FERMILAB-THESIS-2026-14

A MEASUREMENT OF THE PION-ENERGY DEPENDENCE OF MUON NEUTRINO CHARGED-CURRENT SCATTERING TO FINAL STATES WITH ONE CHARGED PION IN NOVA

Erin Ewart

Submitted to the faculty of the Graduate School

in partial fulfillment of the requirements

for the degree

Doctor of Philosophy

in the Department of Physics,

Indiana University

June 2026

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Doctoral Committee

Jon Urheim, PhD

Radovan Dermisek, PhD

Rob de Ruyter, PhD

Mark Messier, PhD

Rex Tayloe, PhD

April 21, 2026

Dedicated to my fiancée Amelia, who has been with me through more than either of us thought we were getting into. Your support will always mean the world to me.

Acknowledgments

Thank you to my thesis advisor, Jon Urheim, whose steady guidance, compassion, and understanding made this possible. I knew going in you were a great scientist— I am eternally grateful you are a great person as well.

Thank you to the rest of my committee: Radovan Dermisek, Rob de Ruyter, Mark Messier, and Rex Tayloe, who have all provided much-needed outside perspectives to my research throughout.

Thank you also to Jon Paley: you went above and beyond to help me when you did not have to. You taught me almost everything I know about neutrino cross-sections, and gave me the tools I needed to learn the rest. Thank you so much for your support throughout this project, it has meant more than I can convey.

Thank you to the Fermilab Accelerator Division, whose tireless work and expertise provided the neutrinos needed to make this work and so much more possible. You all are the unsung heroes of neutrino physics. I would also like to acknowledge Fermilab as a whole for its resources and support, especially everyone who maintains the infrastructure necessary for physics. I would like to acknowledge the Department of Energy for funding support under grant DE-SC0010120.

To my Bloomington friends— Reed, Maria, Becca, Beni, Daniel, Taylor, Luke, Lou, Lauren, and all the rest— you've been the best friends I could ask for, and have kept me sane and happy. Without you I would have turned into the sort of bitter, miserable academic we all hate.

Teresa, Erica, Ashley, and many others beside— thank you for all of your guidance on the details of this field, as well as your personal support through the tough times. I would be a far worse scientist today without all of you demonstrating how to be excellent both at this job and as people.

Amelia, you got the dedication, but I'd be remiss if I didn't mention you again here. I love you so much, and I'm deeply grateful that you support my ambitions even though the life of an

academic is far from ideal.

And finally, to my parents: thank you for raising me with love and kindness, and supporting me through thick and thin after I left the nest. You've supported all of my crazy dreams since I was three, and I know I wouldn't be here today if you weren't willing to take the "why?" of a kid a lot more seriously than most.

Erin Ewart

A MEASUREMENT OF THE PION-ENERGY DEPENDENCE OF MUON NEUTRINO CHARGED-CURRENT SCATTERING TO FINAL STATES WITH ONE CHARGED PION IN NOVA

The study of neutrino oscillations is a main priority for particle physics as the most immediately tractable lever on physics beyond the Standard Model. In particular, more insight into violation of the combined symmetry of charge-conjugation plus parity could yield clues to the origin of matter-antimatter asymmetry, and in some theoretical frameworks the neutrino mass could give insights into dark matter. NOvA is a long-baseline accelerator neutrino experiment with both a near and far detector that seeks to measure several of the parameters of the neutrino mixing matrix, as well as carry out a broad program of additional physics. NOvA has innovated and developed a variety of techniques in the space of neutrino physics, including expanding the use of machine learning techniques in reconstruction. NOvA has also set the stage for the next generation US-based long baseline experiment, DUNE. This dissertation details the creation of a new neutrino interaction vertex reconstruction package for NOvA, which offers enormous improvements in accuracy above the previous vertexer it replaces. This is accomplished using a Convolutional Visual Network trained on large datasets of simulated events. The vertexer is validated thoroughly against data. This dissertation also presents a cross-section measurement for $\nu_\mu + N \rightarrow \mu^- + 1\pi^\pm + X$ (where X does not include additional charged pions) binned in pion kinetic energy, a challenging measurement in general and for NOvA in particular. Measurements like this provide crucial inputs for neutrino interaction models utilized by neutrino oscillation experiments. This measurement in particular is important to the energy ranges relevant for DUNE in a band not well-covered by any other experiment.

Jon Urheim, PhD

Radovan Dermisek, PhD

Rob de Ruyter, PhD

Mark Messier, PhD

Rex Tayloe, PhD

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Neutrinos are both abundant and elusive neutral particles that permeate the universe as a product of all of the most common weak force interactions in the modern universe, ranging from radioactive decay to stellar fusion to nucleosynthesis within supernovae. Their nonzero mass is the most accessible lever on beyond the Standard Model physics available today. For this reason, neutrino physics is considered a top priority of the world particle physics community, with precision measurements of neutrino oscillations forming a large part of this program.

It bears emphasizing how puzzling the nonzero but approximately eV scale neutrino mass is. There is no way to neatly incorporate it into the Standard Model without significant side-effects, and the five to six order of magnitude separation from the suspected mass of the heaviest neutrino and the electron mass begs explanation. It also brings into question whether neutrinos are Dirac and get their mass through coupling to the Higgs like other fermions, or Majorana where they are their own antiparticles and owe their mass to some other mechanism. The former tends to cause more theoretical problems, while the latter is unprecedented among fundamental particles. Many of the explanations of the neutrino mass give a potential window into a "dark sector", which would be cosmologically useful, and the existence of the mass inherently admits the potential for violation of the combined charge-conjugation plus parity symmetry, hereafter referred to as CP violation. This alone would not be sufficient to explain the matter-antimatter problem, but it would add significant additional information towards a resolution.

However, studying neutrinos is not easy. Since they couple only to the weak force, neutrinos interact only at short range and only rarely, necessitating both very large fluxes and very large detectors in order to perform any measurement at all. Moreover, their electrical neutrality means they cannot be focused or otherwise manipulated in the same way we can control charged particles, and therefore every neutrino flux spans a wide range of energies when compared to other particles, and the properties of a neutrino source cannot be monitored in any way that isn't coupled to model assumptions. This exposes neutrino experiments to fairly extreme systematic errors and model dependencies which must be carefully controlled for using a variety of techniques, including and especially the grounding of simulation frameworks in empirical observations and multi-detector setups.

This dissertation lays out the background necessary to understand and present the results of my two main graduate research projects: the development of a new neutrino interaction point reconstruction software package (VertexCVN), and the measurement of the differential cross-section of $\nu_\mu + N \rightarrow \mu^- + 1\pi^\pm + X$ in pion kinetic energy, both within the NOvA experiment. Both of these contributions serve to indirectly improve neutrino oscillation measurements, the former by increasing the accuracy of event reconstruction and thus both recovering measurable events and reducing uncertainties, and the latter by eventually being incorporated into neutrino event generators, which reduces the systematic errors of future measurements.

These seemingly unrelated projects were originally tied together in a coherent narrative— VertexCVN was originally supposed to be a secondary vertexer, which would allow for significantly improved reconstruction in NOvA, especially of charged pions which inelastically scatter within the detector at a high rate. This would have been implemented in a limited production run on selected single charged pion events, enabling accurate charged pion energy reconstruction and high quality differential cross-section measurements in charged pion energy, a measurement which had been attempted and abandoned by several previous analyzers. Unfortunately, it was not to be. The priorities of the experiment required me to focus on the creation of a primary vertexer first, along with a substantially longer than expected validation effort. By the time that was done, it was too late to attempt to develop a secondary vertexer (known to be a hard problem), let alone the downstream reconstruction required to make use of it. Unfortunately, due to further delays in the production, not even VertexCVN could be used in the cross-section analysis I present here though it would have substantially improved my measurement's resolution. Therefore I present the two projects separately.

Chapter 2 of this thesis will present the history of neutrinos, their theoretical framework, and an overview of the current state of the field of neutrino oscillation measurements. Chapter 3 presents an overview of the NOvA experiment and its many interlocking parts. Chapter 4 presents a brief overview of machine learning techniques as well as the construction and results of VertexCVN. Switching attention to the other half of my work, Chapter 5 presents a background on the process of performing a neutrino cross-section measurement and neutrino phenomenology. Finally, Chapter 6 details the cross-section measurement I performed.

Chapter 2

Neutrinos

Neutrinos are an electrically neutral fundamental particle of extremely small mass that have only become detectable relatively recently. They are abundant to cosmological scales due to their production by various stellar processes. However, neutrinos are also defined by their extremely low scattering cross-section, with the mean free path of neutrinos in solid lead being best measured in light-years at the energies of relevance to this thesis. Therefore, neutrinos are both worthwhile to study due to their abundance and their implications for cosmology, and extremely difficult to study due to their inability to be meaningfully controlled and the size of detectors needed to induce even double-digit numbers of interactions from most sources.

2.1 History

Neutrinos were initially proposed as a solution to a very particular problem, that of the continuous energy spectrum of beta decay. The expectation at the time was that beta decay, in which a nucleus gives off a single electron or positron and in the process changes its atomic number, would produce nearly monoenergetic electrons (or positrons) as the product of a two-body decay. However, the actual measured energy spectrum was continuous, and in fact consistent with a three-body decay where the third body was significantly less massive than the electron, then the lightest known massive particle. In fact, the measured spectrum was most consistent with a *massless* third particle, and so when Wolfgang Pauli proposed the neutrino in 1930 it was as a massless particle. While the existence of a massless particle that would be nearly impossible to directly measure due to the very small weak coupling constant was distasteful, it ultimately proved less distasteful than giving up on conservation of momentum or energy. Thus, the neutrino made its way into the Standard Model, with a corresponding neutrino for each generation of charged lepton (e, μ, τ) to conserve lepton number.

2.1.1 Initial detection

Of course, what was nearly impossible to measure in the 1930s eventually became merely extremely difficult to measure as technology improved and the scale at which science was conducted increased. In 1942, Wang Gangchang pointed out that if neutrinos were proposed to be produced as part of beta decay, they then must be measurable through inverse beta decay via $\overline{\nu}_e + p \rightarrow n + e^+$ [1]. All that would be required would be a sufficiently large number of sufficiently high-energy neutrinos, which would shortly be provided by the invention of the nuclear reactor, and a detector capable of distinguishing the signal of inverse beta decay from relevant backgrounds. This last was relatively easy, as the annihilation of the positron followed by the neutron capture would give a distinct signature of two coincident gammas followed after a few microseconds by a third. Capitalizing on this distinct signature, a detector of water sandwiched between three planes of liquid scintillator surrounded by photomultiplier tubes was used to identify this signal and prove the existence of the electron neutrino. [2] Eventually, the muon neutrino would also be directly detected in 1975 [3], and the tau neutrino would later follow in 2000 [4], with the difficult increasing each lepton generation due to the change of source required and the higher mass of the charged lepton involved raising the detection threshold energy.

2.1.2 Discovery of Mass

While nuclear fission reactors were the first manmade high-flux source of neutrinos, the Sun is, among other things, an extremely high-flux source of electron neutrinos via nuclear fusion in its core. However, these solar neutrinos are significantly more difficult to measure than reactor neutrinos given their much lower flux and the inability to characterize backgrounds by turning the flux off, requiring precise, low-background, and very large detectors. The first detector designed to measure this solar neutrino flux was the Homestake experiment [5], which used a 380 cubic meter tank of perchloroethylene located approximately 1500 meters underground in an unused portion of the Homestake Gold Mine, which was then flooded with water to absorb neutrons originating from surrounding radioactivity in the rock. This massive rock overburden allowed for the elimination of nearly all cosmic rays, reducing background to an absolute minimum. The detection mechanism was neutrino capture via $Cl^{37} + \nu_e \rightarrow Ar^{37} + e^-$, where the argon was then periodically purged from the detector with a helium carrier gas, absorbed, and then counted volumetrically. This number of argon atoms could then be converted into an observed electron neutrino flux via the theoretically estimated cross-section.

The major result of this experiment was measuring a flux of electron neutrinos approximately 1/3 of what was expected. This caused a crisis in the fields of solar physics and neutrino physics, as there must have been some serious problem with either the Homestake experiment, the scientific understanding of solar fusion processes, or the scientific understanding of neutrino behavior. This last was held to be unlikely due to the incredible successes of the Standard Model.

Significant work was done on this problem by various experiments and theorists, but the turning point came in 1998 when the Super Kamiokande (Super-K) experiment reported an up-down discrepancy in atmospheric neutrinos, that is, muon neutrinos produced by cosmic rays in the atmosphere. Fewer atmospheric neutrinos coming up through the Earth were detected than expected [6] (neutrino absorption by the Earth is a negligible effect at these energies), strongly implying that the theory of neutrino oscillation via nonzero neutrino mass in 1968 by Pontecorvo [7] was the correct explanation for the solar neutrino problem. His idea was that the enormous mass of the Sun would cause the quantum states of the massive neutrinos to decohere entirely into equal proportions of electron, muon, and tau neutrinos via the Mikheyev–Smirnov–Wolfenstein effect [8] [9] (also simply called the matter effect). This would necessarily also happen to a lesser extent to neutrinos passing through the Earth, leading to the lower imbalance seen by Super-K. This would later be confirmed by the SNO experiment in 2001 [10], which was able to calculate the proportion of electron solar

	Generation					
	I	II	III	I_3	Y	Q
Leptons	$\begin{pmatrix}\nu_e \\ e\end{pmatrix}_L$	$\begin{pmatrix}\nu_\mu \\ \mu\end{pmatrix}_L$	$\begin{pmatrix}\nu_\tau \\ \tau\end{pmatrix}_L$	$+1/2$ $-1/2$	-1 -1	0 1
	e_R	μ_R	τ_R	0	-2	-1
Quarks	$\begin{pmatrix}u \\ d\end{pmatrix}_L$	$\begin{pmatrix}c \\ s\end{pmatrix}_L$	$\begin{pmatrix}t \\ b\end{pmatrix}_L$	$+1/2$ $-1/2$	$+1/3$ $+1/3$	$+2/3$ $-1/3$
	u_R	c_R	t_R	0	$+4/3$	$+2/3$
	d_R	s_R	b_R	0	$-2/3$	$-1/3$

Table 2.1: The fermions of the standard model. I_3 is weak isospin, Y is weak hypercharge, and Q is electric charge. Parentheses are around left-handed isospin partners.

neutrinos to all solar neutrinos with enough precision to ensure statistically significant agreement with the Mikheyev–Smirnov–Wolfenstein effect.

Since then, significant effort has been put forward to measuring the exact form of this oscillation, with hopes for a window into beyond the Standard Model physics.

2.2 Neutrinos in the Standard Model

In the Standard Model, neutrinos are massless and only ever left-handed due to weak force chirality; specifically, they are the weak isospin partners of the left-handed charged leptons (see table 2.1). They couple only through $Z \to \nu_l + \overline{\nu}_l$, called the neutral current interaction, and $W^{\mp} \to \overset{(-)}{\nu_l} + l^{\mp}$ (where $l \in \{e, \mu, \tau\}$), called the charged current interaction. Notably, the coupling is suppressed by the high mass of the W and Z bosons, rendering the weak force both much weaker than the strong and electromagnetic forces as well as much shorter-ranged.

The existence of the neutrino mass presents significant problems for the Standard Model. The weak interaction is empirically chiral, and therefore is formulated in terms of left-handed doublets and right-handed singlets. However, massive particles can necessarily be chirally rotated from left- to right-handed via a change of reference frame. Therefore right-handed neutrinos must exist, but if their mass term is Dirac via the standard Higgs mechanism these right-handed neutrinos would be particles with zero electric charge, no color, and zero weak isospin. They would interact with other particles only through the Higgs, considered by many to be a distasteful proposition. It would also invite the question of why the neutrino's coupling to the Higgs is so much weaker than that of other particles, since the neutrino mass is believed to be around six orders of magnitude lower than the next-lowest particle mass, that of the electron.

However, other possibilities exist. Neutrinos could be Majorana fermions, where they are their own antiparticle. This would make lepton number a non-conserved property and permit neutrinoless double-beta decay, for which searches are currently ongoing. This also leads to a seesaw mechanism [11] which neatly gives the standard neutrinos very low masses at the cost of three corresponding very heavy sterile neutrinos.

2.3 Neutrino Oscillation Physics

The general idea of neutrino oscillations is that while neutrinos are produced in their weak interaction flavor eigenstates, they travel through space in their mass eigenstates (as do all particles in quantum mechanics). Then, if a neutrino undergoes a charged-current scattering event in a detector, this means they are measured in a flavor eigenstate. If the flavor eigenstates do not exactly correspond to the mass eigenstates (which they do not), this then implies a chance of measuring a neutrino in a different eigenstate than it is produced.

I will present here the oscillation framework given by the minimal extension to the Standard Model that admits neutrino oscillations. For simplicity, neutrinos will be assumed to be Dirac, though this is an open question. Other extensions admit interesting features like light sterile neutrinos or heavy right-handed neutrinos which would provide tantalizing levers into proposed "dark sectors", that is, extensions of the Standard Model that seek to explain dark matter.

2.3.1 The PMNS Matrix

The Pontecorvo–Maki–Nakagawa–Sakata (PMNS) matrix provides the simplest framework for examining neutrino oscillations, and has become the null hypothesis for neutrino oscillation experiments after the previous null hypothesis of "no oscillations" was rejected. The PMNS matrix is the unitary mixing matrix between the neutrino mass and flavor eigenstates, and is equivalent in form to the CKM matrix for quark flavor mixing. The general form of the matrix is

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \tag{2.1}$$

Where the ν_i for $i \in \{1, 2, 3\}$ are the mass eigenstates, and U fulfills $U^*U = UU^* = I$ for I the identity matrix. The actual matrix elements are, in general, not a particularly useful representation and cannot be directly measured, and as such the matrix is more typically parameterized using three angles (between each pair of mass eigenstates) and a complex phase:

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta_{23}) & \sin(\theta_{23}) \\ 0 & -\sin(\theta_{23}) & \cos(\theta_{23}) \end{pmatrix} \begin{pmatrix} \cos(\theta_{13}) & 0 & e^{-i\delta_{CP}} \sin(\theta_{13}) \\ 0 & 1 & 1 \\ -e^{i\delta_{CP}} \sin(\theta_{13}) & 0 & \cos(\theta_{13}) \end{pmatrix} \begin{pmatrix} \cos(\theta_{12}) & \sin(\theta_{12}) & 0 \\ -\sin(\theta_{12}) & \cos(\theta_{12}) & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.2}$$

This formulation assumes that neutrinos are Dirac. If they are Majorana, there are two additional phases that do not affect oscillation physics. Small angles indicate relatively little mixing between the mass and flavor eigenstates. Experimentally the angles have been determined to be quite large, especially in comparison to the CKM matrix (see table 2.2). Of particular interest is δ_{CP} as it indicates the level of CP violation occurring in neutrino mixing, with no violation at $\delta_{CP} = \{0, \pi\}$ and maximal violation at $\delta_{CP} = \{\pi/2, 3\pi/2\}$. Unfortunately, δ_{CP} is the most difficult parameter to measure, and its value is currently unknown, with various experiments demonstrating different weak preferences for its value.

Parameter	Value
Δm^2_{21}	$7.50^{+0.19}_{-0.18} \times 10^{-5}\mathrm{eV}^2$
Δm^2_{32}	$2.541 \pm 0.026 \times 10^{-3}\mathrm{eV}^2$ (Normal mass ordering)
Δm^2_{32}	$-2.527 \pm 0.026 \times 10^{-3}\mathrm{eV}^2$ (Inverted mass ordering)
$\sin^2(\theta_{12})$	0.307±0.012
$\sin^2(\theta_{23})$	0.537±0.020 (Normal mass ordering)
$\sin^2(\theta_{13})$	0.0216 ±0.0006
δ_{CP}	$1.21^{+0.19}_{-0.22}$ π rad

Table 2.2: Latest best determinations of the various oscillation parameters per the Particle Data Group [12]. This table is a summary of the information in the "Neutrino Mixing" chapter of that reference. Normal vs inverted mass ordering is determined by whether ν_3 is heavier (normal) or lighter (inverted) than ν_2. These values are an incomplete picture of the state of the field, since the measurements of the various parameters are correlated in ways that simple error statements cannot account for. In particular, δ_{CP} appears more well-specified in this format than it actually is, with its 3σ curve enclosing the vast majority of the parameter's range in current state-of-the-art measurements.

2.3.2 The Oscillation Formula

Standard neutrino oscillation physics is always in the ultrarelativistic limit, as while none of the masses of the mass eigenstates are known with any precision, it *is* known with some (very model- and method-dependent) confidence from both cosmological [12] and some limited direct evidence [13] that their sum is less than tens of eVs. The processes used to detect neutrinos in oscillation experiments have momentum thresholds in the hundreds of keV or higher, and therefore all relevant neutrinos are ultrarelativistic.

In general, there are two types of observations one can make with a source of neutrinos: the disappearance rate, where one measures the number of neutrinos of the same flavor as the source and finds it lower than the no-oscillation case, and the appearance rate, where one measures the number of neutrinos of some other flavor as the source and finds it nonzero. In reality, most neutrino sources are not truly mono-flavor and backgrounds will always exist, somewhat complicating this picture, but for this basic phenomenology we will assume a mono-flavor monoenergetic neutrino source.

The vacuum picture is extremely simple but still illustrative, and will allow us to put things in a form that will prove useful for the more difficult non-vacuum case. In this case we can use the standard formula for the time-evolution of a particle in vacuum, keeping in mind we are interested in detecting the flavor eigenstate at the end:

$$|\nu_l, t\rangle = \sum_j U_{lj}^* |\nu_j, t\rangle = \sum_j U_{lj}^* e^{-iE_j t} |\nu_j, 0\rangle \tag{2.3}$$

where the l index indicates flavor eigenstate, the j index indicates mass eigenstate, and U_{lj} are the elements of the PMNS matrix. Then via further straightforward manipulations, using $\hbar = c = 1$, and taking the ultrarelativistic limit where $t = L$ (L being the distance the neutrino travels from production to detection) yields

$$P\left(\overset{(-)}{\nu}_l \to \overset{(-)}{\nu}_l'\right) = \delta_{ll'} - 4\sum_{i>j}\mathrm{Re}(U_{li}^* U_{l'i} U_{lj} U_{l'j}^*)\sin^2\left(\delta m_{ij}^2 \frac{L}{4E}\right) \tag{2.4}$$

$$\pm 2\sum_{i>j}\mathrm{Im}(U_{li}^* U_{l'i} U_{lj} U_{l'j}^*)\sin\left(\Delta m_{ij}^2 \frac{L}{4E}\right) \tag{2.5}$$

where E is the neutrino energy at production and $\Delta m_{ij}^2 \equiv m_i^2 - m_j^2$ arises from the difference in energies in the exponential, when taking the ultrarelativistic limit. The sign difference for neutrinos vs antineutrinos comes from δ_{CP}, which is what gives U its imaginary part. This dependence on the mass difference is why neutrino oscillations imply nonzero mass for at least all but one of the mass eigenstates.

Most neutrino experiments are not well-modeled by the vacuum approach. Solar neutrino experiments must account for the Sun and, at night, the Earth, long-baseline accelerator neutrino experiments send their neutrinos through the Earth's crust, and while downward-going atmospheric neutrinos are well-modeled as propagating through vacuum, upward-going atmospheric neutrinos have traveled through the entirety of the Earth. As such, it is important to account for the matter effect.

There are two sources of matter effect. The first is the neutral current interaction between neutrinos and (primarily) electrons, where a single Z boson is exchanged (fig. 2.1). However, this interaction is independent of flavor and so introduces only an overall phase, which is not observable. The other is the charged current interaction, which affects only electron (anti)neutrinos, and therefore does produce an observable effect.

This enters into the oscillation framework by way of a potential added to the Hamiltonian

$$V = \pm\sqrt{2}G_F N_e \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{2.6}$$

in the flavor basis where G_F is the Fermi constant and N_e is the number density of electrons and

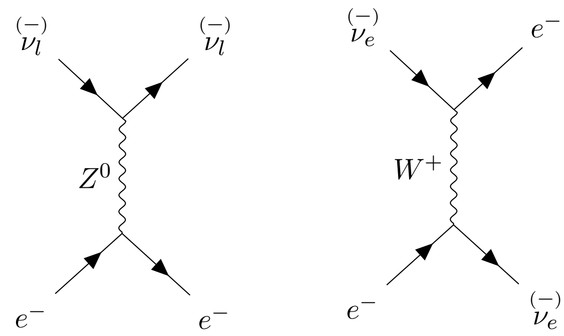

Figure 2.1: The neutral and charged current interactions most relevant to the matter effect.

the potential is positive for neutrinos and negative for antineutrinos. This must then of course be transformed to the mass basis and carried through as before. Predictably, this makes the already somewhat clumsy math quite unwieldy even in the simple case of a constant number density, and so will not be carried out here, especially as the form of the number density depends greatly on the experiment being considered. The most pertinent feature of the matter effect is that whatever correction it introduces to the oscillation will be opposite between neutrinos and antineutrinos in a way that does not depend on the value of δ_{CP} to first order, which can be leveraged to break degeneracies.

We could then of course proceed from here to substitute in the standard parameterization of the PMNS matrix to get particular appearance and disappearance rates in terms of the angles, though this quickly becomes unwieldy unless the conditions of a particular experiment are being examined and this is not relevant to the rest of this thesis. We now see that there are six parameters to be measured in neutrino oscillation experiments: the three angles and one phase of the PMNS matrix, and the two mass splittings (with the third being inferrable from the other two). Oscillation physics alone cannot make a statement on the absolute scale of the neutrino mass, besides that it is indeed very small since the ultrarelativistic limit holds.

2.4 The State of Neutrino Oscillation Measurements

Experimental evidence for neutrino oscillations is at this point robust, with precision determination of the PMNS matrix now the primary focus of ongoing research. It is useful to separate the discussion of neutrino experiments by their primary source of neutrino flux: accelerator-based production, nuclear reactors, the Sun, and cosmic rays impacting Earth's atmosphere. Because neutrino shielding is impossible, an experiment will detect all neutrinos above its detection threshold from any source. However, because such a high flux is needed for a chance at detection most experiments are dominated by some particular source.

This will not be an exhaustive list of neutrino experiments. Only the most representative experiment(s) in any particular category will be summarized, or any that are particularly relevant to the further contents of this thesis.

2.4.1 Accelerators

Accelerator neutrino experiments are costly to operate but have the significant advantage of a relatively controllable flux. They function by colliding a beam of charged particles (typically protons) with some target, and allowing the spray of hadrons from that collision (largely pions and some kaons) to decay in flight. This produces mostly muon neutrinos in a relatively narrow band of energies, typically spanning about an order of magnitude, though positioning detectors off-axis can significantly narrow the energy range at the cost of overall flux. Accelerator neutrino experiments allow for the determination of the baseline (L/E) with relatively few constraints, enabling the targeting of certain oscillation features. Prominent long baseline accelerator-based neutrino experiments include MINOS [14], K2K [15], ICARUS [16], T2K [17], and NOvA [18], with DUNE and Hyper Kamiokande both in development as the next generation of these experiments. Accelerator neutrino measurements are the primary contributor to the measurement of θ_{23} and also provide information on θ_{13}, δ_{CP}, and the mass ordering.

2.4.2 Earth's Atmosphere

Atmospheric neutrino experiments detect neutrinos produced by showering cosmic rays in the upper atmosphere. This flux presents several challenges, with an extremely wide range of energies (starting at around 100 MeV and scaling nearly arbitrarily high), omnidirectional origins leading to travel distance anywhere from tens to tens of thousands of kilometers (including through the Earth), and an appreciable mix of electron neutrinos, muon neutrinos, and their corresponding antineutrinos, with tau neutrinos becoming significant as well at high energies. Fortunately, the cosmic ray flux is well-studied, and the relevant neutrino channels are relatively easy to calculate, which allows for precision calculation of flux ratios. Atmospheric neutrinos are also notable as a background to proton decay searches. As previously discussed, Kamiokande and Super-Kamiokande both measured atmospheric neutrinos as well, and Super-Kamiokande was crucial in the proof of neutrino oscillations via the matter-effect-induced difference between upward-going and downward-going neutrinos. Due to the extremely large range of energies, there are similarly an extremely large number of experiments capable of observing atmospheric neutrinos at various energy ranges. Atmospheric neutrino measurements contribute primarily to θ_{23} and Δm^2_{23} (though not its sign).

2.4.3 The Sun

The solar neutrino flux is complex with a variety of features due to the intricacies of fusion chains, but in general the neutrino energies are relatively low, so they are difficult to detect. Besides the previously discussed Homestake experiment which used chlorine and SNO which used heavy water, there are gadolinium-based detectors such as SAGE [19] GALLEX [20] and GNO [21] which take advantage of the extremely low 233 keV threshold of $\nu_e + {}^{71}\text{Ga} \rightarrow e^- + {}^{71}\text{Ge}$. The Kamiokande experiment and its scaled-up successor Super-Kamiokande are both water Cherenkov detectors that are sensitive to the high-energy end of the solar neutrino flux, with Super-Kamiokande having the same threshold as SNO at 3.5 MeV. Finally, KamLAND [22] and Borexino [23] are both liquid scintillator detectors, with Borexino in particular focusing on extreme radiopurity to enormously reduce backgrounds, allowing it to detect geoneutrinos from radioactivity deep within the Earth and giving it the current lowest threshold of 190 keV. Solar neutrino observations are primary contributors to the measurement of θ_{12} and are important to measuring Δm^2_{12} and θ_{13}.

2.4.4 Nuclear Reactors

Nuclear fission reactors are very high-flux undirected source of $\overline{\nu}_e$ with a fairly narrow energy spectrum at the MeV scale. Despite being a anthropogenic neutrino source, the experiments typically have no control over the production of neutrinos, as the three major medium baseline experiments—Daya Bay [24], Double Chooz [25], and RENO [26]— all use commercial nuclear power plants as their neutrino sources. The neutrino flux produced then must be inferred from the heat output of the reactors. Each of these experiments use gadolinium-doped liquid scintillator as their detection medium. Each experiment makes different choices around where to place their detectors relative to the reactor vessels, though all are on the scale of hundreds of meters. Daya Bay has two detectors at around 500 meters each, while the other two each have a near and far detector. The major next-generation medium baseline reactor neutrino experiment is JUNO, which recently started taking data. Medium baseline reactor neutrino measurements are the primary contributor to measuring the value of θ_{13}. KamLAND is the premier long baseline reactor neutrino experiment, and is currently the dominant contributor to the measurement of Δm^2_{21}.

2.4.5 Other Sources

Geoneutrinos

The decay of various radioisotopes in the Earth, especially uranium and thorium, produces a small low-energy neutrino flux that is nevertheless measurable by some experiments. The discovery of this source was announced in 2005 by KamLAND [27], with work since then by KamLAND [28], Borexino [29], and SNO+ [30] increasing the precision enough to allow limited geophysical inferences to be drawn. This source, however, is generally not useful for oscillation analyses due to the unknown baseline and very low event rates.

Charged Pion Decay-At-Rest

Producing a subset of accelerator-produced neutrinos but with very different characteristics, charged pions decaying at rest are useful for probing coherent elastic neutrino-nucleus scattering (CEvNS). This is a neutral current process with a relatively high cross-section and interesting physics which was first observed by the COHERENT experiment in 2017 [31]. CEvNS is a neutral-current process and therefore gives no direct information on neutrino oscillation. However, it is an excellent probe of other properties of neutrinos that are predicted with great precision by the Standard Model, and therefore also of BSM physics [32].

Chapter 3

The NOvA Experiment

The NOvA (NuMI Off-axis ν_e Appearance) experiment is a long-baseline accelerator neutrino experiment, with the primary goal of measuring θ_{23}, θ_{13}, and Δm^2_{32}, as well as constraining δ_{CP} and the neutrino mass hierarchy, all aspects of the PMNS neutrino mixing matrix. To accomplish this, NOvA has both a near detector (ND) and far detector (FD), which are nearly identical up to size, with the FD being roughly 70 times larger (volumetrically) than the ND [33]. The FD position is optimized to observe the oscillation effects produced by the parameters NOvA seeks to measure, while the ND exists to control the systematic errors on these parameters. The ND also enables a variety of other physics programs, including neutrino cross-section measurements and beyond the Standard Model searches that require a lower cosmic ray background than the FD can provide.

NOvA began development in earnest in the early 2000s, with construction of the FD site beginning in 2009, and the ND following later. A moderately sized prototype, the Near Detector On the Surface (NDOS) constructed at Fermilab provided a technology demonstration and validated the construction techniques to be used for the more logistically difficult ND and FD. NDOS has since been decommissioned and its data is no longer used in analyses. After several rounds of preliminary results, NOvA published its first oscillation measurement in 2016 [34].

3.1 The NuMI Beam

The NuMI (Neutrinos at the Main Injector) beam [35] is produced by Fermilab using the Main Injector, a proton synchrotron so named as it used to fulfill this function for the Tevatron, since shut down. The Main Injector delivers 120 GeV protons in groups (called a “spill”) of approximately 5×10^{13} to NuMI at approximately 1.3 Hz, with each spill lasting about $10\mu s$. The overall power delivered has steadily increased over the lifespan of NuMI, starting at 200 kW in 2005 and increasing to sustained operation exceeding 800 kW in 2021, with a record one hour average power of over 1 MW in 2024.

The spill is directed at a water-cooled graphite target 350m from the extraction point, shaped as a row of 47 twenty millimeter-long fins, each separated by 0.3 mm for a total length of 953.8 mm [35]. This creates a spray of charged mesons, largely pions and kaons, which then enter an intense magnetic field created by an electromagnet consisting of two parabolic horns. These horns act to focus the spray in a charge-dependent way, so that (when running the current forward through the magnets, referred to as Forward Horn Current or FHC) positively-charged particles are focused into a beam while negatively-charged particles are de-focused. The charges are reversed for Reverse Horn Current (RHC) conditions. The now-focused beam of hadrons then proceeds into the decay pipe, a 675m long helium-filled tube, where they are free to decay in-flight and only very rarely interact with the helium. The positively (negatively) charged pions and kaons primarily produce (anti)neutrinos, though $\pi^{\pm} \rightarrow \mu^{\pm} + \overset{(-)}{\nu}_{\mu}$ (branching ratio 99.99%) and $K^{\pm} \rightarrow \mu^{\pm} + \overset{(-)}{\nu}_{\mu}$ (branching ratio 63.56%). However, charged kaons can also decay as $K^{\pm} \rightarrow \pi^{0} + e^{\pm} + \overset{(-)}{\nu}_{e}$ with a branching ratio of about 5%, and more rarely-produced (and unfocused) neutral kaons can produce ν_e as well, leading to some unavoidable ν_e contamination in the beam. At the end of the pipe, most undecayed charged particles are stopped by an absorber composed of aluminum, steel, and concrete. There is then 240m of rock between the absorber and the ND, which ranges out almost all of the remaining muons. A cartoon of this layout is in figure 3.1.

The detectors are placed 14.6 mrad away from the center of the beam. While this does reduce the overall neutrino flux they are exposed to, it significantly narrows the peak of the neutrino energy spectrum and shifts it downward to a baseline that is ideal for its oscillation physics goals (and which would correspond to a feasible Far Detector placement within the continental United States).

It is generally desirable to monitor and report the flux of any given particle beam. Unfortunately,

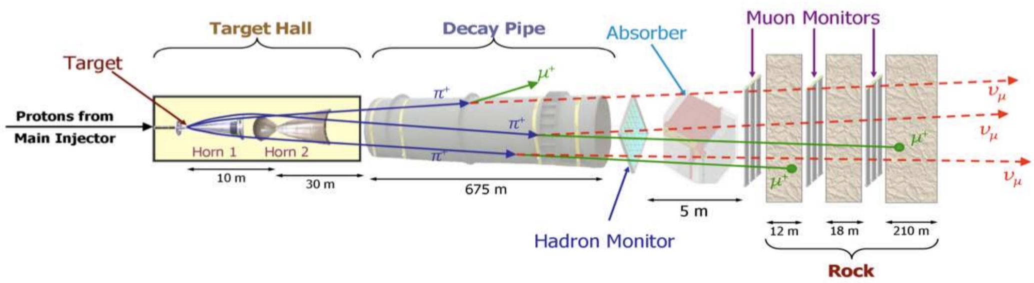

Figure 3.1: A cartoon of the NuMI beamline.

neutrino fluxes are not directly measurable. Therefore, as a proxy for the beam intensity, the Protons On Target (POT) is reported instead as it directly correlates with the resulting integrated neutrino flux. There is also monitoring equipment for the proton beam before the target as well as the interior of the decay pipe to give additional indirect information and help constrain the relationship between POT and neutrino flux.

3.2 The Detectors

NOvA's near and far detectors are constructed from panels (also called "planes") of PVC that have been glued together. Each panel was extruded as a series of 16 tubes of rectangular cross-section (approximately 4cm by 6cm), referred to as "cells", which are themselves glued together to form the panels. Each cell is filled with scintillator fluid (mineral oil doped with additives) and contains a loop of wavelength-shifting optical fiber to collect the scintillation light. Notably, the PVC was mixed with TiO_2 to increase the reflectivity within the cells, ensuring more light reaches the optical fiber. Both ends of this fiber are attached to a single pixel of an avalanche photodiode (APD), which amplifies the signal before it is digitized by a custom continuous readout panel called a Front-End Board (FEB) into units of Analog-to-Digital Converter (ADC) counts (which must then later be calibrated to transform back to photoelectrons and eventually energy deposited). This signal is then sent on to the remaining DAQ (Data AcQuisition) systems, to be

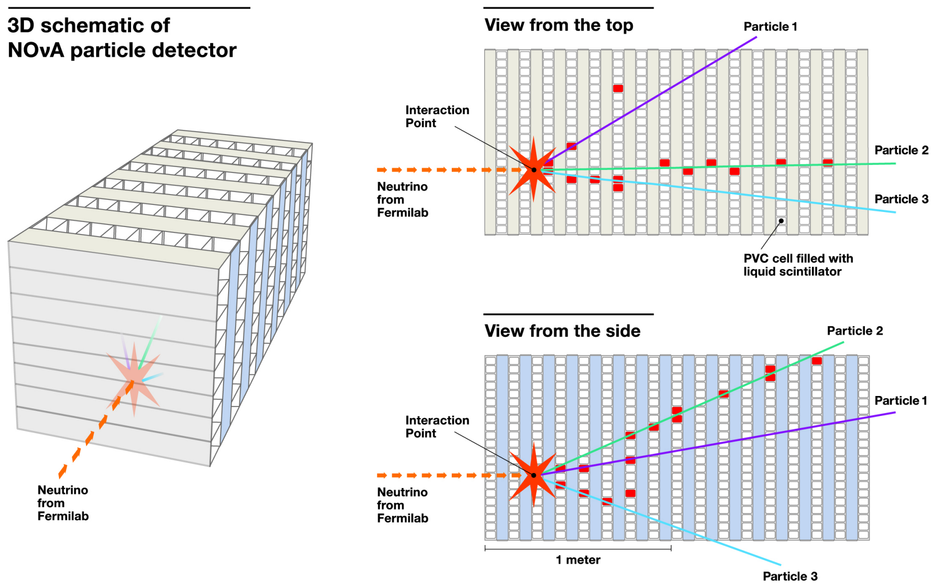

Figure 3.2: A cartoon of the physical layout of the NOvA detectors (left), and how this creates the two "views" (right)

detailed later. Every cell in the plane has the same orientation (horizontal or vertical), giving the plane an overall orientation. The orientations of the planes alternate between horizontal and vertical. Therefore, as a charged particle passes through the detector and ionizes the scintillator, (x, z) and (y, z) coordinate information can be extracted from the alternating planes, where x is horizontal, y is vertical, and z is along the length of the detector (in roughly the beam direction). A visual summary is in figure 3.2. This information is most easily conceived of as two "views" of the detector, with the horizontally-oriented planes giving a side view of events in the detector and the vertically-oriented planes giving a top-down view.

3.2.1 The Near Detector

The Near Detector (ND) is located 105 meters underground located 14.6 mrad off-axis from the NuMI beam center and approximately 1000 meters downstream of the target. The cavern was excavated adjacent and connected to the one used for the on-axis experiments MINOS ND, MIN-

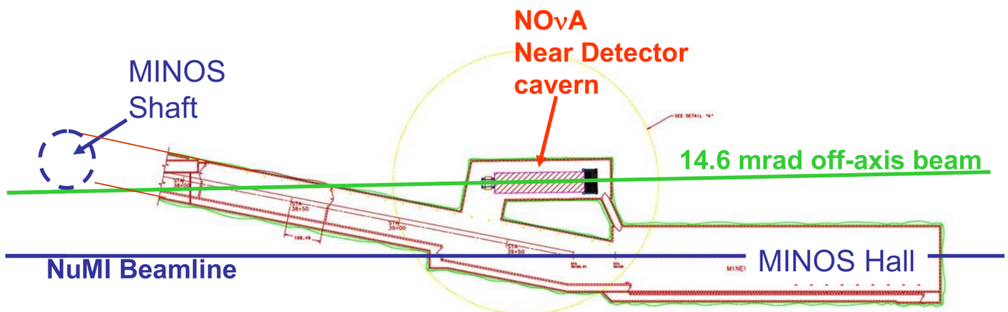

Figure 3.3: A top-down view of the MINOS and NOvA cavern layout at the time of NOvA's planning. From the NOvA TDR figure 5.9 [33]

ERvA, and ArgoNeuT, sharing the same vertical shaft (see figure 3.3). The underground location is for two reasons: the beam itself is pitched downward-going to be able to reach the Far Detector through the Earth's crust, and the rock overburden significantly lessens the number of cosmic rays in the detector. While neutrino detectors are generally large to make up for the small cross-section of neutrinos, the ND can afford to be relatively small due to the sheer intensity of NuMI at this distance. As-such, the ND is "only" 300 tons, approximately 4m wide by 4m tall by 16m long (in the beam direction), with the size considerations driven primarily by event containment and the costs of building a detector underground (including cavern excavation). To further extend the ND's effective energy range without needing to make it even longer, approximately the last three meters of its length is a "muon catcher", consisting of alternating steel plates and the standard PVC detector planes. The density of the plates causes muons to shed their energy much more quickly than they do in the detector medium, while the planes allow the muon's range to be estimated with reasonable precision. However, the trade-off is that non-muon particles cannot be reconstructed in this region.

3.2.2 The Far Detector

The Far Detector (FD) is located 810 km downstream of the target and 14.6 mrad off-axis from the beam center, in northern Minnesota (see figure 3.4). At this distance, the beam spot is very

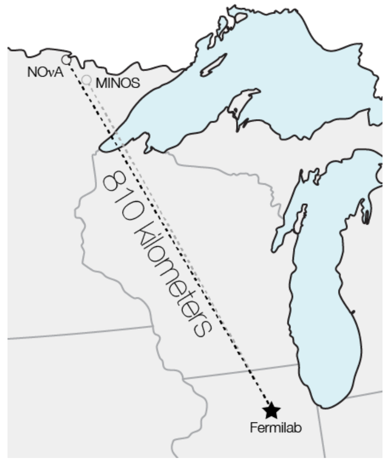

Figure 3.4: The NOvA baseline. Geography both political and physical has enormous impacts on experiment design considerations for these large experiments.

wide, and so the FD's size is more commensurate with what one would expect from a neutrino experiment, at 14 kilotons and approximately 15 m wide, 15 m tall, and 60 m long, subtending only a small fraction of the beam flux's angular distribution. A photograph of the FD is in figure 3.5. Notably, this is of sufficient length to not require a muon catcher, and so the detector is uniformly composed of PVC planes. While the FD sits below grade, it is more at the bottom of a hole than within a cavern, with an overburden of of loose rock and dirt from the excavation on top of a concrete roof which serves as approximately 3.6 m water-equivalent shielding. As such, the FD has a very high rate of cosmic rays, approximately 100 kHz. The FD complex also contains the Shanahan computing center, where initial data processing occurs before being sent to Fermilab.

3.3 Data Acquisition

The data acquisition (DAQ) system of NOvA consists of several subsystems that work together to produce NOvA's data: the readout, timing, triggering, monitoring, and storage systems. A detailed explanation of these systems is well beyond the scope of this thesis, but I will give a brief overview of each.

Figure 3.5: A picture of the NOvA Far Detector. Prominent in the foreground is the device used to tilt the planes into place during construction, now used as a structural support. Some of the readout electronics are visible on top. The walking surface of the top floor is roughly at grade.

3.3.1 Readout

A partial description of the readout has already been provided as part of section 3.2, up to the FEBs. After the FEB has digitized the signal, if it is above threshold, it is sent to a Data Concentrator Module, or DCM. The DCMs pack together signals from up to 64 FEBs into ”millislices” consisting of 5ms of data. The DCMs also timestamp this data using their associated TDU (see 3.3.2). These millislices are then sent to buffer nodes, a collection Unix machines (a few hundred in the FD, a few dozen for the ND, the exact number variable due to failures and replacements) which receive information from the entire detector and assemble it into prospective events based on timing. The underlying data structure of the buffer nodes is a circular buffer, and data will eventually fall off of it if not indicated to be stored by a trigger, further discussed in section 3.3.3.

3.3.2 Timing

A crucial part of assembling events is ensuring that the readouts of every spatial part of the detector are synchronized in time to significant precision to allow for both event generation at all but slicing in particular (see section 3.5). NOvA's timing system allows for precision in the 10s of nanoseconds, with its fundamental unit of time being approximately 8 nanoseconds. This is accomplished through the use of "timing chains", of which each detector has two, an active and a hot-swappable backup. The chains are composed of a string of Timing Distribution Units (TDUs), custom hardware with both FPGA firmware and a Unix kernel. Each chain has a primary TDU at the head of the chain and then several secondaries following, with the primary additionally connected to a GPS unit for a timing reference. Timing calibration is done by bouncing a calibration signal down and then up the chain, allowing each secondary TDU to set its offset from the primary based on the difference between its internal timer and the timestamp included with the signal, and then verified by the return signal. The primary TDUs also connect to web monitoring software, and every TDU can be connected to via Fermilab's network.

I served as both a DAQ expert generally for NOvA as well as a timing expert in particular. As DAQ expert I took shifts responding to calls from the regular shifter to fix various issues with the DAQ as the occurred. As timing expert I was called in to both respond to particularly tricky issues with the timing expert that the on-shift DAQ expert could not solve as well as remote upkeep tasks for the timing system, such as flashing firmware updates.

3.3.3 Triggering and Storage

Triggers on NOvA can be either external or internal. External triggers always read out a certain amount of data around a certain timestamp, with the most important example of this being the beam trigger sent by the accelerator when a spill is delivered to the target. There is also the cosmic trigger, which periodically reads out 500 μs from the FD to allow for cosmic ray callibration and studies, and the Super Nova Early Warning System (SNEWS) trigger, which reads out all data from both detectors for 45 seconds (starting 5 seconds prior to the trigger). Other triggers are the internal, or Data-Driven Triggers (DDTs), which result from processing data in the buffer ring and making a triggering decision from that. DDTs are often designed for rare-process searches, such as fast and slow magnetic monopoles. All triggers are sent to the Global Trigger Processor, which broadcasts the trigger to all buffer nodes and the data logger, which is responsible for reading the data out to local disk and sending information to the monitoring processes (see section 3.3.4). Local data is then picked up by the File Transfer System and sent to long-term storage at Fermilab.

3.3.4 Monitoring

The extreme complexity of NOvA's DAQ inevitably leads to periodic issues as entropy catches up to the various systems. As such, it is important to have a human in the loop continuously monitoring the systems at a high level, as well as having various more specific diagnostic information available for experts to use to fix problems. The Run Control program, DAQ Health Monitor, and other monitoring programs for specific systems all communicate through the Message Passing System, which is documented by the Message Logger, to display the status of each part of the DAQ to the control room. Part of this monitoring is an active event display, which shows both random events from all triggers as well as specifically every beam trigger. There are also "nearline" plots which are updated on the half-hour to provide detailed information about detector function, and "offline" plots which are generated at longer intervals, largely to monitor data quality.

3.4 Calibration

The process of converting from the ADC counts measured by the APDs to energy deposited is extremely important, and a significant source of uncertainty for NOvA. The ND and FD are calibrated separately. Calibration is performed using non-NuMI triggers: in the FD, the cosmic ray trigger (10 Hz steady trigger) and the horizontal muon trigger (throughgoing muon tracks) are used, while the ND, with its much lower cosmic rate, uses an activity trigger that is roughly equivalent to standard quality (but not containment) cuts.

For a given hit to be taken into consideration for calibration, it must be either sandwiched between two other hits in its same view in the X or Y direction or between two hits in its same view in the Z direction. This guarantees that the particle passes all the way through the cell, allowing for very accurate path length reconstruction, especially for the X and Y directions.

With the collection of calibration hits established, the calibration is then performed in stages. First, an approximate calibration is applied globally to take things to units of photoelectrons (PE) from the ADC counts. Then, relative calibration is performed to get the energy distribution flat across the entire detector. This involves compensating for optical attenuation along the fiber to the APD, as well as a threshold effect where a higher energy deposition must occur at the far end of the cell for a hit to be registered at all. Finally, the nonuniform dE/dx across the detector must be accounted for as the muons are slowed through the length and/or width of the detector, especially prominent in the larger FD.

Finally, absolute calibration must be performed to take the photoelectron counts to deposited energy in GeV. This is the largest source of calibration uncertainty, as setting any absolute energy scale is notoriously tricky in general. However, it can be validated against "standard candle" interactions, such as the π^0 invariant mass and the Michel electron spectrum.

3.5 Reconstruction

NOvA's reconstruction begins with the files produced by the DAQ, which consist of a series of events. An event is defined as a group of "hit" information, where a hit is itself a group of information consisting of a cell ID, an amount of charge deposited, and the time at which that deposition was detected. The first step of the reconstruction chain is the slicer, which "slices" the

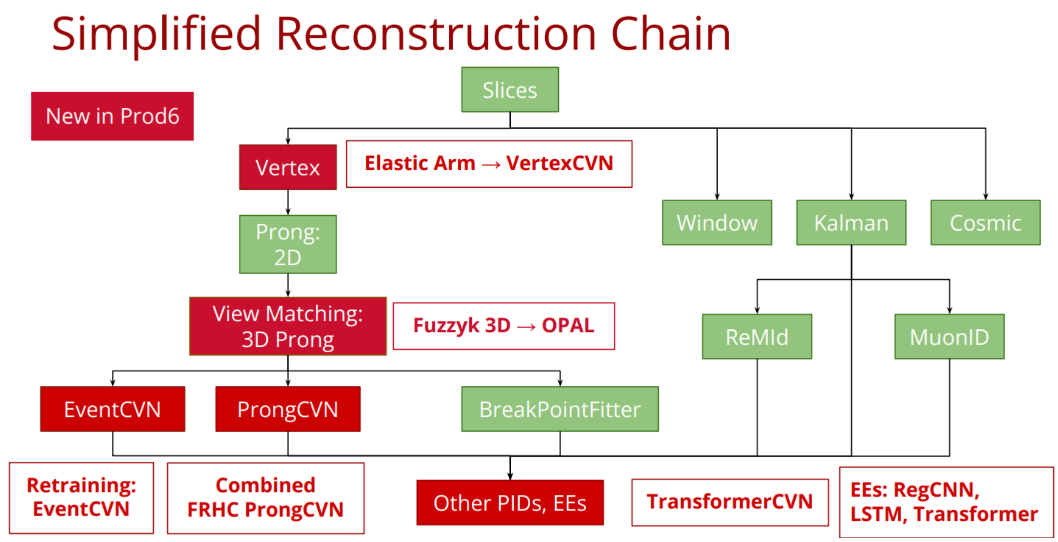

Figure 3.6: A simplified schematic of the NOvA reconstruction chain. Red blocks are changed between productions 5.1 and 6, which will be explained at the end of this section. Credit: Alexander Booth

hits contained in an event by grouping them in space and time, seeking to create slices consisting of the hits corresponding to a single neutrino interaction (or cosmic ray, or muon originating from outside the detector). From here, I will describe the reconstruction chain used in all currently (as of writing) published NOvA analyses. This chain consists of several branches. I will begin with the least complicated and move towards the most. Note that this is a simplified picture. A schematic of the updated chain for future productions is in figure 3.6, which is still largely similar to the old chain.

In the simplest branch, consisting of only one process, cosmic ray rejection is performed. This is currently accomplished using a Convolutional Neural Network (CNN) which is fed only hit-level information, and is run very early in the process to avoid performing additional reconstruction on cosmic slices, shrinking file size and saving time.

In another branch, we start with the Kalman tracker. This is a classical algorithm [36] which seeks to model hits with the physics information of a charged particle moving through the detector with expected trajectory deviations coming from Coulomb scattering. As such, this method excels for particles that create long tracks and do not shower, namely muons and charged pions. This method also inherently uses dE/dx information which is crucial to particle identification, and therefore must be repeated under different particle assumptions. Notably, this process takes place in both views simultaneously, allowing for inherent 3D modeling. This creates 3D track objects, collections of hits with associated information measured by the Kalman algorithm. Two different particle identification algorithms are applied to these tracks, ReMId and MuonID. Both of these seek to identify muons, but MuonID is newer while ReMId is still used in ν_e analyses. ReMId has a particular focus on discriminating muons from charged pions, where charged pions are a product of the significant neutral current background. It is a k-Nearest Neighbor machine learning algorithm that takes as input the dE/dx log-likelihood (LL), scattering LL, scattering length, and fraction of planes used in the dE/dx LL, all of which are outputs from Kalman tracker. MuonID was developed to reduce the model dependencies inherent in Kalman, in particular the track length. It is a Boosted Decision Tree (BDT) that uses only dE/dx information and the scattering LL to select for muons, resulting in less model dependence.

3.5.1 Elastic Arms

In the final significant branch, we start with vertexing. First, a Hough transform is taken of all hits which, in brief, identifies points that bring lines of other points "into focus", i.e., functionally intersections of many lines [37]. This is used to seed an Elastic Arms algorithm [38]. The idea of this algorithm is to model the hits as a vertex from which extends some unknown number of "arms", lines of hits in either a shower or track topology. This algorithm seeks to minimize the penalty function

$$E = \sum_{i=1}^{N}\sum_{\alpha=1}^{M} V_{i\alpha}M_{i\alpha} + \lambda \sum_{i=1}^{N}(\sum_{\alpha=1}^{M} V_{i\alpha} - 1)^2 + \frac{2}{\lambda_\nu}\sum_{\alpha=1}^{M} D_\alpha \tag{3.1}$$

Where N is the number of hits, M is the number of arms (as an adjustable parameter of the fit), $V_{i\alpha}$ is the "association" between hit i and arm α, $M_{i\alpha}$ is the distance between hit i and arm α, and D_α is the distance between the vertex position and the first hit associated with arm α. All terms but the last are standard to the Elastic Arms method, but because the vertex position is not known in this application the third penalty term is required to prevent the vertexer from favoring far upstream positions (with only partial success, as will be discussed in chapter 4). This minimization is started from the seed position of the vertex, and is very sensitive to the seed position, tending to fall into local minima. As such, this method will fail, sometimes spectacularly, if the Hough transform fails to produce a seed close to the true vertex position. It is also computationally costly to perform this minimization procedure on several different seeds. All of these concerns have been addressed by my VertexCVN algorithm (chapter 4), which is being implemented in the next production run.

3.5.2 Prongs

While the Elastic Arms method has relatively good performance in identifying the vertex, it struggles with both the number and direction of the arms themselves. To remedy this, only the vertex position is used going forward, and instead an algorithm based on Fuzzy K-Means [39] [40] is used to cluster the hits in the slice into prongs, by tracing clumps of hits back to the vertex. A prong is a generic term used to refer to any group of hits attributed to a single primary particle. Making prongs is done in each slice separately, so only two-dimensional information is used at this stage. In particular, this method has good performance for the two prong topologies of NOvA: tracks, which are relatively straight lines of charge deposition from a single charged particle undergoing Coulomb scattering, and showers, commonly called electromagnetic showers, in which a photon or electron pair-produces within the detector, leading to a rough cone of charge deposition.

Fuzzy K-Means generally works by minimizing a function to give points (in this case hits) partial membership to each of several clusters (in this case prongs), with each cluster having some cluster center and each point's membership fractions summing to one. This is not suited to NOvA's needs for two reasons, firstly that the number of clusters needed is not known and must be optimized, and secondly that some hits are noise hits and should not be associated with any spatially localized cluster. Therefore, NOvA instead implements a Possibilistic Clustering Algorithm (PCA), which removes the normalization constraint on cluster membership and thus allows hits to be implicitly assigned as noise if they lack strong membership to any cluster.

Because PCA seeks to cluster hits around points in space, a coordinate transform is required to make it work in NOvA's topology of prongs of associated hits radiating out from a vertex. As such, EA's vertex is used to convert each hit into an angle based on its cell center, with associated uncertainty (decreasing the further it is from the vertex). Some predetermined number of cluster centers (in this case one) are seeded based on the angular density profile of the hits. Then, membership is calculated based on distance from the cluster center, the cluster centers are updated to move closer to the center of their associated hits, and if the centers move by more than some threshold then this process is repeated until they are all below threshold. Next, if two cluster centers have converged to the same angle, the duplicate is removed and distances and memberships are recalculated, and the removed duplicate is re-seeded elsewhere in the density profile at some local maximum that has not been used for a seed yet, and the process is repeated. If no such local maximum exists then the number of clusters is simply reduced by one and the process continues. Finally, once the cluster centers are all unique, a check is performed to see if an additional center should be added. If the membership of any hit is below some threshold (1%), then the density profile is re-calculated using only hits below this threshold, a new seed is chosen as the maximum of this profile (if it has not already been used), and then the above is repeated with this added new seed as well as the already-optimized ones. This whole thing is repeated until all cells are members of a cluster to above some threshold or some maximum number of clusters is reached.

3.5.3 View Matching

These 2D prongs then must be matched to produce 3D prongs. This process is significantly more difficult than one might expect, since especially at high energies prongs tend to partially overlap in at least one view and z-coordinate information is never available simultaneously in both views, leading to potential ambiguity. This is done by performing a pairwise Kuiper test between all of the 2D prongs of each view, as follows:

$$K_{ij} = \max\left[E_{X_i}(d) - E_{Y_j}(d)\right] + \max\left[E_{Y_j}(d) - E_{X_i}(d)\right] \tag{3.2}$$

Where the i index goes over all 2D prongs in the X view, the j index goes over all 2D prongs in the Y view, and E is the fractional energy deposition at distance d from the start of the prong. Prong ends must be within one plane of each other. The lowest K_{ij} is deemed a match, and those two 2D prongs removed from further consideration, then the next lowest K_{ij} are matched, and so on. Any remaining prongs are orphaned, which is common in events where two prongs substantially overlap in one of the views.

3.5.4 Break Point Fitter

Break Point Fitter (BPF) [41] is an algorithm to refine the reconstruction of specifically tracks. It is performed on the 3D prongs from the previous subsection. It makes the assumption that the track is primarily in the z-direction, limiting its use to forward- or backward-going particles. The general idea of the algorithm is to assume that the particle will undergo some finite number M scattering events within planes located at z coordinates Z_J along the length of the detector, to be determined, and that we have available to us n measured positions of the particle (x_i, y_i, z_i) along the track with some transverse precision σ_i. For simplicity, only the x_i will be used to represent overall transverse position. Putting this all into math, the trajectory of the particle can be represented as

$$\xi_i = a + bz_i + \sum_{J=1}^{M} \alpha_J (z_i - J_J)\Theta(z_i - Z_J) \tag{3.3}$$

Where ξ_i is the estimated transverse position of the particle between measurements points i and $i+1$, a is the vertex, b is the initial slope, α_J is the scattering angle (assumed to be "small enough" to not cause issues with this representation) at scattering plane J, and Θ is the Heaviside function. The α_J and Z_J are fit using a χ^2 minimization, with

$$\chi^2 = \sum_{i=1}^{n} \frac{(x_i - \xi_i)^2}{\sigma_i^2} + \sum_{J=1}^{M} \frac{\alpha_J^2}{\sigma_J^2} \tag{3.4}$$

Where σ_J is the uncertainty of the scattering angle α_J calculated under a multiple scattering hypothesis. This detailed fit allows for significantly more precise information about tracks to be extrapolated, including and especially a better length estimate than naively comparing the endpoint to the vertex.

3.5.5 Additional Reconstruction

From here, a variety of classifiers and energy estimators are used, both for the overall event and for each prong. These use a variety of machine learning frameworks, from BDTs to Convolutional Visual Networks (CVNs) and even more sophisticated networks. These include EventCVN, which uses a modified version of MobilenetV2 to classify events as $\overset{(-)}{\nu}_\mu$, $\overset{(-)}{\nu}_e$, or neutral current (NC), and ProngCVN, which uses a custom architecture to identify particle ID using either both prong-only views and full-event context or simply prong-only views to reduce model dependence.

3.6 Simulation

In order to interpret its data and develop and test reconstruction techniques, NOvA requires a simulation of its expected conditions. Because of the variety of problems involved, there are several different simulation tools that are used in sequence to provide the final simulation products. All of these tools use Monte Carlo (MC) methods, which seek to model complex random processes by independently or semi-independently choosing different quantities based on known distributions, relying on a large number of simulated events to in aggregate approximate the difficult or impossible to directly calculate underlying distribution through leveraging the law of large numbers. The sequence of simulation tools is NuMI flux simulation, neutrino interaction generation with GENIE, propagating the products of those interactions through the detector with GEANT, conversion of the charge deposition within the detector to scintillation photons, and finally the response of the detector electronics to those photons.

3.6.1 Beamline Simulation

The first piece of the NOvA simulation is shared among the various experiments that use the NuMI beamline to simulate the NuMI flux. A custom combination of GEANT4 [42] and FLUKA [43] [44] called G4NuMI is used to simulate the 120 GeV protons from the main injector interacting with the graphite target and producing hadronic showers. The flux is very sensitive to the exact properties of these hadronic showers. However, the model uncertainty on these showers is relatively high due to the difficulty of simulating strong force interactions in the complex nuclear environment. As such, 100 systematically varied universes are thrown to account for this source of uncertainty, which is not dominant in most analyses but is always significant.

3.6.2 GENIE

The simulated neutrinos from the previous section are then modeled to interact with the detector and the surrounding environment. These interactions are modeled by GENIE [45], a widely-used generator for neutrino interactions. GENIE works by using splines to interpolate between theoretically calculated and/or data-informed differential cross-sections in various variables. It can use any of several models for every step of the process, from nucleon momenta to specific generation processes. Additionally, it can be tuned to better match the actual observations of an experiment by adjusting these splines.

GENIE factorizes event generation by making discrete choices at each step, selecting first the hit nucleon or parton and its momentum, then performing the initial production of particles from the first interaction, then simulating hadronization and from there performing individual hadron transport nucleus, including final state interactions (FSIs). However, this is unphysical, as in reality the physics happens all at once in a wavefunction picture and is affected in difficult-to-model ways by being within the nuclear potential, which is itself not well understood. In particular, hadronization and FSIs are not well-separated in either time or space, so factorizing the two causes systematic issues.

Model uncertainties are encoded as part of the splines used. In order to account for these systematic uncertainties, a multiverse approach is used. Some number of universes are generated by varying the relevant splines within their uncertainty band. This new set of splines can then be used to either generate entirely new events, or to apply weights to existing events to mimic the output that would result from complete re-simulation. GENIE's capability to simplify systematic assessment via reweighting rather than having to run the entire simulation chain from scratch is one of its most important features.

3.6.3 Particle Transport and Detector Response

Once the list of primary particles is generated by GENIE, the particles must be propagated through the detector and surrounding environment. This is accomplished using GEANT4, one of the most popular particle transport models across a variety of energies due to its wide toolbox of available methods. GEANT4 considers each particle as a track which is composed of a list of steps, where each step contains a snapshot of the particle's state at some place and time.

The NOvA detectors and an approximation of their surrounding environment are loaded into

GEANT4 as a geometry, that is, a hierarchy of volumes of specified materials. For the NOvA detectors this geometry is precise, with the detectors being periodically surveyed for their exact position, orientation, and dimensions. As particles pass through volumes, their ionization energy loss and chances of inelastically interacting with the material they are in are simulated at each step. This information is then fed into custom NOvA simulation which simulates the creation of scintillation light, its propagation within the detector cells, and its uptake and transport along the optical fibers. The APD response and digitization are then both simulated, and triggering is skipped as we want to keep all simulated events. At this point the output of the simulation is indistinguishable from real data to the downstream reconstruction, up to true information tracking, and so processing continues identically up to associating reconstructed objects with truth information.

3.7 Productions

The subject of the process by which all of these systems come together to produce analyzable data is worth a brief discussion. NOvA analyzes data in “productions”, with each major production having the title “production X” while minor productions are named for the project for which they were conducted. A major production has two halves: simulation and reconstruction. For the simulation portion, the beam flux, event generation, particle propagation and response are all re-simulated for both nominal and various systematically shifted samples using some frozen configuration of software. For reconstruction, both the re-simulated MC as well as the collected data up to that point are reconstructed using the latest validated methods. Taken together, this produces fully compatible simulation and data with which to perform analyses.

Various parts of this thesis will mention “production 5.1” or “production 6”— production 5.1 is the previous major production configuration, last run for the Neutrino 2024 oscillation results, while production 6 is the latest major production configuration, which changed simulation and substantially changed reconstruction from production 5.1. There are no published production 6 results yet, due to a lengthy validation process needed because of the large number of changes made. Production 5.1 receiving the designation of major production rather than production 5 is best considered a historical accident, the details of which are not relevant to this thesis.

Chapter 4

VertexCVN

Determining the coordinates of the initial interaction of a neutrino with the detector, called the "vertex", is one of the most crucial steps in NOvA's reconstruction due to most of its reconstructed physics observables depending on the accurate identification of this vertex. NOvA's vertexing was implemented very early on in the lifetime of the experiment, and a hand-scan organized by a former Indiana University postdoc [46] revealed that it was due for an update. This coordinated event-by-event comparison of NOvA's reconstruction of simulated events with the simulated truth by human experts led to the realization that in the majority of instances where NOvA's reconstruction significantly failed, it was due to easily identifiable mistakes in vertexing. Because human vision was performing better than the Elastic Arms algorithm, it stood to reason that machine vision methods, which had already seen success in NOvA [47], would likely at least outperform the previous vertexer.

This chapter will start with a brief overview of the history of machine learning techniques, with a particular focus on two classes of techniques that have seen widespread use in particle physics and which will both be relevant to the work presented in this thesis. I will then give a more specific overview of the basic methods of operation of the type of machine learning architecture implemented in NOvA's new vertexer, VertexCVN. Finally, I will describe the creation, training, results, and validation of VertexCVN, and in particular how it compares to the previous vertexer Elastic Arms.

4.1 Machine Learning

Machine learning (ML) in general refers to any technique or algorithm for artificial intelligence that does not rely entirely on human coding but is otherwise semi-independently derived by the computer itself, often involving extrapolating from example data. With preceding theoretical efforts by Turing and many others, the practical form of the field began in 1958-59, with the near-simultaneous implementation of the Perceptron architecture [48] by Frank Rosenblatt in 1958 in the form of the Mark I Perceptron, the first implemented artificial neural network which was capable of rudimentary binary classification of images, and with Arthur L. Samuel's 1959 invention of a program that learned to play checkers [49]. Both neural network and non-neural-network techniques would remain important to the field, though the enormous hardware demands of neural networks would delay their maturation until the development and widespread availability of GPUs.

ML techniques have seen great success in NOvA, and there has been a push in the past several years to improve or supplement various aspects of NOvA's reconstruction with ML. I became involved with ML in NOvA through attempting to develop a secondary vertexer— NOvA's prong-making algorithms struggle with inelastic scatters of the primary particles on the detector medium, and the reconstruction downstream of the prong-making assumes that each prong corresponds to a single particle (or EM shower). As such, supplementing the reconstruction with a secondary vertexer was a long-standing goal of the NOvA reconstruction group, and would in particular allow for much better charged pion reconstruction due to their high cross-section at energies relevant to NOvA.

While I made some technical inroads on the secondary vertexing problem, I will not go into more detail here as fairly early on my efforts were pulled into primary vertexing, as improving NOvA's main vertexer was considered a top priority for the next production run due to it having been identified as the most prevalent weakness in NOvA's reconstruction chain. After I present an overview of relevant ML concepts, I will detail my work on NOvA's new primary vertexer.

4.1.1 Boosted Decision Trees

Boosted Decision Trees (BDTs) are a form of non-neural-network ML that are in widespread use throughout particle physics, due to their simplicity of implementation, ease of training, and broad

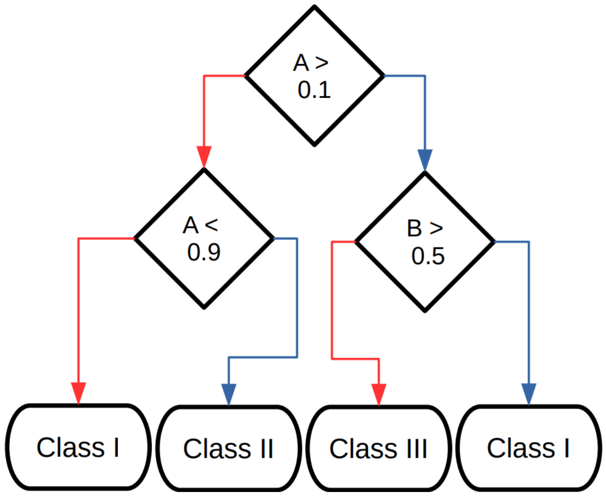

Figure 4.1: An example decision tree with input variables A and B sorting an event into one of three classes. Red arrows (on the left) are followed if the condition in the diamond is true, while the blue arrow is followed if the condition is false. For example, an event with $A = 0.05$ and $B = 0.7$ would follow a blue arrow and then a red arrow to be put in class III.

applicability to a variety of problems. I give an overview of their background here because I will cover the training and implementation of a BDT in my analysis in section 6.5. There are two fundamental concepts here, the "decision tree" and the "boost".

A decision tree is an extremely rudimentary classifier which can, on its own, be optimized using machine learning techniques. It is a binary tree where each node compares some input variable to some threshold and directs the decision down one of two branches, which can lead to either another decision node or to a "leaf" with a classification label (see a simple example in figure 4.1). The variables to decide on for each node, the thresholds in each node, and for non-binary classifiers the labels to assign to the terminal leaves are the parameters to be learned.

A single decision tree is rarely better than what a patient physicist would be able to come up with on their own. However, precisely due to their simplicity, decision trees are a good candidate for boosting [50]. Boosting is the principle by which an ensemble of weak classifiers can be assembled to produce a strong classifier. Because decision trees require minimal time to process and memory to store, it is relatively easy to assemble and train large ensembles of them, while other ML techniques such as neural networks could not practically be created in large enough groups for boosting to be effective. A BDT then is a “forest” of decision trees, where the output of the BDT is the weighted average of the outputs of each tree in the ensemble, where the parameters of each decision tree and the weight on the output of each tree are the parameters to be learned. Typically the binary class labels of -1 and $+1$ (or 0 and $+1$) are used so that the output is a real number in a predetermined (if the weighting is normalized) fixed range.

BDTs have seen widespread usage in particle physics since MiniBooNE [51]. They are generally accepted by the community as reliable and minimally biased, provided training is done carefully. Their adoption has been promoted by their inclusion in the TMVA [26] package of ROOT [52], itself the software package used for data analysis by NOvA and many other experiments.

4.1.2 Neural Networks

Since the field of machine learning often seeks to imitate biological learning, computer scientists have often sought to mimic biological neural networks. To this end artificial neurons are designed. Historically, these were occasionally physical devices, though today they are simulated in software. The typical form of the artificial neuron is a function which takes several inputs, weights them, adds them, and then applies some “activation function” to that result. This mimics biological neurons taking in stimulus through the dendrites, processing it through chemical reactions, and then distributing the output (often binary or nearly binary) through one or more axons.

Notably, the first two steps of an artificial neuron’s function are exactly analogous to matrix multiplication. Therefore, to prevent a neural network from being reducible to a single matrix operation, non-linearity is introduced through the activation function. Neural networks which take the biological metaphor more literally or which seek to mimic a particular biological system as fully as possible have used functions resembling the logistic. However, this is often computationally

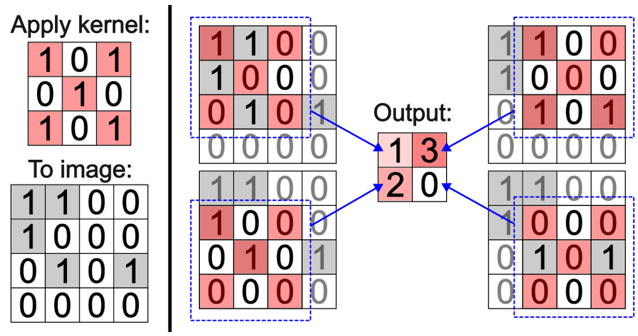

Figure 4.2: Toy illustration of a convolutional layer of a CVN. Only ones and zeroes are used for convenience. In principle both the kernel and image values could contain any real number, in practice they are typically clamped for computational efficiency. The kernel is scanned across the image, in each place multiplying the value in the image by the value in that place of the kernel and summing the results to the output layer. Typically, this process is repeated with several different kernels, producing a stack of outputs.

expensive for little to no (and sometimes negative) gain in performance, and as such threshold activation (the Heaviside step function) or piecewise-linear activation functions such as ReLU have gained significant popularity.

4.2 Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are a class of neural networks which rely less on the biological metaphor and instead aim to bake signal processing schemas directly into the architecture of neural networks. The basic concept behind CNNs is, rather than utilizing computationally expensive all-to-all dense layers, only "nearby" neurons should be processed by using some kernel, where the kernels are the object to be learned. This approach is very general, but the most commonly used subset of CNNs are Convolutional Visual Networks (CVNs), which process images for machine vision. These networks are also easier to explain and understand than general CNNs, and so I will focus my attention on them from here on.

CVNs work on the specific domain of two dimensional images processed by square kernels, which are typically small (three to nine pixels wide). A simplified example of the basic method of operation is shown in figure 4.2. Images are often processed in grayscale, leaving the problem fully two dimensional, which is appropriate when only spatial information is relevant to the problem at hand. Color images can also be processed by treating the red, blue, and green channels separately for the square convolutions, while mixing information between the color layers with depth-wise convolutions. CVNs have proven adept at various classification problems.

This basic defining convolutional operation can then be supplemented by a variety of other techniques, such as restricting the dimensionality of the embedding through bottlenecks, adding in links to aid gradient descent, pooling layers to reduce dimensionality, and more. Development of these supplements and the exact ways to structure the convolutions to best solve various problems is still an are of active research in computer science.

4.3 VertexCVN

Vertexing is one of the first steps in NOvA event reconstruction, taking place after slicer (see section 3.5) and before everything else. NOvA has been using Elastic Arms for its vertexing algorithm up to this point [53]. While it performs well in most cases, it has several known failure modes [46]:

- Events dominated by a single long prong often have the vertex forward-reconstructed, splitting that prong.
- Events that have significant scattering, either hard scattering in the detector medium or tracks substantially bent through soft EM scattering, can cause the vertex to be reconstructed in empty space, often behind the start of the event.
- Hard secondary interactions, especially with charged products perpendicular to the beam direction, can cause the vertex to be placed in the middle of the event.

Given the success of machine learning techniques in NOvA, [54] [55] [56], I have applied a CVN to this problem in order to improve reconstruction accuracy and minimize these failure modes. The

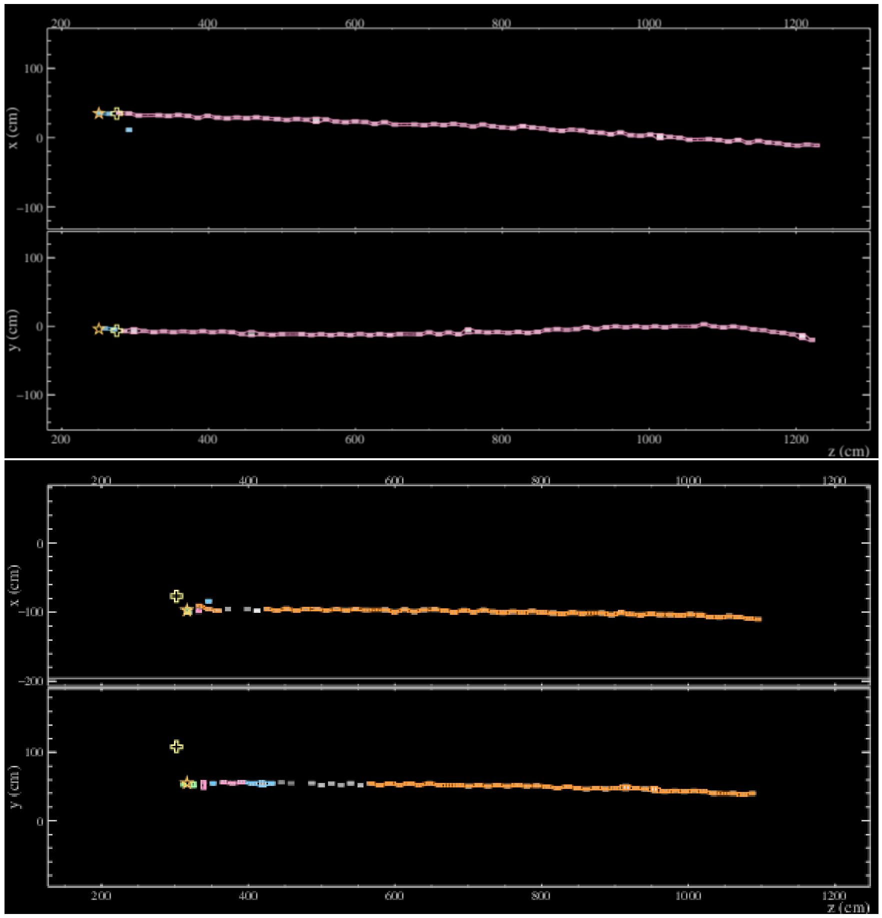

Figure 4.3: Two simulated event displays demonstrating failure modes of Elastic Arms. The orange star is the true interaction point, the yellow cross is the reconstructed vertex, and the colors are the reconstructed prongs. Top: a quasi-elastic event, $\overline{\nu}_\mu + p \to \mu^+ + n$ where the neutron has low momentum. The muon track has been split into two separate prongs, causing the muon energy to be underestimated and an additional particle to be hallucinated. Bottom: a DIS event, $\overline{\nu}_\mu + p \to \mu^+ + p + \pi^-$. All products are visible. Though it's unclear exactly how and there's no clear pattern to which events get reconstructed like this, the hadronic activity has caused EA to place the vertex in empty space away from all activity. Because the vertex is so far off, FuzzyK has mangled the event beyond usability.

figure of merit for optimization was chosen to be the fraction of events where the reconstructed vertex was within 13cm of the true vertex. This was derived from the reconstruction survey handscan [46] which demonstrated a two plane threshold for the resulting prongs to be notably misreconstructed.

4.3.1 Architecture

The architecture for VertexCVN is largely copied from EventCVN, which is itself a modified MobileNetV2 [57]. While this architecture was created to solve classification problems, I will show that it is also very well suited for this location-finding problem with only minor modifications.

Input for the CVN is a pixel map with 100 pixels in Z (i.e. 100 planes in each view) and 80 pixels in X and Y (i.e. 80 cells in each view), for a total size of 100x80x2. The pixel map is created by searching for the most upstream eight-plane window that that contains more than four hits, and then centering the window on the median hit location in the first plane of each view [47]. This pixel map is inherently split into the two views, each of which is processed by identically architectured but otherwise separate sets of layers. Notably, these are *not* processed as different color channels of the same image, as that would imply a spatial relationship between the X and Y coordinates that is not reflective of reality. The first layer provides initial encoding, with a 5x5 kernel and a stride of 2, with 32 filters. Following this is a bottleneck block [57], an average pooling layer, and then two more bottleneck blocks. The two views are then combined by taking the element-wise maximum, with their contents at this point being in some representational space. This result is then sent through a series of average pooling layers and bottleneck blocks, as in figure 4.4.

After the final inverted residual block, there is a global average pooling layer. Following this are two dense layers of 1024 neurons with ReLU activation. These layers are included to essentially allow the network to do the math necessary to go from extracted features to coordinates, with the mathematics required apparently being roughly quadratic. During training, a dropout rate of 0.5

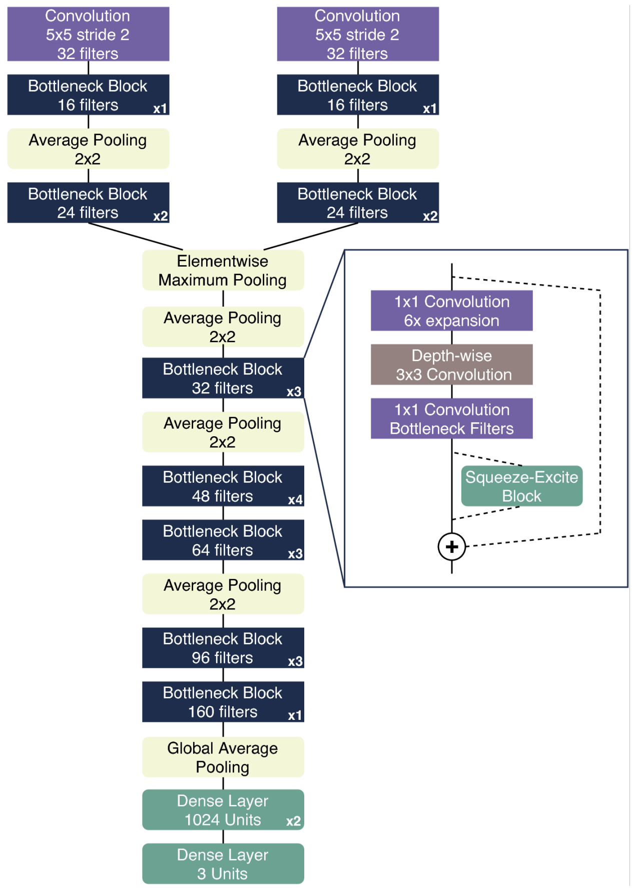

Figure 4.4: The full architecture of VertexCVN. The cutaway is the structure of each bottleneck block. Each view is input separately to the top of each upper branch. The penultimate two dense layers use ReLU activation. The three neurons in the final dense layer are the output vertex coordinates and have no activation or clamping.

is applied to these layers. Dropout in this case is less to avoid overfitting, as the training conditions made this a negligible concern, and more because dropout improves the performance of dense layers in almost all cases [58] [59]. These processing dense layers are followed by a final dense layer with 3 neurons, one for each coordinate, with no activation or clamping. This is the output of the network.

Over the course of development, multiple different output structures were tried, including different activation functions for the dense layers, different numbers of dense layers, and different numbers of neurons in the dense layers. The tested configurations include

- One layer of 512 neurons
- One layer of 1024 neurons
- Two layers of 1024 neurons (used in the final training)
- Three layers of 1024 neurons
- One layer of 2048 neurons
- One layer of 1024 neurons followed by a layer of 2048 neurons
- One layer of 2048 neurons followed by a layer of 1024 neurons

Performance was evaluated for a full training of each of these configurations, with the best-performing one being selected. Other options were briefly attempted but abandoned due to either memory use or processing time.

The final output of the network is in pixel map coordinates. If running at the ART level (the most detailed level of NOvA's software), these are then converted back to detector coordinates using the inverse of the initial linear transformation. Validation work in miniprod 6.1 indicated the need for additional corrections beyond that to fix biases and resolve some correlations in offset between coordinates, detailed in section 4.3.3.

4.3.2 Training

Training Data

Training was performed on two different datasets. Preliminary work was done on a dataset assembled from Near Detector production 5.1 (the latest at the time) simulation for secondary vertexing, but which was easily repurposed for primary vertexing. However, this set had several cuts applied to optimize for training a secondary vertexer which made it unsuitable for the final training of a general primary vertexer. Much of the experimentation with the architecture took place using this dataset.

The second dataset used was assembled from production 5.1 nominal simulation, both FHC and RHC with NOvA-standard quality and containment cuts (for more detail see section 6.2). Two sets were created, one for the Near Detector, and one for the Far Detector, to train two separate vertexers given the differing conditions. For each, further cuts were performed at runtime to ensure the true vertex was not inside of the muon catcher for the ND and that when converted to pixel map coordinates (via simple linear transformation) it was contained in the pixel map- at the edges for X and Y, and between pixels -10 and 90 in Z. The offset in Z was chosen because we expected that a CVN might be able to extrapolate upstream for some distance, while too far forward likely indicated an issue with the pixel map creation.

This dataset was then deterministically split using the standard 80/10/10 ratio into training, testing, and validation sets, where the testing set was used during training to ensure no over-fitting, and the validation set was left boxed until the training was finalized to check that hyperparameters had not been over-fit to the testing set. The order of events within the training and testing sets were randomized per training. The final size of the dataset was 18 million events for the Near Detector and 26 million events for the Far Detector.

Training Process

Due to the extremely large size of the datasets, training converged within a single epoch, and training over an entire epoch was infeasible for time and resource allocation reasons. This has the benefit of making overfitting a negligible concern. A batch size of 64 provided the necessary balance between speed and performance, determined by scanning in powers of 2 from 16 to 1024. Training was performed on the now-decommissioned Fermilab Wilson Cluster.

The loss function used is simply the Euclidean distance squared, in units of pixels. It is squared primarily to allow gradient descent to work properly, but this also avoids the expensive square root operation. Conveniently, this is equivalent to the mean squared error which is commonly used for classification tasks, and therefore a built-in optimized version was used. Pixel map coordinates were used both out of convenience, and because they more properly reflected the relative distance scales between coordinates, with the resolutions of the detectors being slightly finer in the X and Y directions compared to the Z direction.

The optimizer used was the Adam optimizer [60]. The initial learning rate was chosen by first scanning in half-decades from 10^{-1} to 10^{-4}, with $10^{-3.5}$ being the optimal value, and then taking slight variations around that point, all of which performed worse. Performance depended very strongly on learning rate, with next to no convergence if the learning rate was off of this ideal by an order of magnitude or more, but smaller variations around $10^{-3.5}$ showed minimal difference. Extensive studies to determine why the network was so sensitive to learning rate were not carried out, and are a potential starting place for further work on VertexCVN.

A pseudo-epoch size of 2500 batches was chosen to periodically evaluate the network against the testing dataset. This choice was largely arbitrary and was not studied rigorously, except to verify that much larger pseudo-epoch sizes slightly decreased performance. Decreasing the pseudo-epoch was not considered as it would negatively impact training time, which was at a premium.

Training was allowed to run until the optimizer determined convergence. During the final trainings, the network converged in around 90 pseudo-epochs, or around 80% of the Near Detector training dataset and 55% of the Far Detector dataset.

4.3.3 Results and Validation

Initial validation was performed first on the testing dataset to tune hyperparameters as detailed in the previous section, and then on the validation dataset, which was consistent with the testing set. With this confirmation that the model was not over-fit, it was frozen for implementation and validation.

From here, validation proceeded in several stages, corresponding to the wider Reconstruction and Deep Learning working group's plans for validating the various reconstruction changes for production 6. First, unique to vertexing, was a 10% production 5.1 "respin", nicknamed "phase zero"

	Vertexer	fraction r < 13 cm
ND	ElasticArms	0.642
	VertexCVN	0.775
FD	ElasticArms	0.732
	VertexCVN	0.883

Table 4.1: Comparing the vertexers on the figure of merit. Only basic quality and containment cuts have been performed. VertexCVN has been bias-corrected in this table. "r" is the Euclidean distance between the reconstructed and true vertex. This is evaluated over FHC only, but RHC has similar performance.

and referred to as such in some internal documents. This involved re-running the reconstruction chain on existing production 5.1 files with VertexCVN as the vertexer to ensure proper implementation within the ART framework. After clearing up some issues with ART integration, the results of this were largely consistent with the vertexer-only validation, and indicated a 21% relative improvement over Elastic Arms in both detectors as shown in table 4.1 and figure 4.5. Figure 4.6 shows an example of reconstruction improvement on the event level.

The next phase of validation, miniprod6.1, was much more involved. Initial MC-only verification matched previous results, though the higher statistics revealed bias correlations between coordinates, such as the bias in X being strongly dependent on true Z position of the vertex. Comparisons involving data were the first major discrepancy. While the behavior was as expected with no cuts applied, introducing containment cuts and Particle IDentification (PID) cuts caused major shape changes to the raw coordinate distributions.

Eventually, the difference was found to be due to the cross-section weights, the set of weights applied post-hoc to production 5.1 simulation to bring it into line with data. The way these weights were calculated depended on vertexing in an indirect and subtle way, via both vertex-dependent cuts, such as ensuring the vertex is inside the fiducial volume, as well as indirectly vertex-dependent cuts, which are any cuts that use information downstream of the vertexer, that is, almost every

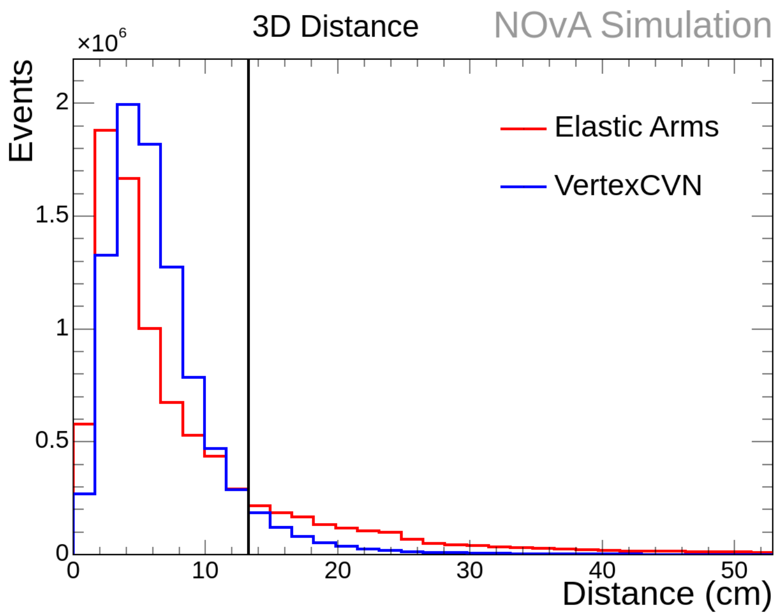

Figure 4.5: A histogram of the Euclidean distance between the reconstructed and true vertex for Elastic Arms and VertexCVN, after standard quality and containment cuts. Note that this is a version of VertexCVN that has not been bias-corrected, so the final performance is slightly better. The vertical black line is at 13cm, the figure of merit. While VertexCVN has slightly lower (about 2cm) precision than Elastic Arms, its accuracy is much greater. Elastic Arm's tail continues out well past the plot cutoff.

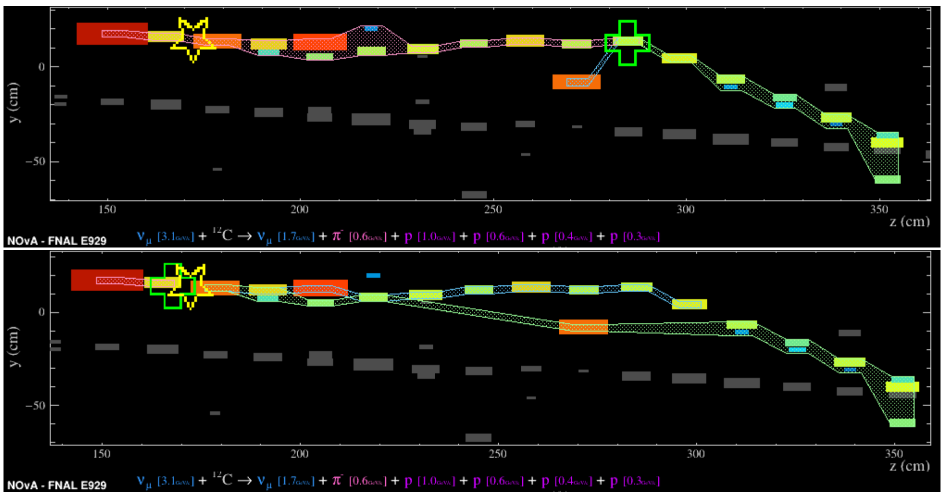

Figure 4.6: A comparison of two reconstructions of the same simulated event. Top: Elastic Arms. Bottom: VertexCVN. Only a single view is shown and the event is zoomed in for clarity. The yellow star is the true vertex, the green cross is the reconstructed vertex, and the colored dotted boxes are the reconstructed prongs. The size and redness of the hit corresponds to energy deposited. This tricky event includes a backward-going proton, a few lower energy protons contributing to near-vertex activity, and a charged pion that undergoes inelastic scattering around the position of the EA reconstructed vertex. VertexCVN gets the vertex location correct, but the pion prong is split in two (green and blue) due to FuzzyK's inability to handle inelastic scattering, demonstrating the need for a secondary vertexer.

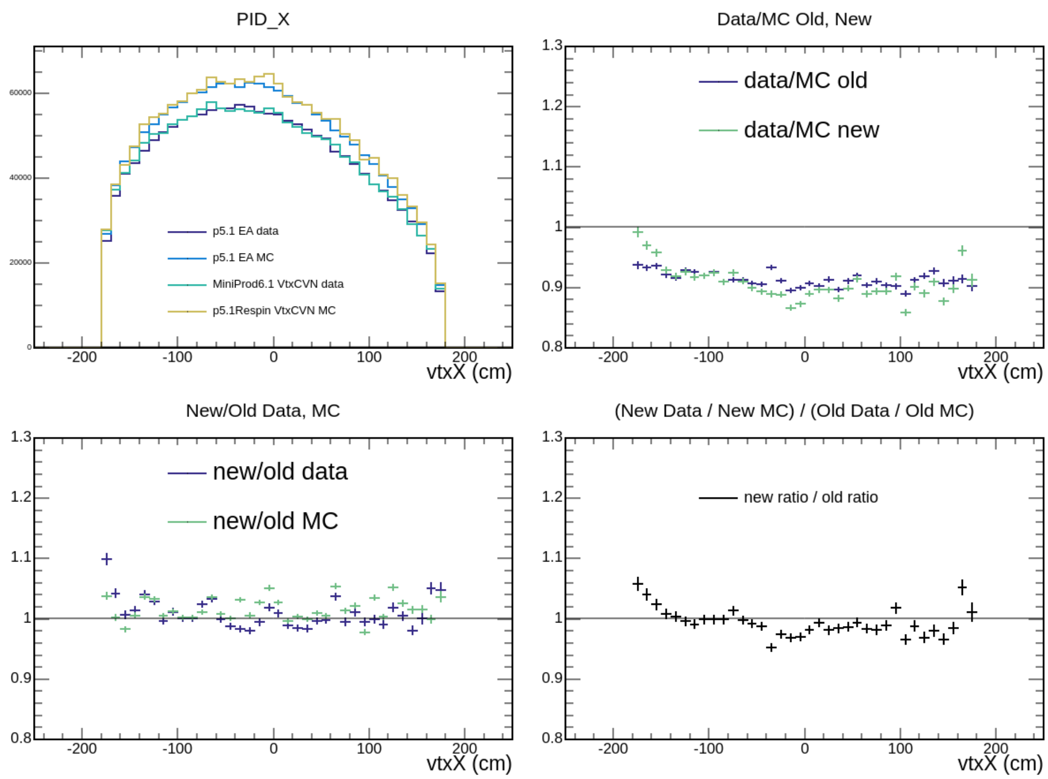

Figure 4.7: "Quad plots" showing the discrepancies between data, simulation, the old vertexer, and the new. Upper left: the raw distributions. p5.1 is production 5.1, the last full production before this validation, p5.1Respin is a partial reprocessing of p5.1 but using VertexCVN. EA is Elastic Arms, the old vertexer, and VtxCVN is VertexCVN, my vertexer. The approximately 90% normalization difference between data and simulation is irrelevant to the validation of VertexCVN, as it results from known effects. Upper right: the ratio of data to simulation for each vertexer. Lower left: the ratio of VertexCVN to EA for data and simulation. Lower right: the ratio of the green from the upper-right to the blue of the upper-right, to evaluate if there has been a notable change in the relationship between the data and simulation. In all ratios, the extreme bin-to-bin variation in certain parts of the detector is because of the binning not precisely following the cell boundaries combined with the plane- and cell- periodic behavior of the Elastic Arms vertexer. The concerning feature is the significant deviations from the mean at the edges of the detector, in both the simple ratios and the ratio of ratios.

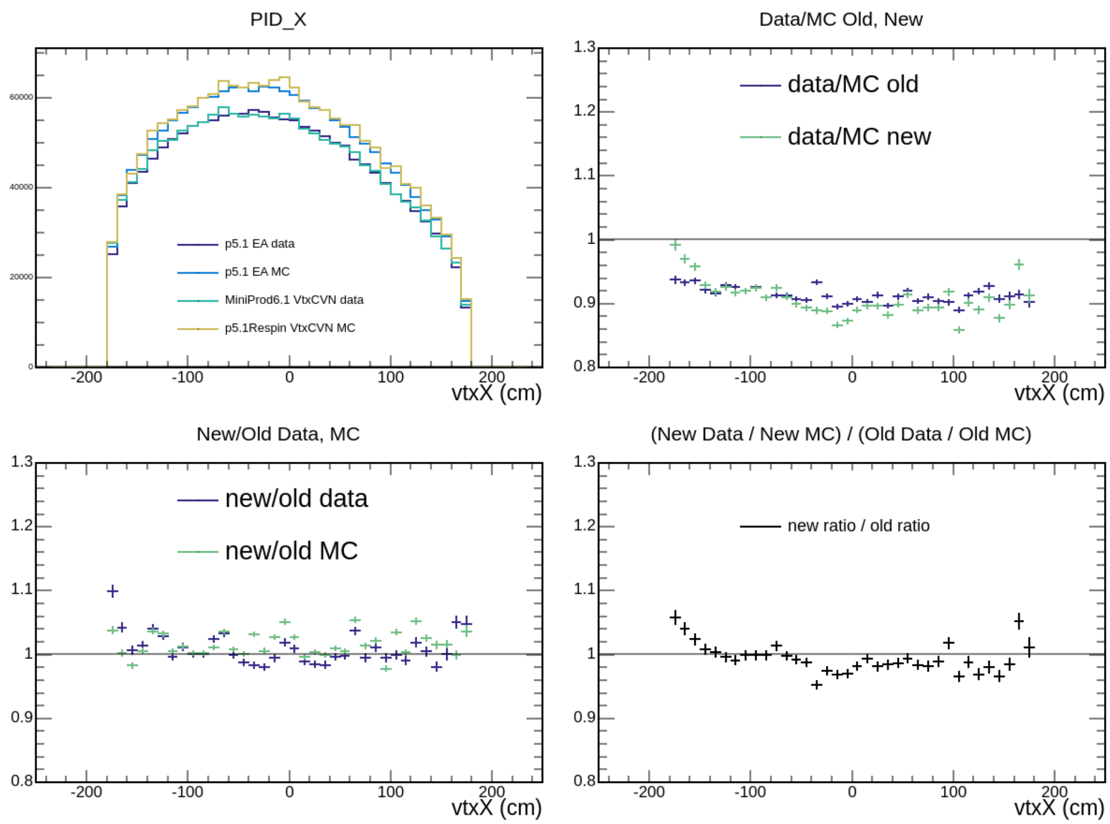

Figure 4.8: As in figure 4.7, but with no cross-section weights applied. Notice that the ratio of ratios is now nearly flat, up to the right edge of the detector (which is in the negative X direction).

cut (see figure 3.6). While it is surprising that this effect is large enough to cause this discrepancy, re-making the quad plots with only the flux weights and no cross-section weights demonstrates that this was the primary issue, as the ratio of ratios is now reasonably flat (see figure 4.8). Unfortunately, the individual ratios are not flat. However, this shows that this difference between data and simulation was already present in production 5.1, and thus the explanation is beyond the scope of the validation effort. Determining the origin of this discrepancy is an ideal spot for further work.

Additionally, there are some final remaining data vs MC discrepancies at the very edge of the detector. However, a hand-scan of hundreds of events at the edge of the detector revealed no visible issues, indicating that this is almost certainly not an issue with VertexCVN.

Finally, the coordinate conversions needed to be finalized. Various plots had indicated that there were also correlations between coordinates, which makes sense because the detector is not a perfect rectangular prism and is additionally tilted slightly off of the coordinate axes, none of which were accounted for in the coordinate conversion used for training. It was decided that only linear corrections would be performed, and that these would be performed separately for the near and far detectors. The process for the near detector is shown in figures 4.9 and 4.10, with the far detector being similar. The near detector fits use combined FHC and RHC data with standard selection cuts, while the far detector fits combine FHC, RHC, fluxswap (changing all ν_μ events to ν_e and vice-versa), and nonswap with standard selection cuts. Notably, linear corrections are insufficient to describe the X-Z correlations in the near detector, however, it was decided that attempting a piecewise linear fit or similar would potentially unfairly bias the data towards the simulation, and could cause vertexing issues if the true detector geometry was to shift even slightly off of the modeled detector geometry.

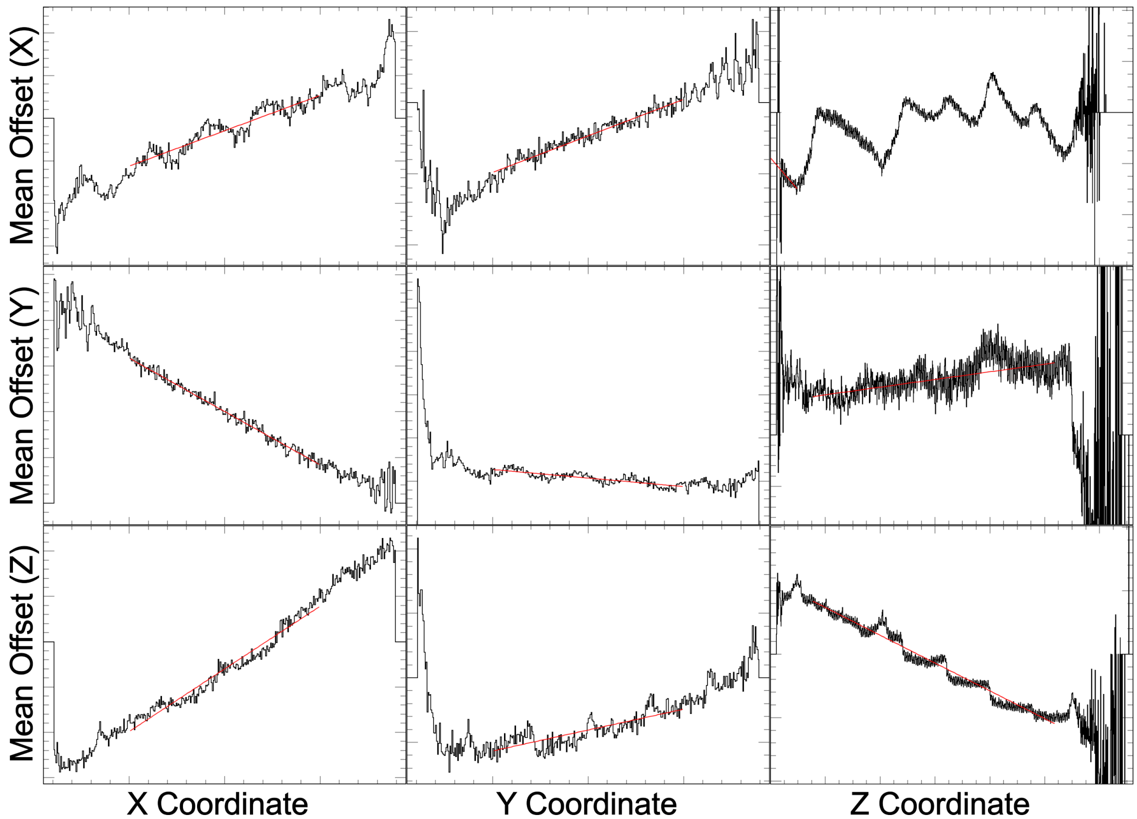

Figure 4.9: The method to determine the bias corrections for each coordinate. These plots are for the ND, FD correction was performed separately. Y-axes are the mean of the difference between the reconstructed vertex coordinate and the true vertex coordinate, X-axes are the reconstructed coordinate within the detector. The red lines are linear fits for each, constrained in the region of the detector that shows linear behavior. The x-offset in the z-coordinate is shown with only the first part of a piecewise-linear fit, which was later changed to a full fit like the rest. Bias correction is taken only from the fits where the offset coordinate matches the absolute coordinate, while slope correction is applied from all nine fits.

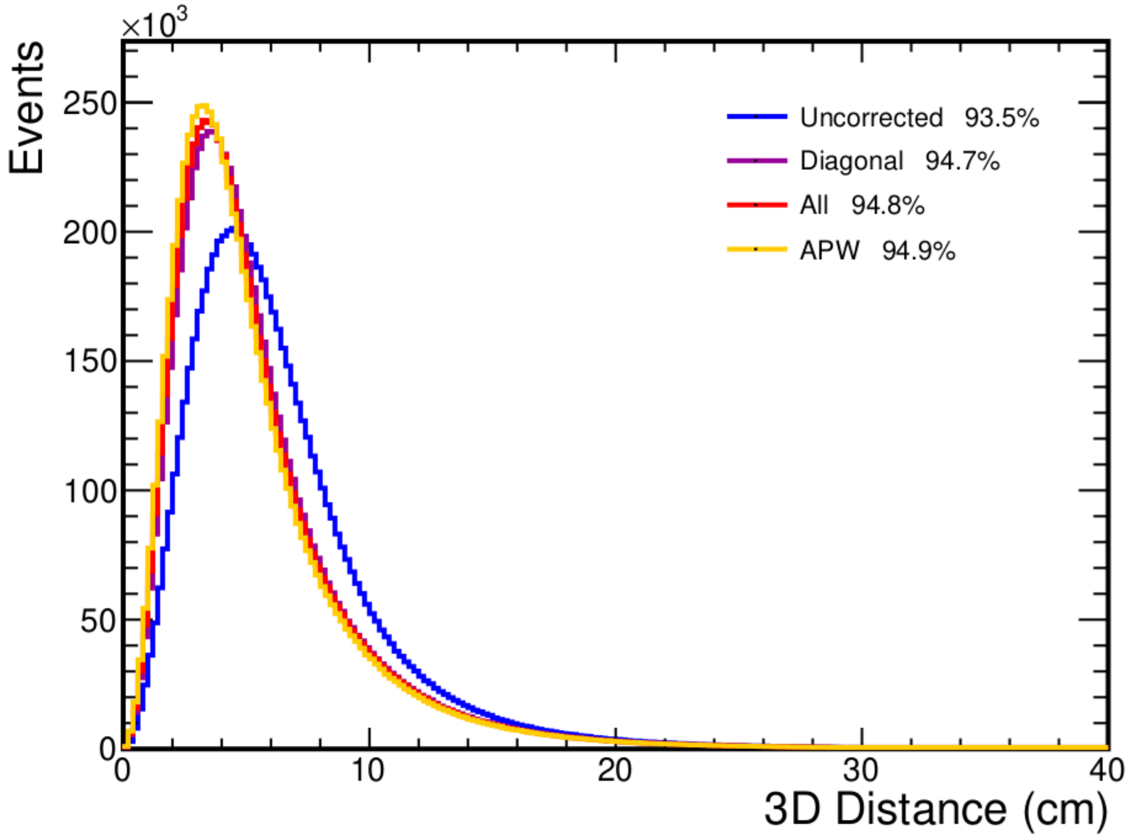

Figure 4.10: Plot showcasing the improvement in vertexer performance with different levels of bias correction in the ND. Blue is pre-corrected, purple applies corrections that only consider the coordinate being corrected, and red uses the information from all nine fits from figure 4.9. Yellow uses the linear piecewise fit for the X offset vs Z plot, which was not used in the final version due to biasing concerns. The percentages next to each label in the legend are the percent of events within 13cm of the true vertex, the figure of merit for the vertexer. Correction provides only a modest improvement in accuracy but a major improvement in precision, though still not to the level of Elastic Arms.

4.3.4 Conclusions

Overall, VertexCVN outperforms Elastic Arms, significantly reducing forms of misreconstruction resulting in split, combined, or otherwise incorrect prong reconstruction, at the cost of a small sacrifice in positional resolution while still being accurate within a cell. Due to the extreme computation efficiency of MobileNetV2 on CPUs, there is actually a speed-up in runtime compared to Elastic Arms, and peak memory usage of the reconstruction chain is minimally affected, as determined during the minprod6.1 production. VertexCVN has also been thoroughly validated. However, in-depth validation work has since revealed that VertexCVN is biased on some less common topologies, such as events containing a single neutral pion and no other particles. Targeting these less common topologies would be the primary goal of a future retraining effort. Impactful secondary goals would include encoding information on the position of the pixel map within the detector into the network which would eliminate the need for post-facto algebraic conversion to detector coordinates, and attempting to increase the resolution to be commensurate with Elastic Arms.

Chapter 5

Neutrino-nucleus scattering phenomenology

It is an unfortunate truth that due to the incredibly low cross-section that neutrinos have to scatter with anything, high target mass in general must be prioritized over target simplicity. While the interaction of a neutrino with a single proton is relatively easy to model, the interactions of neutrinos with even relatively small nuclei like carbon quickly become incredibly difficult to accurately represent. NOvA's successor experiment DUNE will use an even more complex target, argon, making the refinement of this modeling even more urgent.

As such, a great deal of work has been and must continue to be done to constrain models of neutrino-nucleus scattering and explore their rich phenomenology. Some of this work is theoretical, such as teasing out the effects of multi-nucleon correlations on interaction cross-sections and products. However, multiple theoretical models inevitably come into conflict, thus requiring large amounts of experimental data to constrain them— cross-section measurements, both integrated and differential, across as many variables and paradigms as possible.

In this chapter I will first give a brief overview of the phenomenology of the most relevant modes of neutrino-nucleus scattering for NOvA. I will then give an overview of the general process of conducting a neutrino cross-section measurement. Finally, I will summarize the experimental landscape of charged pion production cross-section measurements. It is largely within the context of these measurements that the results of this thesis will add new information and insights for future research.

5.1 Background

In order for neutrino experiments to achieve their physics goals, unavoidable experimental limitations must be addressed so that systematic errors can be substantially controlled. Neutrino experiments face unique difficulty in this, as the incoming neutrino energy cannot be known on an event-by-event basis, which compounds with uncertainties in the knowledge of the energy-dependent flux of the neutrino source. Furthermore, in the most common energy range for accelerator neutrino experiments, the momenta of the struck nucleon(s), multi-nucleon interactions, and final state interactions (FSIs) in which the products of the initial neutrino interaction further interact with the host nucleus are all significant and extremely difficult to model from first principles.

Due to these limitations, large datasets from multiple experimental platforms are needed to constrain models of neutrino-nuclear interactions in order to yield the precision required for the next generation of neutrino experiments, as demonstrated in documents such as the NuSTEC white paper [61] and the DUNE technical design report [62]. It is additionally noteworthy that all of the above factors both motivate a broad program of neutrino cross-section measurements as well as present challenges to the measurements themselves, including the one presented in this thesis.

In this section, I will present a brief review of the various scattering modes that are relevant for the energies of NOvA and DUNE. When discussing the final state products of these modes, FSIs will not be considered unless otherwise noted.

5.1.1 Charged Current Quasi-Elastic Scattering

The simplest interaction mode, Charged Current Quasi-Elastic (CCQE) events are in the regime where the scattering is well-modeled by a single interaction with an isolated nucleon moving at some momentum, with the only source of inelasticity being the particles changing type and therefore mass through W boson exchange. Specifically, the reaction for this mode is $\overset{(-)}{\nu}_\mu + n(p) \rightarrow \mu^\mp + p(n)$. These interactions are easy to identify due to their simple final state topology of a single charged lepton with a single nucleon (though this can be changed by FSIs), and dominate at neutrino energies in the hundreds of MeV, though they still contribute significantly at higher energies as well.

This type of interaction is of particular interest due to the ease of neutrino energy reconstruction.

As a two-body scattering problem where the momenta of the outgoing lepton and hadron can usually both be measured precisely, the primary loss of precision is due to the unknown initial momentum of the nucleon, allowing detailed nuclear models to be probed if the neutrino flux is well-understood, or detailed constraints to be placed on of the flux if the nuclear model is well-understood (for example, on a hydrogen target).

However, even in this simple mode complications arise. Observations of both neutral current and charged current quasi-elastic scattering at MiniBooNE [63][64] revealed an excess of events. This excess was resolved by refining the phenomenology of meson exchange currents (MEC) [65], which, to be extremely brief, involve the exchange of virtual mesons between nucleons which correlate those nucleons in such a way that significantly increases the neutrino-nucleus scattering cross-section, with the effect peaking around 1 GeV.

5.1.2 Resonant Production

Of particular interest to my analysis is the resonant production regime. Here, the interaction is modeled as the neutrino exciting the nucleon it hits to a baryonic resonance, most commonly the $\Delta(1232)$ state though higher resonances are relevant to DUNE and to some degree NOvA. The subset of these interactions most relevant to charged pion production in NOvA can be represented as

$$\nu_\mu + p \to \mu^- + \Delta^{++} \to \mu^- + p + \pi^+ \tag{5.1}$$

$$\nu_\mu + n \to \mu^- + \Delta^{+} \to \mu^- + n + \pi^+ \tag{5.2}$$

$$\overline{\nu}_\mu + p \to \mu^+ + \Delta^{0} \to \mu^+ + p + \pi^- \tag{5.3}$$

$$\overline{\nu}_\mu + n \to \mu^+ + \Delta^{-} \to \mu^+ + n + \pi^- \tag{5.4}$$

The multiple hadronic products increase the relevance of FSIs to this interaction type, and in particular the charged pion can frequently be fully re-captured by the nucleus, causing resonant production to be a major background to CCQE measurements. One of the major theoretical challenges of modeling resonant production is dinucleon or multinucleon correlations, which are believed to have a significant impact on this scattering mode. Unfortunately, they are also extremely difficult to model due to the requirement of relaxing the nucleon factorizability assumption that makes most computational nuclear physics tractable. The most studied form of these multinucleon correlated scattering modes is the "2p2h" (two particle two hole) mode, involving the ejection of two nucleons from the target nucleus. Even this is a relatively broad class which overlaps with the MEC processes discussed in the previous section.

5.1.3 Coherent Production

Another important process for charged pion production is coherent production, defined as processes where the target nucleus is left in its ground state rather than having a nucleon ejected like in QE or being in an excited state such as resonant production. The specific process relevant for this analysis is shown in equation 5.1.3:

$$\nu_\mu + A \rightarrow \mu^- + \pi^+ + A \tag{5.5}$$

Constraining this process is particularly important since it can mimic CCQE if the pion is confused for a proton, and vice versa.

5.1.4 Inelastic Scattering

Shallow and deep inelastic scattering dominate at high neutrino energies. Shallow inelastic scattering is poorly understood theoretically and is the transition region between resonant production and deep inelastic scattering (DIS). DIS is substantially better understood, and is the regime best modeled by a neutrino interacting with a parton rather than an entire nucleon. These events tend to have very complicated final topologies due to hadronization, and are therefore the least appropriate for precision measurement in the typical energy ranges of accelerator neutrino experiments, while still being extremely relevant for the energies of DUNE as DIS becomes the dominant mode in the few-GeV range. Past the energy range of accelerator neutrino experiments the hadronic showers become predictable enough for precision measurement. These high energy ranges are therefore the subject of the most study at, for instance, the various LHC experiments, resulting in a modeling gap in the few-GeV transition region.

5.2 Neutrino Cross-Section Measurements

Due to the importance of neutrino cross-section measurements as outlined above, several programs and even dedicated experiments have engaged vigorously performing these measurements. MicroBooNE, NOvA, T2K, and others all have robust cross-section measurement programs in addition to their primary oscillation physics goals. MINERvA, meanwhile, is unique as a fully-dedicated neutrino scattering experiment, designed to measure neutrino cross-sections on a variety of materials using the NuMI beamline.

Neutrino cross-section measurements are complex, requiring several intermediary steps to proceed from some initial spectrum measured by the detector through to a cross-section that can be used by phenomenologists to update event generators. To understand this process, we will go step by step using the cross-section formula as a guide. The formula for measuring a differential flux-integrated cross-section, using T_π as the differential variable, is

$$\left(\frac{\partial\sigma}{\partial T_\pi}\right)_\alpha = \frac{\sum_j U_{j\alpha} M_j P_j}{\epsilon_\alpha \Phi N (\Delta T_\pi)_\alpha} \tag{5.6}$$

where σ is the cross-section, T_π is the kinetic energy of the single charged pion, α indexes over true pion kinetic energy bins, j indexes over reconstructed pion kinetic energy bins, U is the unfolding matrix, M_j is the total number of events passing selection cuts in bin j, P_j is the purity in bin j, ϵ_α is the efficiency in bin α, Φ is the flux, N is the number of targets (typically in nucleons), and $(\Delta T_\pi)_\alpha$ is the bin width of bin α.

The determination of each of the quantities on the right hand side of this equation and their uncertainties is the task of the cross-section measurement, and gives a guide to the steps of the measurement. First is the determination of the purity. In a "cuts-based" analysis, this is accomplished using simulation alone. This has the benefit of being straightforward, but typically exposes the measurement to model bias and/or a great deal of uncertainty. More favored methods are data-driven purity estimates, typically through a template fit. This standard technique used in many fields across the sciences uses the separation between signal and background on some axis to estimate their relative proportions, and includes robust uncertainty analysis. Systematic errors are accounted for via an input covariance matrix which is used in the usual way to calculate the χ^2 metric that the fit optimizes. The output covariance matrix of the fit can also be determined via fluctuations around the fit minimum, which determines the fit errors and can be carried forward through the rest of the analysis.

5.2.1 Unfolding

While not all physics measurements perform unfolding, it is standard in the field of neutrino cross-section measurements. This is because the ultimate goal of these measurements is to be incorporated into event generators to further oscillation physics, and as such the measurements must be made compatible with each other. Therefore, transforming from the reconstructed space of the differential variable to the corresponding true space by correcting for detector response is necessary. The naïve approach would be to simply compute the response matrix (also called the migration matrix), which is essentially the two-dimensional histogram between the reconstructed space and the true space binnings subject to either row- or column-normalization (depending on the formulation), and then invert it. However, this tends to amplify noise and produce large bin-to-bin fluctuations in the unfolded result, and theoretically speaking the response matrix need not be invertible. As such, the unfolding must be regularized through one of several methods.

A complete overview of such methods is beyond the scope of this dissertation, but I will give a brief overview of the method I use in my own measurement, called variously D'Agostini unfolding [66], iterative unfolding, and (misleadingly) Bayesian unfolding. In this method, rather than inverting the response matrix, the true bins are treated as "causes" and the observed events in the reconstructed binning as "effects", and therefore the response matrix can be viewed as giving us information on $P(E|C_i)$, that is, the probability of an effect E (such as measuring an event in a given reconstructed bin) given the cause C_i. Then to back out the true distribution from the reconstructed we need the $P(C_i|E)$, and rather than obtaining these via matrix inversion we instead use Bayes formula:

$$P(C_i|E) = \frac{P(E|C_i) \cdot P(C_i)}{\sum_l P(E|C_l) \cdot P(C_l)} \tag{5.7}$$

where $P(C_i)$ is the initial probability of cause C_i. These are initially set to the true distribution, but are updated iteratively using the actual observed number of events in each bin, roughly corresponding to updating the prior for the expected number of events in that bin.

While Bayes' theorem is used to update the distribution, the method does not use strictly Bayesian statistics, hence why "Bayesian unfolding" is potentially misleading terminology. The number of iterations used is the regularization parameter of this method, determined by minimizing the χ^2 between unfolded fluctuated distributions and the true distribution from simulation, or alternatively by looking by-eye at when the distribution stabilizes but the errors begin to increase.

5.2.2 Efficiency, Flux, and Targets

Efficiency correction is the process of taking an observed number of events and extrapolating to the theoretical true number of events a perfect detector with the corresponding fiducial volume would observe. Unfortunately, there is usually no data-driven way to do this, and so the modeling of the detector must be relied on, typically using a simulation-based approach.

Target counting methods are unique to the experiment, but also rely on effective detector modeling. So long as the detector is made with reasonably pure materials (or the impurities have been measured) and the fiducial volume is relatively uniform in composition, target counting is typically a negligible source of uncertainty.

Flux estimation is always extremely difficult, though again the exact issues and the methodologies used to address them are experiment-specific. For a discussion of the NuMI beamline flux estimation used later in this thesis, see section 3.1.

5.3 Charged Pion Production Measurements

Cross-sections that include charged pions in their final states are particularly difficult measurements to perform across experiments, due to the large cross section of pions on both detector media and within the target nucleus as part of FSIs. There are also significant contributions from 2p2h processes, which are an area of active ongoing research both theoretical and experimental. Finally, since resonant production is a major mode of pion production, modification of the properties of $\Delta(1232)$ when bound in a nucleus are another major effect that must be considered [67] [68].

Charged pion production cross-sections are a relatively small field, with contributions mainly from the Kamiokande experiments, MiniBooNE and its successor MicroBooNE, ArgoNeuT, and MINERvA. Of these, MINERVA is by far the best positioned for high-quality results as a dedicated neutrino cross-section experiment with a particular focus on how cross-sections scale with target A. This is to allow extrapolation from current low-A target experiments such as NOvA and T2K to future high-A target experiments, most significantly DUNE.

A summary of charged pion production cross-section measurements is presented in table 5.1. While MINERvA and T2K have both covered part of the 150-330 MeV energy range, both of their

Experiment	Target	Sgn	Measurements	T_π	Ref.
MiniBooNE	CH_2	π^+	$\frac{d\sigma}{dX} : X \in \{T_\mu, T_\pi, Q^2\}$ $\frac{d\sigma}{dX}(E_\nu) : X \in \{T_\mu, T_\pi, Q^2\}$ $\frac{d^2\sigma}{dXdY} : (X,Y) \in \{(T_\mu, T_\pi), (T_\pi, \theta_\pi)\}$ $\sigma(E_\nu)$	0-400 MeV	[69]
MINERvA	CH	$\pi^\pm$	$\frac{d\sigma}{dX} : X \in \{\theta_\pi, T_\pi\}$	35-350 MeV	[70]
T2K	H_2O	π^+	$\frac{d\sigma}{dX} : X \in \{p_\pi, p_\mu, \theta_\mu, \theta_\pi, \theta_{\mu\pi}\}$ $\sigma(E_\nu)$	0.10-1.9 GeV	[71]
	CH		$\frac{d\sigma}{dX} : X \in \{p_\pi, p_\mu, \theta_\mu, \theta_\pi, \theta_{\mu\pi}\}$ $\frac{d^2\sigma}{dp_\mu \cos(\theta_\mu)}$	0-3 GeV	[72]
MicroBooNE	Ar	$\pi^\pm$	σ $\frac{d\sigma}{dX} : X \in \{\theta_\mu, p_\mu, \theta_\pi, p_\pi, \theta_{\mu\pi}\}$	30-50 MeV	[73]
MINERvA	CH C H_2O Fe Pb	π^+	$\frac{d\sigma}{dX} : X \in \{p_\mu, p_{\mu,T}, p_{\mu,\parallel},$ $p_\pi, p_{\pi,T}, p_{\pi,\parallel},$ $\theta_\mu, T_\pi, \theta_\pi, Q^2, W_{exp}\}$	50-350 MeV	[74]
K2K	CH	π^+	$\frac{\sigma_{CC1\pi^+}}{\sigma_{CCQE}}(E_\nu)$	—	[75]
MiniBooNE	CH	π^+	$\frac{\sigma_{CC1\pi^+}}{\sigma_{CCQE}}(E_\nu)$	—	[76]
MINERvA	CH	$\pi^\pm$	$\sigma(E_\nu)$ $\frac{d\sigma}{dX} : X \in \{p_\mu, \theta_\mu, Q^2\}$	—	[77]
ArgoNeuT	Ar	$\pi^\pm$	σ $\frac{d\sigma}{dX} : X \in \{p_\mu, \theta_\mu, \theta_\pi, \theta_{\mu\pi}\}$	—	[78]
T2K	H_2O	$\pi^\pm$	σ	—	[73]
NOvA	CH	$\pi^\pm$	$\frac{d\sigma}{dT_\pi}$	150+ MeV	—

Table 5.1: All relevant $\nu_\mu CC1\pi^{\pm(+)}$ measurements. The first section has the papers that include differential measurement in T_π (or a related variable), the second section are all other measurements. The bottom row is the analysis presented in section 6. With thanks to Alesha Devitt's doctoral thesis [79] for the formatting and citations from 2022 and earlier.

results either use coarse binning or have large uncertainties in this range. This is the gap that NOvA is uniquely well-placed to fill, and which my analysis seeks to address.

Chapter 6

Analysis

As discussed previously in section 5.1, neutrino cross-section measurements are a crucial input to Monte Carlo generators, which are themselves necessary to perform neutrino oscillation measurements. Semi-exclusive measurements that include charged pions in their allowed final states are considered particularly difficult to measure due to charged pions tending to re-scatter in the detector medium and generally being easily confused with protons. However, these measurements, especially at the energies of NOvA, are particularly crucial for the energy range that DUNE will operate in.

NOvA has had a variety of charged pion analyses in development for quite some time, producing several theses [80] [81] [82] [83]. A notable gap in the coverage of these measurements is a differential measurement in the energy of the leading charged pion. This is a particularly challenging measurement for NOvA, as its standard reconstruction tools for the pion energy have a high failure rate. This analysis seeks to plug that gap.

6.1 General Approach

This analysis seeks to measure a single-differential cross-section measurement in pion kinetic energy for the signal $\nu_\mu + N \to \mu^- + 1\pi^\pm + X$, where X has no charged pions.

Measuring a cross-section for final states containing a charged pion, particularly in pion energy, is a hard problem for NOvA. NOvA's reconstruction does not have a mechanism for handling primary particles inelastically re-scatting in the detector medium, which causes significant underestimation of the energy of higher-energy pion prongs. Prior to production 6, NOvA also did not account for overlapping prongs, causing issues for several common topologies of pion production. Finally, due, to the relatively low spatial resolution of the detectors, hadronic energy estimation is very difficult for short tracks, as is particle identification.

To compensate for these factors, we separate the signal into two different categories: "golden" pions that do not significantly scatter inside the detector, and "non-golden" pions that do. Additionally, there are misreconstructed pions, where the identified single pion prong in the signal event is actually a non-pion, typically a proton, and therefore does not give useful information about the pion energy.

Accounting for all of these effects would be complex in a cuts-based analysis, or require extremely strict cuts that would result in a very low efficiency. Therefore, we use a template fitting methodology with four templates: golden signal, non-golden signal, misreconstructed, and background.

Separating these four categories is done using a BDT, to be described in section 6.5. The GoldenBDT result acts as the fitting variable. Binning must be optimized both in T_π and in the BDT variable to provide the best fit results.

Finally, the purity-corrected spectra for the golden and non-golden signal categories are separately unfolded and efficiency-corrected, while the misreconstructed and background categories are discarded since they are only useful for determining the purity (in particular, the misreconstructed category cannot be usefully unfolded). A cross-section is then calculated for both categories, and then these results are combined using the covariance information between the two measurements in each analysis bin for a final reportable cross-section.

6.2 Selection

Selection is minimal and limited largely to quality and containment cuts, with the philosophy that the BDT should be able to sort background from signal well enough. Investigating tighter initial selection cuts to reduce background and hopefully help the BDT perform better is a good subject of additional work for this analysis.

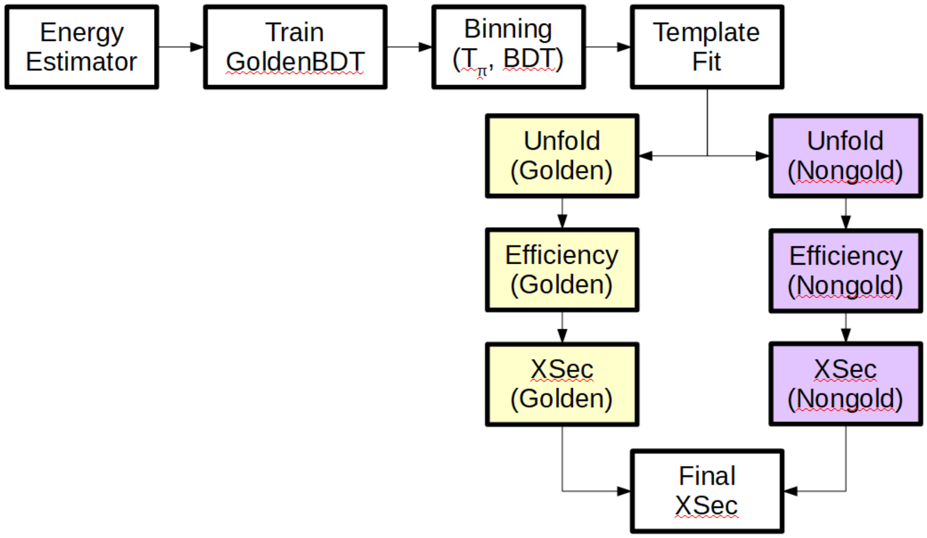

Figure 6.1: Outline of the analysis procedure. The analysis bifurcates after the template fit, with separate purity corrections for the golden and nongold categories. Once separate cross-section measurements are reached for both categories, these measurements are combined, taking into account correlations.

The first cuts applied are the standard NOvA quality and containment cuts. These ensure that there are at least two reconstructed prongs in the event, that the event contains at least 20 hits and at least four contiguous planes have hits, that the vertex and all Kalman tracks start in the fiducial volume, and that no prongs contain hits outside the fiducial volume.

The next cuts are copied from NOvA's ν_μ charged-current inclusive cross-section measurement, to ensure future compatibility if desired. These include a cut on the MuonID of the leading muon prong, and cuts on the kinetic energy and angle of the muon, as detailed in that paper [84].

Finally, there are two more cuts for charged pion reconstruction. First, I require that the prong with the highest pionID that is not the muon prong (called the "leading pion prong" from here on) has at least three hits per view in both views of the detector. This is also implicitly an angle-dependent length cut. Additionally, I cut on the reconstructed angle of the leading pion prong to make sure it is not within $\pm 20°$ of perpendicular to the detector Z-axis. This is because insufficiently horizontal prongs are not reliably reconstructed in the variables that are important to the analysis, including and especially prong length.

6.3 Energy Estimator

A range-based energy estimator is used. This will work best on the non-scattering sample. The scattering sample will be biased downwards, with the amount of bias increasing with pion kinetic energy in a way that is predictable in aggregate, though not for each individual event. This bias will be accounted for in the response matrix for that sample, at the cost of an increased dependence on the GEANT systematic. The energy estimator is determined using a particularly clean golden sample with extra cuts on prong purity and efficiency. This sample is binned in track length (as measured by Break Point Fitter [41]) vs true pion kinetic energy (see figure 6.2). The lower threshold of pion kinetic energy is determined from where the event rate drastically decreases due to the cut on number of hits in each plane. This 2D histogram is then sliced along the length variable, and each slice has a gaussian fitted to it in the range of the full width at half maximum, creating a dataset of central value points with associated errors. A chi-squared linear fit is then performed through these points, which is used as the energy estimator. For the range of kinetic energies considered (130 MeV threshold, 310 MeV as the lower edge of the unbounded bin), this linear relationship is a very good description. However, it is non-linear at both lower and higher energies.

6.4 Binning in T_π

For a physically meaningful cross-section, we need our kinematic bins to be wider than the resolution. To determine this, I started with a 2D histogram of true T_π vs Reco T_π − True T_π, filled with both golden and non-golden events. I proceeded from the low T_π threshold of 130 MeV to fit a gaussian to the full width at half max to determine the resolution. If the resolution was greater

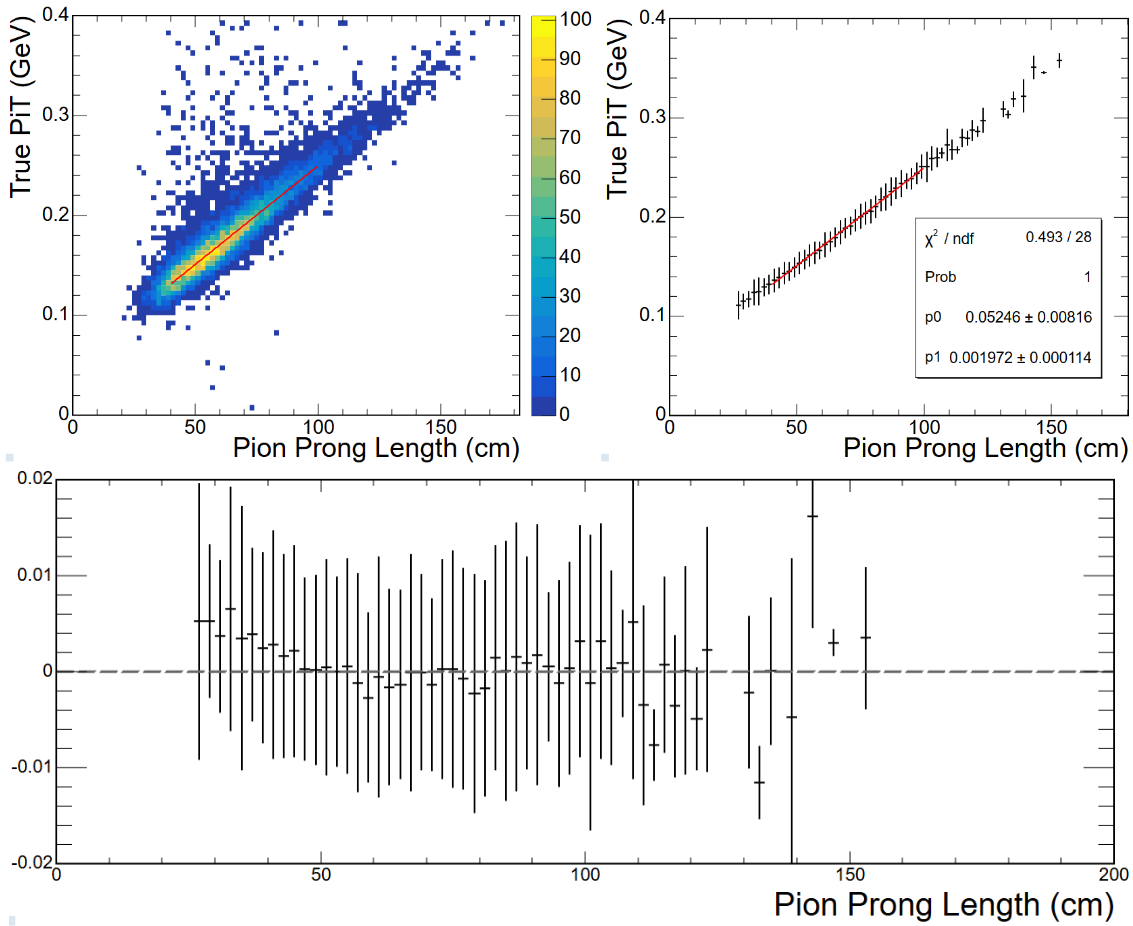

Figure 6.2: Left: Plot of the pared-down golden pion sample. Z-axis is number of events. The linear relationship in this range is clear. Right: the data from the previous plot has been sliced and had a gaussian fit to each slice to give the data points. The ROOT fitting utilities are then used to perform a linear fit to these data. Bottom: residual plot of the fit. There is a slight upward bias at the bottom end of the range, but still in agreement with zero. This bias was left instead of trying other functional forms to avoid overfitting to the simulation.

than the bin width times an additional (semi-arbitrary) factor of 1.5, I combined that bin with the bin immediately to its right and repeated until the condition was met. I then start over with the next bin to the right. This process is guaranteed to give an "optimal" binning, in the sense that it will yield the maximum number of bins fulfilling the condition, with a final unbounded bin. Because the energy estimator is only fully valid to 310 MeV, the penultimate bin's upper boundary was moved to that value, which still results in a bin wider than the resolution in that bin, just by a factor less than 1.5. I then checked the resulting resolution, bias, and unmigrated fractions (figure 6.3). The unmigrated fraction was relatively low compared to normal standards, however, this is expected due to the presence of non-golden and misreconstructed events.

The unbounded bin does also technically meet the resolution requirement, though using that terminology is somewhat misleading. Rather, reporting the cross-section in that bin is valid because while the energy reconstruction is known to be inaccurate within that bin, the reconstructed energy is strictly lower than the true energy, and therefore any event reconstructed into the final bin definitely belongs there. It is also worth noting that for technical reasons a finite upper bound must be chosen for the "unbounded" bin. In this case, that is 120 GeV, so chosen because it is the energy of the main injector protons used to create the neutrinos for NuMI, and therefore no downstream product could have a higher energy, barring some truly spectacularly high-momentum nucleon and very "lucky" momentum distributions for both the neutrino production and the following pion production, a statistically insignificant situation.

6.5 GoldenBDT

The template fit must be performed in some variable that demonstrates good separation between the various categories under consideration. A comprehensive review of available variables revealed that none of them provided the separation necessary. As such, I turned to the standard tool in particle physics for this task: the Boosted Decision Tree (BDT). For a brief overview of BDTs refer

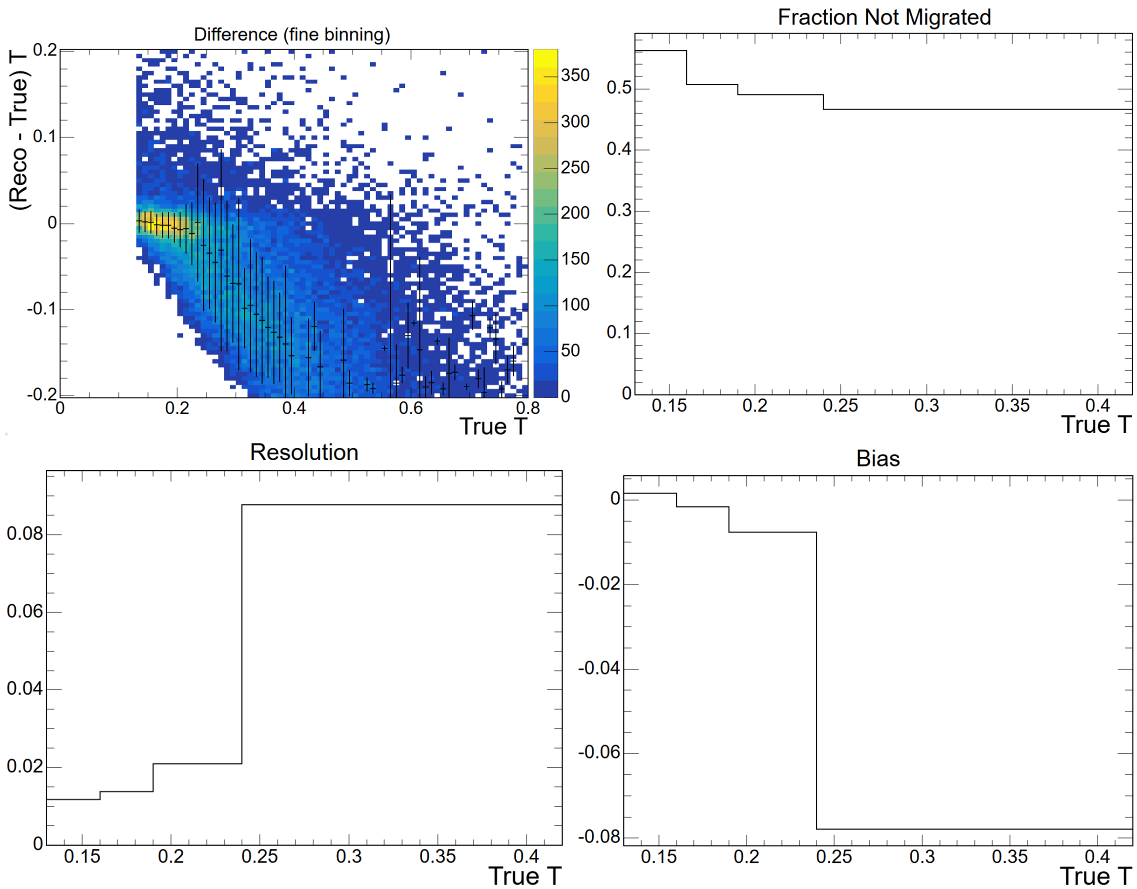

Figure 6.3: All axes are in units of GeV. Upper-Left: the initial histogram from which binning is calculated. The marked points with errors are from the gaussian full width at half max (FWHM) fit. Upper-Right: After binning, the unmigrated fraction in each bin. Normally a value of 0.6 is the target, but due to the inclusion of the nongolden events, this is good enough. Bottom-Left: Resolution in each bin after binning, from a gaussian FWHM fit over the (reco - true) axis within the bin. Lower-Right: Bias in each bin, from a gaussian FWHM fit over the (reco - true) axis within the bin.

to section 4.1.1.

GoldenBDT was trained on the four classes (golden, non-golden, misreconstructed, and background) using TMVA [26], on 10% of production 5.1 (see section 3.7) ND FHC nominal MC, with 75% of that being the training set and the remaining 25% as the testing. The inputs are in table 6.5 (the particle ID scores are derived from ML methods trained only on single prongs to avoid bias), and the distributions of the most important variables can be seen in figure 6.4. These input variables were selected by putting training on all potentially relevant variables and then eliminating low-contribution or half of a pair of highly correlated variables until all remaining variables were only moderately correlated and all had relatively high importance. Multiclass BDTs in TMVA work by training a different BDT for each class, with that class as the "signal" for that BDT. Of the four resulting BDTs, the background-class BDT was the one with good shape distinction between all four classes, and would become the BDT I would use. The final configuration used 1200 trees of depth 3, which maximized performance while minimizing overfitting, which was a serious issue. A depth of 3 is deeper than the standard 1 or 2 used in most BDTs, however, the drop in performance between from depth 3 to depth 2 makes it clear that three-variable correlations are important to separating the categories. I have not investigated exactly which correlations are at play, as determining the details of how any ML system functions is time-consuming and requires thorough work to prevent jumping to conclusions.

Name	Description	Importance (pre-fit)
PionID	pionID score for selected pion prong	0.1985
ProtonID	protonID score for selected pion prong	0.1825
MuonID	muonID score for selected pion prong	0.1361
Width	Shower width of selected pion prong	0.1327
Second PionID	pionID score of second-best prong	0.1231
PhotonID	photonID score for selected pion prong	0.1162
Plane Gap	Max plane gap of selected pion prong	0.1109

The outputs can be seen in figure 6.5. The background and misreconstructed categories numer-

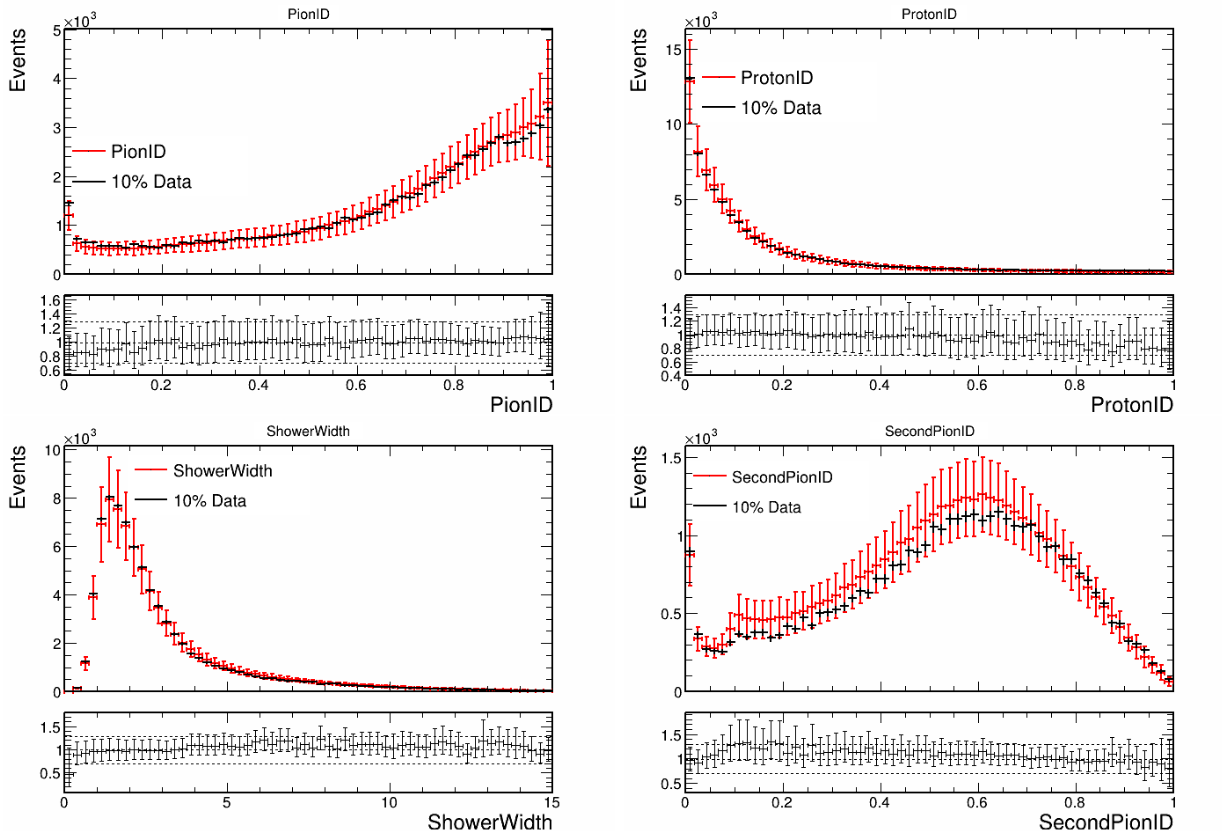

Figure 6.4: Four of the five most important BDT inputs. Muon ID was skipped because it is not graphically illustrative. Data and simulation agree within error for all variables.

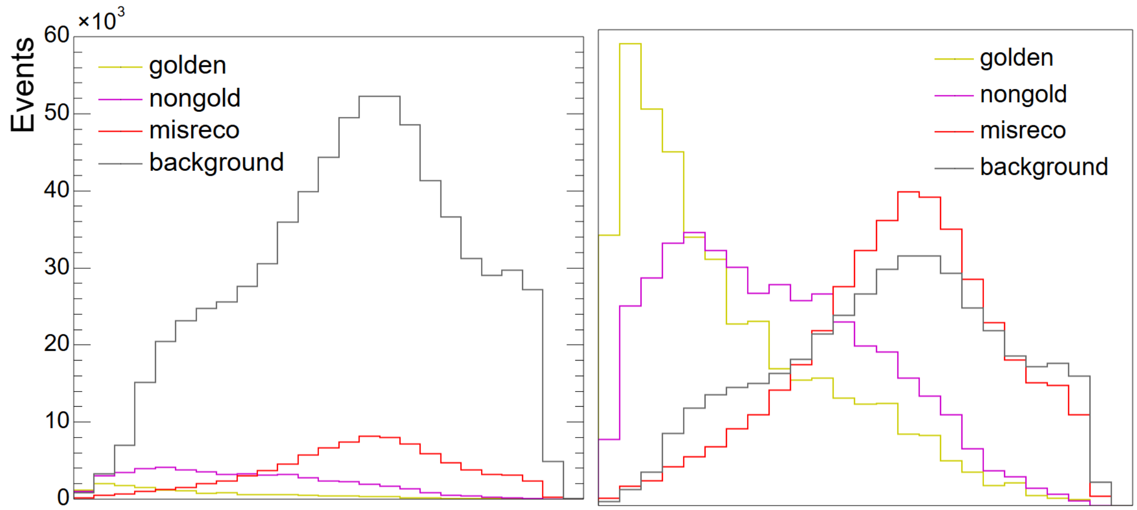

Figure 6.5: Left: final templates in even binning. Right: the same, but area-normalized to show shape

ically dominate, but with careful binning to be described in the next section the shape differences can still be exploited for productive measurement.

6.6 Binning in GoldenBDT

To optimize binning in the fit variable a metric for optimization must be selected. The choice was somewhat arbitrary, but I decided to use the sum of the uncertainties on the fit parameters in the golden and non-golden templates, for a Poisson-fluctuated fake data universe (or the mean error on the fit parameter of an ensemble of universes) with statistical errors only (since systematics slow the fit process untenably). To optimize, I made templates in 120 even bins, with the intent to have a final binning that unevenly combines those bins. This combination can be represented as a list of dividers $\{x_0, x_1, ..., x_n\}$ for some fixed n, such that the list is strictly increasing. Each x_i represents the index of the bin whose low edge becomes the low edge of the new bin. Experimenting with naïve binning showed that 12-14 bins was reasonable. I decided to use 12 bins to guard against there being fewer signal events in the data than expected. To optimize, I started with a fixed Poisson-fluctuated universe, and tried about 100 different initial lists of bin dividers. I wrote a code to pass back and forth across this list, adjusting each entry in turn (except the ends) by $\{-5, -2, -1, 0, +1, +2, +5\}$, keeping whichever had the lowest in the metric, until the list had not changed for an entire pass.

This method is extremely sensitive to getting stuck in local minima. This turns out to be a good thing— the global minimum for most if not all Poisson-fluctuated fake data universes is severely overfit to that universe. To account for this, starting with the lowest metric for the test universe and working towards the highest, I tested each candidate binning against 200 Poisson-fluctuated universes and found the metric computed on the mean error. For the earliest entries on this list, that metric was higher than the average of the entries on the initial list. It then rapidly dropped to below the average before starting to rise again as I went down the list. The binnings near the minimum were fairly similar to each other, so I simply took the lowest one (in metric on the ensemble) that I found. I verified it was reasonable both by-eye and by testing it on an even larger ensemble of universes. I then produced templates for this binning that included full systematics, and repeated the analysis on an ensemble of 200 Poisson-fluctuated fake data universes. The results were stable, and so I decided to move on to more sophisticated fake data studies using this binning.

6.7 Purity correction

Purity correction is done via global template fit, using the process designed by Cathal Sweeney and described in his thesis [80]. The gist is that each template is fit in each analysis bin simultaneously, seeking the χ^2 minimum between the summed scaled templates and the data via some covariance matrix constructed by systematic inputs, shown in figure 6.6 with the component-wise breakdown in appendix A.1. The MINUIT package (included in ROOT) is used to optimize the fit. The main modification here is that I have four templates, and thus four purity corrections. However, I discard both the background and the misreconstructed sample, in the latter case because it cannot be unfolded due to the reconstructed and true T_π values being largely uncorrelated. Therefore, I carry forward two purity-corrected spectra in reconstructed pion kinetic energy, one for the golden sample and the other for the nongolden (per figure 6.1).

The errors on the fit are determined by fluctuating the input spectrum around its central value and creating a covariance matrix out of the results. This output covariance matrix for the nominal MC is shown in figure 6.7.

6.8 Template Fitter Fake Data Studies

I conducted a variety of fake data studies to validate the template fitter's validity across the potential range of systematics and physical results that it might encounter in the data. The first of these was on a GENIE multiverse. I generated 300 GENIE universes independent of the ones that were used to construct the covariance matrix, and ran the template fitter on all of them. The results were reasonable, with only a couple failures. The main criterion I used for evaluation was agreement between the fit-scaled contents of the analysis bin and the true contents of that bin in the fake data.

Finally, I created a custom reweight function that directly impacted the quantity to be measured, the pion kinetic energy. This was done by reweighting signal events according to the true kinetic energy of the pion, starting with a weight of 1 at the threshold value of 130 MeV and scaling to a maximum of 1.2 at the upper bound of the final bounded bin, 310 MeV. The template fitter also performed successfully here, as seen in figure 6.9

6.9 Generator Studies

In order to assess the distinguishing power of the analysis and get a sense of the range of what to expect from the data, it is a good idea to see what various event generators, essentially substitutes for theoretical models, predict for the cross-section. Four generators were used for this: the most up-to-date version of GENIE (a more recent version than was used for the production that this

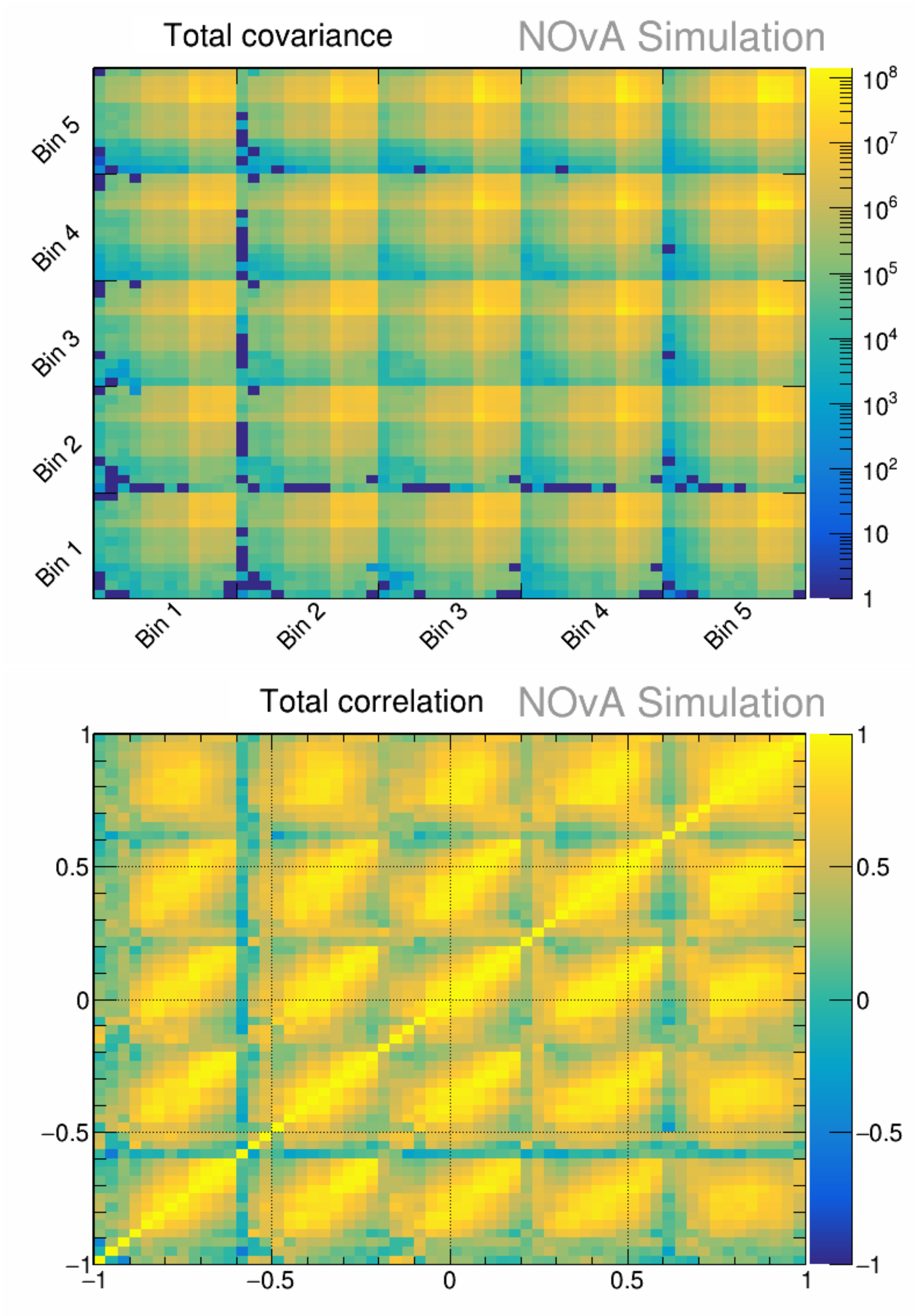

Figure 6.6: Total input covariance matrix and its corresponding correlation matrix. The five large divisions are the bins in T_{π}, while the subdivisions are in the BDT binning. The various energy bins are highly correlated, mostly from GENIE systematics.

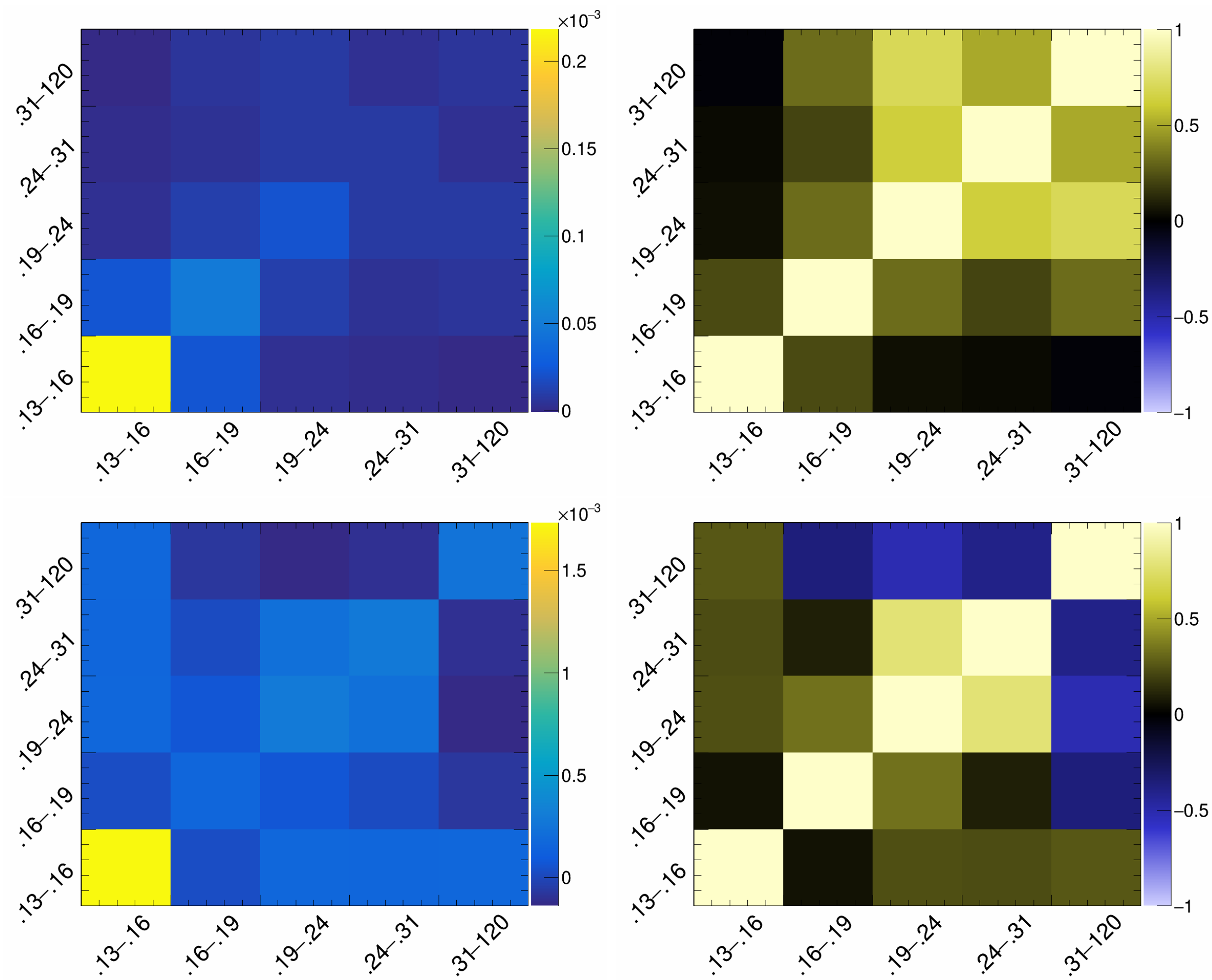

Figure 6.7: The output purity covariance matrices of the global template fit for the nominal sample. The global template fit has managed to remove the correlations between most bins. Top: golden. Bottom: nongold. Left: covariance. Right: correlation.

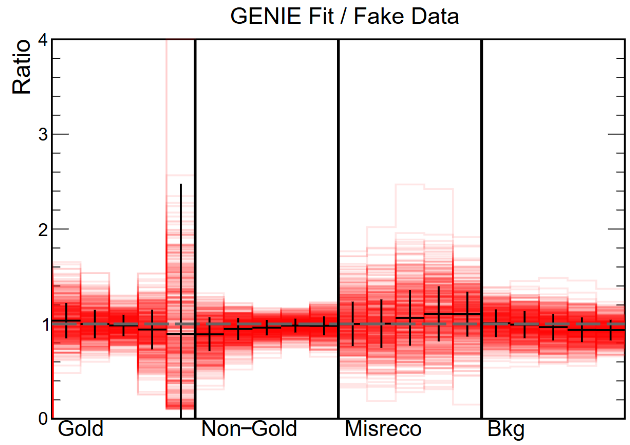

Figure 6.8: Per-universe ratios of the fit-scaled templates to that universe's fake data. The black is the median value, and its error bars are the standard deviation of the values within that bin determined by fitting a gaussian.

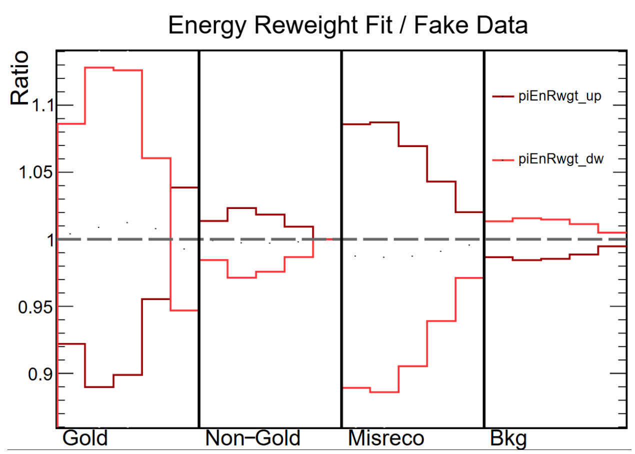

Figure 6.9: The results of the energy reweight fitting. Agreement is within error.

analysis uses), NuWro, NEUT, and GiBUU.

GENIE's basic functionality and configuration for production 5.1 nominal was reviewed in section 3.6. The latest GENIE version used here updates the nuclear, quasi-elastic, and 2p2h models to SuSA [85], uses the effective internuclear transport model from INTRANUKE, and is untuned.

NuWro [86] [87] is a neutrino event generator developed by the University of Wroclaw in 2004 and continually updated since. NuWro uses the Local Fermi Gas model for calculating nucleon states, the Llewellyn-Smith model [88] with dipole form factors for quasi-elastic scattering, the NSV 2p2h model [89], the PYTHIA [90] hadronization model, and an in-house developed resonant production and FSI models.

NEUT [91] is the event generator developed, used, and maintained by the Kamiokande experiments, starting with the original Kamiokande atmospheric neutrino experiment in the 1980s. Like production 5.1 GENIE it uses the local Fermi gas model for the nucleons and Berger-Sehgal for resonant events, but it uses the Valencia [92] model for quasi-elastic and MEC events and a custom semi-classical intranuclear cascade model for final state interactions.

GiBUU (Giessen Boltzmann-Uehling-Uhlenbeck transport model) [93] takes the approach of propagating phase space distributions rather than particles, aiming for greater physical accuracy than cascade models (which all the other generators listed here are). In particular, it accounts for nucleons being bound and uses the same potentials for both the initial interaction as well as FSIs. However, it also has the major weakness of not describing coherent production. As such, its only similarity with the other generators is using PYTHIA for its hadronization model.

Generator studies were conducted using CONVENIENT, a currently internal to NOvA framework that puts the four major generators in the same file format for direct comparison. Implementation was straightforward, as most of the heavy lifting is done by the framework. Notably, in order to get agreement with the version of GENIE used for production 5.1, the spectra had to

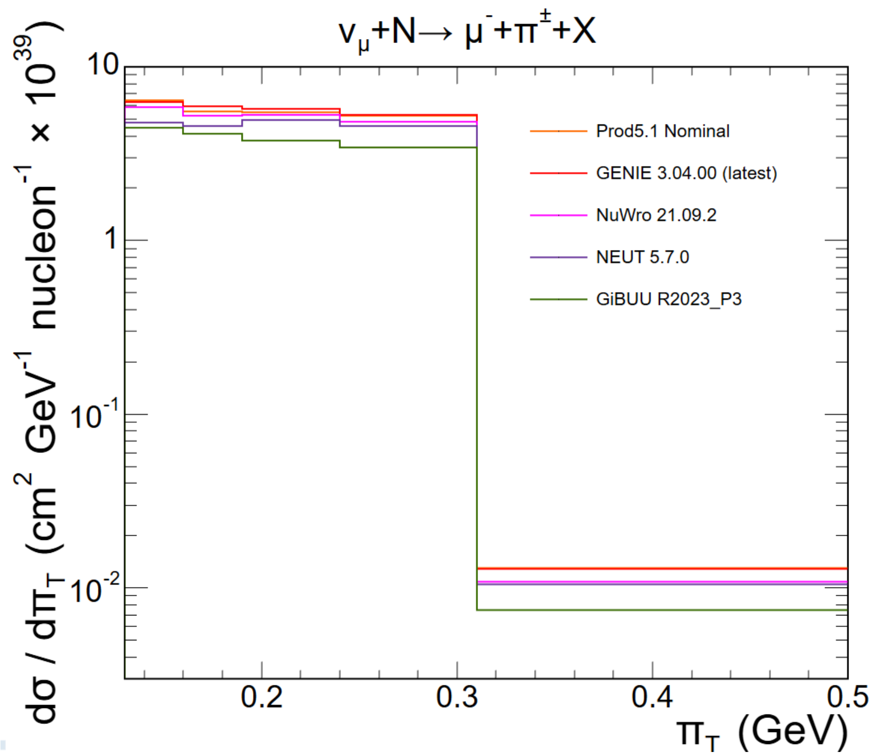

Figure 6.10: A comparison of generator results for four major event generators. The primary differences visible in this binning are of magnitude, though some shape discrimination is possible with NEUT. The log scale is used only to be able to include the last bin cleanly.

be calculated with no weights other than flux central value, but closure was achieved with this configuration. The spread of the generators is within expectations, and there are no unexpected features. They are also around what one would expect from the MINERvA results [94].

6.10 Unfolding

As explained in section 5.2, it is necessary to perform unfolding on the purity-corrected data to take it from the reconstructed space to the true space of the differential variable. The unfolding technique I use for this analysis is the D'Agostini method [66], also called iterative unfolding or (misleadingly) Bayesian unfolding. In this method, the inverted response matrix is repeatedly applied to the measurement until the chi-squared difference from the predicted nominal is minimized.

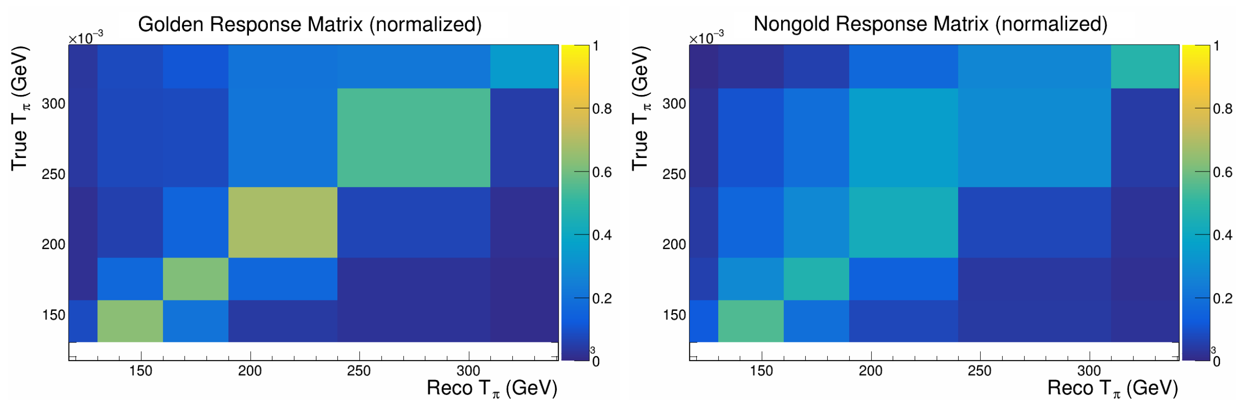

Figure 6.11: Response matrices for the gold and nongold samples, row-normalized. The bias in the nongold is expected due to the length-based energy reconstruction systematically underestimating the track length as well as the energy lost to hard scattering being unaccounted for.

The number of iterations used is the regularization parameter of this method. This is handled using RooUnfold [95], a standard tool. D'Agostini unfolding was chosen due to its simplicity, standardized use in NOvA, and good performance across fake data.

The response matrix is calculated for the nominal and every systematically varied sample, to be used in the construction of the final cross-section covariance matrix. However, when treating the systematic universes as fake data, the nominal response matrix is used. The nominal response matrices are shown in figure 6.11.

To avoid bias it is standard to pre-commit to a number of iterations. Therefore I measure the chi-squared for 1-6 iterations of unfolding for both GENIE multiverses and poisson-fluctuated fake data sets (see figures 6.12 and 6.13).

For most of these samples, unfolding either converges at infinity (presumably, that or a very large value) or it converges after one iteration. As such, I take one iteration as my value for both the golden and non-golden samples. It is not uncommon in cross-section measurements to have significant variance around different methods to determine the regularization parameter for unfolding, and it is generally considered a judgment call.

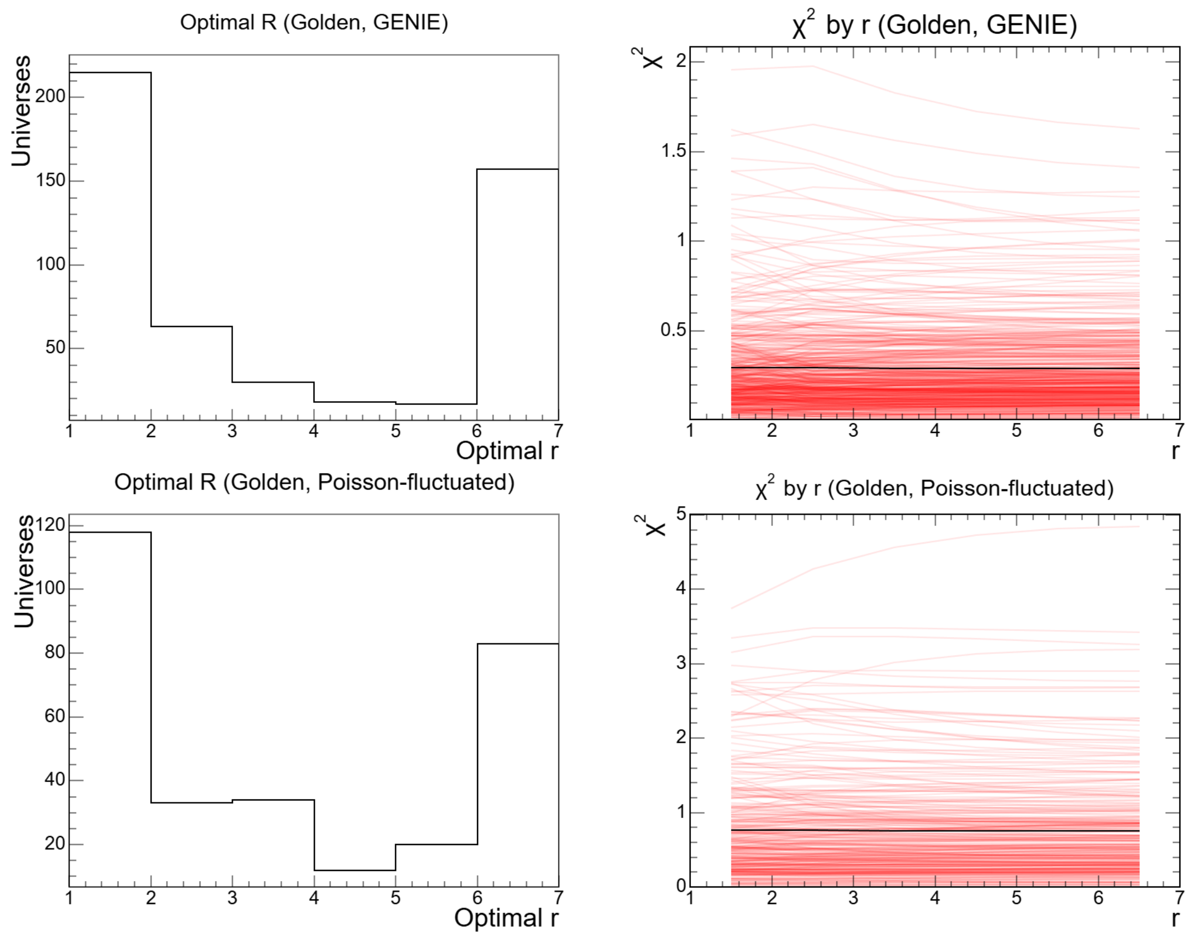

Figure 6.12: Unfolding optimization for the golden sample. Top-left: optimal number of iterations (R) for GENIE universes. Top-right: trend lines of the χ^2 for GENIE universes, note how they are almost flat. Bottom-left: optimal R for poisson-fluctuated fake data. Bottom-right: trend lines of the χ^2 for poisson-fluctuated fake data.

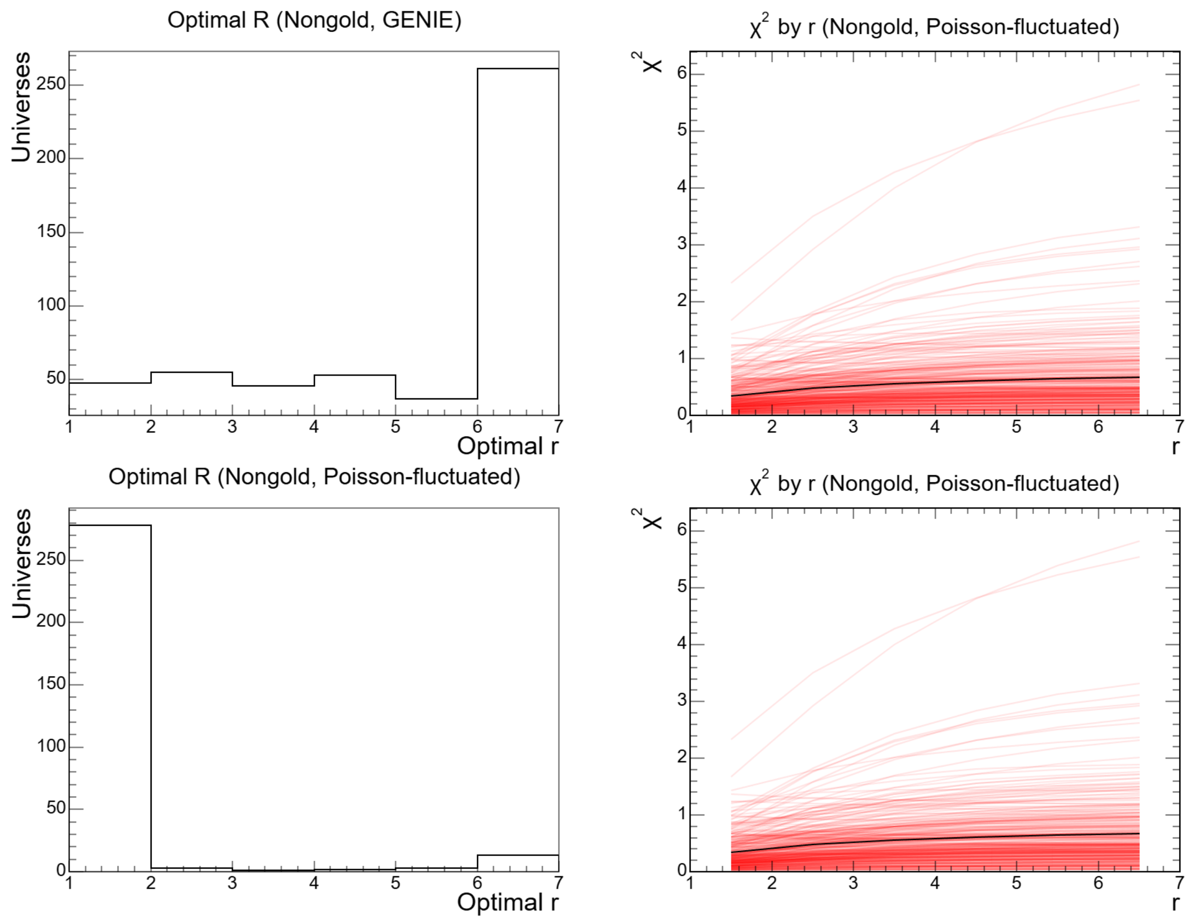

Figure 6.13: Unfolding optimization for the nongold sample. Top-left: optimal number of iterations (R) for GENIE universes. Note non-convergence, see figure 6.14. Top-right: trend lines of the χ^2 for GENIE universes. Bottom-left: optimal R for poisson-fluctuated fake data. Observe how clearly it favors one iteration. Bottom-right: trend lines of the χ^2 for poisson-fluctuated fake data.

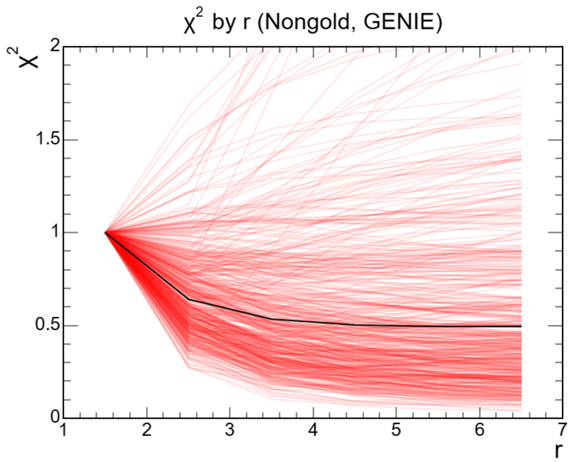

Figure 6.14: Further examination of the GENIE fake data for the nongold sample. Here, each universe line has been divided by its value at one iteration to reveal per-universe trends. Almost all of the change happens between one and two iterations. This gives a weak argument for using two iterations for the nongold that I ultimately decided against.

6.11 Measurement

The final steps to make the measurement are flux, target counting, and efficiency. Target counting is handled in the standard way for NOvA, by sampling the GEANT4 geometry of the fiducial volume with 10^8 points, randomly selecting an atomic Z-value at each point with probabilities determined by the material composition of the detector at that point, and taking the volume-weighted sum. The flux is calculated using the NOvA-standard DeriveFlux tool, for nominal, beam systematics, and PPFX multiverse systematics. The nominal and beam systematics are calculated using only the PPFX central value weight, while the PPFX universes are calculated only with their corresponding universe weight.

The efficiency is estimated with the simulation, separately for each systematic universe as well as the nominal for the purposes of constructing the final covariance matrix, though again only the nominal efficiency correction is used for fake data studies. Efficiency here is defined with the denominator being the total number of signal events in the fiducial volume, while the numerator is the selected signal subcategory events. Therefore, there are separate efficiencies for the golden and non-golden samples.

Then, using the standard cross-section formula as detailed in section 5.2 the cross-section is computed separately for the golden and non-golden samples across each systematically varied sample. The results for all of the systematically varied samples are then combined into two final covariance matrices, one for the golden sample and one for the non-golden. Then, using the BLUE method illustrated by Valassi [96], these dependent and correlated measurements are combined into a final measurement. Notably, this takes into account information from all measurements of all bins to determine the combination in each individual bin, subject to normalization conditions. From the same, the combined covariance matrix can also be computed, as well as its parts from the statistical and various systematic contributions.

The BLUE method is robust in that it contains several checks on its own validity, including the compatibility of the distributions to be combined (accounting for the covariance matrix), as well as the ability to evaluate the coefficients of combination which, if too large, exposes the combination to errors from small fluctuations in the covariance matrix. If either of these conditions is violated, there is the option to fall back to a combination that considers only measurements within the same bin for calculating that bin's combined value, in this case block-diagonalizing the covariance matrix into 2×2 blocks.

6.12 Cross Section Fake Data Studies

Before applying this methodology to actual data, fake data studies were carried through all the way to the cross section measurement. This includes the previously discussed energy reweight fake data, but also 20% increases or decreases in the potentially tricky single proton and multiple charged pion backgrounds (figure 6.15). All of these were handled well. The energy reweight results are in figure 6.16. The others agree closely, and can be seen in appendix A.2.

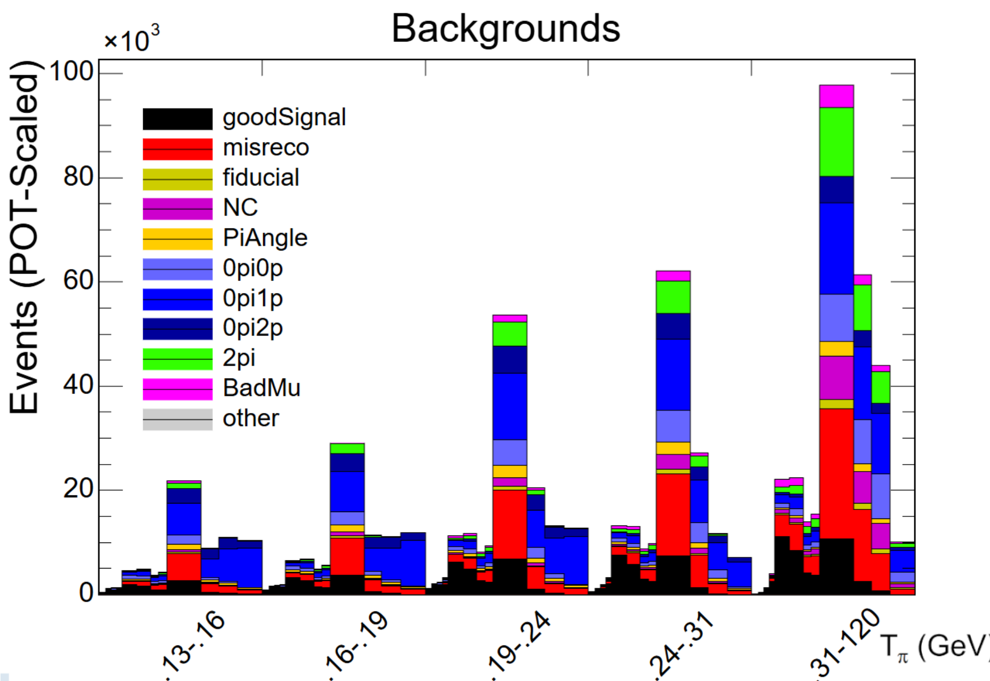

Figure 6.15: A breakdown of the various background components. "goodSignal" combines both gold and nongold, "PiAngle" are events failing the pion angle quality cut, "0pi0p" and "0pi1p" both have zero charged pions and either zero or one proton, while "0pi2p" has zero charged pions and two or more protons. "2pi" is two or more charged pions, and "BadMu" are events where the muon fails any of its relevant cuts. Category membership is tested in order, so an event will appear in exactly one category, the highest in the legend if it qualifies for multiple. Note the very large one proton and multi-pion backgrounds.

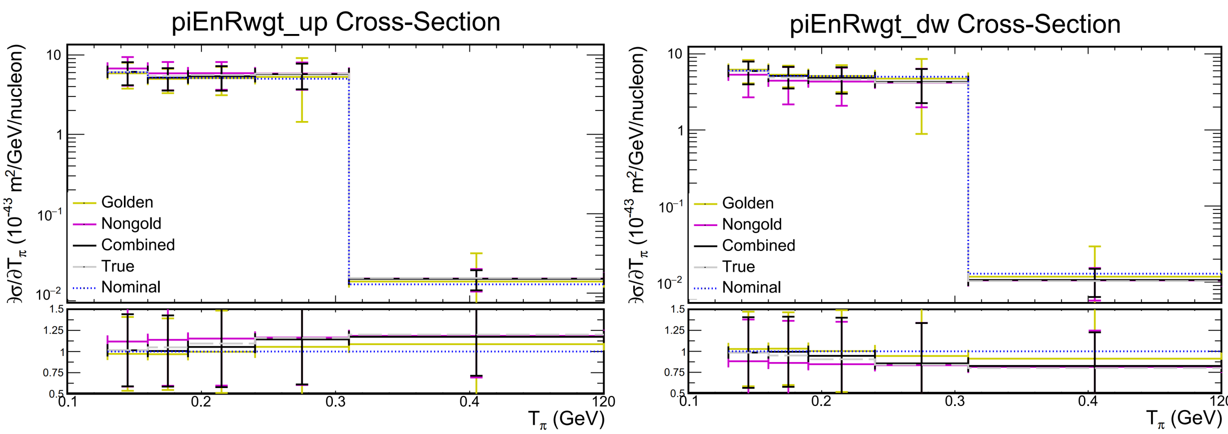

Figure 6.16: The most extreme fake data examples. The combined central value tracks the (calculated) true value of the fake data. It is difficult to distort the data significantly from the error bars in a physical way, which is why the weighting is capped at ±20%.

6.13 Results

Data was unboxed in two stages. First, 10% data was run for both the BDT inputs (shown previously) and the analysis. The event counts per bin are very low for the golden template, and so it was unsurprising that the fit failed on it, however, the nongold results looked reasonable. As such, I then looked at the full data.

Unfortunately, it was immediately clear there were significant but fixable issues. The final few BDT bins (corresponding to the background-dominated portion) were wildly off in each analysis bin, and the fit was massively reducing the background template and enhancing the misreconstructed template. Furthermore, the templates looked strange even in the low-BDT region— it seemed like I could do a better fit by-eye. When carried through to measurement and combination, the cross-sections were completely incompatible with the golden being very low and the nongold very high. Variation to this degree— especially the non-monotonic behavior in the first four bins of analysis bin two and the uptick in the last bin in analysis bin three— was not seen in any of the fake data studies.

Closer investigation revealed that the input covariance matrix to the fit was the culprit of the unintuitive fitting behavior, forcing the templates to be pulled in strange tensions between each other and the analysis bins. I therefore re-ran the fit and all subsequent steps with statistical errors only, and that is the result I will be presenting here. I unfortunately ran out of time to diagnose and correct whatever specific issue was happening with the systematics.

Figure 6.17 shows the data against the stacked templates before fitting, and figure 6.18 shows the same after fitting. Again, this is using statistical errors only, so the input covariance matrix to the fit is diagonal. While there is some clear tension in the fit, the results are not unreasonable, especially with the large systematic errors on the templates we know are present (but unutilized).

The resulting purity correction is shown in figure 6.19, followed by the unfolded spectrum in figure 6.20. The first bin is somewhat questionable for both golden and nongold, but not enormously far from expectation when considered individually.

Finally, the measurement is carried through with efficiency correction and dividing by flux, target count, and bin width. The results are shown in figure 6.21, with the combined result (via BLUE with only the covariance matrix derived from the poisson-fluctuated fits) shown against generator predictions in figure 6.22. Here the large tension between the golden and nongold measurements in the first bin can be seen, while the other bins look more reasonable, especially the third and fourth (the golden is always unreliable in the fifth bin and the combination is therefore always dominated by the nongold).

6.14 Discussion

No concrete conclusions can be drawn from the analysis as-is due to systematic uncertainties not being included. A more thorough investigation into the failure modes of this analysis structure is required. In particular, a better validation framework for the template fit should be devised. It would likely also be useful to attempt to further reduce backgrounds before fitting, and it

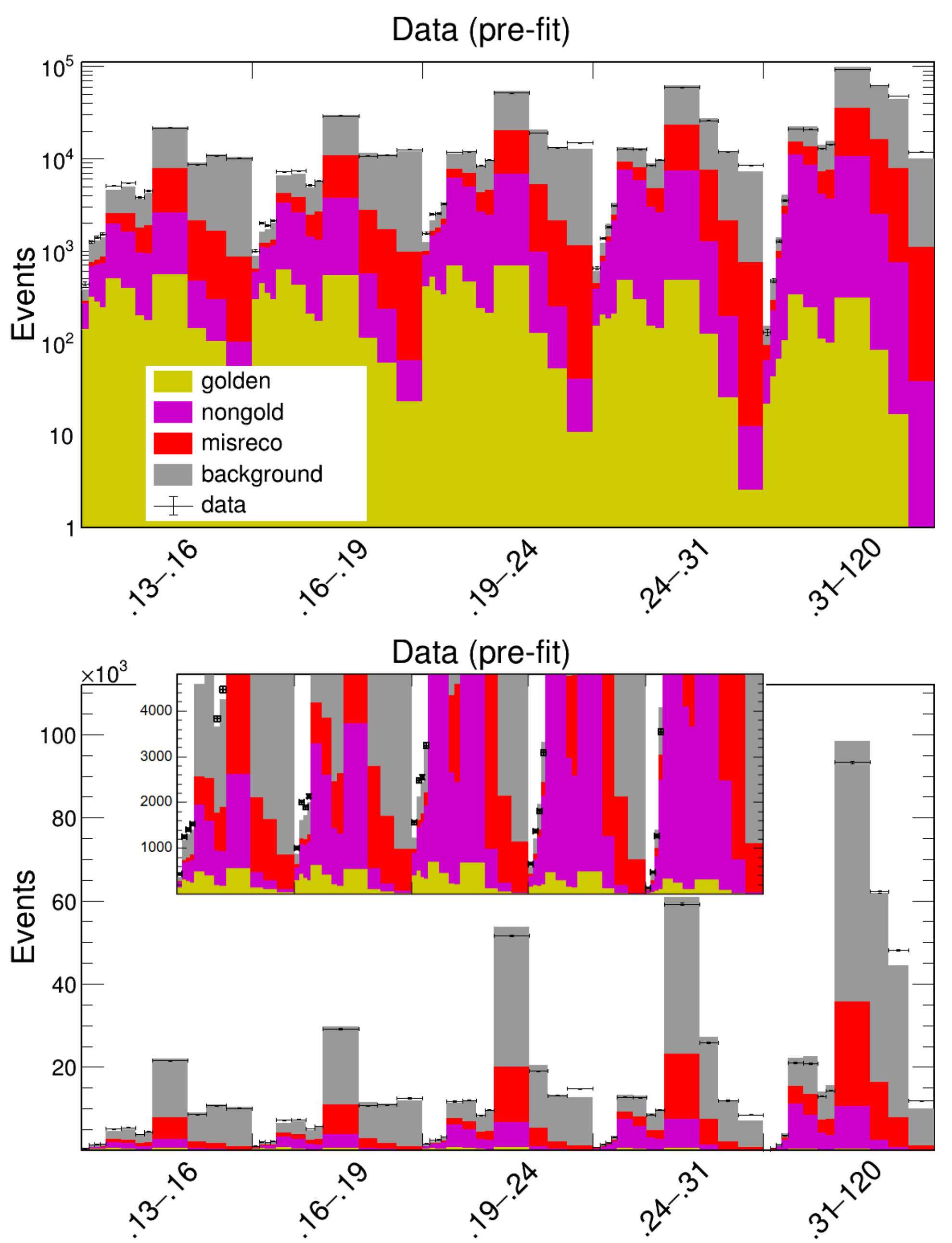

Figure 6.17: Pre-fit data spectrum against nominal templates (POT-scaled). The x-axis labels are T_π in units of GeV, and the sub-binning is in the BDT axis. Top: log scale. Bottom: linear scale, the legend is the same as log but is omitted for space. Note the 10^3 scaling for number of events. Zoomed inset to show first four bins, which are the most important for the golden template and also important for nongold.

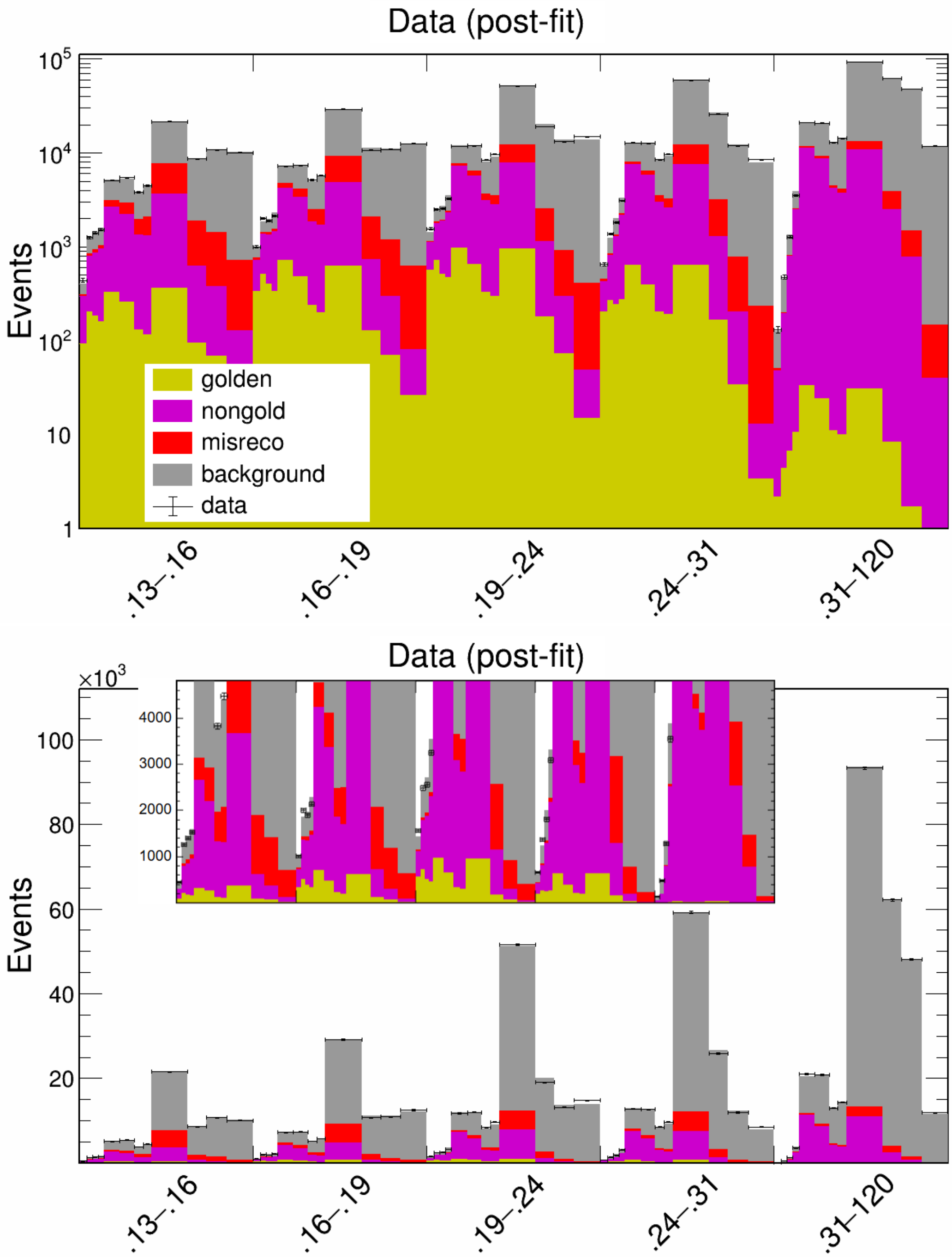

Figure 6.18: The same as the pre-fit spectra, (figure 6.17), but the templates are now scaled to the values determined in the stats-only template fit. Note the tension in the inset zoom, especially in the second energy bin— unfortunately, I have not been able to determine the source of this shape difference, as the background breakdown (see figure 6.15) does not suggest any particular source.

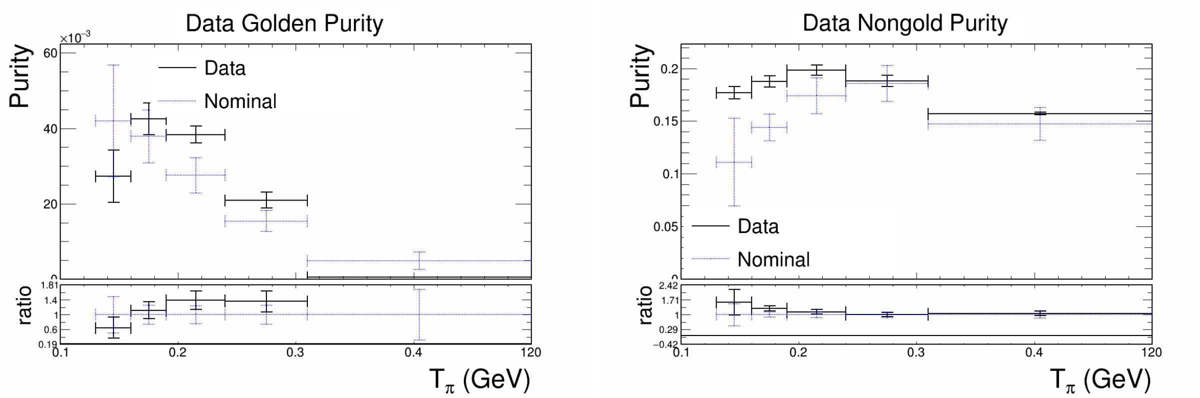

Figure 6.19: Purity results for the data compared to nominal. We do not necessarily expect them to track well, but nominal is the only reasonable point of comparison available. The final bin is shortened for legibility, but does extend to 120 GeV. The error bars for the nominal are for full systematics. Left: golden sample. Note the major shape difference stemming from the first bin. Because this is from a stats-only fit, the error bars are artificially small and should not be trusted. Right: nongold. Again, the error bars are much smaller than they should be with systematics included.

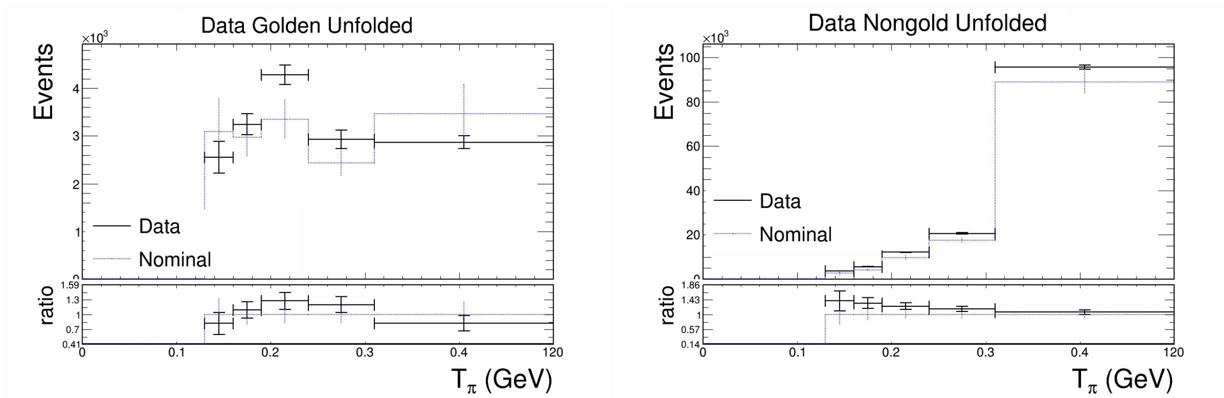

Figure 6.20: The data spectrum purity corrected and then unfolded (one iteration of D'Agostini unfolding). The nominal is included for reference, but we do not necessarily expect the data to track it. The final bin is shortened for legibility, but does extend to 120 GeV. The error bars on the nominal are for full systematics. Left: golden sample. Because this is propagated from a stats-only fit, the error bars are artificially small and should not be trusted. Right: nongold. Again, the error bars are much smaller than they should be with systematics included.

would likely simplify things to combine the misreconstructed and background templates, as their separation is at this point a historical artifact of the analysis.

The nature of the systematic errors should also be assessed, particularly which of the GENIE pion-production knobs it is appropriate to include. These are likely the source of the large correlations in the input covariance matrix to the fit.

It may also be worth investigating other unfolding methodologies to see if they are more stable and yield more consistent regularization parameters between assessment schemes. Singular Value Decomposition (SVD) unfolding would be easy to implement, but its theoretical validity in this case, given the very small number of analysis bins, should be checked.

The use of a BDT to separate the categories was largely out of convenience and due to a shortage of time. It is almost certain that a more sophisticated ML method would provide better separation and enable a superior template fit.

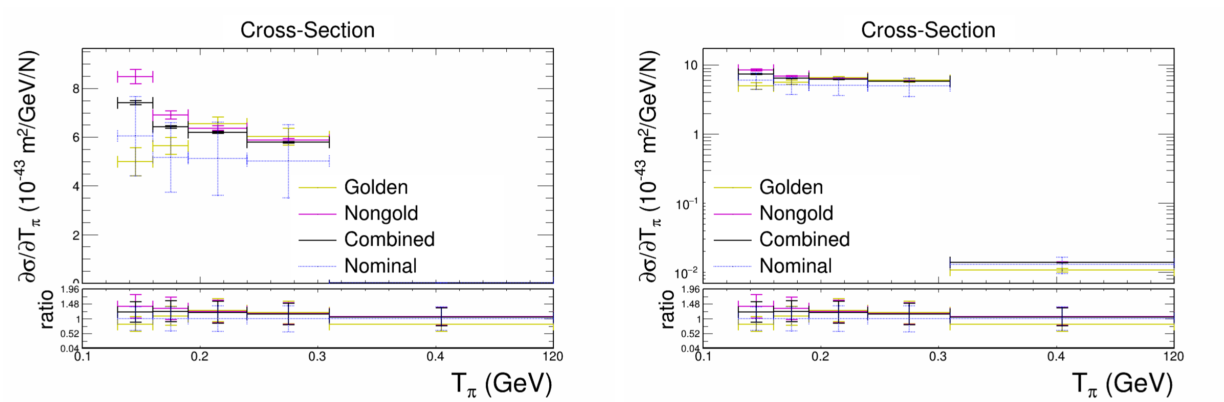

Figure 6.21: The final cross-section measurements. "N" refers to the number of nucleons in the fiducial volume. The covariance matrix used for the combination was derived only from the poisson-fluctuated fits, and therefore does not capture systematic errors, only the fit and statistical errors. As such the error bars are artificially small for the golden, nongolden, and their combination. The combination was performed under full BLUE. Compatibility failed for both full BLUE and partial BLUE, but this is unsurprising due to the artificially small error bars. Due to the significant tension in the first bin, that result should be disregarded. The subsequent bins are more reasonable. Left: linear scale. The final bin is too small to be visible. Right: log scale, to show final bin.

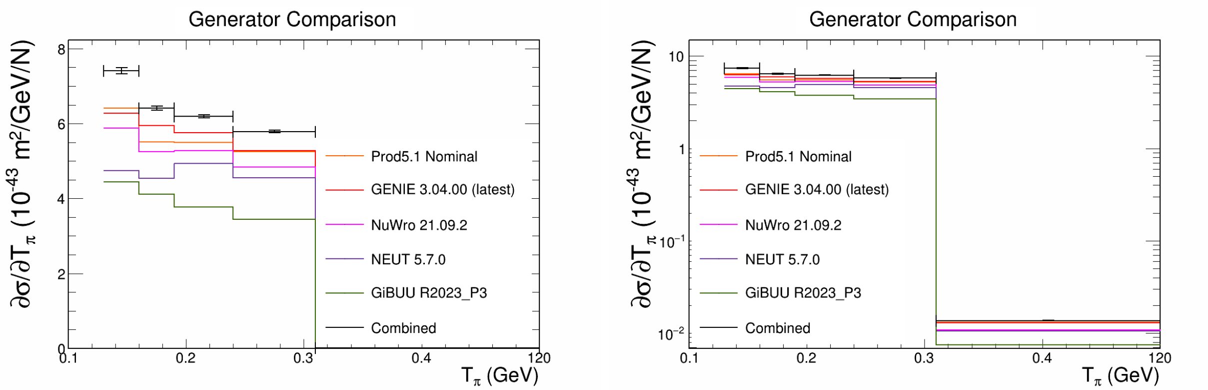

Figure 6.22: A comparison to the generators. "N" refers to the number of nucleons in the fiducial volume. Generator details are in section 6.9 and figure 6.22. Because of the incomplete error bars, the ability to draw conclusions is limited. However, the general shape agrees with the generators, and the magnitude disfavors GiBUU. Left: linear scale. The final bin is too small to be visible. Right: log scale, to show final bin.

Finally, more thorough fake data studies need to be carried out, given that the issues borne out by the real data were not caught by the studies I performed. It is evident that I erred too far on the side of keeping the studies physical, and in doing so neglected to properly stress-test the analysis machinery.

Chapter 7

Conclusions

In this chapter, I will briefly review the conclusions of my work. This chapter is speculative and less formal, and is primarily aimed at members of NOvA who want to improve on my work.

7.1 VertexCVN

VertexCVN performs very well in comparison to Elastic Arms, and has been adopted as NOvA's new primary vertexer. However, there are still some ways in which it is not a strict upgrade, offering opportunities for further improvement.

Firstly there is the precision. In theory, VertexCVN has access to nearly as much information as ElasticArms does, and so it should be able to resolve the vertex to better sub-cell precision. I suspect that the use of coordinate conversions is what is limiting this, as the ways the geometry of the detectors is not a perfect grid effectively smear the idealized "pixel". While it is not straightforwardly clear how to augment VertexCVN with better detector geometry information, this would allow for major improvements. An option is including the pixel center information in color-equivalent layers. Another is feeding the coordinates of the center cells of the pixel map into the network somehow, perhaps immediately before the dense layers. However, this would implicitly require the network to independently learn the intricacies of the detector geometry, which is likely too difficult for the current architecture and would negatively impact the required memory profile.

Besides the precision, VertexCVN has been found during validation to perform in a biased way with certain event types, including and especially neutral current interactions that create one π^0 with no other visible detector activity. Accounting for this is relatively simple conceptually: re-train VertexCVN with a data-set that has been properly weighted to more evenly represent different event topologies. Unfortunately, any re-training incurs an enormous validation burden, so this would only be worth it if a streamlined vertex validation framework is introduced for the potential NOvA production 7.

Overall, VertexCVN is a highly performant tool, especially given the operating constraints of NOvA's analysis framework. Due to its relatively low memory profile and high evaluation speed even on CPUs, I believe that it is an excellent candidate for adaptation to other experiments with image-like output and relatively low access to GPUs.

7.2 Cross-Section Analysis

While it is disappointing I could not get the measurement to converge in time, the results of the diagnostics I was able to run are promising. The template fit input covariance matrix simply is not sufficient for the task at hand. It is likely that it is in some way overspecified, locking the fit into obeying the ways GENIE expects the cross-section to vary. Removing the knobs that impact overall normalization might be all that is required, though there are likely other subtleties at play. It could also be the case that rather than removing portions of it, the GENIE systematic needs to be augmented somehow, presumably in a way that lessens correlations.

I suspect that GENIE is the culprit both because it is the largest systematic and because I at one point performed a fit excluding only the GENIE multiverse component of the input covariance matrix. Unfortunately, I lost this plot in the scramble of debugging, and recreating it is nontrivial.

Throughout chapter 6 I detailed various ways the analysis framework needs improved to be able to catch the cases that came up in the real data. I stand by those suggestions for the next analyzer to tackle this problem, but I do believe that the core issue I encountered was with the fit.

As such, I hold that this analysis has successfully proved the feasibility of a differential cross-section measurement in charged pion energy using the NOvA Near Detector, even if I was not able to fully complete that measurement.

Finally, I offer some speculation on the potential future results of this analysis if it is completed. While the uncertainty of the eventual final result is expected to be large, likely precluding shape-based discrimination between the tightly clustered GENIE, NuWro, and NEUT models, the statistics-only fit is highly suggestive of excluding the GIBUU model by overall normalization. If efforts to refine the selection and/or GoldenBDT pay off substantially, it might even be possible to slightly disfavor NEUT given its much lower value in the first analysis bin.

Bibliography

[1] Wang K C. A suggestion on the detection of the neutrino. Phys. Rev., 1942, 61: 97-97.

[2] Cowan C L, Reines F, Harrison F B, et al. Detection of the free neutrino: a confirmation. Science, 1956, 124(3212): 103-104.

[3] Danby G, Gaillard J M, Goulianos K, et al. Observation of high-energy neutrino reactions and the existence of two kinds of neutrinos. Phys. Rev. Lett., 1962, 9: 36-44.

[4] Kodama K, Ushida N, Andreopoulos C, et al. Observation of tau neutrino interactions. Physics Letters B, 2001, 504(3): 218-224.

[5] Davis R. A review of the homestake solar neutrino experiment. Progress in Particle and Nuclear Physics, 1994, 32: 13-32.

[6] Super-Kamiokande Collaboration, Fukuda Y, Hayakawa T, Ichihara E, et al. Evidence for oscillation of atmospheric neutrinos. Phys. Rev. Lett., 1998, 81: 1562-1567.

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

[8] Wolfenstein L. Neutrino oscillations in matter. Phys. Rev. D, 1978, 17: 2369-2374.

[9] Mikheyev S P, Smirnov A Y. Resonance Amplification of Oscillations in Matter and Spectroscopy of Solar Neutrinos. Sov. J. Nucl. Phys., 1985, 42: 913-917.

[10] SNO Collaboration, Ahmad Q R, Allen R C, Andersen T C, et al. Measurement of the rate of $\nu_e + d \to p + p + e^-$ interactions produced by ^8B solar neutrinos at the sudbury neutrino observatory. Phys. Rev. Lett., 2001, 87: 071301.

[11] Mandal S, Miranda O G, Garcia G S, et al. Toward deconstructing the simplest seesaw mechanism. Phys. Rev. D, 2022, 105: 095020.

[12] Particle Data Group, Navas S, et al. Review of particle physics. Phys. Rev. D, 2024, 110(3): 030001.

[13] KATRIN Collaboration, Aker M, Batzler D, Beglarian A, et al. Direct neutrino-mass measurement based on 259 days of KATRIN data. Science, 2025, 388(6743): 180-185.

[14] MINOS Collaboration, Adamson P, Andreopoulos C, Armstrong R, et al. Measurement of the neutrino mass splitting and flavor mixing by MINOS. Phys. Rev. Lett., 2011, 106: 181801.

[15] K2K Collaboration, Ahn M H, Aliu E, Andringa S, et al. Measurement of neutrino oscillation by the K2K experiment. Phys. Rev. D, 2006, 74: 072003.

[16] Abratenko P, Aduszkiewicz A, Akbar F, et al. ICARUS at the fermilab short-baseline neutrino program: initial operation. The European Physical Journal C, 2023, 83(6): 467.

[17] Abe, K., Akhlaq, N., Akutsu, R., et al. Measurements of neutrino oscillation parameters from the T2K experiment using 3.6×10^{21} protons on target. Eur. Phys. J. C, 2023, 83(9): 782.

[18] The NOvA Collaboration, Abubakar S, Acero M A, Acharya B, et al. Precision measurement of neutrino oscillation parameters with 10 years of data from the NOvA experiment. Phys. Rev. Lett., 2026, 136: 011802.

[19] Abdurashitov J N, Veretenkin E P, Vermul V M, et al. Solar neutrino flux measurements by the soviet-american gallium experiment (SAGE) for half the 22-year solar cycle. Journal of Experimental and Theoretical Physics, 2002, 95(2): 181-193.

[20] Hampel W, Handt J, Heusser G, et al. GALLEX solar neutrino observations: results for GALLEX IV. Physics Letters B, 1999, 447(1): 127-133.

[21] Altmann M, Balata M, Belli P, et al. GNO solar neutrino observations: results for GNO I. Physics Letters B, 2000, 490(1): 16-26.

[22] KamLAND Collaboration, Gando A, Gando Y, Hanakago H, et al. ^{7}Be solar neutrino measurement with KamLAND. Phys. Rev. C, 2015, 92: 055808.

[23] Borexino Collaboration, Basilico D, Bellini G, Benziger J, et al. Final results of borexino on CNO solar neutrinos. Phys. Rev. D, 2023, 108: 102005.

[24] Daya Bay Collaboration, An F P, Bai W D, Balantekin A B, et al. Measurement of electron antineutrino oscillation amplitude and frequency via neutron capture on hydrogen at daya bay . Phys. Rev. Lett., 2024, 133: 151801.

[25] Double Chooz, Soldin P. Precision Neutrino Mixing Angle Measurement with the Double Chooz Experiment and Latest Results. PoS, 2024, TAUP2023: 228.

[26] RENO Collaboration, Jeon S, Kim H I, Choi J H, et al. Measurement of reactor antineutrino oscillation parameters using the full 3800-day dataset of the RENO experiment. Phys. Rev. D, 2025, 111: 112006.

[27] Araki T, Enomoto S, Furuno K, et al. Experimental investigation of geologically produced antineutrinos with KamLAND. Nature, 2005, 436(7050): 499-503.

[28] Gando A, Gando Y, Ichimura K, et al. Partial radiogenic heat model for earth revealed by geoneutrino measurements. Nature Geoscience, 2011, 4(9): 647-651.

[29] Borexino Collaboration, Agostini M, Altenmüller K, Appel S, et al. Comprehensive geoneutrino analysis with borexino. Phys. Rev. D, 2020, 101: 012009.

[30] SNO + Collaboration, Abreu M, Albanese V, Allega A, et al. Measurement of reactor antineutrino oscillation at SNO+. Phys. Rev. Lett., 2025, 135: 121801.

[31] Akimov D, Albert J B, An P, et al. Observation of coherent elastic neutrino-nucleus scattering . Science, 2017, 357(6356): 1123-1126.

[32] Barranco J, Miranda O G, Rashba T I. Probing new physics with coherent neutrino scattering off nuclei. Journal of High Energy Physics, 2005, 2005(12): 021.

[33] Ayres D, Drake G, Goodman M, et al. The NOvA Technical Design Report. Fermilab, 2007: 600.

[34] NOvA Collaboration, Adamson P, Ader C, Andrews M, et al. First measurement of muon-neutrino disappearance in NOvA. Phys. Rev. D, 2016, 93: 051104.

[35] Adamson P, Anderson K, Andrews M, et al. The NuMI neutrino beam. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 2016, 806: 279-306.

[36] Frühwirth R. Application of kalman filtering to track and vertex fitting. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 1987, 262(2): 444-450.

[37] Fernandes L, Oliveira M. Real-time line detection through an improved hough transform voting scheme. Pattern Recognition, 2008, 41: 299-314.

[38] Ohlsson M, Peterson C, Yuille A L. Track finding with deformable templates — the elastic arms approach. Computer Physics Communications, 1992, 71(1): 77-98.

[39] Krishnapuram R, Keller J. A possibilistic approach to clustering. IEEE Transactions on Fuzzy Systems, 1993, 1(2): 98-110.

[40] Yang M S, Wu K L. Unsupervised possibilistic clustering. Pattern Recognition, 2006, 39(1): 5-21.

[41] Baird M D. An Analysis of Muon Neutrino Disappearance from the NuMI Beam Using an Optimal Track Fitter. Indiana U., 2015.

[42] Allison J, Amako K, Apostolakis J, et al. Recent developments in Geant4. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 2016, 835: 186-225.

[43] Ahdida C, Bozzato D, Calzolari D, et al. New capabilities of the FLUKA multi-purpose code . Frontiers in Physics, 2022, Volume 9 - 2021.

[44] Battistoni G, Boehlen T, Cerutti F, et al. Overview of the FLUKA code. Annals of Nuclear Energy, 2015, 82: 10-18.

[45] Andreopoulos C, et al. The GENIE Neutrino Monte Carlo Generator. Nucl. Instrum. Meth. A, 2010, 614: 87-104.

[46] Smith E S. Technical note: identification of necessary reconstruction improvements via hand-scan. Nova Internal Document 54684, 2022.

[47] Groh M. Constraints on Neutrino Oscillation Parameters from Neutrinos and Antineutrinos with Machine Learning. Indiana U., Bloomington (main), 2021.

[48] Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain.. Psychological review, 1958, 65 6: 386-408.

[49] Samuel A. Some studies in machine learning using the game of checkers. II—recent progress. Annual Review in Automatic Programming, 1969, 6: 1-36.

[50] Drucker H, Cortes C. Boosting decision trees//Touretzky D, Mozer M, Hasselmo M. Advances in Neural Information Processing Systems: Vol. 8. MIT Press, 1995.

[51] Roe B P, Yang H J, Zhu J, et al. Boosted decision trees as an alternative to artificial neural networks for particle identification. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 2005, 543(2): 577-584.

[52] Brun R, Rademakers F. ROOT: An object oriented data analysis framework. Nucl. Instrum. Meth. A, 1997, 389: 81-86.

[53] Messier M D. Vertex reconstruction based on elastic arms. Nova Internal Document 7530, 2012.

[54] Radovic A. CVN for the second electron neutrino appearance analysis. Nova Internal Document 15386, 2016.

[55] Elkins M, Doyle D. Prod 5 prong CVN validation. Nova Internal Document 42229, 2020.

[56] Yu S, Djurcic Z. Improving energy reconstruction with 2D prong convolutional neural network - technical note. Nova Internal Document 42295, 2019.

[57] Sandler M, Howard A, Zhu M, et al. MobileNetV2: inverted residuals and linear bottlenecks// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2018.

[58] Hinton G E, Srivastava N, Krizhevsky A, et al. Improving neural networks by preventing co-adaptation of feature detectors. arXiv (2012). https://arxiv.org/abs/1207.0580.

[59] Salehin I, Kang D K. A review on dropout regularization approaches for deep neural networks within the scholarly domain. Electronics, 2023, 12(14).

[60] Kingma D P, Ba J. Adam: A Method for Stochastic Optimization. arXiv (2014). https://arxiv.org/abs/1412.6980.

[61] Alvarez-Ruso L, Sajjad Athar M, Barbaro M, et al. NuSTEC white paper: status and challenges of neutrino–nucleus scattering. Progress in Particle and Nuclear Physics, 2018, 100: 1-68.

[62] Abi B, Acciarri R, Acero M A, et al. Deep underground neutrino experiment (DUNE), far detector technical design report, volume II: DUNE physics. arXiv (2020). https://arxiv.org/abs/2002.03005.

[63] MiniBooNE Collaboration, Aguilar-Arevalo A A, Anderson C E, Bazarko A O, et al. First measurement of the muon neutrino charged current quasielastic double differential cross section . Phys. Rev. D, 2010, 81: 092005.

[64] MiniBooNE Collaboration, Aguilar-Arevalo A A, Brown B C, Bugel L, et al. First measurement of the muon antineutrino double-differential charged-current quasielastic cross section. Phys. Rev. D, 2013, 88: 032001.

[65] Megias G D, Donnelly T W, Moreno O, et al. Meson-exchange currents and quasielastic predictions for charged-current neutrino-^{12}C scattering in the superscaling approach. Phys. Rev. D, 2015, 91: 073004.

[66] D'Agostini G. A multidimensional unfolding method based on bayes' theorem. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 1995, 362(2): 487-498.

[67] Kim H, Schramm S, Horowitz C J. Delta excitations in neutrino-nucleus scattering. Phys. Rev. C, 1996, 53: 2468-2473.

[68] Singh S, Vicente-Vacas M, Oset E. Nuclear effects in neutrino production of Δ at intermediate energies. Physics Letters B, 1998, 416(1): 23-28.

[69] MiniBooNE Collaboration, Aguilar-Arevalo A A, Anderson C E, Bazarko A O, 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, 2011, 83: 052007.

[70] MINERvA Collaboration, Eberly B, Aliaga L, Altinok O, et al. Charged pion production in ν_μ interactions on hydrocarbon at $E_\nu = 4.0$ GeV. Phys. Rev. D, 2015, 92: 092008.

[71] T2K Collaboration, Abe K, Andreopoulos C, Antonova M, 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, 2017, 95: 012010.

[72] The T2K Collaboration, Abe K, Akutsu R, Ali A, et al. Measurement of the muon neutrino charged-current single π^+ production on hydrocarbon using the T2K off-axis near detector ND280. Phys. Rev. D, 2020, 101: 012007.

[73] MicroBooNE Collaboration, Abratenko P, Aldana D A, Arellano L, et al. Measurement of single charged pion production in charged-current ν_μ-ar interactions with the MicroBooNE detector. Phys. Rev. D, 2026, 113: 032007.

[74] The MINERvA Collaboration, Bercellie A, Kroma-Wiley K A, Akhter S, et al. Simultaneous measurement of muon neutrino ν_μ charged-current single π^+ production in CH, C, H_2O, fe, and pb targets in MINERvA. Phys. Rev. Lett., 2023, 131: 011801.

[75] K2K Collaboration, Rodriguez A, Whitehead L, Alcaraz J L, et al. Measurement of single charged pion production in the charged-current interactions of neutrinos in a 1.3 GeV wide band beam. Phys. Rev. D, 2008, 78: 032003.

[76] MiniBooNE Collaboration, Aguilar-Arevalo A A, Anderson C E, Bazarko A O, et al. Measurement of the ratio of the ν_μ charged-current single-pion production to quasielastic scattering with a 0.8 GeV neutrino beam on mineral oil. Phys. Rev. Lett., 2009, 103: 081801.

[77] MINERvA Collaboration, McGivern C L, Le T, Eberly B, et al. Cross sections for ν_μ and $\overline{\nu}_\mu$ induced pion production on hydrocarbon in the few-GeV region using MINERvA. Phys. Rev. D, 2016, 94: 052005.

[78] The ArgoNeuT Collaboration, Acciarri R, Adams C, Asaadi J, et al. First measurement of the cross section for ν_μ and $\overline{\nu}_\mu$ induced single charged pion production on argon using ArgoNeuT . Phys. Rev. D, 2018, 98: 052002.

[79] Devitt A. A measurement of single charged pion production in MicroBooNE. Lancaster University, 2023.

[80] Sweeney C. Measurement of the muon neutrino charged-current single charged pion production cross-section in the NOvA Near Detector. University College London, University Coll. London, University Coll. London, 2023.

[81] Kuruppu K D C D. Measurement of the Total Cross-Section of Muon Neutrino Charged-Current Coherent Pion Production in NOvA Near Detector. South Carolina U., University of South Carolina, South Carolina U., 2023.

[82] Rojas P N. A measurement of muon neutrino charged-current interactions with a charged pion in the final state using the NOνA near detector. Colorado State U., Fort Collins, 2023.

[83] Tripathi J. Charged Pion semi inclusive charged current cross section measurement in the NOvA experiment at Fermiilab. Panjab U., 2020.

[84] NOvA Collaboration, Acero M A, Adamson P, Aliaga L, et al. Measurement of the double-differential muon-neutrino charged-current inclusive cross section in the NOvA near detector. Phys. Rev. D, 2023, 107: 052011.

[85] González-Jiménez R, Megias G D, Barbaro M B, et al. Extensions of superscaling from relativistic mean field theory: the SuSAv2 model. Phys. Rev. C, 2014, 90: 035501.

[86] Juszczak C, Nowak J A, Sobczyk J T. Simulations from a new neutrino event generator. Nucl. Phys. B Proc. Suppl., 2006, 159: 211-216.

[87] Golan T, Sobczyk J, Żmuda J. NuWro: the wrocław monte carlo generator of neutrino interactions. Nuclear Physics B - Proceedings Supplements, 2012, 229-232: 499.

[88] Llewellyn Smith C. Neutrino reactions at accelerator energies. Physics Reports, 1972, 3(5): 261-379.

[89] Nieves J, Simo I R, Vacas M J V. Inclusive charged-current neutrino-nucleus reactions. Phys. Rev. C, 2011, 83: 045501.

[90] Bierlich C, Chakraborty S, Desai N, et al. A comprehensive guide to the physics and usage of PYTHIA 8.3. arXiv (2022). https://arxiv.org/abs/2203.11601.

[91] Hayato Y, Pickering L. The NEUT neutrino interaction simulation program library. The European Physical Journal Special Topics, 2021, 230(24): 4469-4481.

[92] Nieves J, Amaro J E, Valverde M. Inclusive quasielastic charged-current neutrino-nucleus reactions. Phys. Rev. C, 2004, 70: 055503.

[93] Buss O, Gaitanos T, Gallmeister K, et al. Transport-theoretical description of nuclear reactions . Physics Reports, 2012, 512(1): 1-124.

[94] MINERvA Collaboration, Eberly B, Aliaga L, Altinok O, et al. Charged pion production in ν_μ interactions on hydrocarbon at $E_\nu = 4.0$ GeV. Phys. Rev. D, 2015, 92: 092008.

[95] Brenner L, Balasubramanian R, Burgard C, et al. Comparison of unfolding methods using RooFitUnfold. International Journal of Modern Physics A, 2020, 35(24): 2050145.

[96] Valassi A. Combining correlated measurements of several different physical quantities. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 2003, 500(1): 391-405.

Appendix A

Analysis Supplementary Plots

This appendix is a collection of plots to supplement the content of chapter 6.

A.1 Purity

This section presents the breakdown of the input covariance matrix to the golden template fit, by systematic uncertainty source. GENIE, by far the largest contributor, is in figure A.1. The other non-beam non-target systematics are presented left to right in order of decreasing magnitude in figure A.2. Finally, all of the non-PPFX beam systematics are in figure A.3.

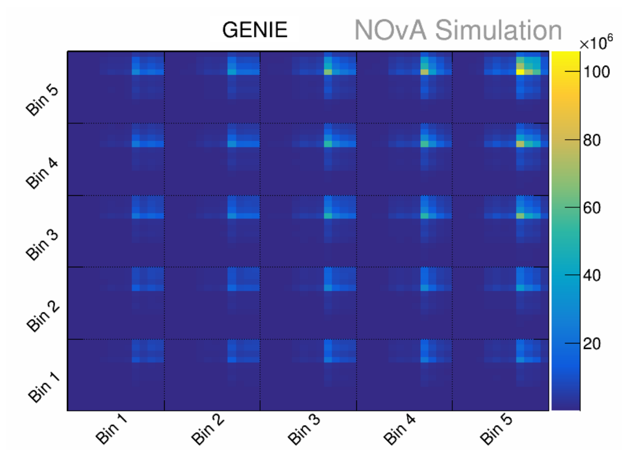

Figure A.1: The GENIE component of the template fit input covariance matrix. The superbins are listed by their index, which correspond to ranges of T_π. The subbins are in the BDT variable.

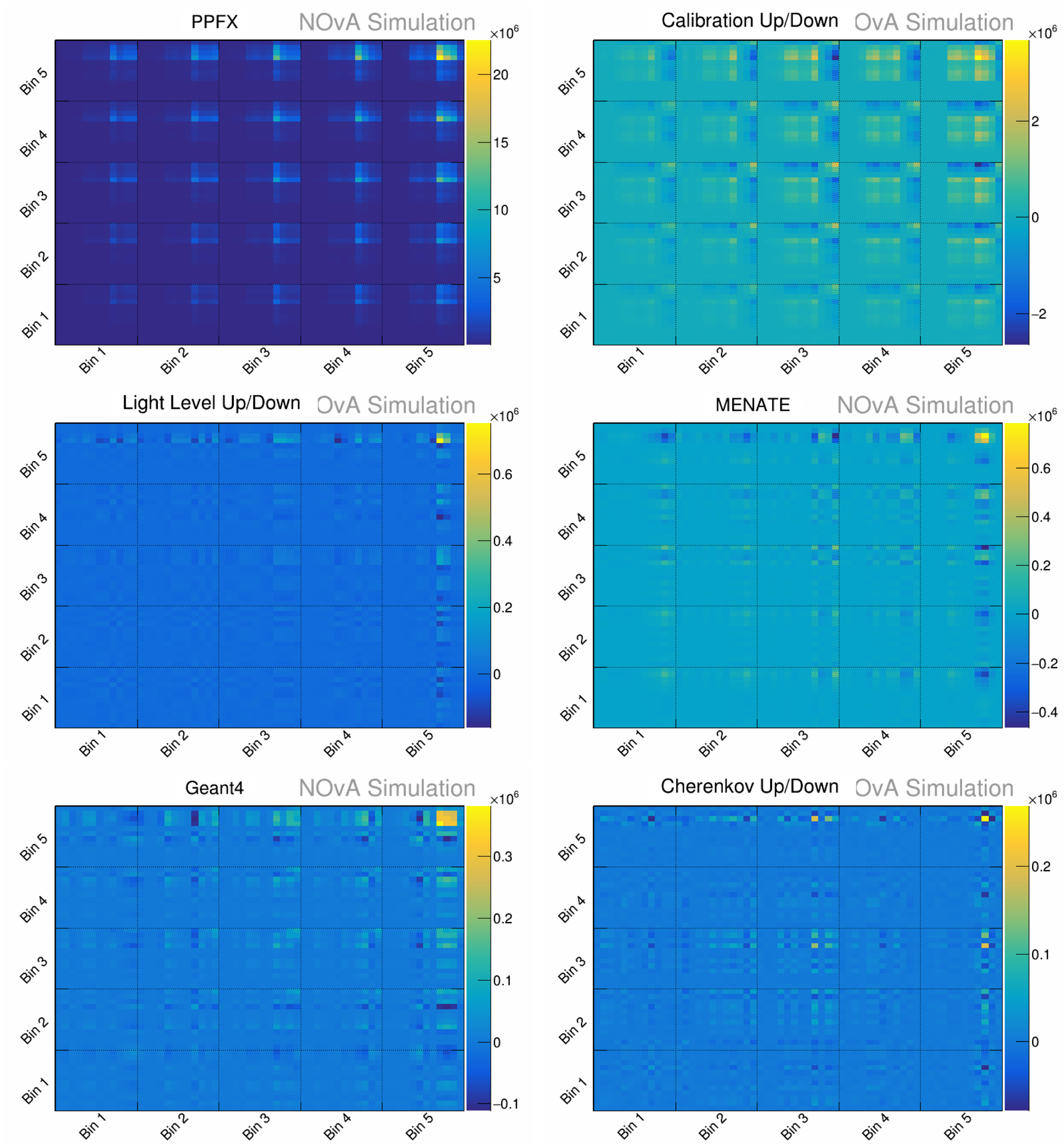

Figure A.2: The non-beam systematic components of the input covariance matrix. Menate is the neutron systematic.

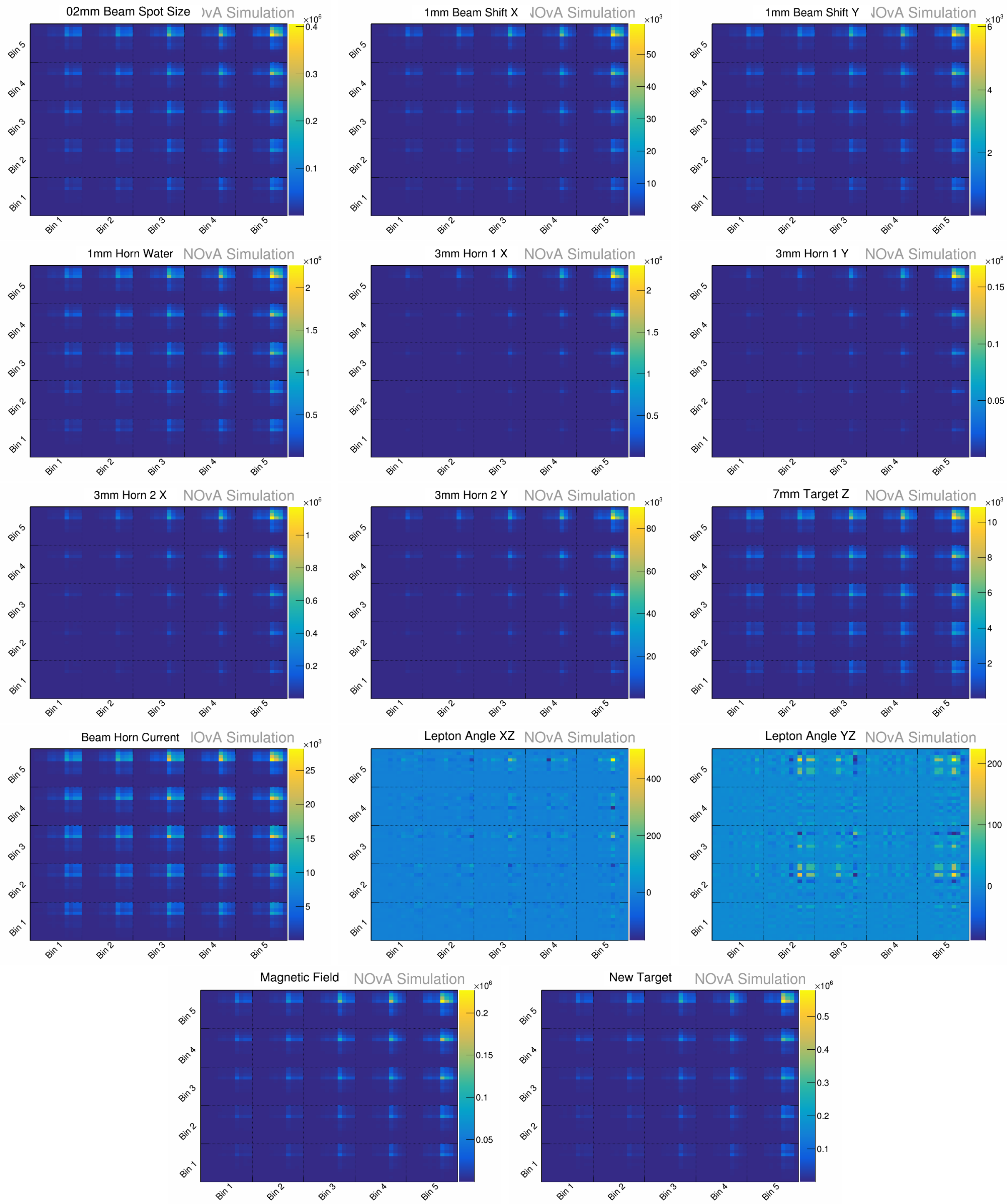

Figure A.3: The beam systematic components of the input covariance matrix.

A.2 Fake Data Cross-Sections

The other two fake data studies conducted were done by enhancing and suppressing the most relevant backgrounds to this analysis. I had wished to conduct more, but ran out of time. Selecting and performing additional studies to better test the analysis structure and code is imperative for improving this analysis.

The first background enhanced was the multi-charged-pion background. This one is especially significant as charged pion identification is always difficult for NOvA, and the only lever GoldenBDT was given on this background was the second-highest prong pionID score. Fortunately, the analysis handles it well, as shown in figure A.4. It might be reasonable to re-do this study at a 100% shift for a more thorough stress-test.

The second background enhanced was the single proton background. I decided to not restrict this to just the background, and instead allow the weight to enhance or suppress *all* events with exactly one proton above the kinetic energy threshold (adjusted to account for the different mass), leading to a major enhancement of the misreconstructed sample as well. Since single protons are easily confused for charged pions, it is important to make sure the analysis is robust to this effect. The results, shown in figure A.5, demonstrate that the analysis is capable of handling this situation. While exaggerating this effect to 100% would likely cause serious physicality issues, a weight higher than 20% might still be worth checking.

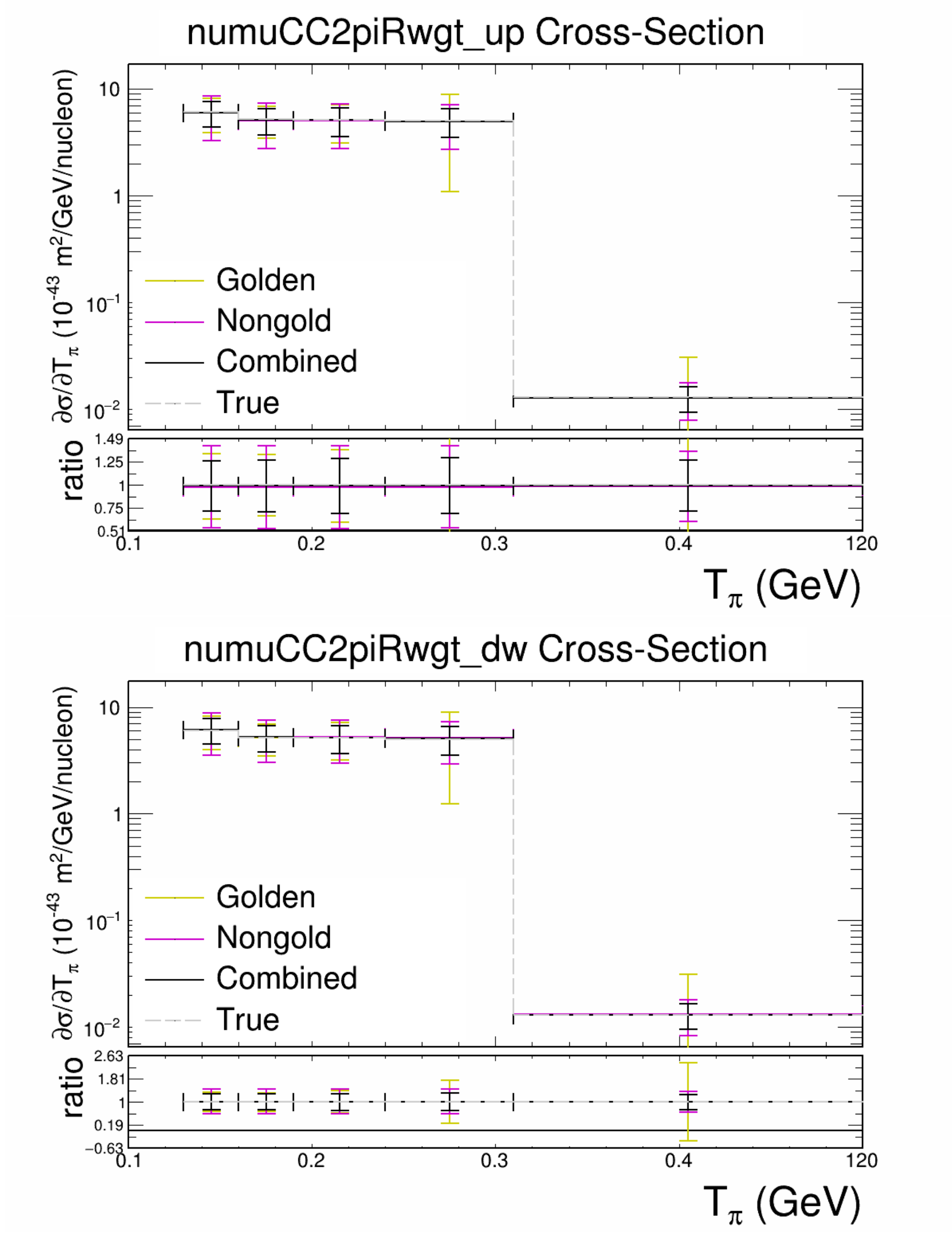

Figure A.4: Fake data study for enhancing/suppressing the multi-charged-pion background by 20%. Effects are minimal, demonstrating that GoldenBDT is able to handle this background.

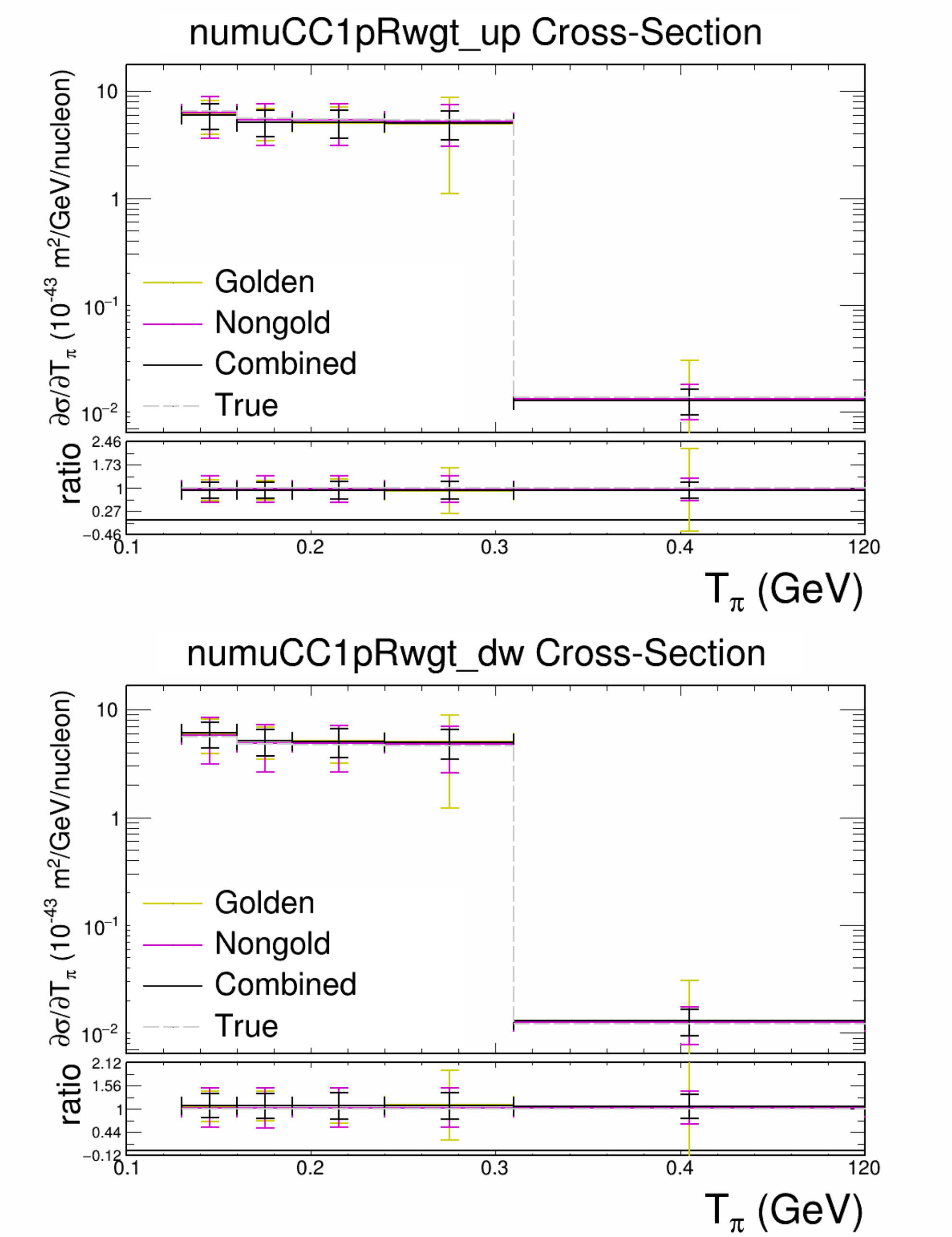

Figure A.5: Fake data study for enhancing/suppressing single proton events by 20%. Effects are minimal, demonstrating that GoldenBDT is able to handle this situation.

CURRICULUM VITAE

ERIN EWART

email eeewart.work@gmail.com

EDUCATION

Ph.D. Physics 2018-June 2026

Indiana University

Experimental high-energy particle physics, with a focus on neutrinos. Member of the DUNE and NOvA collaborations.

B.S. Physics, Math 2014-2018

Rice University

B.S. in physics and mathematics, concentration general physics. President's Honor Roll Fall '14, Fall '15.

WORK EXPERIENCE

Graduate RA 2019-2026

Indiana University, Part-time

Research work funded by the Department of Energy via Dr. Jon Urheim, on the DUNE / ProtoDUNE and NOvA experiments.

- Worked on a more sophisticated model of wireplane shadowing for ProtoDUNE
- Attempted to develop Michel electron identification using the photon detection system leading to improvements in the photon simulation framework

- DAQ and timing expert for NOvA for two years
- Developed a machine learning based vertexer for NOvA with much better accuracy than the previous vertexer, which will be the one used in future production campaigns
- Currently working on a Near Detector analysis in NOvA that aims to measure the cross section for ν_μ charged-current interactions that result in a single charged pion (plus anything else) in the final state, binned in pion energy

Lab Instructor 2018-2019

Indiana University, Part-time

Taught undergraduate lab sections, which involved explanations of the principles to be explored in the experiment of the week, answering student questions and aiding their understanding of the materials, providing office hours, and grading lab reports.

Undergraduate Researcher Summer 2017

Fermilab, Full-time

Worked with Dr. Minerba Betancourt on the feasibility of using electron scattering data to constrain the nuclear models used in the event generator GENIE for neutrino research. Worked with Dr. Richard Gran and Dr. Daniel Ruterbories on validating a version of GENIE for use by the MINERνA collaboration.

Grading Fall 2016

Rice University, Part-time

Graded homework for MATH 355 Linear Algebra. Approximately 3 hours per week.

Undergraduate Researcher Summer 2016

Rice University, Full-time

Worked with Dr. Kaden Hazzard on the many-bodied behavior of ultracold non-reactive molecules in an optical lattice, supported by a scholarship from the Physics and Astronomy Department at Rice University.

SELECTED PUBLICATIONS

Jeremy Wolcott (for the NOvA collaboration). *New neutrino oscillation results from NOvA with 10 years of data*. June 2024. DOI: `10.5281/zenodo.12704805`. URL: `https://doi.org/10.5281/zenodo.12704805`

B. Abi et al. “First results on ProtoDUNE-SP liquid argon time projection chamber performance from a beam test at the CERN Neutrino Platform”. In: *JINST* 15.12 (2020), P12004. DOI: `10.1088/1748-0221/15/12/P12004`. arXiv: `2007.06722 [physics.ins-det]`

Erin Ewart, Michael L. Wall, and Kaden R. A. Hazzard. “Bosonic molecules in a lattice: Unusual fluid phase from multichannel interactions”. In: *Phys. Rev. A* 98 (1 July 2018), p. 013611. DOI: `10.1103/PhysRevA.98.013611`. URL: `https://link.aps.org/doi/10.1103/PhysRevA.98.013611`

Erin Ewart and Minerba Betancourt. “Validating nuclear models: electron data to interpret neutrino physics”. In: *Fermilab Internal Documenatation* (Aug. 2017). DOI: `10.2172/1462232`

TALKS

NOvA Collaboration Meeting Summer 2024 July 2024

Minneapolis, Minnesota. Reconstruction & Deep Learning Plenary. *Alexander Booth (presenter), Wenjie Wu, Barbara Yaeggy, Abdul Wasit Yahaya (presenter), Erin Ewart (presenter)*

APS 2023 April Meeting April 2023

Minneapolis, Minnesota. Convolutional Visual Networks for Secondary Vertexing in NOvA. *Erin Ewart*

POSTER PRESENTATIONS

Neutrino 2024 June 2024

Milan, Italy. Recent Advancements in machine learning techniques utilised by NOvA. *Ashley Back, Alexander Booth (presenting), Erin Ewart, Alexander Shmakov, Alejandro Yankelevich*

Fermilab undergraduate poster session August 2017

Batavia, Illinois. Validating nuclear models for neutrino scattering using electron data. *Erin Ewart and Minerba Betancourt*

Workshop on Frontiers of Quantum Materials November 2016

Houston, Texas. Phase diagram of ultracold bosonic molecules in lattice with many-channel collisions. *Erin E. Ewart (presenter), Michael L. Wall, Kaden R. A. Hazzard*

APS DAMOP meeting May 2016

Providence, Rhode Island. Ultracold nonreactive molecules in an optical lattice: connecting chemistry to many-body physics. *Rick Mukherjee (presenter), Erin Ewart, Shah S. Alam,* *Michael L. Wall, Andris Doçaj, and Kaden R. A. Hazzard*

AWARDS

URA Travel Grant 2019

Fermilab Users Meeting

$800 travel grant to attend the summer 2019 Fermilab user's meeting.

OUTREACH

Indiana University Bloomington Science Fest 2018-2019, 2024

Bloomington, Indiana. Science demos for children and other community members to generate interest in physics.

ADDITIONAL PUBLICATIONS

All publications with authorship not listed under selected publications.

Adam Abed Abud et al. "Neutrino Interaction Vertex Reconstruction in DUNE with Pandora Deep Learning". In: (Feb. 2025). arXiv: `2502.06637 [hep-ex]`

M. A. Acero et al. "Measurement of the double-differential cross section of muon-neutrino charged-current interactions with low hadronic energy in the NOvA Near Detector". In: (Oct. 2024). arXiv: `2410.10222 [hep-ex]`

M. A. Acero et al. "Measurement of d2sigma/d—q—dEavail in charged current neutrino-nucleus interactions at <Ev> = 1.86 GeV using the NOvA Near Detector". In: (Oct. 2024).

arXiv: 2410.05526 [hep-ex]

Adam Abed Abud et al. "The track-length extension fitting algorithm for energy measurement of interacting particles in liquid argon TPCs and its performance with ProtoDUNE-SP data". In: *JINST* 20.02 (2025), P02021. DOI: 10.1088/1748-0221/20/02/P02021. arXiv: 2409.18288 [physics.ins-det]

M. A. Acero et al. "Dual-Baseline Search for Active-to-Sterile Neutrino Oscillations in NOvA". In: *Phys. Rev. Lett.* 134.8 (2025), p. 081804. DOI: 10.1103/PhysRevLett.134.081804. arXiv: 2409.04553 [hep-ex]

Adam Abed Abud et al. "DUNE Phase II: scientific opportunities, detector concepts, technological solutions". In: *JINST* 19.12 (2024), P12005. DOI: 10.1088/1748-0221/19/12/P12005. arXiv: 2408.12725 [physics.ins-det]

Adam Abed Abud et al. "First measurement of the total inelastic cross section of positively charged kaons on argon at energies between 5.0 and 7.5 GeV". In: *Phys. Rev. D* 110.9 (2024), p. 092011. DOI: 10.1103/PhysRevD.110.092011. arXiv: 2408.00582 [hep-ex]

Adam Abed Abud et al. "Supernova Pointing Capabilities of DUNE". In: (July 2024). arXiv: 2407.10339 [hep-ex]

M. A. Acero et al. "Search for CP-Violating Neutrino Nonstandard Interactions with the NOvA Experiment". In: *Phys. Rev. Lett.* 133.20 (2024), p. 201802. DOI: 10.1103/PhysRevLett.133.201802. arXiv: 2403.07266 [hep-ex]

Adam Abed Abud et al. "Performance of a Modular Ton-Scale Pixel-Readout Liquid Argon Time Projection Chamber". In: *Instruments* 8.3 (2024), p. 41. DOI: 10.3390/

instruments8030041. arXiv: 2403.03212 [physics.ins-det]

Adam Abed Abud et al. "Doping liquid argon with xenon in ProtoDUNE Single-Phase: effects on scintillation light". In: *JINST* 19.08 (2024), P08005. DOI: 10.1088/1748-0221/19/08/P08005. arXiv: 2402.01568 [physics.ins-det]

Adam Abed Abud et al. "The DUNE Far Detector Vertical Drift Technology. Technical Design Report". In: *JINST* 19.08 (2024), T08004. DOI: 10.1088/1748-0221/19/08/T08004. arXiv: 2312.03130 [hep-ex]

M. A. Acero et al. "Expanding neutrino oscillation parameter measurements in NOvA using a Bayesian approach". In: *Phys. Rev. D* 110.1 (2024), p. 012005. DOI: 10.1103/PhysRevD.110.012005. arXiv: 2311.07835 [hep-ex]

A. Abed Abud et al. "Impact of cross-section uncertainties on supernova neutrino spectral parameter fitting in the Deep Underground Neutrino Experiment". In: *Phys. Rev. D* 107.11 (2023), p. 112012. DOI: 10.1103/PhysRevD.107.112012. arXiv: 2303.17007 [hep-ex]

Adam Abed Abud et al. "Highly-parallelized simulation of a pixelated LArTPC on a GPU". In: *JINST* 18.04 (2023), P04034. DOI: 10.1088/1748-0221/18/04/P04034. arXiv: 2212.09807 [physics.comp-ph]

A. Abed Abud et al. "Identification and reconstruction of low-energy electrons in the ProtoDUNE-SP detector". In: *Phys. Rev. D* 107.9 (2023), p. 092012. DOI: 10.1103/PhysRevD.107.092012. arXiv: 2211.01166 [hep-ex]

Adam Abed Abud et al. "DUNE Offline Computing Conceptual Design Report". In:

(Oct. 2022). arXiv: 2210.15665 [physics.data-an]

M. A. Acero et al. “Monte Carlo method for constructing confidence intervals with unconstrained and constrained nuisance parameters in the NOvA experiment”. In: *JINST* 20.02 (2025), T02001. DOI: 10.1088/1748-0221/20/02/T02001. arXiv: 2207.14353 [hep-ex]

Adam Abed Abud et al. “Reconstruction of interactions in the ProtoDUNE-SP detector with Pandora”. In: *Eur. Phys. J. C* 83.7 (2023), p. 618. DOI: 10.1140/epjc/s10052-023-11733-2. arXiv: 2206.14521 [hep-ex]

M. A. Acero et al. “Measurement of the ν_e-Nucleus Charged-Current Double-Differential Cross Section at $\langle E_\nu \rangle = 2.4$ GeV using NOvA”. In: *Phys. Rev. Lett.* 130.5 (2023), p. 051802. DOI: 10.1103/PhysRevLett.130.051802. arXiv: 2206.10585 [hep-ex]

Adam Abed Abud et al. “Separation of track- and shower-like energy deposits in ProtoDUNE-SP using a convolutional neural network”. In: *Eur. Phys. J. C* 82.10 (2022), p. 903. DOI: 10.1140/epjc/s10052-022-10791-2. arXiv: 2203.17053 [physics.ins-det]

Adam Abed Abud et al. “Scintillation light detection in the 6-m drift-length ProtoDUNE Dual Phase liquid argon TPC”. In: *Eur. Phys. J. C* 82.7 (2022), p. 618. DOI: 10.1140/epjc/s10052-022-10549-w. arXiv: 2203.16134 [physics.ins-det]

A. Abed Abud et al. “Snowmass Neutrino Frontier: DUNE Physics Summary”. In: (Mar. 2022). arXiv: 2203.06100 [hep-ex]

A. Abed Abud et al. “A Gaseous Argon-Based Near Detector to Enhance the Physics

Capabilities of DUNE". In: (Mar. 2022). arXiv: 2203.06281 [hep-ex]

M. A. Acero et al. "Measurement of the double-differential muon-neutrino charged-current inclusive cross section in the NOvA near detector". In: *Phys. Rev. D* 107.5 (2023), p. 052011. DOI: 10.1103/PhysRevD.107.052011. arXiv: 2109.12220 [hep-ex]

A. Abud Abed et al. "Low exposure long-baseline neutrino oscillation sensitivity of the DUNE experiment". In: *Phys. Rev. D* 105.7 (2022), p. 072006. DOI: 10.1103/PhysRevD.105.072006. arXiv: 2109.01304 [hep-ex]

M. A. Acero et al. "Improved measurement of neutrino oscillation parameters by the NOvA experiment". In: *Phys. Rev. D* 106.3 (2022), p. 032004. DOI: 10.1103/PhysRevD.106.032004. arXiv: 2108.08219 [hep-ex]

A. Abed Abud et al. "Design, construction and operation of the ProtoDUNE-SP Liquid Argon TPC". In: *JINST* 17.01 (2022), P01005. DOI: 10.1088/1748-0221/17/01/P01005. arXiv: 2108.01902 [physics.ins-det]

A. Abed Abud et al. "Searching for solar KDAR with DUNE". In: *JCAP* 10 (2021), p. 065. DOI: 10.1088/1475-7516/2021/10/065. arXiv: 2107.09109 [hep-ex]

M. A. Acero et al. "Extended search for supernovalike neutrinos in NOvA coincident with LIGO/Virgo detections". In: *Phys. Rev. D* 104.6 (2021), p. 063024. DOI: 10.1103/PhysRevD.104.063024. arXiv: 2106.06035 [hep-ex]

M. A. Acero et al. "Search for Active-Sterile Antineutrino Mixing Using Neutral-Current Interactions with the NOvA Experiment". In: *Phys. Rev. Lett.* 127.20 (2021), p. 201801. DOI:

10.1103/PhysRevLett.127.201801. arXiv: 2106.04673 [hep-ex]

M. A. Acero et al. "Seasonal variation of multiple-muon cosmic ray air showers observed in the NOvA detector on the surface". In: *Phys. Rev. D* 104.1 (2021), p. 012014. DOI: 10.1103/PhysRevD.104.012014. arXiv: 2105.03848 [hep-ex]

V. Hewes et al. "Deep Underground Neutrino Experiment (DUNE) Near Detector Conceptual Design Report". In: *Instruments* 5.4 (2021), p. 31. DOI: 10.3390/instruments5040031. arXiv: 2103.13910 [physics.ins-det]

B. Abi et al. "Experiment Simulation Configurations Approximating DUNE TDR". In: (Mar. 2021). arXiv: 2103.04797 [hep-ex]

B. Abi et al. "Prospects for beyond the Standard Model physics searches at the Deep Underground Neutrino Experiment". In: *Eur. Phys. J. C* 81.4 (2021), p. 322. DOI: 10.1140/epjc/s10052-021-09007-w. arXiv: 2008.12769 [hep-ex]

B. Abi et al. "Supernova neutrino burst detection with the Deep Underground Neutrino Experiment". In: *Eur. Phys. J. C* 81.5 (2021), p. 423. DOI: 10.1140/epjc/s10052-021-09166-w. arXiv: 2008.06647 [hep-ex]

B. Abi et al. "Neutrino interaction classification with a convolutional neural network in the DUNE far detector". In: *Phys. Rev. D* 102.9 (2020), p. 092003. DOI: 10.1103/PhysRevD.102.092003. arXiv: 2006.15052 [physics.ins-det]

B. Abi et al. "Long-baseline neutrino oscillation physics potential of the DUNE experiment". In: *Eur. Phys. J. C* 80.10 (2020), p. 978. DOI: 10.1140/epjc/s10052-020-08456-z.

arXiv: 2006.16043 [hep-ex]

Babak Abi et al. "Deep Underground Neutrino Experiment (DUNE), Far Detector Technical Design Report, Volume III: DUNE Far Detector Technical Coordination". In: *JINST* 15.08 (2020), T08009. DOI: 10.1088/1748-0221/15/08/T08009. arXiv: 2002.03008 [physics.ins-det]

Babak Abi et al. "Deep Underground Neutrino Experiment (DUNE), Far Detector Technical Design Report, Volume I Introduction to DUNE". In: *JINST* 15.08 (2020), T08008. DOI: 10.1088/1748-0221/15/08/T08008. arXiv: 2002.02967 [physics.ins-det]

Babak Abi et al. "Deep Underground Neutrino Experiment (DUNE), Far Detector Technical Design Report, Volume II: DUNE Physics". In: (Feb. 2020). arXiv: 2002.03005 [hep-ex]

Babak Abi et al. "Deep Underground Neutrino Experiment (DUNE), Far Detector Technical Design Report, Volume IV: Far Detector Single-phase Technology". In: *JINST* 15.08 (2020), T08010. DOI: 10.1088/1748-0221/15/08/T08010. arXiv: 2002.03010 [physics.ins-det]