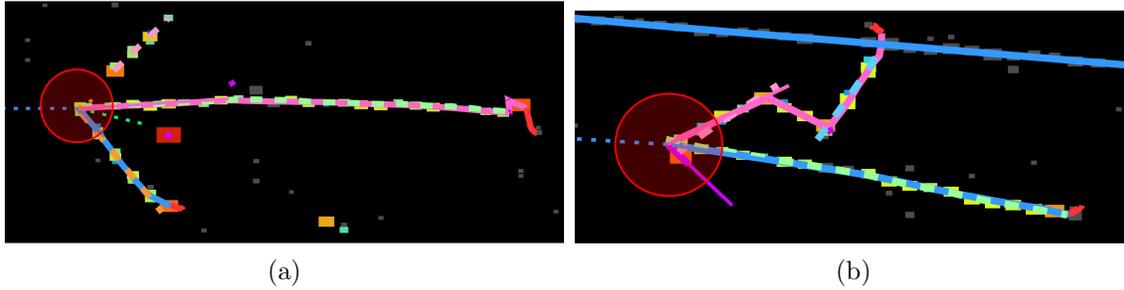


FIGURE 5.11. Zoomed event displays of simulated events with a red circle showing the  $<15$  cm track start requirement. The event in the left pane is a signal event where the neutral pion was not selected. The event in the right pane shows a signal event with a pion that scattered multiple times creating multiple reconstructed objects.

5.3.2. TRACK REQUIREMENTS. The CC Pion measurement required the identification of both a muon and a charged pion. This means that at a minimum of two 3D tracks needed to be fully reconstructed in order to allow identification otherwise the event must be rejected. Both of these tracks were required to have their start position within 15 cm of each other. As pions in this energy range often reinteract and when they scatter, additional tracks can result. Additionally, this helps to exclude selecting track-like neutral pion decays. Neutral pions travel some distance before decaying, producing a gap between the start of the neutral pion track and the start of the muon track. Figure 5.11 shows zoomed-in event displays illustrating these effects. A red circle highlights the 15 cm region that a track must start within to be considered a candidate charged pion track. The right-hand display shows a pion that scattered multiple times creating more than one reconstructed track and the left-hand display shows a signal DIS event where the neutral pion was not selected as the pion track due to the  $>15$  cm gap between the start of the muon track and the start of the neutral pion track.

5.3.3. MUON THRESHOLD. The last selection performed was a muon energy threshold requirement that the muon track have at least 500 MeV of Kinetic Energy and a maximum of

2500 MeV. Also, an angle threshold was applied that required the muon to have a minimum  $\cos \theta_\mu$  of 0.5. These kinematic requirements helped to ensure a well reconstructed muon.

5.3.4. PURITY. Purity is a ratio of a specific signal to the overall dataset. This is a useful metric when optimizing selection. As more requirements are created and applied to the selection, it is desired for the purity to increase for a given sample. Mathematically, purity is defined as,

$$(10) \quad Purity = \frac{N^{signal}}{N^{total}}$$

where  $N^{signal}$  is the number of true signal events in the dataset, and  $N^{total}$  is the number of total events, signal and background, in the dataset after the selection is applied.

5.3.5. SUMMARY OF SELECTED EVENTS. Table 5.1 shows how the simulated signal was sculpted from the initial dataset where the sample was scaled to  $8.09 \times 10^{20}$  POT. Each row includes a requirement to those accumulated in the previous row. Each column is a specific metric for the sample.

TABLE 5.1. Table showing the summary of simulated events from the final selection (scaled to  $8.09 \times 10^{20}$  POT)

	Signal	Total	Purity	Sig. Eff.	Rel. Sig. Eff.
No Cuts	$1.11 \times 10^6$	$1.38 \times 10^8$	0.8%	N/A	N/A
Quality	$1.11 \times 10^6$	$9.84 \times 10^7$	1.1%	100%	100%
Containment	198027	$4.84 \times 10^6$	4.1%	17.8%	17.8%
Fiducial	191052	$2.58 \times 10^6$	7.4%	17.2%	96.5%
MuonID	178717	$1.42 \times 10^6$	12.6%	16.1%	93.5%
$\geq 2$ Tracks + 15 cm Req.	51187	190665	26.8%	4.6%	28.6%
Muon Threshold	47280	111815	42.3%	4.3%	92%

## 5.4. PION IDENTIFICATION

It is necessary to identify the charged pion track in events to distinguish signal events from background events. A multivariate approach was chosen to help identify events with these charged pion tracks. Specifically, a Boosted Decision Tree was employed, using the TMVA library[29], as it has been used by many other measurements and is well understood. Additionally BDT network<sup>2</sup> have relatively small runtimes and scale quite well. Another benefit of using a BDT is that it is easier to understand what the algorithm is doing compared to more complicated machine learning techniques, such as Deep Learning. Using a somewhat less sophisticated approach also provides a good baseline to compare against future iterations of the CC Pion measurement that will use these Deep Learning techniques.

5.4.1. BDT INPUT VARIABLE SELECTION. Many features were considered as inputs to be used in the BDT, however five were chosen based based on their shape differences between charged pions and other particles. Greater shape difference can yield higher discriminating power in the BDT network. The distribution for each feature described below can be seen in Figure 5.12, broken out by particle type.

- (1) Average  $\frac{dE}{dx}$  in the last 10 cm of the track

This variable is the energy deposited at the last 10 cm of the track divided by 10 cm.

- (2) Average  $\frac{dE}{dx}$  in the Last 40 cm of the Track

This variable is the energy deposited at the last 40 cm of the track divided by 40 cm.

- (3)  $\frac{dE}{dx}$  Log-Likelihood

This variable is the difference between the  $\frac{dE}{dx}$  log-likelihood using a muon assumption and the  $\frac{dE}{dx}$  log-likelihood using a pion assumption.

---

<sup>2</sup>Hereby referred to as simply “BDT”

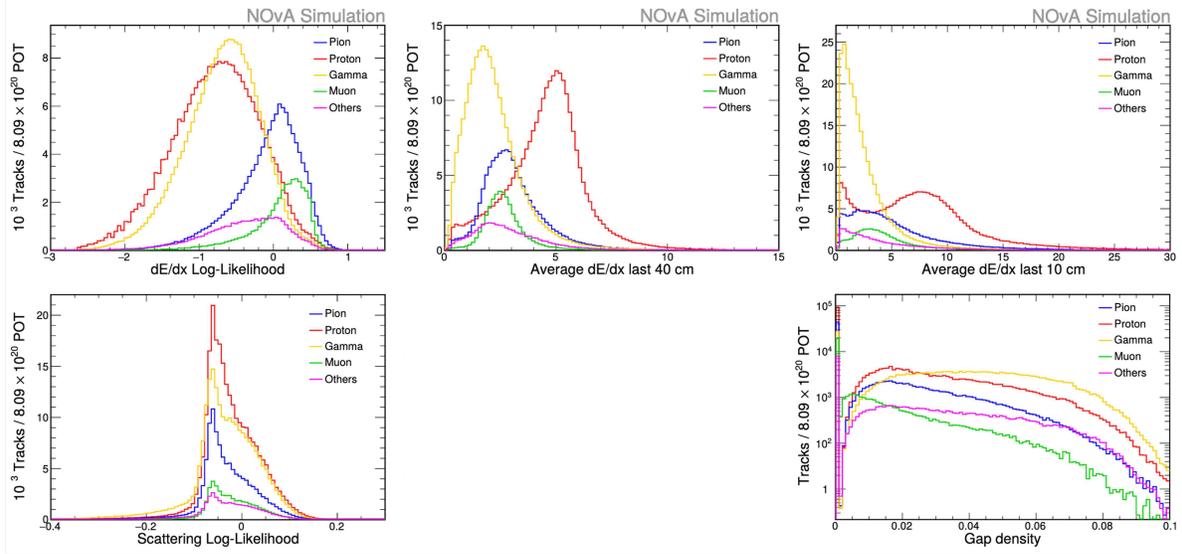


FIGURE 5.12. The five input parameters used in the BDT, broken out by particle type.

#### (4) Scattering Log-Likelihood

This feature is the scattering log-likelihood with a muon assumption subtracted by the scattering log-likelihood with a pion assumption. The scattering is calculated as  $\frac{\theta^2}{d}$  where  $\theta$  is defined as the scattering angle and  $d$  is the distance from the last scatter.

#### (5) Track Gap Density

This value is calculated by taking the number planes within the track where no hits were recorded and dividing by the length of the track.

The BDT response was optimized by minimizing the fractional uncertainty on the final measurement while also using the Receiver Operating Characteristic (ROC) curve to determine discrimination performance and utilizing the Kolmogorov-Smirnov (KS) score to optimize against overtraining. Events may contain multiple pion tracks, however, only the track most likely to be a pion is used for this measurement. Therefore the track with the highest pion score in the event, excluding the tracked selected by MuonID, is considered the

pion track. This highest value is called the best pion score. Figure 5.13 shows best pion score in the event broken out by the true type of particle that created the track for all selected events in the simulated sample. There is a strong response in the high score region from pions, which is desired, however there is also a noticeable response from muons and a slight peak in that region for the proton distribution as well. The peak in the muon distribution is a combination of three effects: mismatched 3D reconstruction, incorrect muon selection, and pion decay. The 3D misreconstruction occurs from events where the pion track and the muon track have approximately the same length in the z-direction and the reconstruction incorrectly matches the 2D tracks together when forming the 3D track. This mismatch effect can also occur between protons and pions. The incorrect muon selection most often occurs when the selected muon track from MuonID wasn't actually a muon, but instead a particle with similar properties, such as a charged pion. Lastly, pion decay can cause the track to be identified as a muon since the most common pion decay product are muons. If the pion decays into a muon and the muon produces a long enough track, the reconstruction software may consider that to be a muon track, even though the muon came from a final state pion. As seen in Figure 5.13, the BDT does a good job of assigning low PionID scores to gammas as a majority of the gammas have scores below zero. Figure 5.14 shows the distribution of the best pion score, broken out by the type of event. The signal distribution in orange is strongly peaked in the high pion score region. However, there is also a peak in the low score region. This is primarily due to events where the pion was not reconstructable, such as events where the pion did not appear in the final state or if the pion was poorly reconstructed. The only background events with a strong peak in the signal region are from events with energy below the reconstruction threshold.

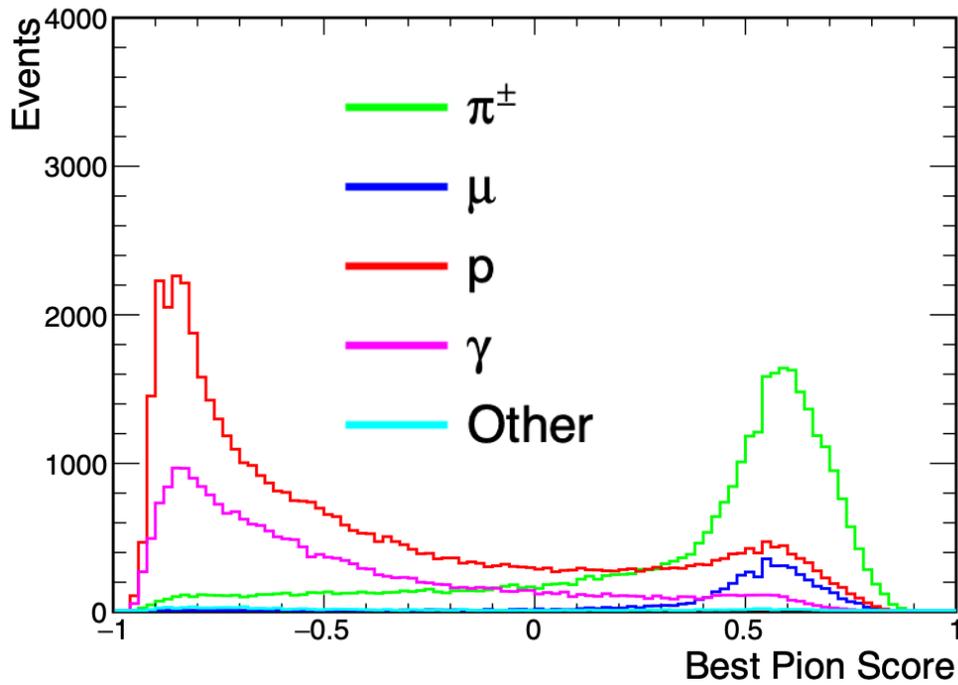


FIGURE 5.13. The highest PionID value among all tracks in an event, excluding the track selected by MuonID, broken out by particle type.

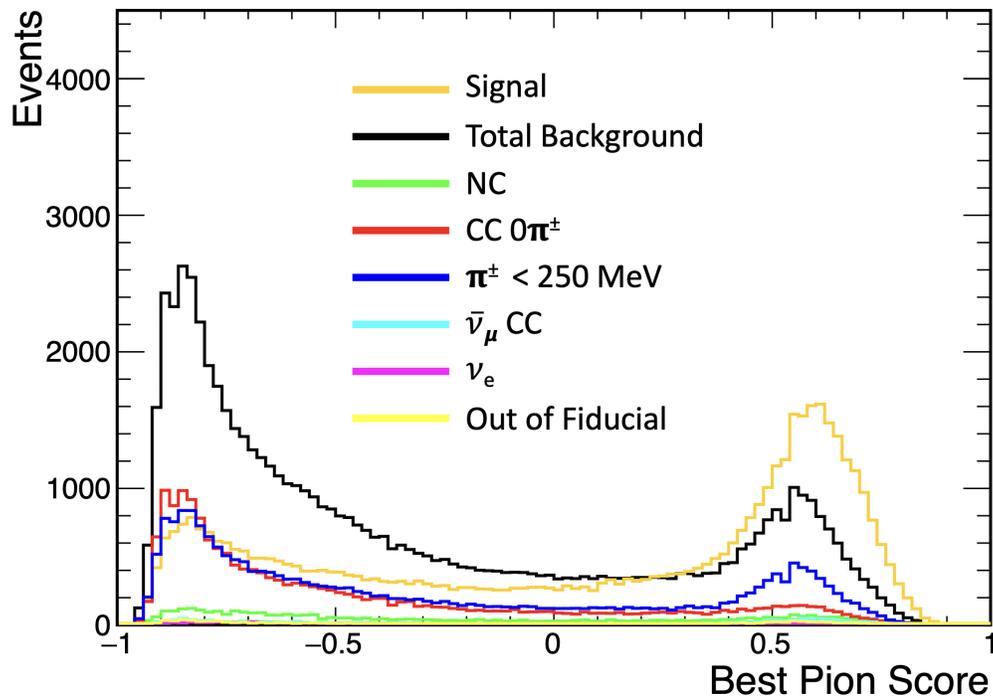


FIGURE 5.14. The highest PionID value among all tracks in an event, excluding the track selected by MuonID, broken out by the topology of the event.

5.4.2. OPTIMIZATION OF THE PION EVENT SELECTION. As with the other selection optimizations described so far, this selection was optimized by minimizing the total fractional uncertainty on the cross section. The total fractional uncertainty is calculated using Equation 7. The total fractional uncertainty on the cross section comprises four parameters: the statistical uncertainty on the number of selected events, the statistical uncertainty on the number of background events, the systematic uncertainty on the number of background events, and the systematic uncertainty on the efficiency. Adding these four uncertainties in quadrature yields the cross section uncertainty. The statistical uncertainties are taken as the  $\sqrt{N}$  normalized by the square of the difference between the number of selected and background events. To calculate the systematic uncertainty in the nominal background, the difference between the nominal and shifted samples is calculated. It is assumed that all of the systematic parameters vary independently such that the total systematic uncertainty can be calculated by adding them in quadrature. Figure 5.15 shows each of the four terms as function of the minimum accepted value for the Best Pion Score. Figure 5.16 shows the total fractional uncertainty on the cross section as a function of the minimum accepted value for the Best Pion Score. From this figure, it can be seen that requiring a minimum Best Pion Score of 0.6 results in the lowest uncertainty on the final measurement.

## 5.5. MEASUREMENT BINNING

The previous  $\nu_\mu$  CC inclusive measurement performed by the NO $\nu$ A collaboration utilized a 2D binning scheme using muon kinematic variables:  $T_\mu$  and  $\cos\theta_\mu$ , as described in the previous chapter. The same preselection, a similar selection, and the utilized binning scheme can be adopted for this analysis. A check on the resolution of the muon kinematics, which was used to determine the binning for the  $\nu_\mu$  CC inclusive measurement, yields similar

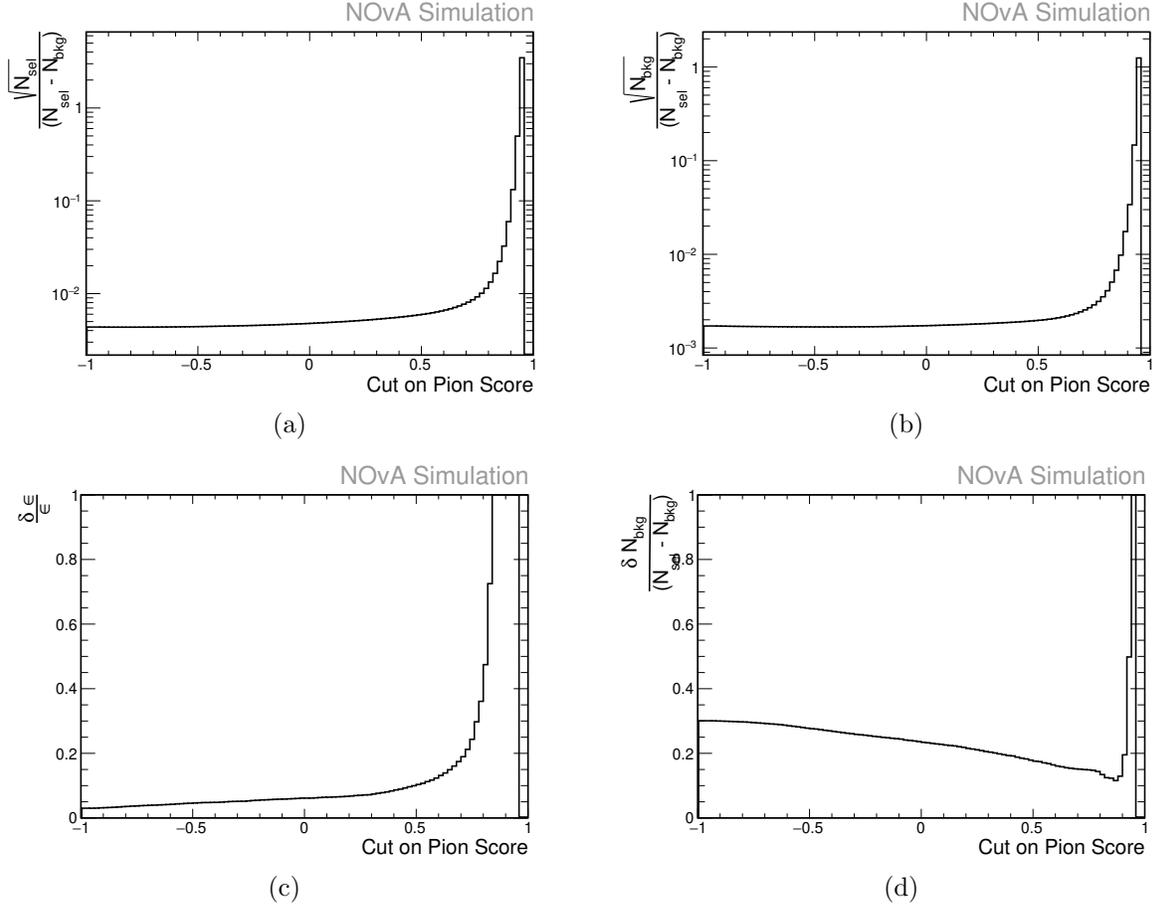


FIGURE 5.15. The four components contributing to the calculation of the fractional uncertainty on the cross section as a function of the cut value used in the analysis, (a) is the statistical uncertainty on the number of selected events, (b) is the statistical uncertainty on the number of background events, (c) is the total systematic uncertainty on the efficiency, the quadratic sum of all the model parameter uncertainties, (d) is the total systematic uncertainty on the number of background events.

results between the two signal definitions. The absolute resolution as a function of the muon kinematic parameters is shown in Figure 5.17. This figure suggests that using a binning scheme similar to the  $\nu_\mu$  CC inclusive measurement was a good starting point for this measurement. Based on the resolution of the kinematics alone, the finest binning scheme possible is shown in Figure 5.18 for the CC Pion measurement on the left, and the binning used in the  $\nu_\mu$  CC inclusive measurement[27] for comparison on the right. These plots show

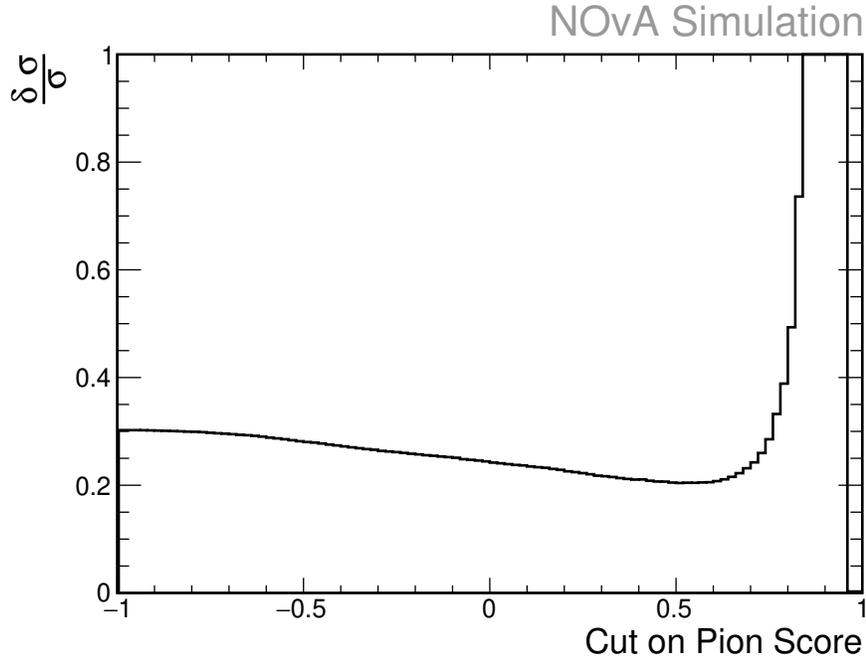


FIGURE 5.16. Total cross section uncertainty as a function of the cut value on the BDT response.

the total number of selected events, signal plus background, in each bin. The difference in analysis space and event rate between the two samples is due to additional cuts made for the pion analysis that were not made in the  $\nu_\mu$  CC inclusive measurement. Therefore, the final binning used in this measurement is more coarse than the  $\nu_\mu$  CC inclusive measurement. Additionally, the template fit that is performed on the data and simulation, which is a different approach than that used by the  $\nu_\mu$  CC inclusive measurement, necessitates an even more coarse binning scheme than that of Figure 5.18, which is described in Section 5.7.

## 5.6. BACKGROUND ESTIMATION

To measure the cross section using Eq. 3, the number of background events contained in the selection must be estimated. This can be done by investigating the sidebands, or events outside of the signal region in the BDT response, however, this approach was not used for this analysis due to large contributions from the signal in the sideband region.

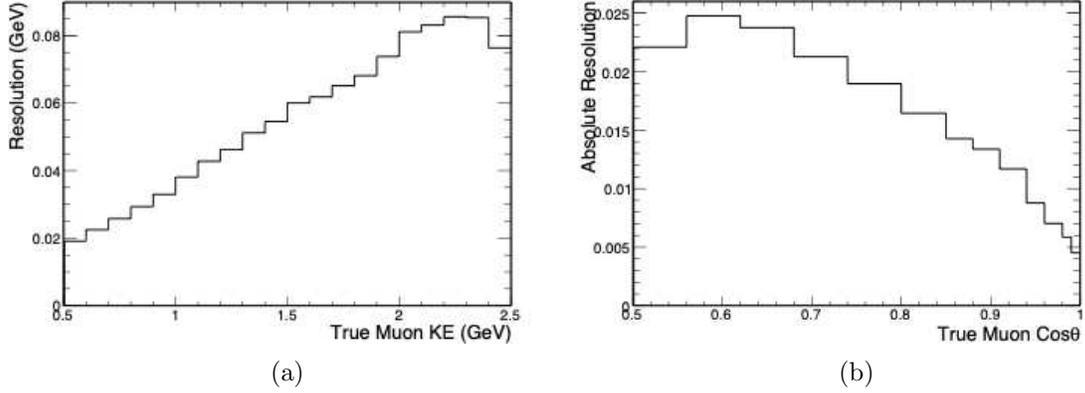


FIGURE 5.17. Absolute resolutions as a function of the true muon kinematic parameter using the semi-inclusive pion signal definition. The left pane shows the absolute energy resolution of the muon and the right pane shows the absolute resolution on the muon direction,  $\cos\theta$ , relative to the direction of the neutrino beam.

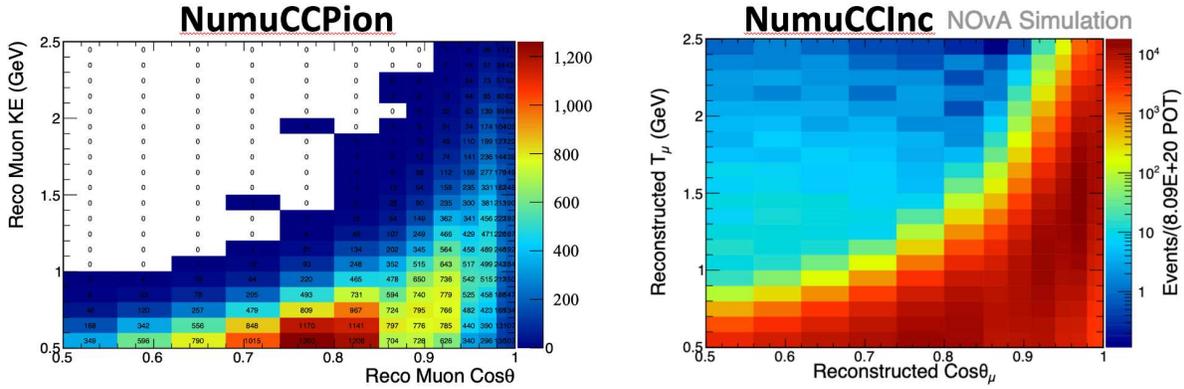


FIGURE 5.18. Distribution of events as a function of the reconstructed kinetic energy and reconstructed direction of the muon track, after the selection was applied to the simulated dataset using the finest possible binning based on the resolution of the kinematic distributions. In the left pane is the described distribution using the Inclusive CC Charged Pion selection and binning scheme. In the right pane is a described distribution using the Inclusive CC selection and binning scheme.[27]

5.6.1. BACKGROUND EVENTS. The three largest backgrounds in this analysis are CC  $0\pi^\pm$ , below threshold events, and NC DIS. The Best Pion Score Distributions for these events were shown previously in Figure 5.14. CC  $0\pi^\pm$  events can be further split into CCQE-like and CC  $\pi^0$ -like events. CCQE-like events are events with a muon and at least one proton, typically producing two or more tracks, however no pion is produced. CC  $\pi^0$ -like are events that

produce a muon, at least one proton, and at least one neutral pion, however, no charged pions are present. Typically, with these types of events, the muon is selected correctly and then the proton is selected as the pion track as it is the only other track in the event. If the proton isn't in the final state, the decay products of the neutral pion may be selected instead. Below-threshold events are background events where charged pions, with energy below where they can be reconstructed, were produced. There are also below threshold events where the pion did meet the energy requirements to be considered a signal pion, but the muon produced was outside of the accepted energy range or had too large of an angle with respect to the neutrino beam. NC DIS events have no primary muon but can produce multiple pions which can sometimes be miscategorized as a muon track. The miscategorization primarily comes if a produced pion has a similar  $dE/dx$  profile to a muon. Less commonly, this miscategorization can also occur if the pion decays quickly after the interaction, as the resultant muon can produce a track. There are additional interaction channels that contribute to the background, however at a much reduced rate, such as  $\nu_e$  and  $\bar{\nu}_e$  CC events and out of fiducial events, where events that occurred outside of the fiducial volume but were reconstructed with in the fiducial volume.

## 5.7. SIGNAL ESTIMATION USING A DATA-DRIVEN TEMPLATE FITTING PROCEDURE

Due to the large number of background events in the signal region and the large amount of signal events in the sideband region, a different approach was needed to estimate and constrain the signal. Two previous measurements in NO $\nu$ A [30][31] faced similar challenges and used a novel technique of fitting simulation to data utilizing the shape differences present between simulated background and signal distributions called templates. This technique was employed for this measurement using the PionID.

5.7.1. THE TEMPLATE FITTING PROCEDURE. The template fitting is performed for PionID templates in each of the 2D kinematic bins, in which the analysis will be presented. The fitting procedure was a global fit, where all kinematics bins are simultaneously fit, while minimizing the reduced  $\chi^2$ , given by:

$$(11) \quad \chi^2 = (x_i - \bar{x}_i)^T V_{ij}^{-1} (x_j - \bar{x}_j)$$

where  $i$  and  $j$  are template bins and  $x_i$  is the measured number of selected events from data in bin  $i$ ,  $\bar{x}_i$  is the expected number of selected events from simulation in that bin, nominal in MC, and  $V_{ij}$  is the total covariance matrix calculated as the sum of the total systematic covariance matrix and the statistical covariance:

$$(12) \quad V_{ij} = V_{ij}^{syst} + V_{ij}^{stat}$$

where  $V_{ij}^{stat}$  is the statistical variance and  $V_{ij}^{syst}$  is the total systematic covariance defined as the sum of the covariance matrices for each of the individual systematic samples indexed by  $z$ :

$$(13) \quad V_{ij}^{syst} = \sum_z V_{ij,z}$$

For the systematic samples that provide a single shift representing a change of one standard deviation of a parameter, such as the calibration, light, and Cherenkov samples, the covariance is calculated as

$$(14) \quad V_{ij,syst} = (s_i - \bar{s}_i)(s_j - \bar{s}_j)$$

For the multiverses systematic samples, such as PPFX and GenieMV, the average covariance was taken over all the universes using the Bessel correction:

$$(15) \quad V_{ij,syst} = \frac{\sum_{n=1}^U (s_{n,i} - \bar{s}_i)(s_{n,j} - \bar{s}_j)}{U - 1}$$

$\bar{x}_j$ , from Equation 11, can be written as a sum of the templates:

$$(16) \quad \bar{x}_j = a_i * N_i^{Signal} + b_i * N_i^{Background}$$

where  $a_i$  and  $b_i$  are the normalization free parameters in bin  $i$ , for the signal and background templates respectively.  $N_i^{Signal}$  and  $N_i^{Background}$  are the number of signal and background events in each template bin  $i$ .

The fitting and minimization procedure is outlined in Figure 5.19. First, templates were generated from the simulation and data for each kinematic bin, an illustration of a typical kinematics bin is shown in Figure 5.20. Next, a covariance matrix was calculated using all of the systematic samples. A global fit was then performed multiple times, each with a different seed value for the normalization parameters. The fit was performed using the CERN Minuit package[32], initially using fast fit function. The normalization seed values were varied from 0.2 to 1.8 using 0.2 increments. This fit was performed 64 times using these varying seeds to ensure that the smallest  $\chi^2$  was found. Best fits with large correlations between the normalization parameters, unphysical results, or errors were discarded, while other fit results were saved. If no convergence was found with the fit, the fit was discarded. Once the fast fits were performed on all 64 seeds, a slow, more accurate, fit was performed using the seed that resulted in the lowest  $\chi^2$  from the saved fast fits. This slower fit function yielded the most accurate minimum compared to the fast fit method which used a coarser minimization

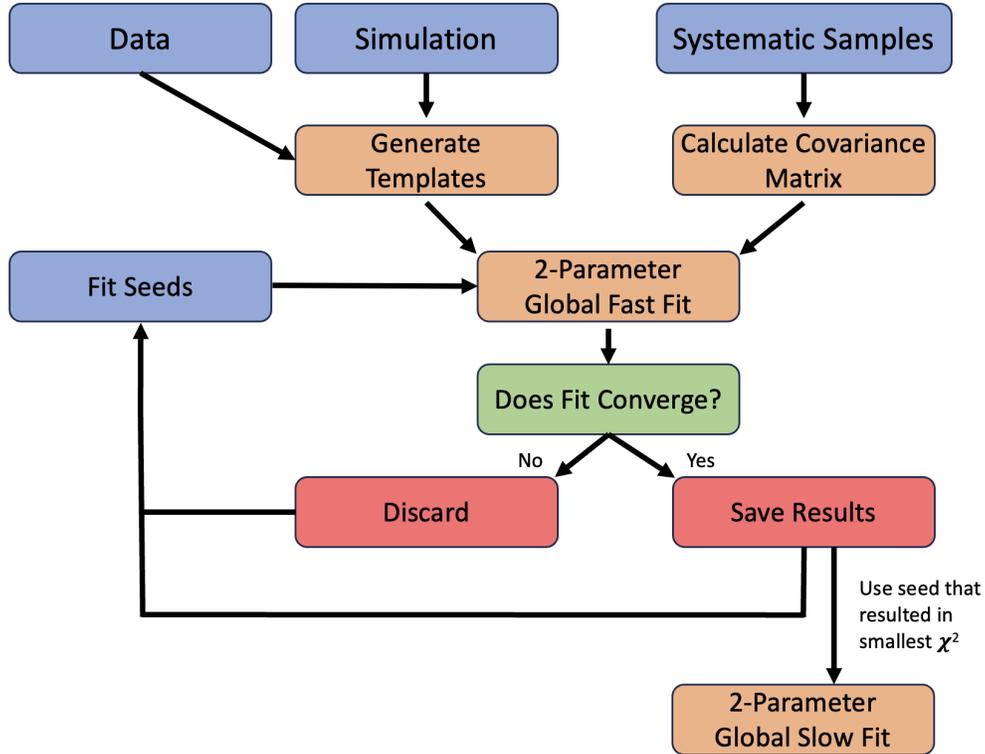


FIGURE 5.19. Cartoon illustrating the template fitting procedure workflow.

method. Once the slow fit was finished, the normalization parameters were saved along with the uncertainty in each bin from the fit.

In order to ensure sufficient statistics, and hence fit convergence, two requirements were imposed on each kinematic bin in order to be used in the global fit. Each kinematic bin must contain at least 400 signal events to ensure that the bin will not be statistically limited. Additionally, the signal to noise ratio in each kinematic bin must be at least 0.4. This requirement was necessary to ensure the signal template wasn't overwhelmed by the background template during fitting.

5.7.2. DETERMINING THE BINNING SCHEMES. Two binning schemes had to be determined, one for the templates and the other for the kinematics. To determine the best binning scheme for the templates, the fitting procedure was performed on different template binning

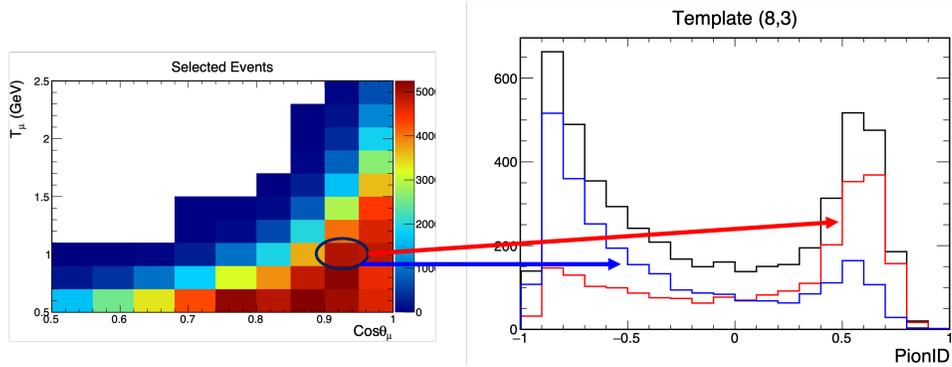


FIGURE 5.20. Example illustrating the signal and background templates coming from a kinematic bin.

schemes, while keeping the kinematic binning fixed. The template shape varies based on granularity of the binning. If the binning is too coarse, the shape distinction between signal and background can be lost. If the template bins are too fine, then the fit becomes sensitive to statistical fluctuations in the data. Four different binning schemes were tested, 10 bins, 20 bins, 50 bins, 100 bins, all using equal bin widths. The 20 bin scheme was found to be the optimal choice.

To determine the kinematic binning scheme, the template binning scheme was kept fixed at the optimal 20 bins, and the binning of the kinematic variables was varied. Initially, the kinematic binning scheme from Figure 5.18 was tested, but resulted in poor fit convergence. Coarser binning schemes were tested and the optimal 2D binning scheme was found. Figure 5.21 shows the number of selected signal events in each kinematic bin using the finalized binning scheme, after the application of the full selection on the simulated dataset.

5.7.3. TESTING FIT ROBUSTNESS. In order to ensure that the fitting procedure would function correctly on data samples, a series of robustness tests were performed. In each test, the nominal templates were fit to a fake data sample. These fake samples were: Poisson varied dataset, a flat 20% shift applied to the signal portion of the nominal simulation, and

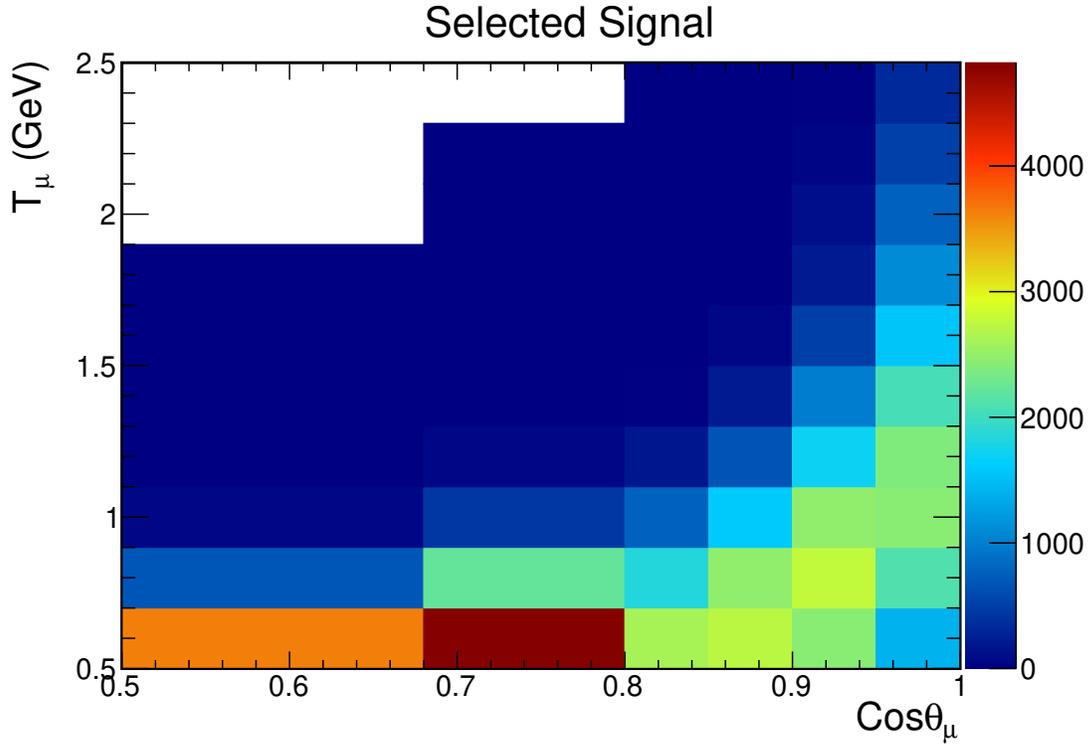


FIGURE 5.21. Distribution of selected signal events as a function of the true kinetic energy and scattering angle of the final-state muon, after the full selection is applied to the simulated dataset.

datasets with  $\pm 1\sigma$  of deviation for the signal portion of the nominal simulation by varying GENIE model parameters. Each of these tests resulted in fit convergence and produced fit uncertainties between 10% - 20% in most kinematic bins.

## 5.8. UNFOLDING

In an ideal detector, any measured physical value would match to the true physical value. In practice, what the detector measures will differ from what truly occurred for a few reasons, such as: measurement resolution, detection thresholds, non-linear detector response, and imperfect reconstruction algorithms. These effects are said to “smear” the distributions of the true physical quantities into the distributions the detector measures. Unfolding is the act of reversing this smearing effect from the measured distributions to obtain the true physical

values. The true physical quantities are of the most importance for the reported result as that allows comparisons between different experiments and theory. For the measurement described in this thesis, the physical quantities of interest are the discrete distributions of the kinetic energy and scattering angle of the final state muon. The initial folded measurement can be described by a matrix transformation of the true values of the discrete kinematic distributions into the reconstructed values of the kinematic distributions[33]:

$$(17) \quad \mu = Ax + b$$

where  $\mu$  is the observed distribution by the detector,  $x$  is the true underlying signal distribution,  $b$  is the estimated background distribution, and  $A$  is the transformation matrix, also called the smearing matrix. Using simulation, where both the observed distributions and the true underlying distributions are known, this smearing matrix can be estimated. In data however, only the observed distribution can be measured. Using Equation 17, and assuming  $A$  is both square and not singular, the true distribution can be written as:

$$(18) \quad x = A^{-1}(\mu - b)$$

Unfortunately, this alone is not enough to recover the true distributions as this method assumes simple linear effects from the transformation, while in reality detector effects can be non-linear. Instead, a probabilistic approach is taken, using Bayes' theorem to iteratively unfold the distribution in multiple steps[34]. This technique is most commonly called either the D'Agostini method or iterative Bayesian unfolding. Using this technique, the iterative improvement on the measured result in each bin of the distribution,  $x_j^{(n+1)}$ , relative to the

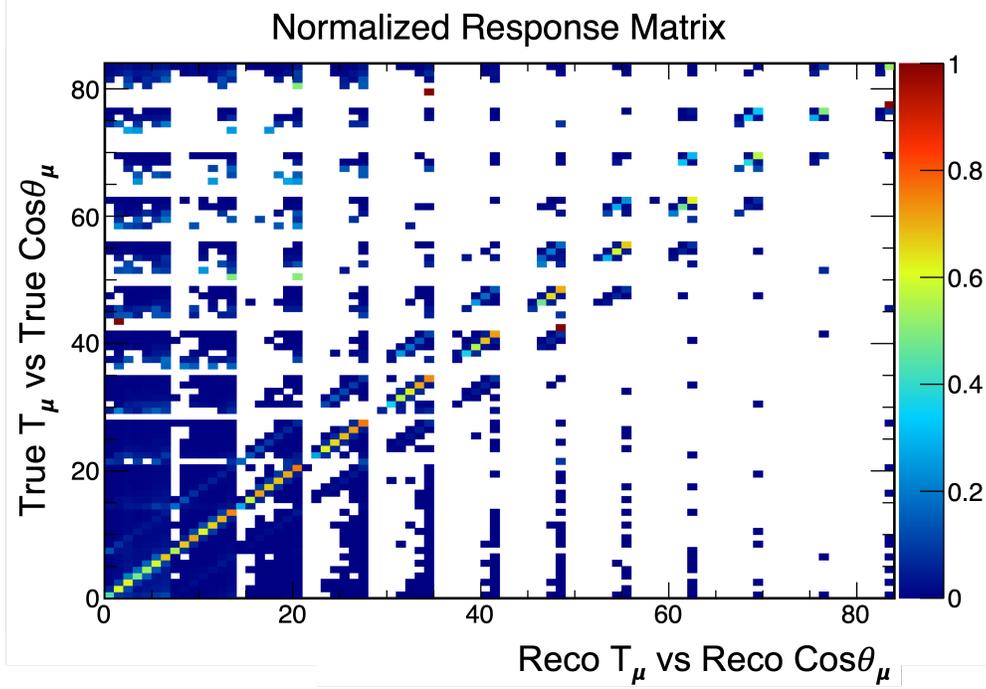


FIGURE 5.22.

previous iterative improvement,  $x_j^n$ , can be written as:

$$(19) \quad x_j^{(n+1)} = x_j^n \sum_{i=1}^M \frac{A_{ij}}{\varepsilon_j} \frac{y_i}{\sum_{k=1}^N A_{ik} x_k^{(n)}}$$

where  $\varepsilon_j$  is the selection efficiency of the  $j$ th bin. A plot of the normalized response matrix,  $A$ , for this measurement is shown in Figure 5.22, with the true kinematic quantities on the y-axis and the reconstructed quantities on the x-axis.

5.8.1. OPTIMIZING THE NUMBER OF UNFOLDING ITERATIONS. Because this method is recursive, the optimal number of iterations used in the unfolding procedure must be determined. With D’Agostini’s method, small numbers of iterations bias the unfolding towards the true distribution, while large numbers of iterations result in large variances in the measured distribution. Therefore, to determine the optimal number of iterations, one must consider the effects of bias and variance. The mean squared error (MSE) was chosen as the

optimization metric, with the iteration having the smallest MSE being the desired number of iterations to use for the final measurement. This metric has been used in other cross section measurements on  $\text{NO}\nu\text{A}$  successfully[35][27]. The MSE is defined as:

$$(20) \quad \text{MSE} = \text{Var} + \text{Bias}^2$$

where Var is the variance of the unfolded sample, and the bias is the difference between the unfolded sample and the true distribution. The MSE can be calculated using simulation as the true and reconstructed distributions are available. To test the unfolding and calculate the MSE, a statistically independent simulated sample with varied GENIE model parameters was used. This was to ensure that the unfolding procedure would still perform well despite any differences between data and simulation. Three-hundred simulated datasets were produced by randomly varying all of the GENIE model parameters. Six specific datasets were chosen for this procedure: the dataset with the largest bin-to-bin variation relative to the standard simulation, the dataset with the largest integral difference relative to the standard simulation, and 4 datasets that represented a  $+1\sigma$ ,  $+2\sigma$ ,  $-1\sigma$ , and  $-2\sigma$  of deviation from the nominal simulation. The variance, squared bias, and MSE as a function of the number of unfolding iterations are shown in Figure 5.23 for the  $+1\sigma$  varied sample. The variance is shown in the top left, the bias on the top right. The results from each of these tests are shown in Table 5.2. Most of the tests yielded a minimum MSE value at the same number of iterations, 2, except for one test, the  $+2\sigma$  deviation which yielded a minimum at 3 iterations, however, the MSE found at 2 and 3 iterations for that sample were similar in value. Therefore two iterations was chosen as the optimal number of iterations to perform the unfolding. This is

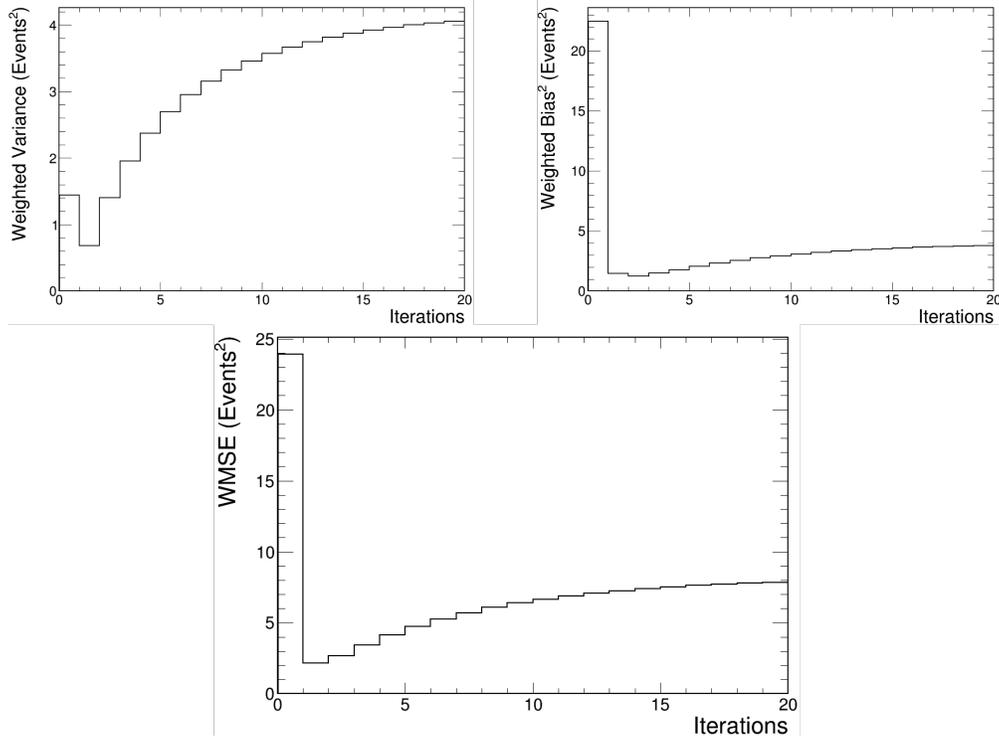


FIGURE 5.23. Variance, Bias<sup>2</sup>, and MSE as a function of the number of iterations for the +1 $\sigma$  varied sample.

similar to the number of iterations used in other NO $\nu$ A measurements, where typically 2-4 iterations grant the smallest MSE and are therefore optimal.

TABLE 5.2. Optimized number of unfolding iterations.

Simulation Dataset	Optimized # of Iterations
Largest Bin-to-Bin Difference	2
Largest Integral Difference	2
+1 $\sigma$	2
+2 $\sigma$	3
-1 $\sigma$	2
-2 $\sigma$	2

## 5.9. EFFICIENCY CORRECTION

In order to correctly evaluate a cross section, corrections must be made to account for inefficiencies in the measurement selection. The efficiency, along with the associated uncertainty, can be estimated using simulation. The efficiency is applied after the data has been

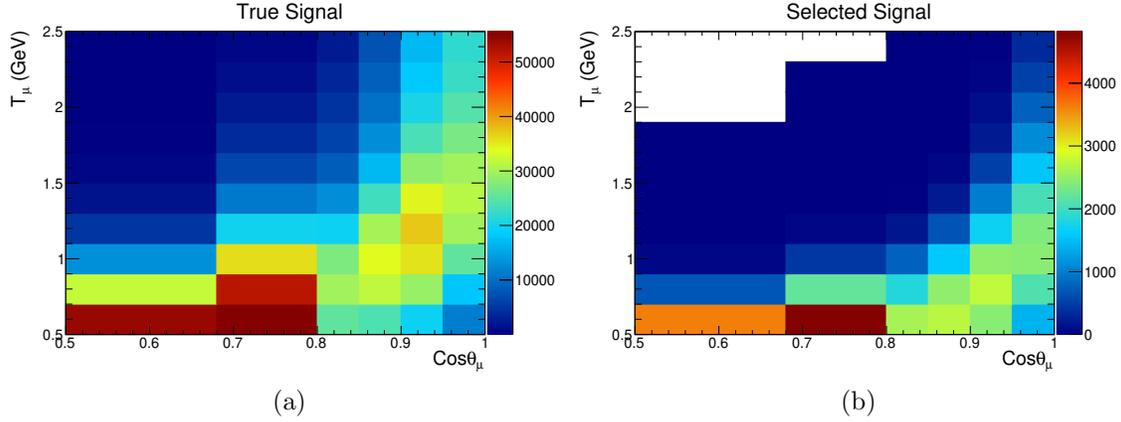


FIGURE 5.24. Distribution of signal events as a function of the true kinetic energy and direction of the produced muon. This plot contains all simulated signal events prior to the application of the selection to the simulated dataset.

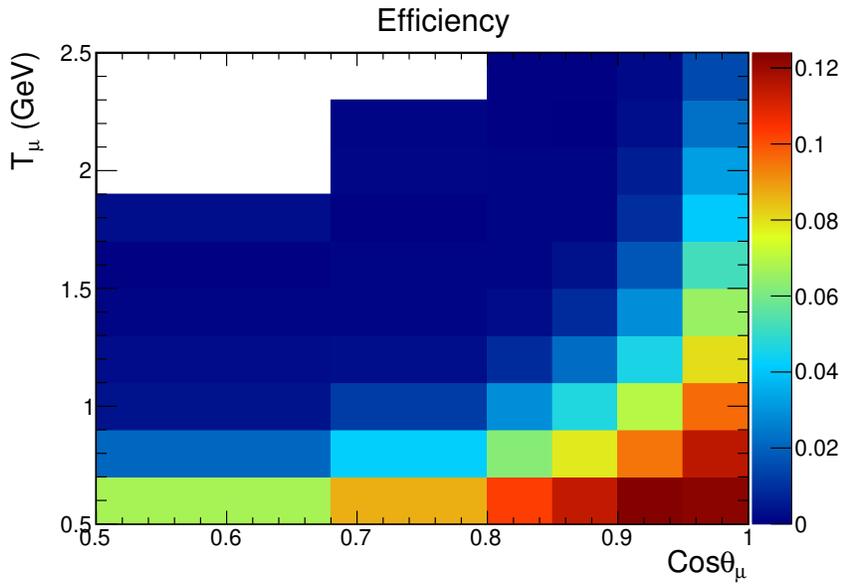


FIGURE 5.25. Efficiency as a function of the kinetic energy and scattering angle of the final-state muon.

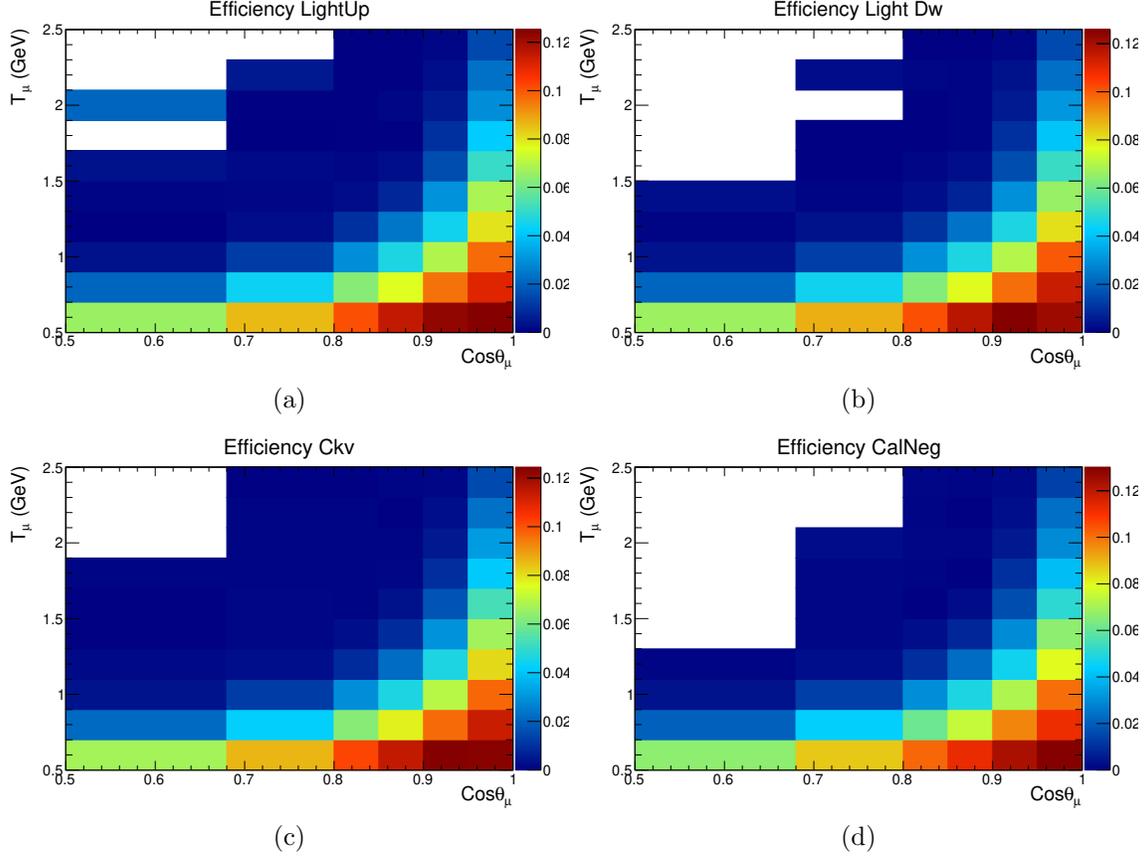
unfolded into the true distributions. To estimate the efficiency, Equation 4 is used. The left pane of Figure 5.24 shows the number of selected events in each bin from the full selection using the simulation, while the right pane shows the true number of signal events in each bin prior to selection using the simulation. Figure 5.25 shows the total estimated efficiency calculated from simulation.

5.9.1. UNCERTAINTY ON THE EFFICIENCY CORRECTION. There are three primary sources of uncertainty in the calculation of efficiency: the statistical uncertainty coming from the number of events in the measurement; systematic uncertainties from the beam and detector modelling; and the systematic uncertainties from parameters in the physics models used in the simulation. The total uncertainty assigned to the efficiency is obtained by adding the individual uncertainty contributions in quadrature. Correlations are ignored. The individual uncertainties in each bin coming from detector effects are shown in Figure 5.26. The uncertainty from GENIE and PPFX modelling is shown in Figure 5.27. The top panes show the upper and lower bounds of the uncertainty due from GENIE model parameters and the bottom panes show the upper and lower bounds on the uncertainty due to PPFX model parameters. The total fractional uncertainty assigned to the efficiency is shown in Figure 5.28.

## 5.10. FLUX CALCULATION

The integrated  $\nu_\mu$  flux is needed for the cross section measurement. The beam flux, in general, describes how many particles are incident on the detector per area as a function of neutrino energy. Therefore, the integrated flux is the total number of  $\nu_\mu$  that passed into the fiducial volume per unit time and integrated over the data collection period corresponding to the data set. The measurement described in this thesis uses the value calculated for the inclusive CC measurement. Using Equation 5, the integrated flux is defined as:

$$(21) \quad \phi(E) = \frac{N^{Signal}(E)}{N^{Target}\sigma(E)}$$



where  $N^{Signal}(E)$  is the number of signal events as a function of the neutrino energy,  $N^{(Target)}$  is the number of targets in the fiducial volume, and  $\sigma(E)$  is the cross section as a function of the neutrino energy. To estimate the integral flux with this equation, GENIE is used, simulating quasi-elastic  $\nu_\mu$  CC events interacting on carbon, as the detector is largely composed of carbon[36]. This is used in conjunction with an internal toolkit called PPFX (Package to Predict the Flux) which contains the uncertainties from the NuMI beam modelling and data from other sources such NA61 experiment[37] at CERN. The uncertainty associated with the predicted flux is calculated by producing 100 multiverses and measuring the  $1\sigma$  deviation from nominal. Figure 5.29 shows the estimated flux as a function of neutrino energy with the associated uncertainties. From this method, the estimated integral  $\nu_\mu$  flux is  $7.801^{37} \text{ m}^{-2} \pm 11.4\%$ .

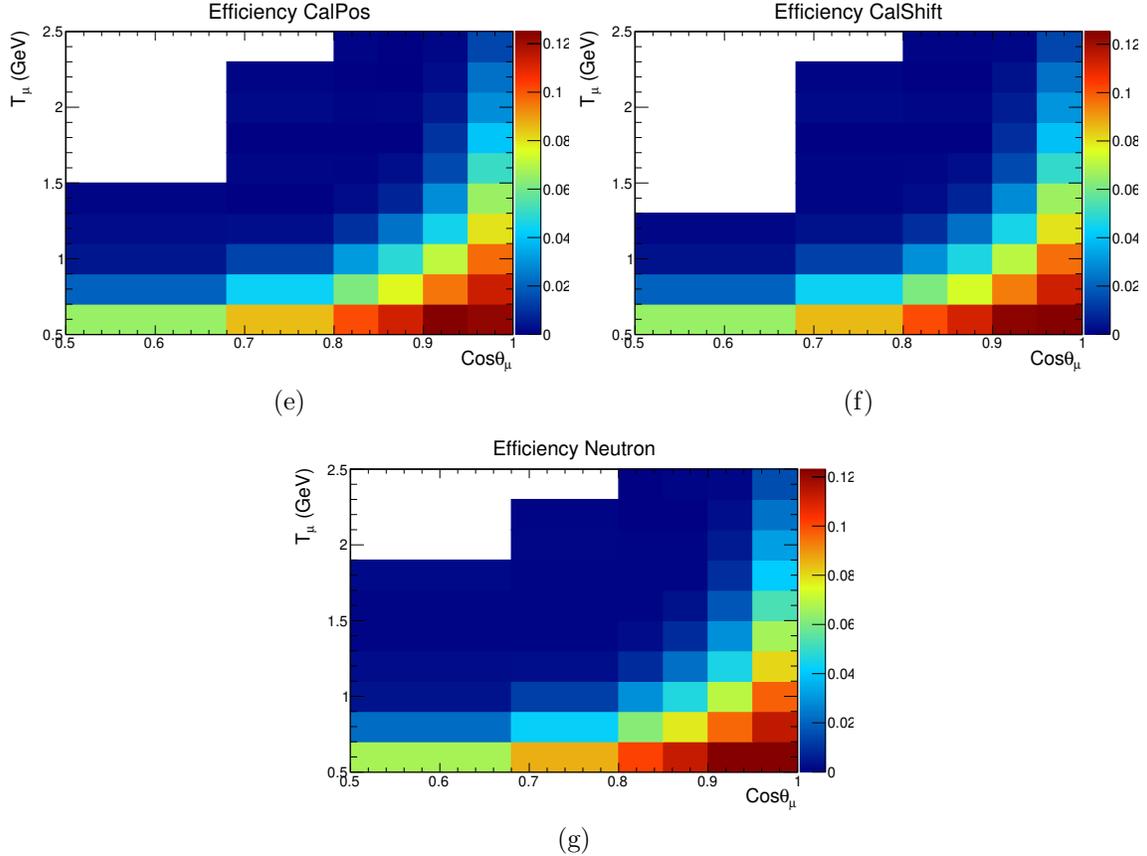


FIGURE 5.26. The individual systematic uncertainties on the efficiency per bin associated with detector model parameters.

### 5.11. DETERMINATION OF TARGET NUCLEON NUMBER

The total number of nucleons was calculated for the cross section estimation, allowing for comparison with other experiments. The number of target nucleons in the fiducial volume was calculated. The number of nucleons estimated in the previous  $\nu_\mu$  CC inclusive measurement was used for the measurement described in this thesis, as both measurements used the same fiducial volume. This number was estimated by randomly sampling the simulated detector geometry to determine the element at a specific point, and thus the number of nucleons, one million times. This process was repeated one million times and a Gaussian curve was fit to the distribution[35]. Figure 5.30 shows the distribution of the measured number of

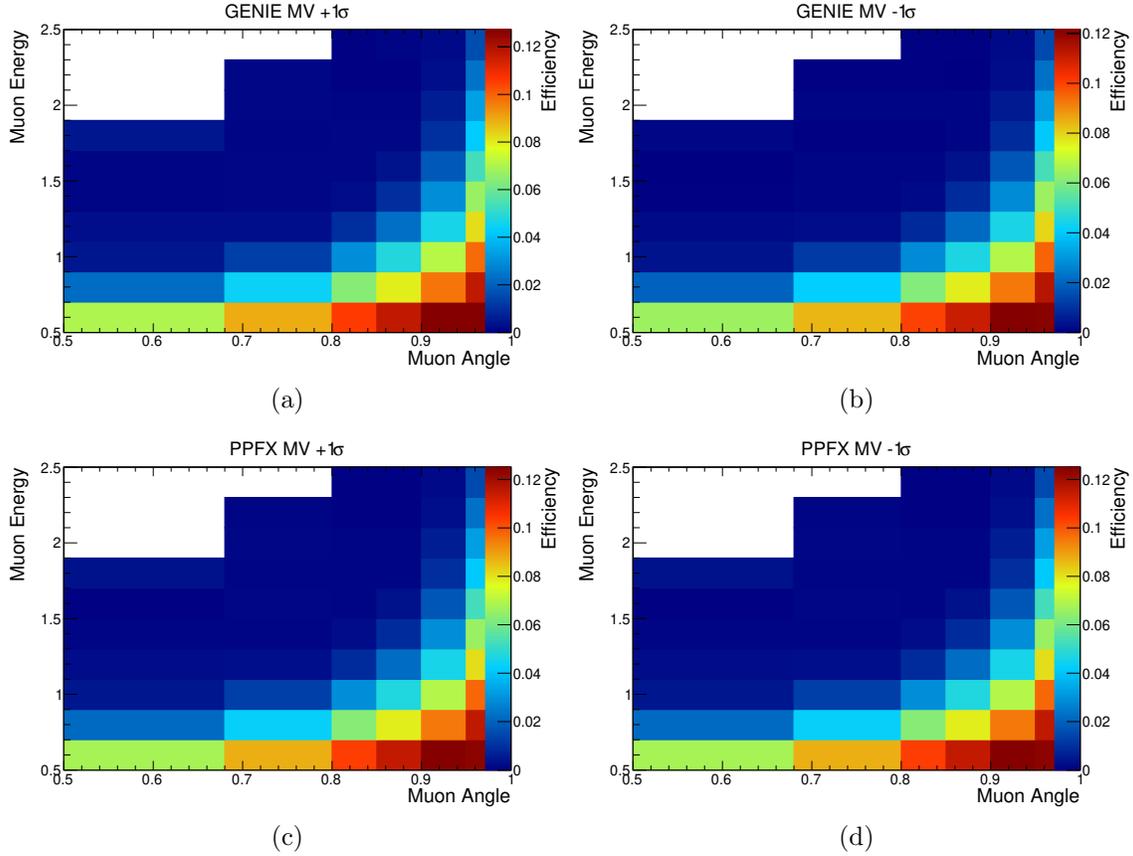


FIGURE 5.27. Uncertainty per bin associated with GENIE and PPFX model parameters.

nucleons from one million trials along with the mean and uncertainty from the fit. The total estimated number of target nucleons in the fiducial volume is  $3.890 \times 10^{31} \pm 0.33\%$  with the error being statistical in nature due to the random sampling process.

## 5.12. UNCERTAINTY CALCULATION

Each term in the cross section equation, Equation 6, has an associated uncertainty that needs to be calculated in order to determine the total uncertainty on the cross section measurement. The uncertainties for each of those quantities were given in the relevant sections above. To calculate the total uncertainty on the final measurement each quantity will be treated as independent and uncorrelated to one another, therefore the uncertainties

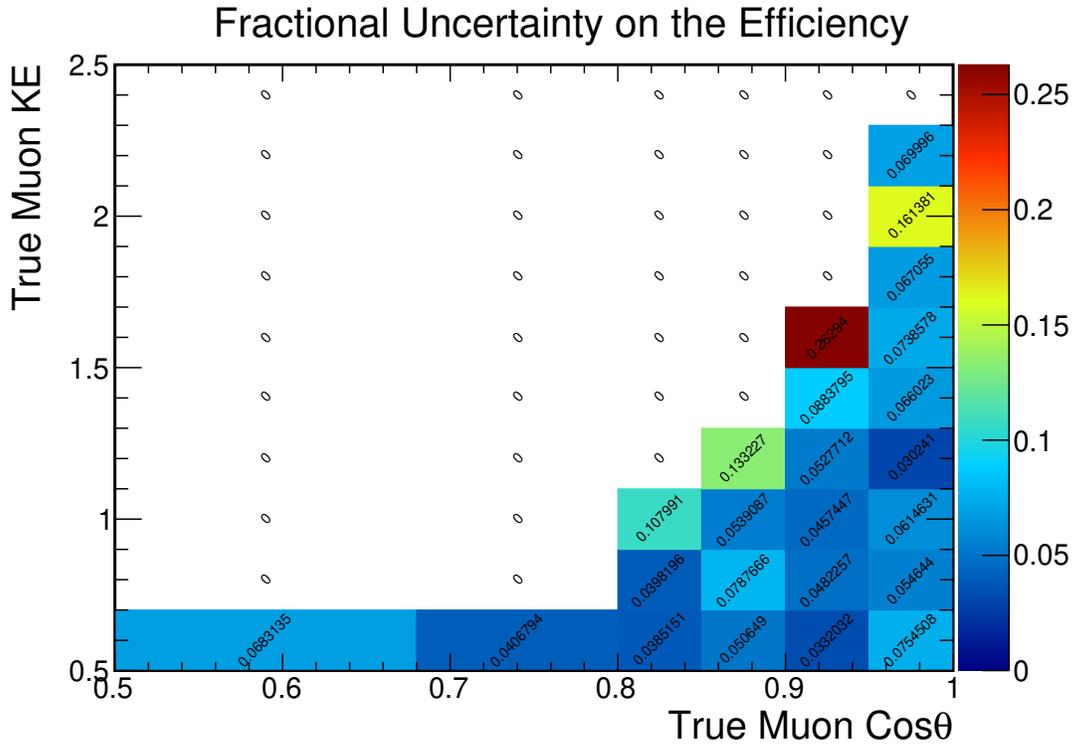


FIGURE 5.28. The total uncertainty on the efficiency per kinematic bin. This includes the uncertainties from detector, beam, flux, and physics model parameters, and the statistical uncertainty.

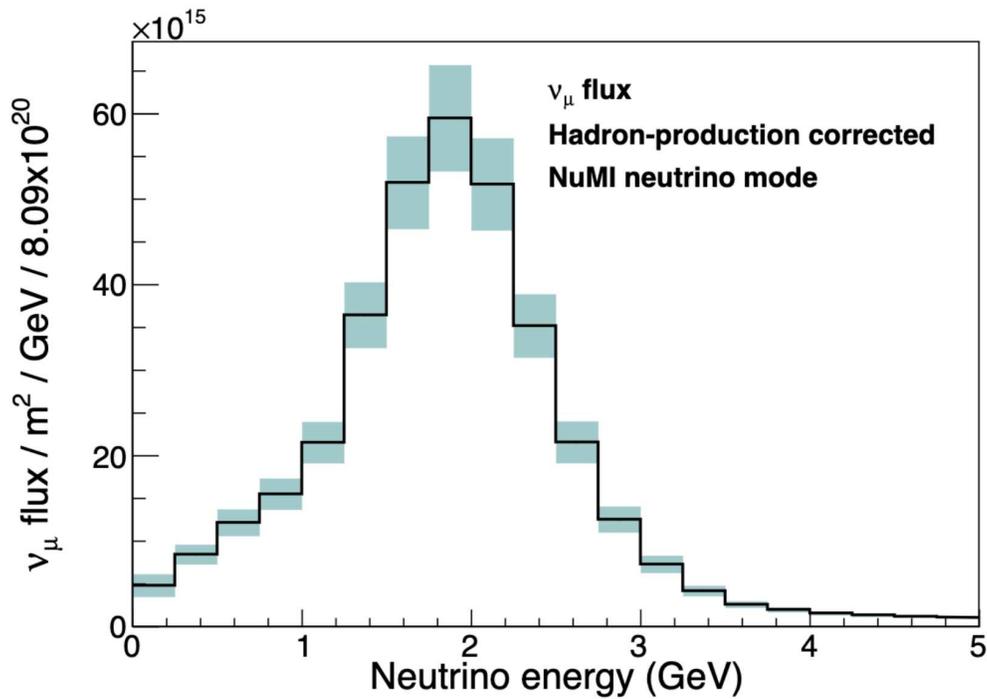


FIGURE 5.29.  $\nu_\mu$  flux as a function of neutrino energy.[35]

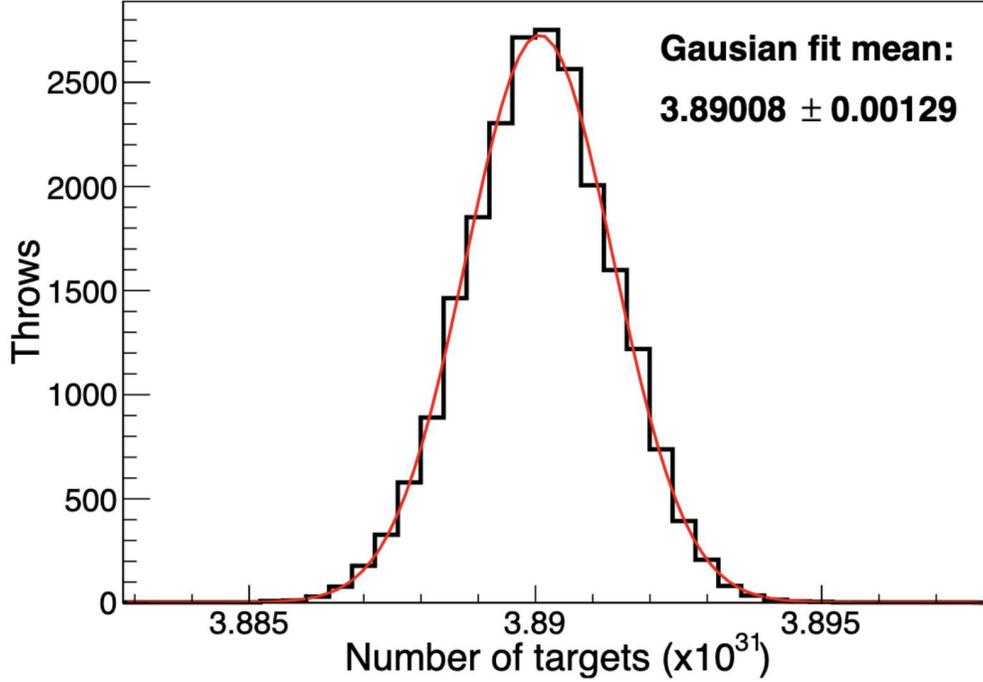


FIGURE 5.30. Distribution of the estimated number of targets in the fiducial volume from the simulation.[35]

will be added in quadrature using the following equation:

$$(22) \quad \frac{\delta_\sigma}{\sigma} = \sqrt{\sum \frac{\delta_{x_i}}{x_i}}$$

where  $\frac{\delta_\sigma}{\sigma}$  is the fractional uncertainty on the cross section and  $\frac{\delta_{x_i}}{x_i}$  is the fractional uncertainty of parameter  $x_i$  from the cross section equation.

### 5.13. MOCK DATA TEST

Before unblinding the analysis, it is imperative to test the entire analysis framework using a statistically independent dataset of Monte Carlo simulation. This sample, called “fake data”, was made using the nominal simulation, but was 25% the size of the full simulation dataset.

5.13.1. **MOCK DATA CROSS SECTION.** The procedure used for the fake data measurement was identical to the procedure used on the data for the final measurement. First, the templates from the nominal simulation were fit to the fake data. The fit converged, as can be seen in the extracted signal distribution in Figure 5.31. The extracted distribution of signal events was then unfolded into true space, yielding Figure 5.32. This unfolded distribution was then efficiency-corrected, by dividing each bin in the unfolded distribution by the corresponding bin in the efficiency plot shown in Figure 5.25, yielding the signal distribution in Figure 5.33. Each bin is then divided by the integrated flux, the number of targets, and the bin width, giving the double-differential cross section, shown in the right pane of Figure 5.34. For comparison, the double-differential cross section, as calculated from the nominal simulation is shown in the left pane of Figure 5.34. Figure 5.35 shows the ratio of the calculated cross section extracted from the fake data relative to the cross section used in the nominal simulation in each bin. The bins in the forward-going region,  $\cos\theta_\mu > 0.95$ , match the nominal simulation quite well, however, at larger muon scattering angles, there is some difference between the two samples. This is because the shape difference between signal and background in PionID is diminished in those regions as compared to the forward-going bins.

5.13.2. **MOCK DATA CROSS SECTION UNCERTAINTY.** The resultant fractional uncertainty per bin on the fit is shown in Figure 5.36. Each bin was then added in quadrature with the uncertainties from: the efficiency, the flux, and the number of targets. The fractional uncertainty on the extracted cross section from the fake data sample is shown in Figure 5.37.

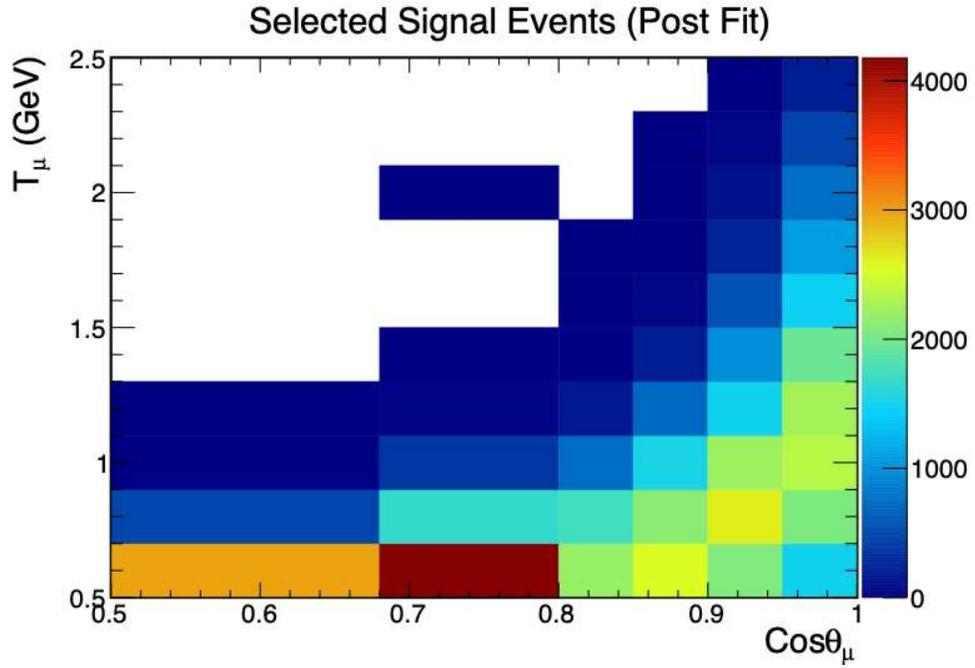


FIGURE 5.31. Distribution of signal events from the fake data sample using the template fitting procedure.

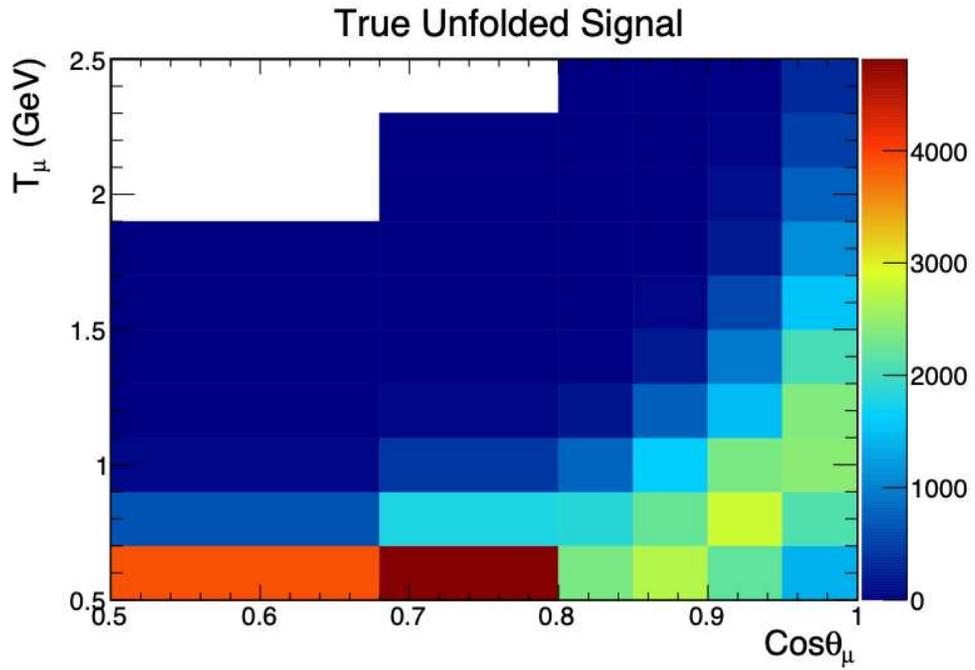


FIGURE 5.32. Distribution of unfolded signal events from the fake data sample.

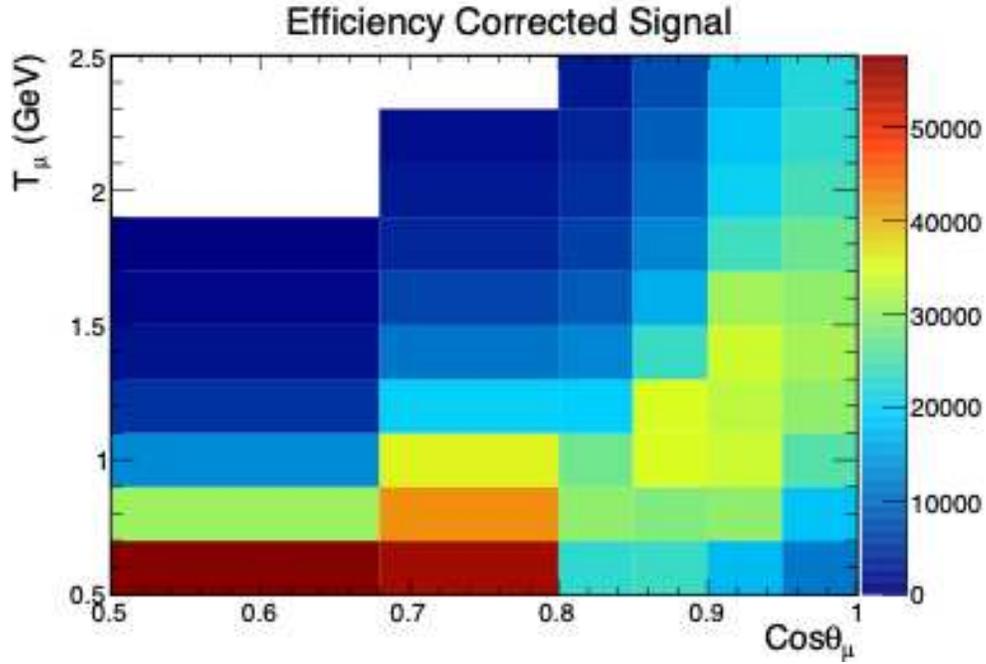


FIGURE 5.33. Distribution of signal events from the fake data sample after the efficiency correction had been applied.

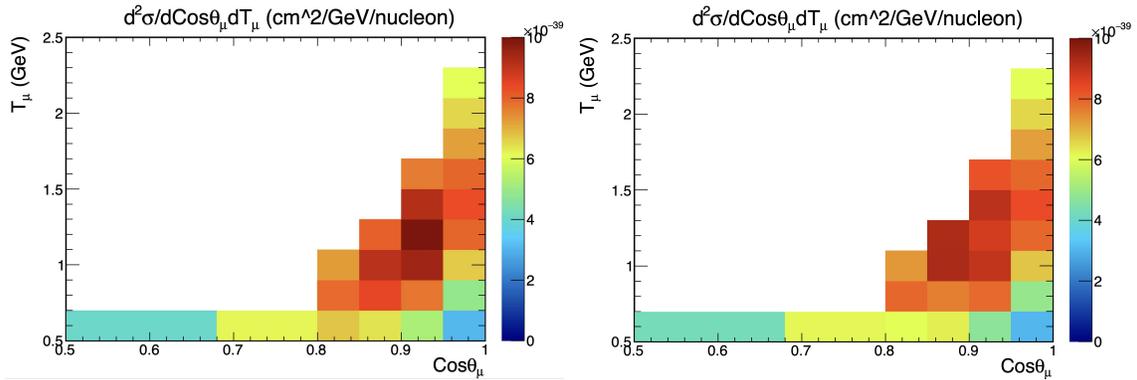


FIGURE 5.34. The double differential cross section from the nominal simulation in the left pane and the double differential cross section calculated from the mock data test on the right.

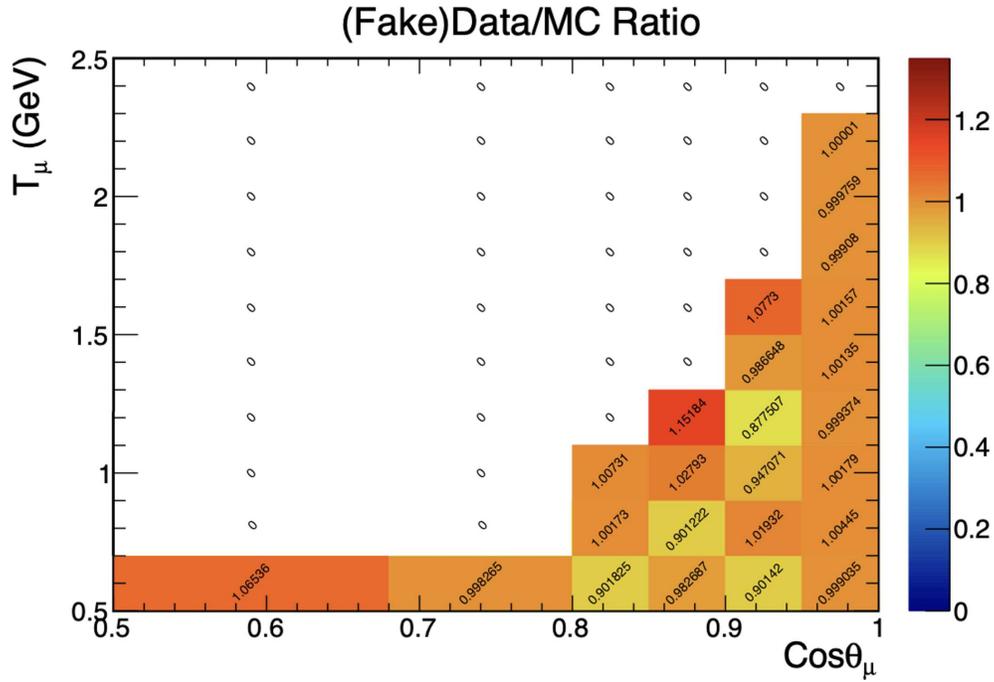


FIGURE 5.35. Ratio of the fake data cross section divided by the nominal simulation cross section in each bin.

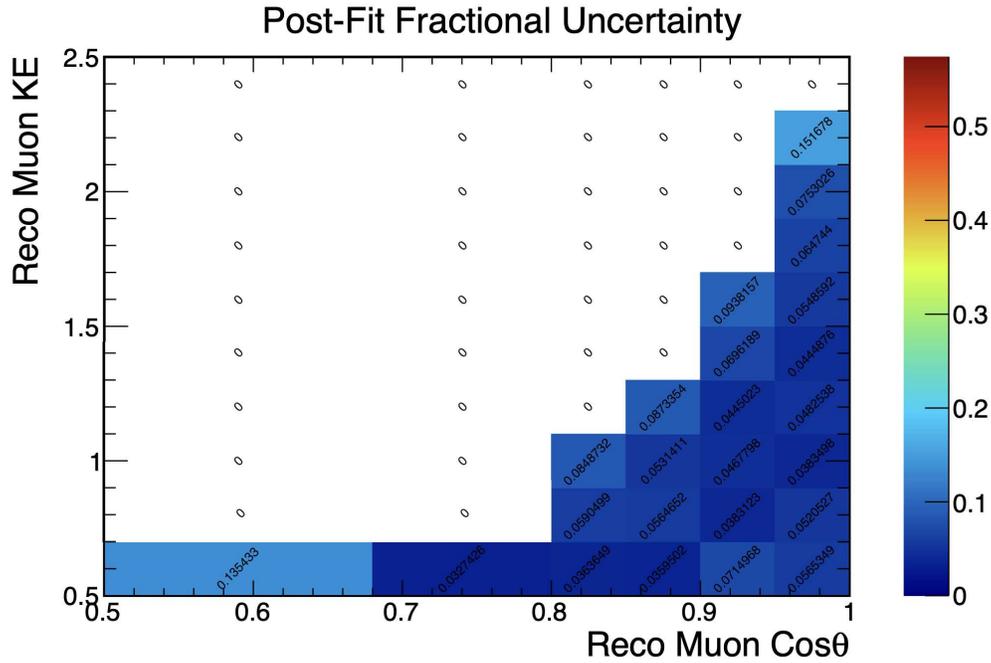


FIGURE 5.36. Fractional uncertainty in each bin from the template fitting procedure performed on the fake data sample.



## CHAPTER 6

# RESULTS

The cross section measurement process described in the previous chapter has been applied to NO $\nu$ A 's Near Detector dataset, with data collected from August 2014 to February 2017, while the NuMI beam was running in the FHC configuration. Prior to this point, the analysis and measurement had been blinded from data.

### 6.1. FIT RESULTS

Figure 6.1 shows the distribution of the number selected events on the data. The templates for three bins used in the fitting procedure can be seen in Figure 6.2, containing the PionID distributions of the pre fit templates, the post fit templates with normalization parameters applied, as well the data distribution for each bin. Only the reported bins, bins passing the 400 signal events and 0.4 signal to background ratio requirements, are shown in this figure. Each plot is labelled with a bin number in (x,y) coordinates, with (1,1) corresponding to the bottom left bin in Figure 6.1 and (6,10) corresponding to the top right bin in that same plot. These figures also contain the sum of the background and signal templates, in black, which is the equivalent to the total predicted number of events in each bin. The complete set of templates, for all bins, is in Appendix A.

### 6.2. MEASURED CROSS SECTION

Following the fitting procedure, the cross section was calculated. The cross section was calculated by applying the normalization parameters from the fit to the templates, unfolding that distribution, and applying the efficiency correction, flux, and number of targets to calculate the cross section. In detail, this procedure was:

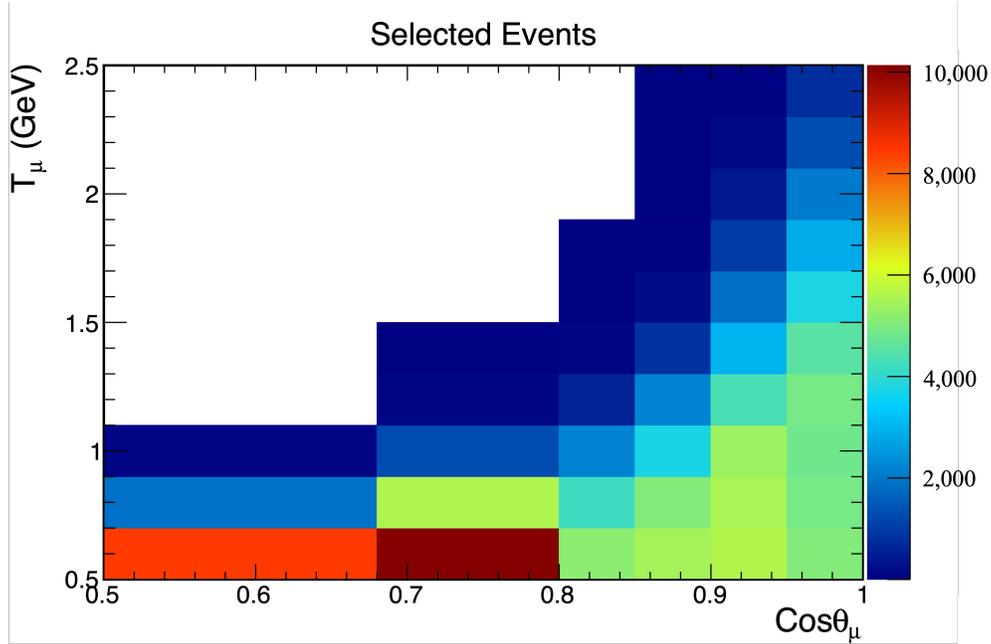
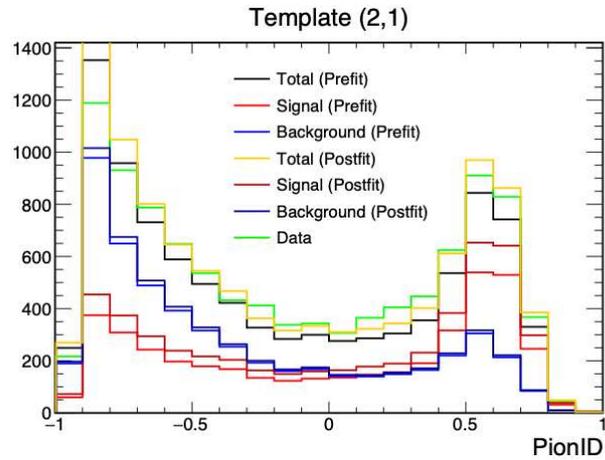
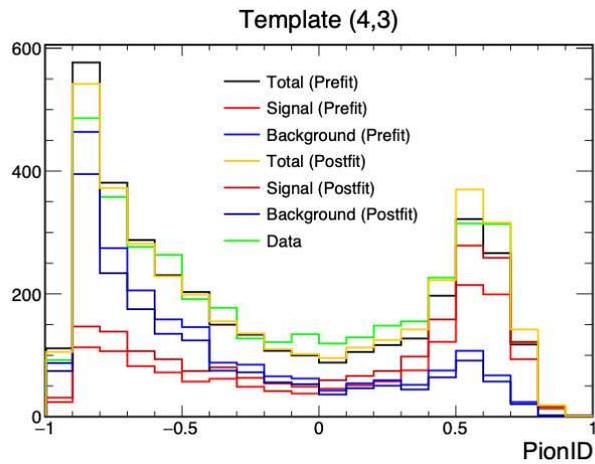


FIGURE 6.1. Distribution of events from the data after the full selection has been applied.

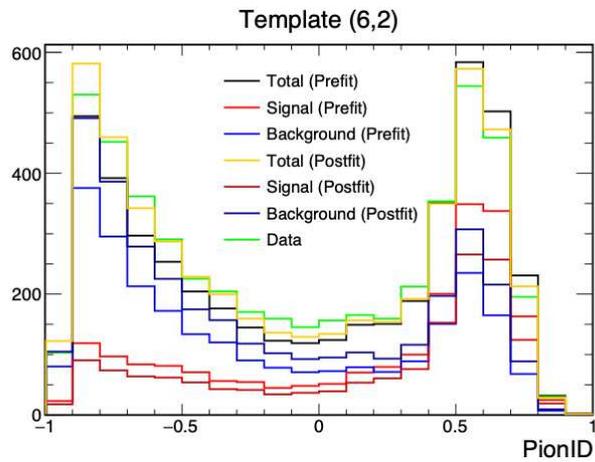
- (1) The normalization parameters from the template fitting procedure on the data were applied to the signal templates. This yielded the estimated number of signal events per kinematic bin in the data, shown in Figure 6.3.
- (2) The unfolding process was applied to the distribution of signal events from the previous step, using two iterations. The unfolded distribution is shown in Figure 6.4.
- (3) The efficiency correction was applied. Each bin of the unfolded distribution was divided by the efficiency in that bin from 5.25. The efficiency correction distribution of signal events is shown in Figure 6.5.
- (4) Each bin of efficiency corrected distribution was divided by: the number of targets, the integrated flux, the kinetic energy bin width, and the scattering angle bin width. This step yielded the complete double-differential cross section.



(a)



(b)



(c)

FIGURE 6.2. Data template, in orange, along with the pre and post fit templates for signal and background. The simulated estimation of the data is shown in black.

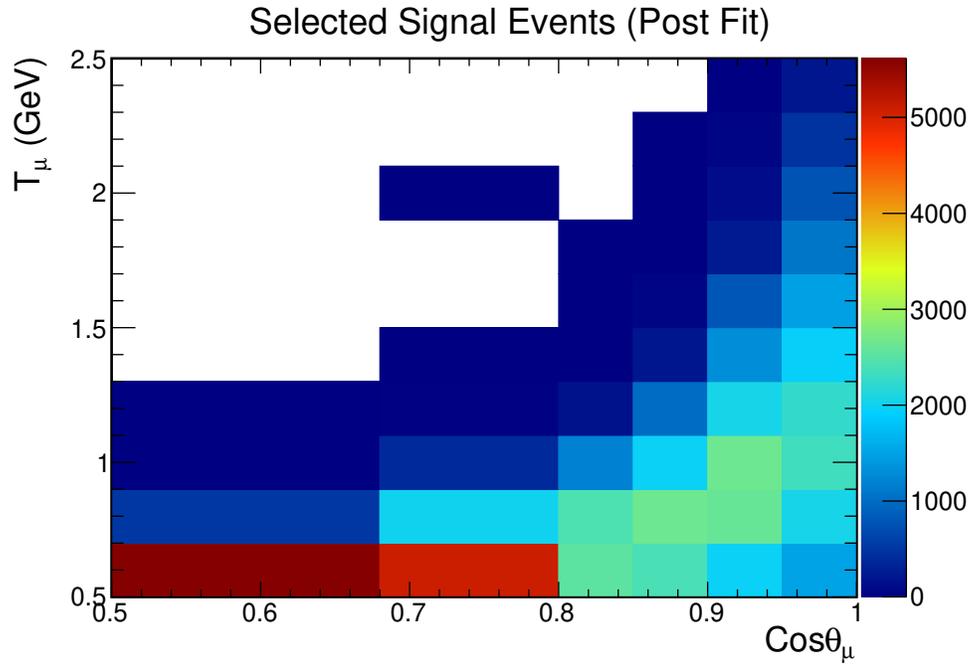


FIGURE 6.3. Distribution of the number of signal events extracted from the template fitting procedure performed on the data.

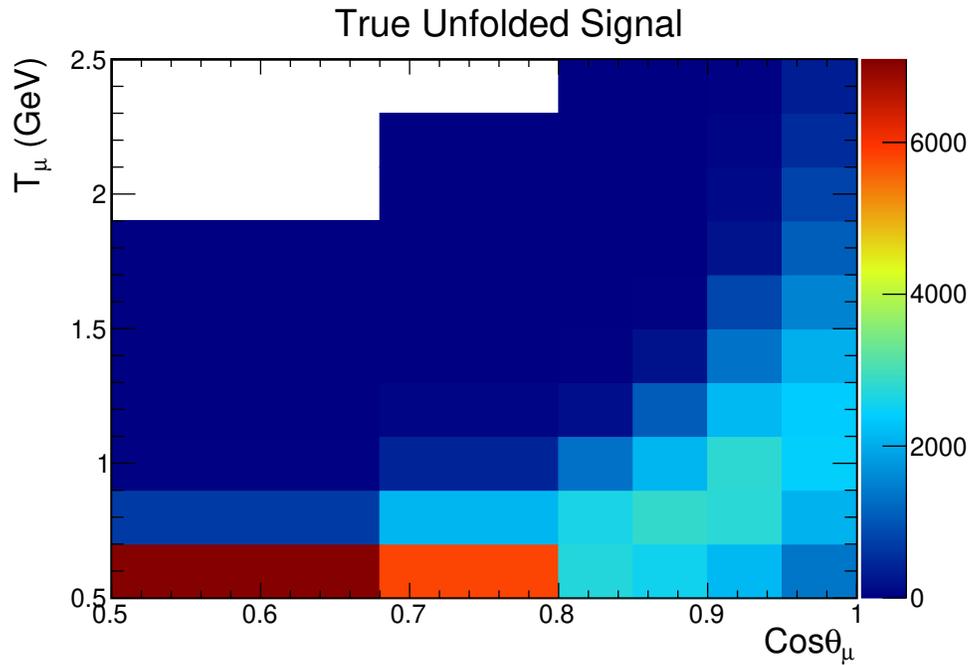


FIGURE 6.4. Distribution of the measured number of signal events after performing the unfolding procedure.

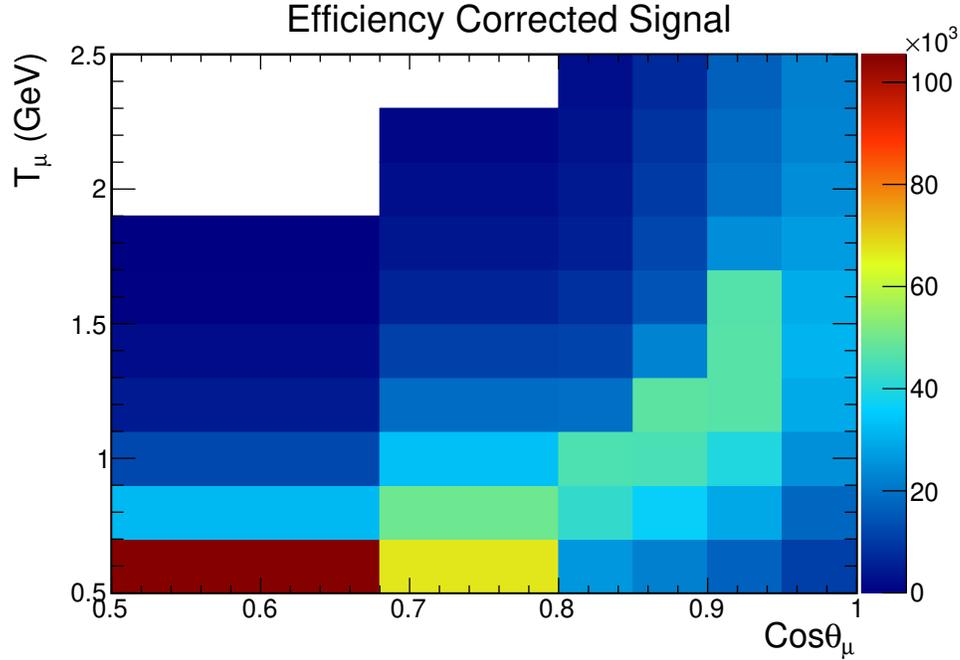


FIGURE 6.5. Distribution of the measured signal events after the efficiency corrections have been applied.

The final measured double-differential cross section for the  $\nu_\mu$  CC  $\pi^\pm$  channel is shown in Figure 6.6 (only the reported bins). Figure 6.7 shows the ratio of the measured cross section from data relative to the simulation. The measured double-differential cross section is similar to the simulated prediction in the forward-going region,  $\cos\theta_\mu > 0.95$ . At larger scattering angles, an excess in the measured cross section relative to the simulated prediction was observed. The excess is larger than the uncertainty in most bins. Events in the forward-going region tend to have simple topologies, such as coherent processes or one muon and one charged pion final state events which are well understood processes due to previous measurements. Events containing muons with larger scattering angles,  $\cos\theta_\mu < 0.95$ , tend to have more complex topologies, such as multi-pion DIS, which are difficult to measure and aren't as well understood. Future NOvA analyses of this final state will be performed with

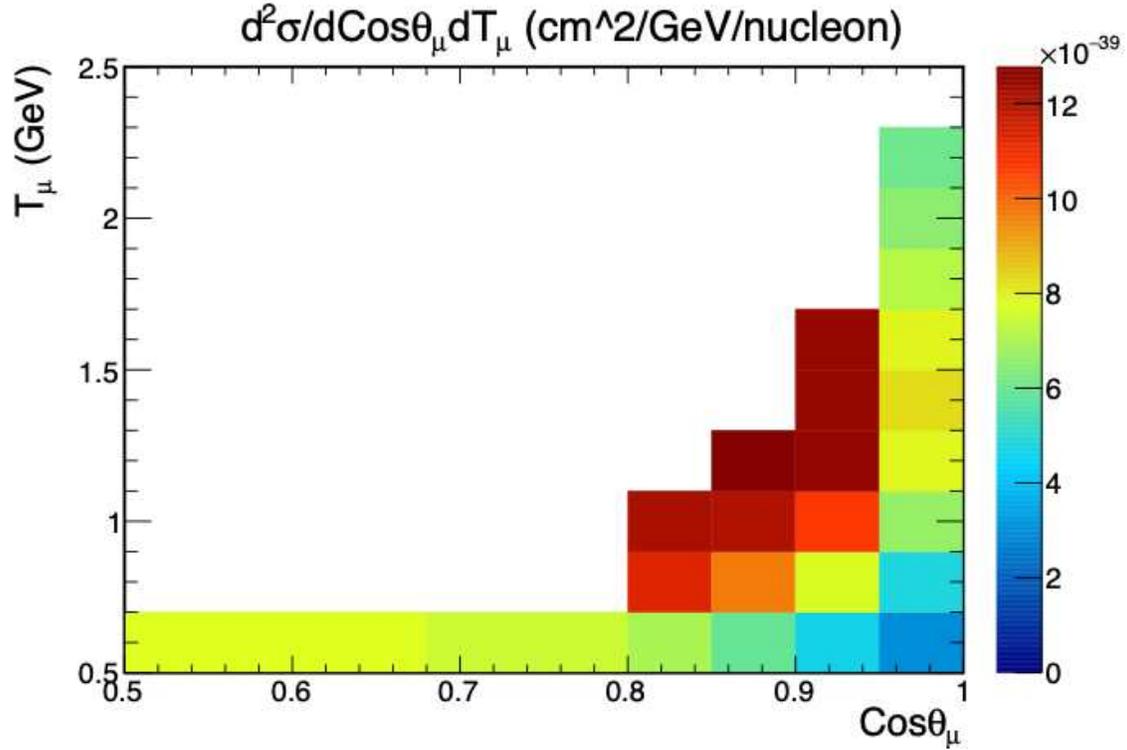


FIGURE 6.6. The measured double differential cross section of the  $\nu_\mu$  CC  $\pi^\pm$  channel.

more sophisticated pion reconstruction and energy estimation tools, which will shed light on the observed differences.

### 6.3. UNCERTAINTY ON THE MEASUREMENT

The resulting uncertainty from the fit is shown in Figure 6.8. This was used in combination with the uncertainties described in the previous chapter to calculate the total uncertainty on the cross section, shown in Figure 6.9. Most total fractional uncertainties were between the values of 10% to 20%. There were two bins with values outside of that range. The bin with the largest scattering angle had the largest fractional uncertainty which came from a large fit uncertainty. The large fit uncertainty in that bin had also been seen in the mock data test and is due to the discrimination between signal and background not being as strong as the rest of the bins. The second largest uncertainty is in the kinematic space

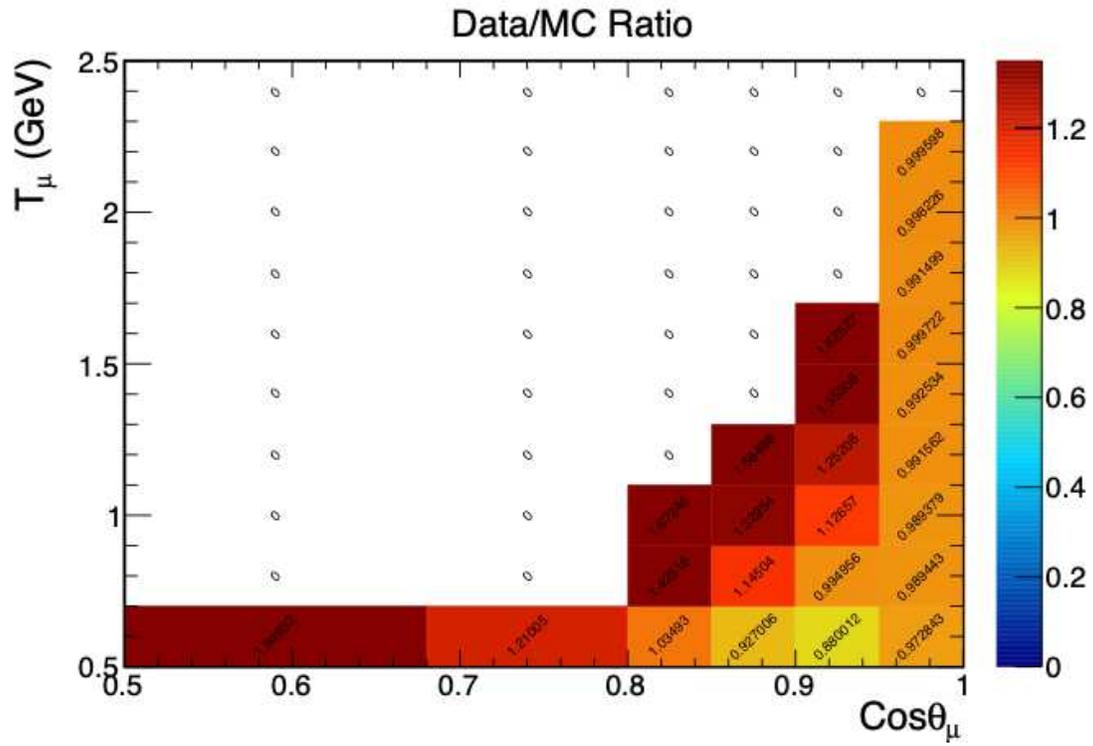


FIGURE 6.7. The Data/MC cross section ratio.

where  $0.90 < \text{cos}\theta_\mu < 0.95$  and  $1.5 \text{ GeV} < T_\mu < 1.7 \text{ GeV}$ . This came from the uncertainty on the efficiency and was seen in the mock data test as well.

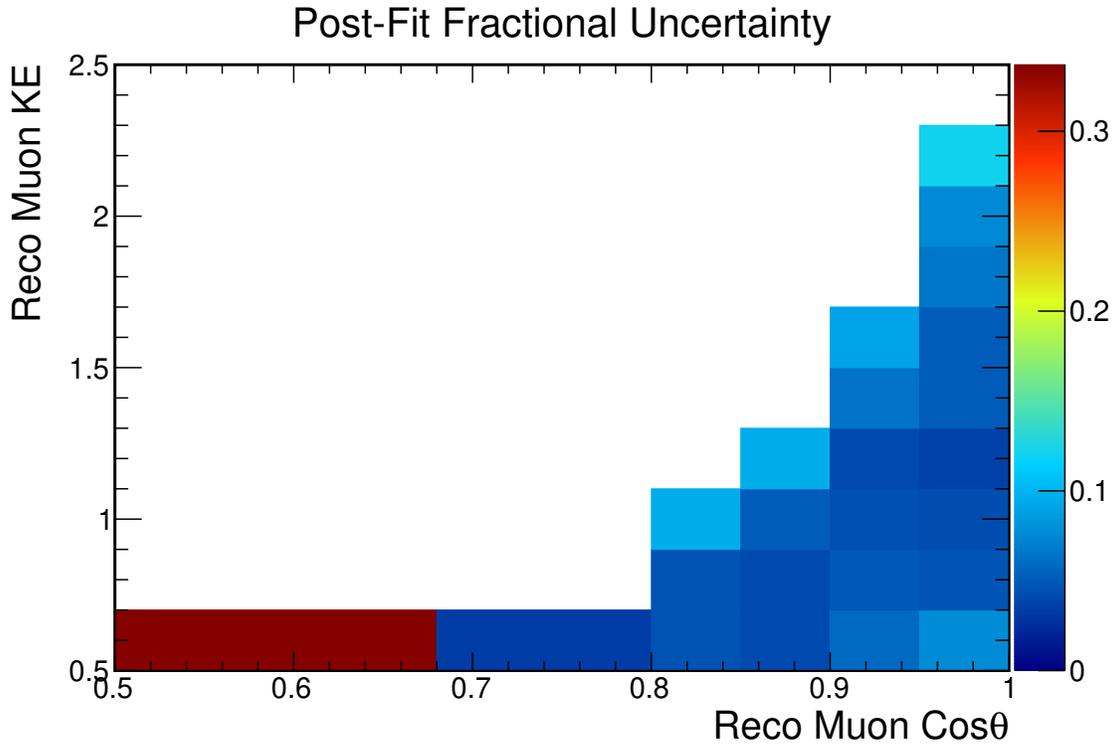


FIGURE 6.8. Fractional uncertainty on the signal from the fit

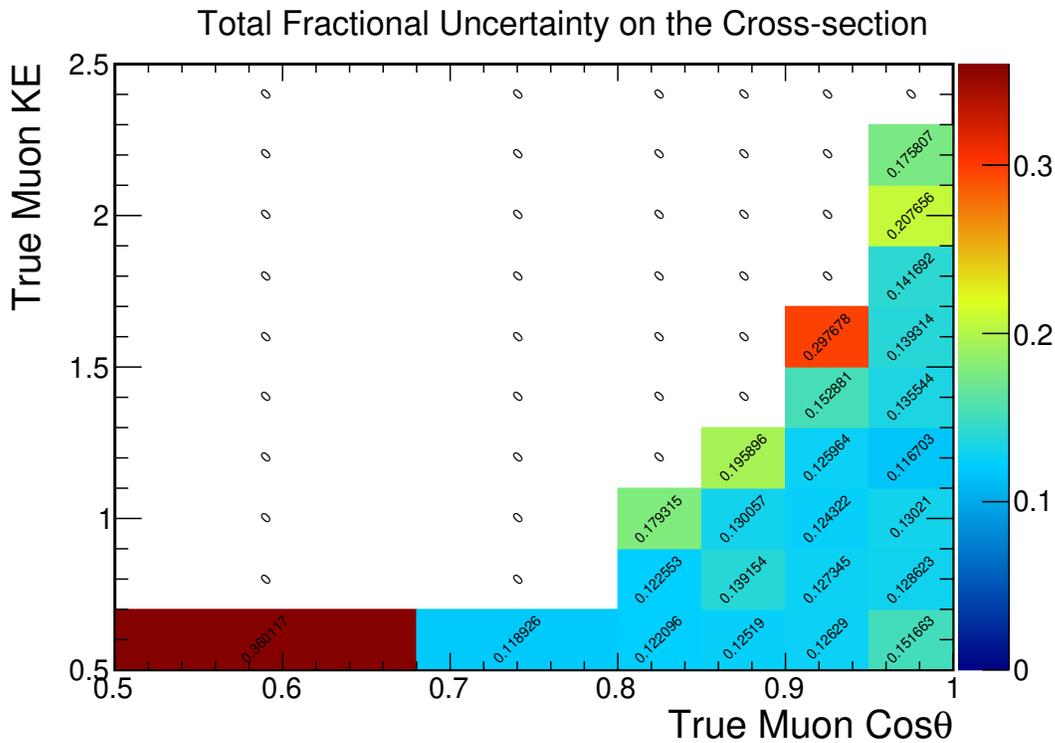


FIGURE 6.9. Total fractional uncertainties on the final cross section measurement.

## CHAPTER 7

# CONCLUSIONS AND FUTURE WORK

### 7.1. CONCLUSIONS

A measurement of muon neutrino-nucleus interactions with a charged pion in the final state was made. The double-differential cross section was measured and was within uncertainty of the simulated prediction in the forward going region,  $\cos\theta_\mu > 0.95$ . Excesses in the extracted signal (greater than 25%), relative to the simulation, were found at large scattering angles. These excesses were greater than the estimated uncertainties ( $\sim 15\%$ ). This discrepancy may come from either the model parameters used in simulation or from the difficulty of selecting events with a pion using existing tools.

### 7.2. FUTURE WORK

This analysis was the first CC  $\pi^\pm$  measurement using the NOvA experiment and used very basic pion identification and energy estimation tools. This measurement will serve as a baseline to future analyses that will have the benefit of more sophisticated and mature analysis and reconstruction tools. Future analyses that are able to better reconstruct pions and better measure the angle and energy of the resultant pions will have the ability present the measurement in pion kinematics. Those measurements will be quite useful to the neutrino interaction modelling community. Similar techniques to what have been used in this analysis could be applied to those analyses.

7.2.1. USING THE LATEST VERSION OF NOVASOFT AND DATA. Newer versions of NovaSoft, and the included analysis software CAFAna, have been released since the conclusion of this analysis. These updated tools include more kinematic variables that can be used to

improve PionID and energy estimation. Additional data has been collected and processed using the newer versions of NovaSoft that will be used in future iterations of this analysis.

7.2.2. MOVING TO GENIE 3. This analysis was performed using events simulated with an earlier version of the interaction modeling software than current analyses use. While GENIE 2 used the Rein-Sehgal model to describe pion production in neutrino events, GENIE 3 uses the improved Berger-Sehgal model, which parameterizes the Rein-Sehgal model using recent pion measurements on carbon.

7.2.3. USE OF DEEP LEARNING TO IDENTIFY THE PION. PionID was effective at selecting pions, however, newer techniques involving deep learning exist, such as convolved neural networks (CNN), which have shown greater efficiency at correctly selecting pions.

## BIBLIOGRAPHY

- [1] A. Himmel, “New oscillation results from the nova experiment,” July 2020.
- [2] J. A. Formaggio and G. P. Zeller, “From  $\nu_e$  to  $\bar{\nu}_e$ : Neutrino cross sections across energy scales,” *Rev. Mod. Phys.*, vol. 84, pp. 1307–1341, Sep 2012.
- [3] D. Griffiths, *Introduction to Elementary Particles*. WILEY, 1987.
- [4] L. M. Brown, “The idea of the neutrino,” *Physics Today*, vol. 31, pp. 23–28, 09 1978.
- [5] E. Fermi, “Tentativo di una teoria dei raggi  $\beta$ ,” *Il Nuovo Cimento (1924-1942)*, vol. 11, pp. 1–19, Jan 1934.
- [6] F. Reines and C. L. Cowan, “The neutrino,” *Nature*, vol. 178, pp. 446–449, Sep 1956.
- [7] R. Davis, D. S. Harmer, and K. C. Hoffman, “Search for neutrinos from the sun,” *Phys. Rev. Lett.*, vol. 20, pp. 1205–1209, May 1968.
- [8] M. A. Ramírez *et al.*, “Measurements of neutrino oscillation parameters from the T2K experiment using  $3.6 \times 10^{21}$  protons on target,” Mar 2023.
- [9] M. A. Acero *et al.*, “Improved measurement of neutrino oscillation parameters by the NOvA experiment,” *Physical Review D*, vol. 106, Aug 2022.
- [10] F. An *et al.*, “Neutrino physics with JUNO,” *Journal of Physics G: Nuclear and Particle Physics*, vol. 43, p. 030401, Feb 2016.
- [11] I. Esteban, M. Gonzalez-Garcia, M. Maltoni, T. Schwetz, and A. Zhou, “The fate of hints: updated global analysis of three-flavor neutrino oscillations,” *Journal of High Energy Physics*, vol. 2020, Sep 2020.
- [12] D. Scully, *Neutrino Induced Coherent Pion Production*. PhD thesis, University of Warwick, 2013.

- [13] D. Rein and L. M. Sehgal, “Neutrino-excitation of baryon resonances and single pion production,” *Annals of Physics*, vol. 133, no. 1, pp. 79–153, 1981.
- [14] B. Eberly *et al.*, “Charged pion production in  $\nu_\mu$  interactions on hydrocarbon at  $E_\nu = 4.0$  GeV,” *Phys. Rev. D*, vol. 92, p. 092008, Nov 2015.
- [15] A. A. Aguilar-Arevalo *et al.*, “Measurement of neutrino-induced charged-current charged pion production cross sections on mineral oil at  $E_\nu \sim 1$  GeV,” *Phys. Rev. D*, vol. 83, p. 052007, Mar 2011.
- [16] K. Abe *et al.*, “First measurement of the muon neutrino charged current single pion production cross section on water with the t2k near detector,” *Phys. Rev. D*, vol. 95, p. 012010, Jan 2017.
- [17] D. S. Ayres, G. R. Drake, M. C. Goodman, J. J. Grudzinski, V. J. Guarino, R. L. Talaga, A. Zhao, /Argonne, P. Stamoulis, E. Stiliaris, G. Tzanakos, and M. Zois, “The nova technical design report,” 10 2007.
- [18] G. J. Feldman, J. Hartnell, and T. Kobayashi, “Long-baseline neutrino oscillation experiments,” *Advances in High Energy Physics*, vol. 2013, p. 475749, Mar 2013.
- [19] NO $\nu$ A Collaboration, “NO $\nu$ A approved plots.” NO $\nu$ A Internal Figure.
- [20] C. Andreopoulos, C. Barry, S. Dytman, H. Gallagher, T. Golan, R. Hatcher, G. Perdue, and J. Yarba, “The genie neutrino monte carlo generator: Physics and user manual,” 2015.
- [21] T. Bhlen, F. Cerutti, M. Chin, A. Fass, A. Ferrari, P. Ortega, A. Mairani, P. Sala, G. Smirnov, and V. Vlachoudis, “The fluka code: Developments and challenges for high energy and medical applications,” *Nuclear Data Sheets*, vol. 120, pp. 211–214, 2014.

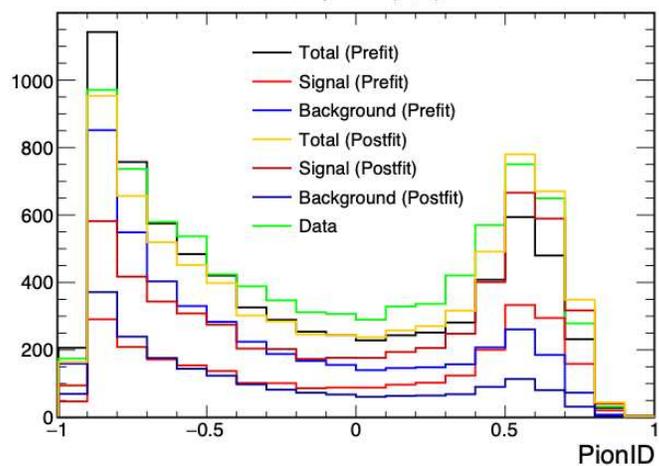
- [22] L. Aliaga *et al.*, “Neutrino flux predictions for the numi beam,” *Phys. Rev. D*, vol. 94, p. 092005, Nov 2016.
- [23] M. Baird, J. Bian, M. Messier, E. Niner, D. Rocco, and K. Sachdev, “Event reconstruction techniques in nova,” *Journal of Physics: Conference Series*, vol. 664, p. 072035, Dec 2015.
- [24] R. E. Kalman, “A new approach to linear filtering and prediction problems,” *Trans. ASME, D*, vol. 82, pp. 35–44, 1960.
- [25] M. Baird, J. Bian, M. Messier, E. Niner, D. Rocco, and K. Sachdev, “Event reconstruction techniques in nova,” *Journal of Physics: Conference Series*, vol. 664, p. 072035, Dec 2015.
- [26] L. Breiman, *Classification and Regression Trees*. Routledge, 1984.
- [27] L. Aliaga *et al.*, “Measurement of the  $\nu_\mu$  cc inclusive cross section in the NO $\nu$ A near detector.” NO $\nu$ A Internal Document, DocDB 32688, 2019.
- [28] M. Muether, D. Doyle, and J. Paley, “Derivation of the relative cross-section uncertainty figure of merit for selection criteria in cross section analyses.” NO $\nu$ A Internal Document, DocDB 49166, 2019.
- [29] K. Albertsson *et al.*, “Tmva 4, toolkit for multivariate data analysis with root, users guide,” 2020.
- [30] D. Pershey, “A  $\nu_\mu n \rightarrow \mu\pi^0 x$  cross section analysis with the NO $\nu$ A near detector.” NO $\nu$ A Internal Document, DocDB 18768, 2017.
- [31] M. A. Acero *et al.*, “Measurement of the  $\nu_e$  nucleus charged-current double-differential cross section at  $e_\nu = 2.4$  gev using nova,” vol. 130, Feb 2023.

- [32] F. W. James, “Minuit function minimization and error analysis: Reference manual version 94.1,” 1994.
- [33] S. Schmitt, “Data unfolding methods in high energy physics,” *EPJ Web of Conferences*, vol. 137, p. 11008, 2017.
- [34] G. D’Agostini, “Improved iterative bayesian unfolding,” 2010.
- [35] L. Aliaga, “Analysis of the  $\nu_\mu$  cc with low hadronic activity events in the NO $\nu$ A near detector.” NO $\nu$ A Internal Document, DocDB 49881, 2022.
- [36] M. Strait, “Geometry changes tech note.” NO $\nu$ A Internal Document, DocDB 23132, 2018.
- [37] L. Aliaga, “Numi beam prediction for the nova 2017 analyses.” NO $\nu$ A Internal Document, DocDB 23441, 2017.

APPENDIX A

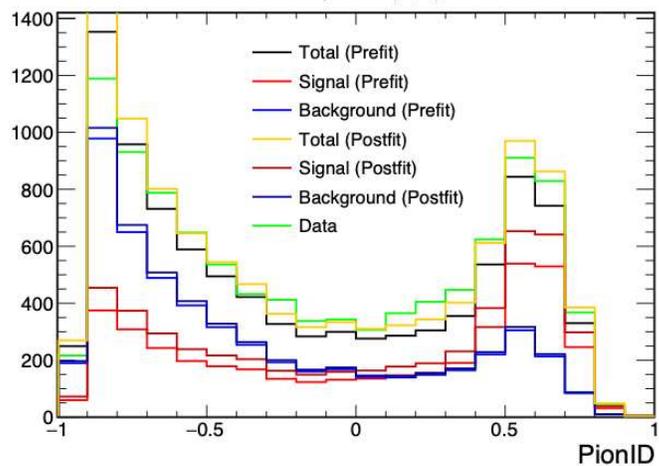
ADDITIONAL FIGURES

Template (1,1)



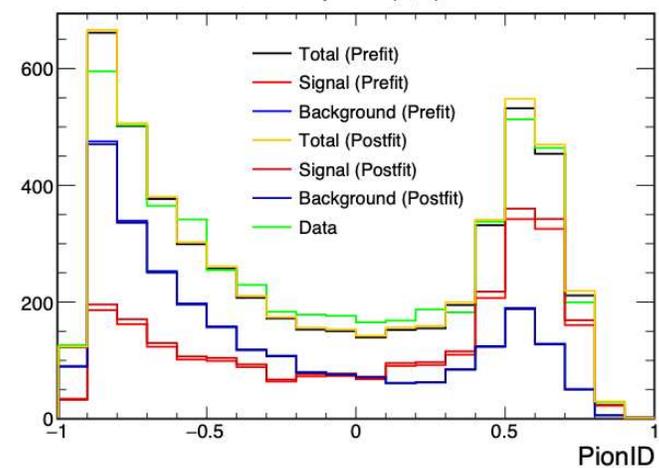
(a)

Template (2,1)



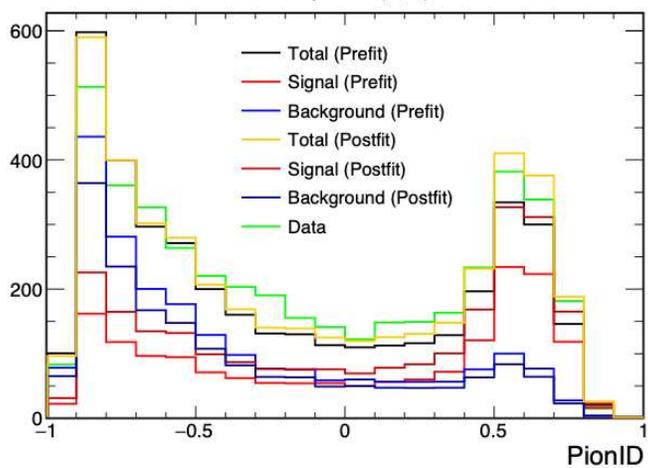
(b)

Template (3,1)



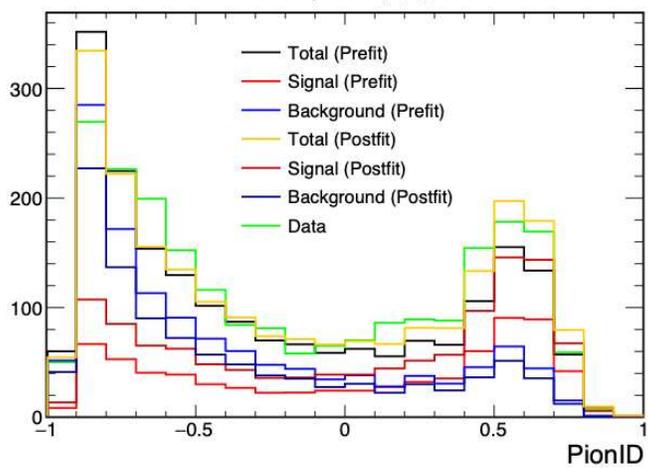
(c)

Template (3,2)



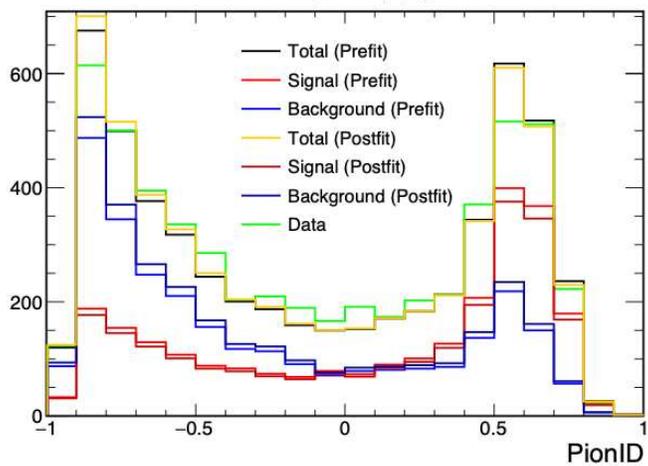
(d)

Template (3,3)



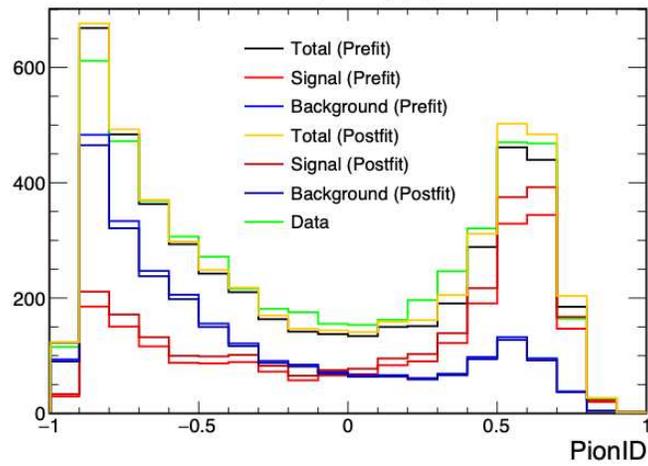
(e)

Template (4,1)



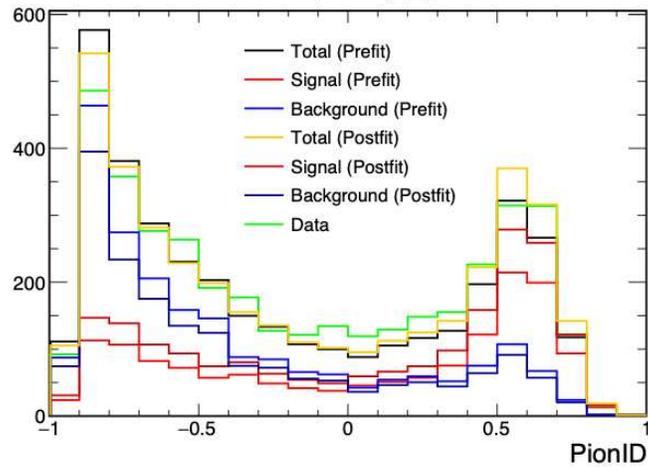
(f)

Template (4,2)



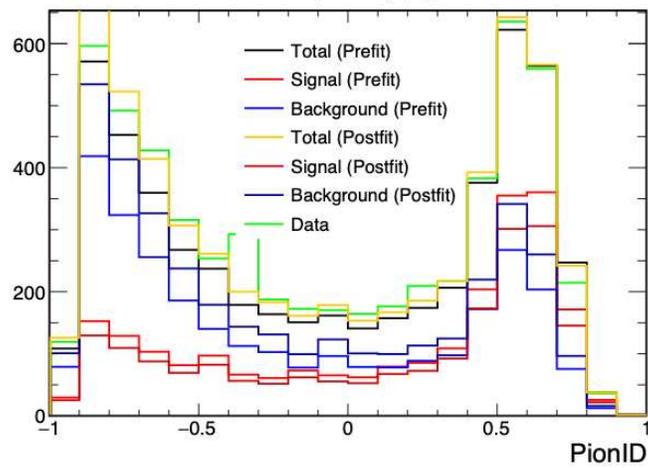
(g)

Template (4,3)



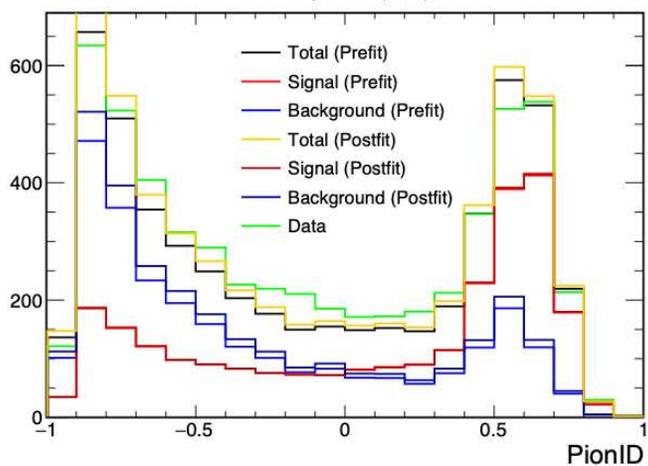
(h)

Template (5,1)



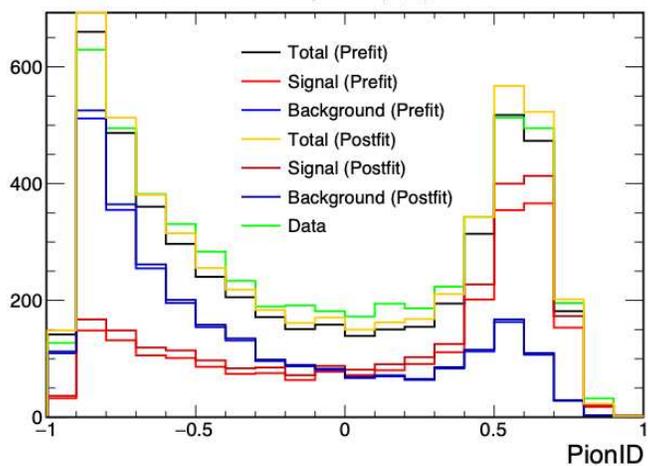
(i)

Template (5,2)



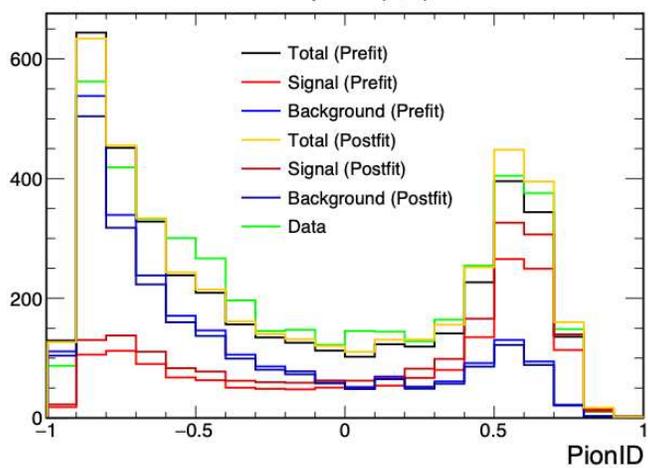
(j)

Template (5,3)



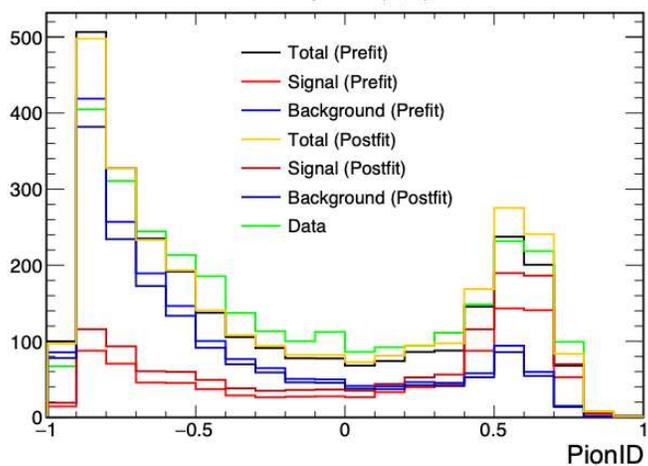
(k)

Template (5,4)



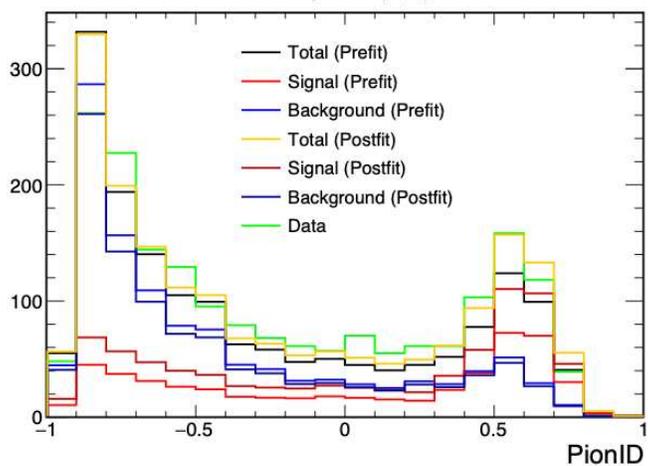
(l)

Template (5,5)



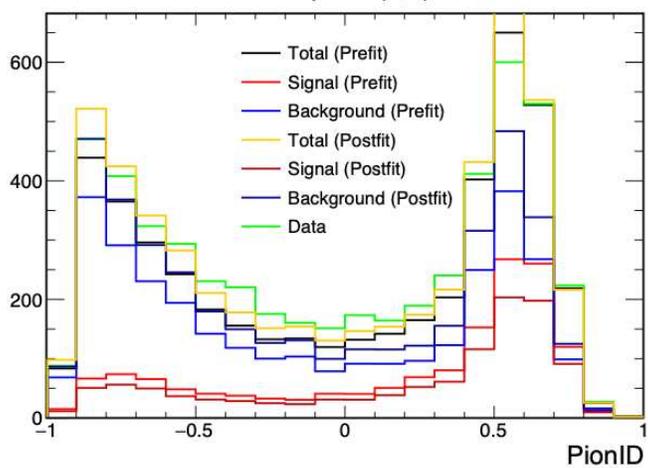
(m)

Template (5,6)



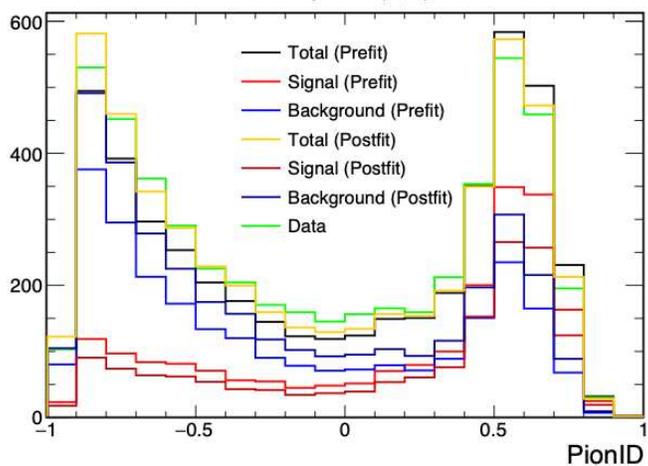
(n)

Template (6,1)



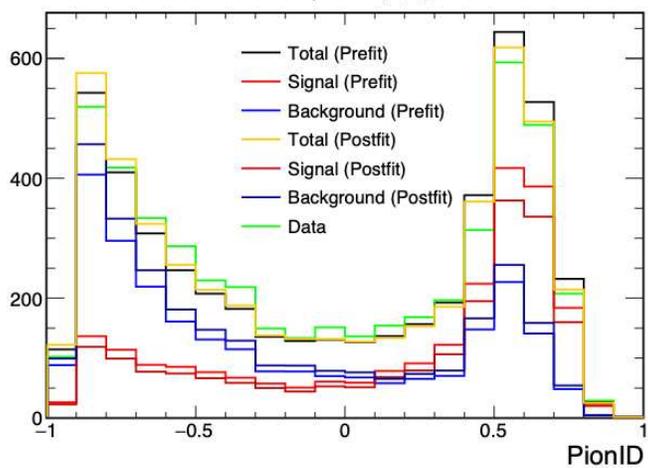
(o)

Template (6,2)



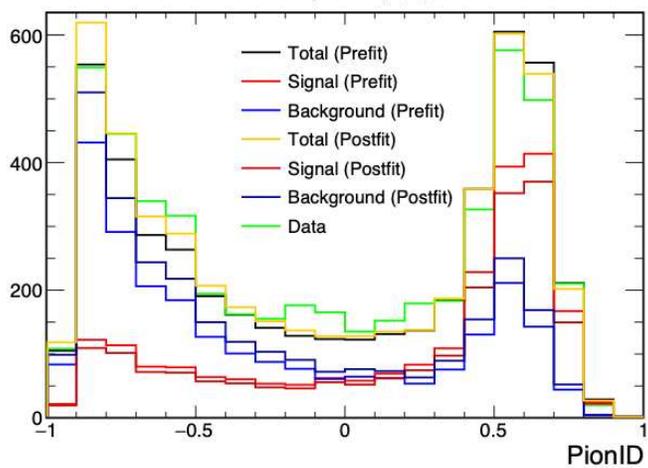
(p)

Template (6,3)



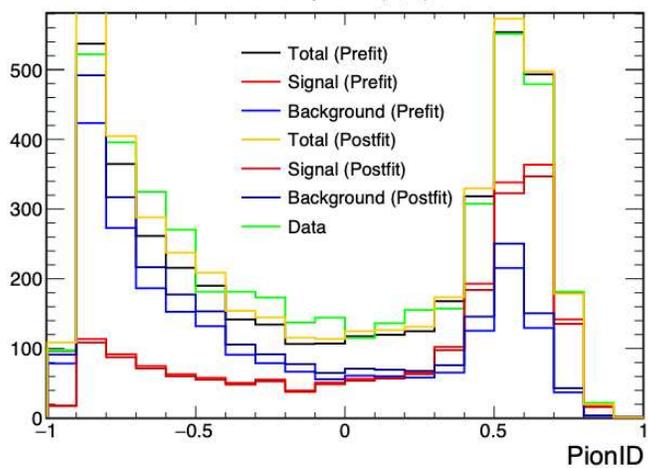
(q)

Template (6,4)



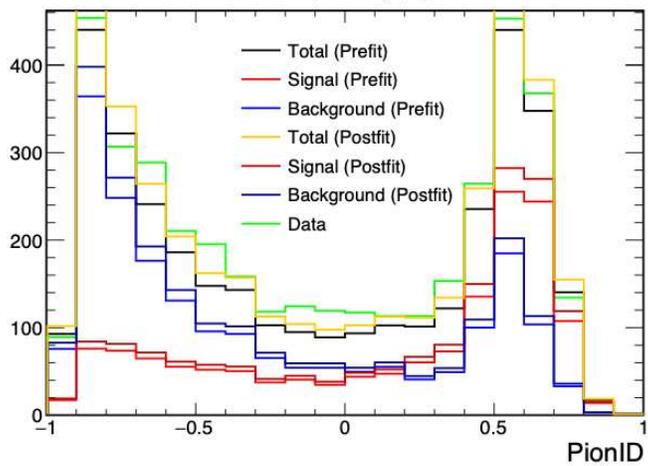
(r)

Template (6,5)



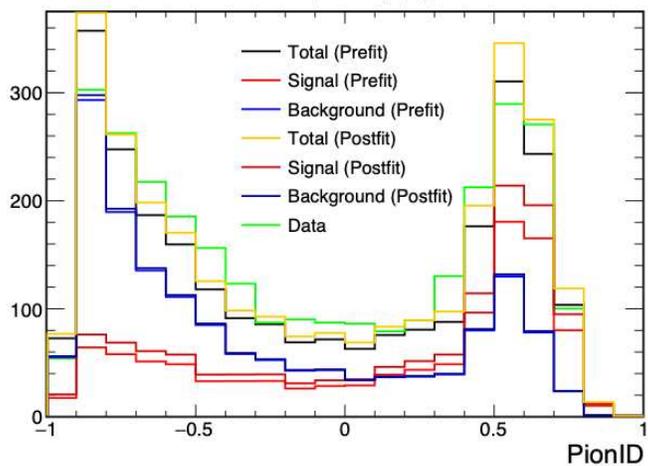
(s)

Template (6,6)

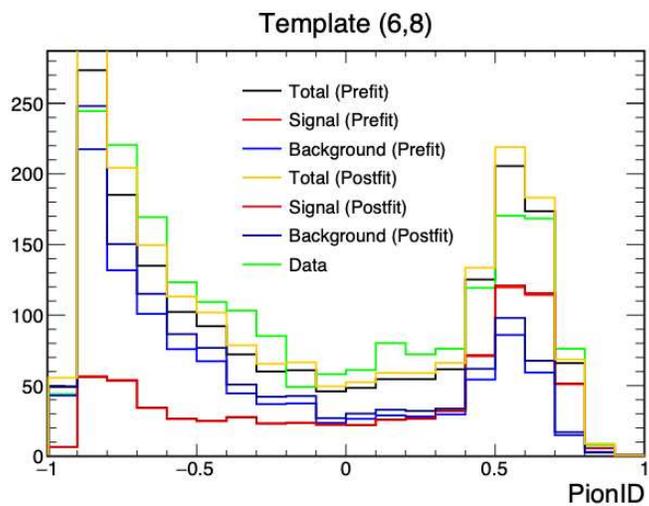


(t)

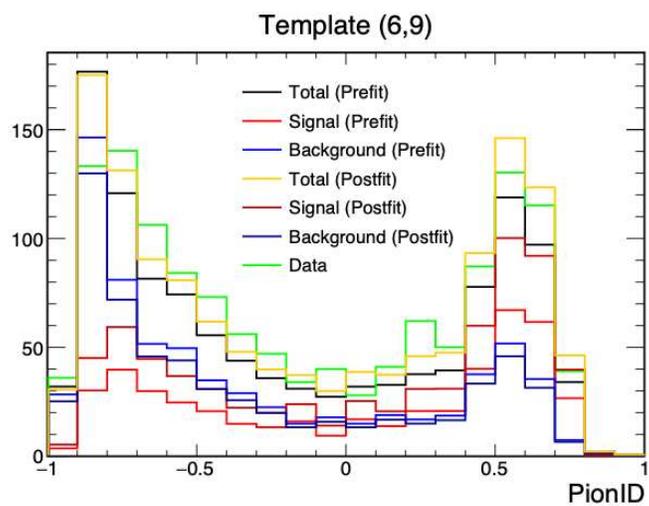
Template (6,7)



(u)



(v)



(w)

FIGURE A.1. The complete template set from the data. The data template, in orange, along with the pre and post fit templates for signal and background. The simulated estimation of the data is shown in black.

## APPENDIX B

### LIST OF ABBREVIATIONS

**APD:** Avalanche Photo Diode

**AWS:** Amazon Web Services

**BDT:** Boosted Decision Tree

**CAF:** Common Analysis Format/File

**CC:** Charged Current

**CNN:** Convolutional Neural Network

**COH:** Coherent

**CVN:** Convolutional Visual Network

**DAQ:** Data Acquisition System

**DIS:** Deep Inelastic Scattering

**FD:** Far Detector

**FEB:** Front End Board

**FHC:** Forward Horn Current

**FNAL:** Fermi National Accelerator Laboratory

**KE:** Kinetic Energy

**MC:** Monte Carlo

**MPPC:** Multi Pixel Photon Counter

**MSE:** Mean Square Error

**NC:** Neutral Current

**ND:** Near Detector

**NO $\nu$ A :** NuMI Off-Axis  $\nu_e$  Appearance Experiment

**NuMI:** Neutrinos from the Main Injector

**OoF:** Out of Fiducial

**PMNS:** Pontecorvo-Maki-Nakagawa-Sakata

**PPFX:** Package to Predict the Flux

**PVC:** Polyvinyl Chloride

**QE:** Quasi-Elastic

**RES:** Resonant

**RHC:** Reverse Horn Current

**ROC:** Receiver Operating Characteristic

**SM:** Standard Model