

UNIVERSITY OF CALIFORNIA,
IRVINE

Optimizing Neutrino Flavor Conversion Measurements through Machine Learning

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Physics

by

Alejandro Jaime Yankelevich

Dissertation Committee:
Professor Jianming Bian, Chair
Professor Henry Sobel
Professor Juan Pedro Ochoa-Ricoux

2026

TABLE OF CONTENTS

	Page
LIST OF FIGURES	v
LIST OF TABLES	ix
ACKNOWLEDGMENTS	xi
VITA	xii
ABSTRACT OF THE DISSERTATION	xiii
1 Introduction to Neutrino Physics	1
1.1 Neutrinos and the Standard Model	1
1.2 Sources of Neutrinos	2
1.2.1 Accelerator Neutrinos	2
1.2.2 Atmospheric Neutrinos	4
1.2.3 Solar Neutrinos	4
1.2.4 Reactor Neutrinos	6
1.2.5 Supernova Neutrinos	6
1.2.6 Astrophysical and Cosmogenic Neutrinos	7
1.2.7 Cosmic Neutrino Background	7
1.3 Neutrino Flavor Conversion	8
1.4 The MSW Effect for Solar Neutrinos	11
2 The NOvA Experiment	15
2.1 Neutrino Beam	15
2.2 Detector Design	18
2.3 Event Topologies	20
3 NOvA Event Reconstruction and Simulation	24
3.1 Simulation	28
3.2 Event and Particle Classification	28
3.3 TransformerCVN	31
3.3.1 Network Performance	33
3.3.2 Interpretability	41

4	The 3-Flavor Oscillation Analysis	47
4.1	Sensitivity to Oscillation Parameters	47
4.2	Event Selection	50
4.3	Energy Estimation	55
4.4	Systematic Uncertainties	59
4.4.1	Calibration	59
4.4.2	Detector Response	59
4.4.3	Lepton Reconstruction	60
4.4.4	Cross Sections	60
4.4.5	Beam Flux	60
4.4.6	Neutron Model	61
4.4.7	ND/FD Uncorrelated Errors	61
4.4.8	Near-to-Far Extrapolation	61
4.4.9	Error Summary	62
5	Bayesian Analysis Framework	64
5.1	Markov Chain Monte Carlo	65
6	Results with 10-Year Dataset	68
6.1	Mass Ordering and Octant	74
6.2	CP Violation	75
6.3	Comparisons to Other Experiments	78
7	Summary and Future Measurements	81
8	The Super-Kamiokande Experiment	84
8.1	Detector Design	84
8.2	Data Acquisition	89
8.3	Wideband Intelligent Trigger	91
8.4	Calibration Sources	92
9	Low Energy Event Reconstruction	95
10	The Solar Analysis	101
10.1	Sources of Background	101
10.2	Event Selection	103
10.3	Signal Extraction	104
10.3.1	Solar Distribution Shapes	104
10.3.2	Extended Likelihood Method	105
11	Extension to 3.49 MeV	107
11.1	Event Selection	107
11.1.1	Development of Machine Learning Methods	108
11.1.2	Dataset Generation	114
11.1.3	Final Selection	116
11.2	Signal Extraction	117

11.3	Systematic Uncertainties	120
11.3.1	Background Shape	120
11.3.2	Signal Extraction	121
11.3.3	Trigger Efficiency	122
11.3.4	BDT Input Variables	123
11.3.5	Energy-Related Uncertainties	125
11.3.6	Error Summary	127
12	Solar Energy Spectrum Results	128
12.1	Interaction Rate Measurement in 3.49 MeV	128
12.2	Survival Probability Fit	131
13	Summary and Future Sensitivity	133
	Bibliography	136

LIST OF FIGURES

	Page
1.1 Illustration of relativistic aberration	3
1.2 The solar fusion proton-proton chain with branching ratios	5
1.3 The solar neutrino flux from various processes as a function of energy	5
1.4 Graph illustrating the effective mass of the ν_e and $\bar{\nu}_e$ states as a function of electron density	12
1.5 ν_e survival probability as a function of neutrino energy	13
2.1 Baselines of the Fermilab long-baseline neutrino oscillation experiments	16
2.2 The Fermilab accelerator complex	17
2.3 The NuMI beam	17
2.4 The neutrino energy spectrum for several potential off-axis angles	18
2.5 NOvA plastic extrusions each with one wavelength shifting fiber	18
2.6 Schematic of NOvA detector and generation of top and side views from vertical and horizontal planes respectively	19
2.7 NOvA ND front view and back view	20
2.8 NOvA FD front view	20
2.9 Example NOvA event displays	21
2.10 Feynman diagrams illustrating QE, Res, and DIS	22
2.11 Content of interaction modes as a function of hadronic energy	23
3.1 Overview of the NOvA reconstruction chain.	25
3.2 An example event display of a full 550 μ s window	26
3.3 An example event display with reconstruction products shown	27
3.4 An example pixel map for a ν_e CC event with the XZ view and YZ view	29
3.5 The 2-view Siamese MobileNet architecture of EventCVN	30
3.6 The 4-view Siamese GoogLeNet architecture of ProngCVN	30
3.7 TransformerCVN network architecture	31
3.8 Event network output score distributions for ν_e and $\bar{\nu}_e$	35
3.9 Prong network output score distributions for ν_e and $\bar{\nu}_e$ prongs	35
3.10 TransformerCVN event confusion matrices	35
3.11 EventCVN event confusion matrices	36
3.12 TransformerCVN full prong confusion matrices	36
3.13 TransformerCVN 5-class prong confusion matrices	37
3.14 ProngCVN 5-class prong confusion matrices	37

3.15	Event classification confusion matrices for retrained TransformerCVN	39
3.16	Event output score distributions separated by true class for retrained TransformerCVN	39
3.17	Prong classification confusion matrices for retrained TransformerCVN	40
3.18	5-class prong classification confusion matrices for retrained TransformerCVN	40
3.19	The average attention score for each prong type for a given event or prong type	42
3.20	Saliency map for one example event	43
3.21	Grid of aggregated saliency maps and difference maps for every pair of prong types	44
3.22	Integrated salience map demonstrating the profile of different prong types	46
4.1	Time series showing the daily FHC neutrino beam or RHC antineutrino beam POT recorded by NOvA	48
4.2	Illustration of NOvA's sensitivities to θ_{12} and θ_{13} through the energy spectrum of the $\bar{\nu}_\mu$ disappearance channel using toy MC	48
4.3	Illustration of NOvA's sensitivities to δ_{CP} , the sign of θ_{13} (the mass ordering), and the octant of θ_{23} by comparing the rate of ν_μ vs $\bar{\nu}_\mu$ appearance between FHC and RHC modes through the $\bar{\nu}_\mu$ appearance channel	49
4.4	Cutflow chart illustrating the $\bar{\nu}_\mu$ selection	51
4.5	Cutflow chart illustrating the 2020 and 2024 $\bar{\nu}_\mu$ selections	52
4.6	The Core PID cut optimization projected onto 2D Core BDT vs EventCVN CC score	53
4.7	The Core PID cut optimization projected onto 2D EventCVN Cosmic score vs EventCVN CC score	54
4.8	The Peripheral PID cut optimization projected onto 2D EventCVN Cosmic score vs EventCVN CC score	54
4.9	Histogram of the weighted average true E_{ν} energy as a function of E_{EM} and E_{Had}	56
4.10	Histograms of the true quantity vs the reconstructed quantity for the primary muon and hadronic part in the $\bar{\nu}_\mu$ CC events.	57
4.11	The $\bar{\nu}_\mu$ selection spectra	57
4.12	The ν_μ selection spectra	58
4.13	The $\bar{\nu}_\mu$ selection spectra separated by hadronic energy quartile	58
4.14	Illustration of the extrapolation procedure	62
4.15	Summary of all errors on the integral number of $\bar{\nu}_\mu$ and ν_μ signal and background event	63
5.1	Illustration of MCMC	66
5.2	Comparison of posterior distributions of θ_{12} and θ_{13} vs δ_{CP} between Stan and ARIA MCMC methods	67
6.1	Illustration of the 1D and 2D Daya Bay reactor constraint.	70
6.2	Triangle plot of marginalized posterior probability densities	71
6.3	Triangle plot of marginalized posterior probability densities with the 1D Daya Bay constraint	72

6.4	Triangle plot of marginalized posterior probability densities with the 2D Daya Bay constraint	73
6.5	Marginalized posterior probability densities for from a fit to data with different options of the Daya Bay constraint	75
6.6	Posterior probability densities for θ_{CP}	76
6.7	Posterior probability densities for the Jarlskog invariant	77
6.8	NOvA 2024 results for vs θ_{CP} compared with NOvA 2020 results and results from other experiments	78
6.9	NOvA 2024 results for – compared with NOvA 2020 results and results from other experiments	79
6.10	Comparison of central values and 1σ intervals for absolute value of in the normal ordering and inverted ordering from various accelerator, reactor, atmospheric experiments, and the two joint fits	80
8.1	Diagram of Super-K	85
8.2	Pictures of Super-K ID and bottom of OD	85
8.3	Illustrations of Cherenkov cones in Super-K	86
8.4	Diagram of the Super-K water system	87
8.5	Schematic view of the water circulation paths in the Super-K tank	87
8.6	Illustration of the Super-K DAQ since SK-IV from the PMT readout to processing for long-term storage through both the software trigger and WIT system	90
8.7	The time variation of the SLE trigger efficiency estimated by the MC simulation of solar neutrinos	90
8.8	The trigger efficiency of the WIT system as a function of reconstructed electron kinetic energy	92
8.9	Schematic of the LINAC geometry feeding into the detector	93
8.10	An overview of DT data taking	93
9.1	The PDF of the timing residuals based on LINAC events used for the vertex reconstruction maximum likelihood fit	96
9.2	The distribution of the opening angle between the direction of the MC electron and the vector from the reconstructed vertex to each hit-PMT's position	97
9.3	The vertex, direction, and energy resolution in SK-IV as a function of the true electron kinetic energy	98
9.4	Illustration of the MSG calculation	100
10.1	SK-IV vertex distribution in 3.49 MeV - 3.99 MeV	102
10.2	Solar angle distribution for the full energy range of 3.49–19.49 MeV using the SK-IV dataset	106
11.1	An example 3.85 MeV event display	108
11.2	An example graph structure	110
11.3	A 2D polynomial fit to the background vertex distribution	111
11.4	Distributions of the most important reconstructed variable inputs to the BDT in the WIT dataset for signal and background following preselection cuts	112
11.5	Architecture of the hybrid network	113

11.6	“Rejection” ROC curve showing the background rejection as the inverse of the false positive rate (FPR) vs the true positive rate (TPR, or signal efficiency) for the WIT 2.49 MeV - 3.49 MeV testing dataset	113
11.7	Illustration of the “triple-fit” BDT training procedure	114
11.8	Rejection ROC curve showing the effect of the “triple-fit” procedure for training BDTs	115
11.9	Solar angle distributions for the WIT dataset in all MSG bins for data selection, calculated background shape, and calculated signal shape added to background shape all with statistical error bands	118
11.10	Solar angle distributions for the SLE dataset in all MSG bins separated by trigger threshold period	118
11.11	Solar angle distributions aggregated over MSG bins for the WIT dataset . .	119
11.12	Solar angle distributions aggregated over MSG bins for the SLE datasets in 2.99 MeV - 3.49 MeV separated by trigger threshold period	119
11.13	Comparison of background shape calculation methods for systematic error calculation in each MSG bin for WIT and SLE 34-hit period in 2.99 MeV - 3.49 MeV	121
11.14	Ni calibration positions and their corresponding regions in the TFV	122
12.1	Solar angle distribution in 2.99 MeV - 3.49 MeV aggregated over the WIT and SLE 34-hit datasets	130
12.2	The measured electron kinetic energy spectrum in SK-IV with the addition of the lowest energy bin	130
12.3	survival probability spectra as a function of neutrino energy obtained from the exponential (left) and quadratic (right) fits to SK recoil electron data . .	132
13.1	Predicted confidence level contours for and	135

LIST OF TABLES

	Page
3.1 Aggregated metrics accuracy, precision, and area under a standard ROC curve (ROC AUC) for event classification and prong classification	34
3.2 ROC AUC values for individual labels in a one-vs-rest manner for either event classification or prong classification	34
3.3 ROC AUC values for individual labels in a one-vs-rest manner for either event classification or prong classification on the production 6 dataset	38
4.1 Observed and predicted numbers of ν_e , $\bar{\nu}_e$ and LowE ν_μ events in the neutrino beam, and $\bar{\nu}_\mu$ and $\bar{\nu}_\tau$ events in the antineutrino beam	55
4.2 Systematic and statistical uncertainties on the oscillation parameters θ_{12} , θ_{13} , and δ	63
6.1 Summary of the oscillation parameters measured by NOvA for various reactor constraint options	69
6.2 Bayes factors, with percentage posterior probability preferences, illustrating the preference for the normal ordering over the inverted ordering and for the upper octant over the lower octant	74
8.1 List of Super-K operations phases, their duration, number of ID PMTs, photocoverage of those PMTs, and the concentration of gadolinium dissolved in % by weight	89
11.1 Variables input to the BDT	110
11.2 Livetimes of the datasets used in this selection	117
11.3 Trigger efficiency of the online WIT cuts for solar MC	123
11.4 MSG systematic error in signal MC efficiency for the WIT dataset using plus and minus shifts under various methods	124
11.5 Correlated energy error table extended to 2.49 MeV	125
11.6 Energy scale error in signal MC efficiency for all datasets using plus and minus shifts under various methods	126
11.7 Systematic error summary with significant error sources	127
12.1 Expected number of unoscillated MC signal events, fit number of signal events, and their ratio in each energy and MSG bin	129
12.2 Resulting coefficients a , b , and c comparisons for the quadratic parameterization fit. .	131

12.3	Resulting coefficients and comparisons for the exponential parameterization	
	fit.	132

ACKNOWLEDGMENTS

I would like to thank my PhD advisers Jianming Bian and Henry Sobel for their guidance and support throughout my PhD research and for helping me build the skills necessary for a career in particle physics. I would also like to thank the other members of both my Advancement and Dissertation committee: Juan Pedro Ochoa-Ricoux, Pierre Baldi, and Jonathan Feng. I greatly appreciate Michael Smy's extensive guidance on the Super-K experiment and the framework for the solar analysis. Thank you to Bruce Cortez for all his work on helping to optimize the solar selection BDT. I would like to thank Jeff Griskevich for all his support on DUNE purity monitor and cold electronics projects, all the valuable hardware training necessary for this work, and for maintaining our lab in the Physical Sciences High Bay. Thank you to all the previous students and postdoctoral researchers who helped me get started with working on the NOvA, DUNE, and Super-K experiments: Nitish Nayak, Scott Locke, Wenjie Wu, and Benjamin Jargowsky, as well as all of the current members for their help and collaboration: Yiwen Xiao, Zhenyang Chen, Larry Zhao, and Jiaxi Liu. Thank you to everyone in Pierre Baldi's group in the computer science department including Alexander Shmakov, Junze Liu, Edgar Robles, Dikshant Sagar, and Kaiwen Yu who developed tools for NOvA and DUNE neutrino event reconstruction using state-of-the-art technology in machine learning. Thank you to the Water Cherenkov Machine Learning Group for their help in researching machine learning tools for the Super-K solar analysis. Thank you to all the undergraduate students in our group who helped us develop tutorials and train other students in machine learning tools and especially for their work to help test the many thousand DUNE ASIC chips. I would like to thank all the leaders, conveners, and analyzers in the NOvA reconstruction and 3-flavor groups including Alex Himmel, Erika Catano-Mur, Zoya Vallari, Liudmila Kolupaeva, Artur Sztuc, Ashley Back, Erica Smith, Alex Booth, and Barbara Yaeggy, the Super-K low energy group including Yasuo Takeuchi and Yuuki Nakano, and DUNE cold electronics consortium including David Christian, Cheng-Ju Lin, and Roger Huang for their dedication to their roles in these analyses and all their advice for improving the work in my PhD research. I would especially like to thank my family for their support throughout the PhD program.

The text and figures of Chapter 6 and portions of Chapter 4 of this dissertation are adaptations of the material as it appears in Physics Review Letters 136:011802, used with permission from the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license.

This work was supported by U.S. Department of Energy Office of Science grant DE-SC0009920 as well as the Office of Science Graduate Student Research (SCGSR) program under contract number DE-SC0014664. This work was supported by the U.S. National Science Foundation award 2013073.

VITA

Alejandro Jaime Yankelevich

EDUCATION

Doctor of Philosophy in Physics

2026

University of California, Irvine

Irvine, California

Master of Science in Physics

2022

University of California, Irvine

Irvine, California

Bachelor of Science in Physics

2019

California Institute of Technology

Pasadena, California

RESEARCH EXPERIENCE

Graduate Student Researcher

2021–2026

University of California, Irvine

Irvine, California

Graduate Student Research Fellowship

2023–2024

Fermi National Accelerator Laboratory

Batavia, Illinois

TEACHING EXPERIENCE

Teaching Assistant

2020

University of California, Irvine

Irvine, California

ABSTRACT OF THE DISSERTATION

Optimizing Neutrino Flavor Conversion Measurements through Machine Learning

Alejandro Jaime Yankelevich

Doctor of Philosophy in Physics

University of California, Irvine, 2026

The phenomenon of neutrino flavor conversion whereby the flavor of a neutrino particle can change between its time of production and later detection was the first definitive evidence of physics beyond the Standard Model. Some of the oscillation parameters used to describe this conversion are not yet well measured, leaving important questions still open regarding flavor conversion both in vacuum and as neutrinos travel through matter. NOvA is a long-baseline neutrino oscillation experiment that uses Fermilab’s predominantly ν_μ NuMI beam. A 14 kton oil-based liquid scintillator far detector 810 km away is used to measure neutrino oscillation through the $\nu_\mu \rightarrow \nu_\tau$ disappearance and $\nu_\mu \rightarrow \nu_e$ appearance channels. Super-K is a 50 kton water Cherenkov detector, which measures the disappearance of ν_μ produced during solar fusion. The high density environment of the sun decreases the ν_μ survival probability at higher energies observable in Super-K compared to the vacuum-dominated oscillations at lower energies. However, the transition region is overshadowed by radioactive background in the detector. In both of these experiments, the separation of neutrino detection events from background and classification of neutrino flavor are crucial tasks that benefit from the introduction of machine learning. Chapter 1 gives an overview of neutrinos and the context under which flavor conversion is measured in this dissertation. Chapters 2–7 present the results of a Bayesian sampling approach for the latest NOvA 3-flavor oscillation analysis with ν_μ protons on target in neutrino mode and $\bar{\nu}_\mu$ in antineutrino mode collected over 10 years. Chapters 8–13 present the results of extending the Super-K solar analysis to lower energies during its fourth phase with 2970 days of livetime.

Chapter 1

Introduction to Neutrino Physics

1.1 Neutrinos and the Standard Model

The Standard Model of particle physics is currently the best description of the fundamental particles of nature. With the notable exception of gravity, the model describes fermions, which make up the matter in the universe, and their interactions mediated by force-carrying bosons. The massless photon (γ) mediates the electromagnetic force, the massless gluon (g) mediates the strong force, and the massive $W^\pm$ and Z^0 bosons mediate the weak force. Since the $W^\pm$ is charged while the Z^0 is neutral, interactions involving an exchange of a $W^\pm$ is called the charged current (CC) and interaction involving the Z^0 is called the neutral current (NC). Among the fermions there are six quarks with charge and color; three charged leptons: the electron (e), muon (μ), and tau (τ); and three neutrinos with flavors corresponding to the charged leptons: electron neutrino (ν_e), muon neutrino (ν_μ), and tau neutrino (ν_τ). Since neutrinos have neither charge nor color, they cannot interact via the electromagnetic or strong force and can therefore only interact via the weak force. This property, along with their very small mass, results in very small cross sections for their interactions with other particles except at very high energies. This makes their detection much more difficult

compared to other particles in the Standard Model.

The neutrino was first postulated by Wolfgang Pauli in 1930 to solve the energy conservation problem in beta decay theorizing that missing energy is carried away by a light, neutral particle [1]. The neutrino was later discovered by Clyde Cowan and Frederick Reines in 1956 through the “Project Poltergeist” experiment in which a detector with 200 L of water with dissolved cadmium chloride sandwiched between tanks of liquid scintillator detected neutrinos from a nuclear reactor at the Savannah River Plant [2] (see Section 1.2.4). For most of the history of the Standard Model, neutrinos were assumed to be massless.

1.2 Sources of Neutrinos

Neutrinos play a fundamental role in nuclear reactions due to the law of lepton number conservation. Lepton number refers to the number of leptons corresponding to each flavor before and after a reaction: L_e . Both e^- and ν_e have electron number $L_e = 1$, and their anti-particles, e^+ and $\bar{\nu}_e$, have electron number $L_e = -1$. This is clearly seen in beta decay, where a nucleon transforms and must emit e^- to conserve charge as well as $\bar{\nu}_e$ to conserve electron number. The total lepton number $L = L_e + L_\mu + L_\tau$ must also be conserved. Flavor-specific lepton numbers have only been found to be violated in neutrino flavor conversion (Section 1.3), but there is currently no evidence for total lepton number violation.

1.2.1 Accelerator Neutrinos

Many major neutrino experiments, particularly long-baseline neutrino oscillation experiments, are designed around a neutrino beam. The technology for the neutrino beam originates from the Brookhaven Alternating Gradient Synchrotron, which was used to discover the muon neutrino in [3]. Long baseline neutrino oscillation experiments include MINOS and K2K in the past, NOvA and T2K in the present, and DUNE and Hyper-K in the future. The starting point for neutrino beams is proton accelerators. Protons are accelerated

to 10 GeV – 100 GeV before colliding with a graphite target, producing hadrons such as pions and kaons. Focusing horns then steer the hadrons in the desired direction. The polarity of the horn current determines the direction of the magnetic field, which in turn determines whether positive hadrons or negative hadrons are steered in the desired direction. The pions then decay via $\pi^+ \rightarrow \mu^+ \nu_\mu$ in the case of forward horn current (FHC) or $\pi^- \rightarrow \mu^- \bar{\nu}_\mu$ for reverse horn current (RHC) with a branching ratio of $> 99\%$. While pions make up $\sim 20\%$ of the hadrons depending on energy, the remaining are primarily kaons. Kaons also primarily decay to $\mu \bar{\nu}$, but can also decay via $K^0 \rightarrow \pi^+ \pi^-$ with a branching ratio of $\sim 1\%$. Therefore, neutrino beams are predominately $\mu \bar{\nu}$ but with a contamination of $\pi^+ \pi^-$.

In the frame of rest of the decaying hadron, the neutrinos are emitted isotropically. However, due to relativistic aberration, the neutrino beam becomes a cone in the lab frame (Figure 1.1). For pion frame-of-rest angle relative to the beam θ^* , lab frame angle θ , and $\beta = v/c$,

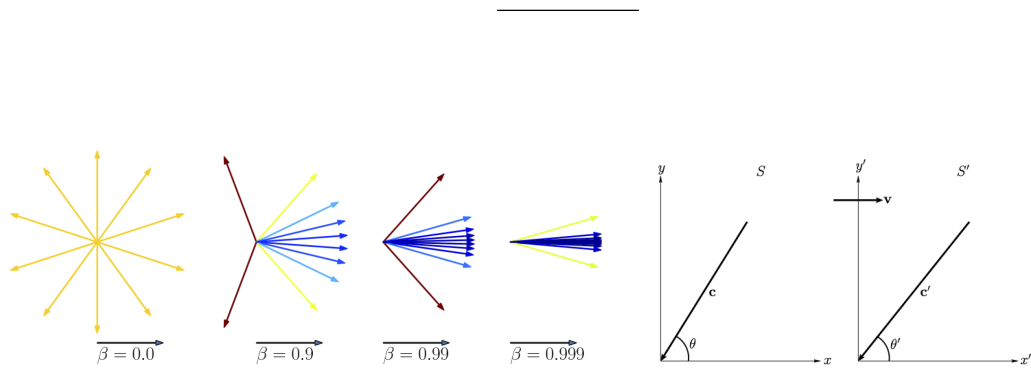


Figure 1.1: Illustration of relativistic aberration [4, 5].

The degree of aberration is dependent on the pion energy such that neutrinos from higher energy pions are more concentrated in the forward direction. This effect makes it possible to select neutrinos at desired energies by placing detectors at a certain “off-axis” angle away from the center of the beam at the cost of event rate. The distribution of neutrino energy

at angle as a function of pion energy is

1.2.2 Atmospheric Neutrinos

Atmospheric neutrinos are produced in a similar but natural process. Cosmic rays, such as protons and heavy ions, collide with the atmosphere and create large hadronic showers spanning several kilometers. The pions and kaons in these showers similarly decay to produce μ^+ or μ^- , but without human-controlled magnetic fields to separate signs. The muons then decay to e^+ or e^- . Together, these decay processes produce approximately twice as many ν_μ than $\bar{\nu}_\mu$.

1.2.3 Solar Neutrinos

The sun is the largest source of neutrinos on Earth by sheer number. Neutrinos are produced during fusion in the core of the sun. Since fusion fundamentally involves the production of neutrons in order to support the nuclei of heavier elements requiring either positron emission or electron capture, fusion produces ν_e to balance electron number. Figure 1.2 shows the proton-proton chain of fusion reactions in the sun in which hydrogen is fused into helium. The CNO cycle can also fuse hydrogen into helium but is nondominant in the sun due to its low mass. These processes result in the energy spectrum shown in Figure 1.3.

Neutrino detectors with sizable solar neutrino datasets include Super-Kamiokande, SNO, and Borexino. As water-Cherenkov detectors, Super-Kamiokande and SNO can only observe neutrinos from the β^+ step in the chain, which produce electrons above Cherenkov threshold. *hep* neutrinos have not yet been observed due to their low flux. As a liquid scintillator detector, Borexino has measured the flux from various solar neutrino sources including *pp*, $\bar{\nu}_e$, and *pep* in addition to ν_e .

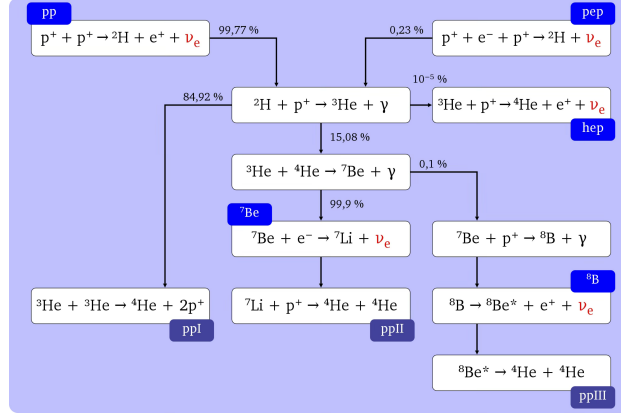


Figure 1.2: The solar fusion proton-proton chain with branching ratios. The blue and purple boxes show the name of the process. Reactions emitting neutrinos are highlighted in red [6].

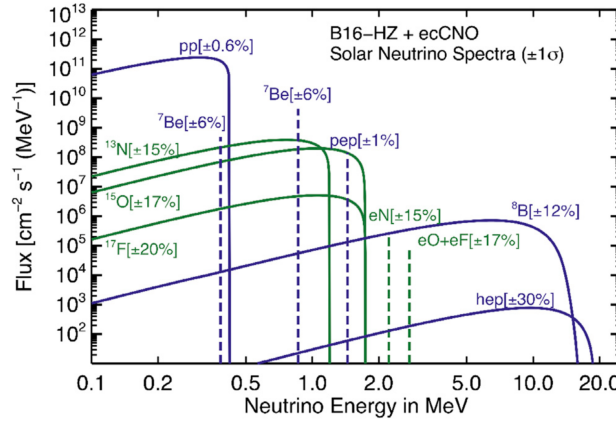


Figure 1.3: The solar neutrino flux from various processes as a function of energy. Proton-proton chain sources are in blue while CNO cycle sources are in green [7].

An important interaction method for solar neutrinos is elastic scattering on electrons.

This interaction is open to all neutrino flavors, but $\bar{\nu}_e$ has a greater cross section. Since elastic scattering produces a recoil electron whose direction preserves the direction of the incoming neutrino, the solar direction can be used to identify solar neutrino events.

1.2.4 Reactor Neutrinos

Conversely, as neutrons in heavy nuclei undergo beta decay during fission, e^- must be produced to balance electron number: $n \rightarrow p + e^- + \bar{\nu}_e$. These $\bar{\nu}_e$ can be found from nuclear reactors on Earth and are also the easiest of all flavors to detect at the ~ 1 MeV energies of reactor neutrinos due to the “inverse beta decay” (IBD) interaction channel available only to $\bar{\nu}_e$.

–

The neutron is then captured by an atom, e.g. hydrogen for water-based detectors, and the excited nucleus then emits one or more photons as it deexcites. This delayed coincidence signature of a positron followed by a burst of photons $\sim 100 \mu\text{s}$ later makes it easy to separate IBD events from background such as decay of radioactive impurities in the detector. Reactor neutrinos were therefore the first source of neutrinos to be discovered, and there have been many reactor neutrino experiments in the recent past including Daya Bay, KamLAND, Double Chooz, PROSPECT, and RENO, culminating in JUNO, the largest operating today.

A small flux of “geoneutrinos” from the decay of radioactive elements within the Earth is also detectable.

1.2.5 Supernova Neutrinos

Core-collapse supernovae emit an astronomical number of neutrinos ($\sim 10^{58}$) in ~ 10 s. Fusion of elements lighter than iron is energetically favorable. Once these elements run out, the iron core collapses under its gravity. If the mass of this core is large enough, the electron degeneracy pressure cannot maintain the structure either and “neutronization” occurs in which protons and electrons combine to form neutrons and $\bar{\nu}_e$ just as in fusion but at a much larger rate. The neutrinos from this burst then collide with the falling outer layers of the star and ejects them, creating neutrinos of all flavors that carry away $> 99\%$ of the supernova’s

energy.

The only supernova burst ever to be detected was the 1987A supernova. Three neutrino experiments detected neutrinos from this supernova: 11 by Kamiokande [8], 8 by IMB [9], and 5 by Baksan [10], all within the same 13 s interval with energies 7.5 MeV – 40 MeV. If a core-collapse supernova were to occur in the galactic center today, Super-Kamiokande is expected to detect events, 95% being IBD events and 5% being elastic scattering events for neutrinos of all flavors pointing to the supernova with 3° – 7° total accuracy [11].

Only a supernova within the Milky Way galaxy or neighboring Large and Small Magellanic Clouds would create an intense enough flux to detect a neutrino burst, but a supernova occurs within the universe at a rate of 1 Hz. This creates a constant background known as the Diffuse Supernova Neutrino Background (DSNB), and should result in a few detections per year. The DSNB has not yet been discovered beyond a slight indication [12].

1.2.6 Astrophysical and Cosmogenic Neutrinos

Astrophysical sources also produce ultra-high energy neutrinos with energies of 10^{13} eV – 10^{15} eV produced in active galactic nuclei and gamma ray bursts [13]. Neutrino experiments searching for these ultra-high energy neutrinos include IceCube and KM3NeT. “Cosmogenic” neutrinos with energies $>10^{17}$ eV are also predicted from the interaction of cosmic rays with the cosmic microwave background photons but have not yet been discovered beyond one such potential event by KM3NeT [14].

1.2.7 Cosmic Neutrino Background

The cosmic neutrino background consisting of neutrinos decoupled from matter one second after the Big Bang is also postulated. These neutrinos would have much lower energy than previously mentioned sources at 10^{-6} eV – 10^{-4} eV.

1.3 Neutrino Flavor Conversion

Measurements of the solar neutrino interaction rate by the Homestake experiment disagreed with the flux predicted by the Standard Solar Model (SSM). Homestake used perchloroethylene to observe solar ν_e CC interactions on chlorine, and the observed rate disagreed with the SSM by a factor of one third [15]. This discrepancy was exacerbated by later experiments such as Kamiokande, which observed solar neutrinos through elastic scattering on electrons. Since this channel has a small but nonzero component from ν_μ and ν_τ , the observed elastic scattering interaction rate disagreed by a factor of one half with the SSM. The directionality of elastic scattering also confirmed that the observed neutrinos came from the sun. This discrepancy was called the “solar neutrino problem”. A similar “atmospheric neutrino ” problem developed between Kamiokande’s and IMB’s measurements of atmospheric ν_μ and ν_τ and their predicted ratio.

The problem was eventually resolved through the discovery of neutrino oscillations. In 1998, Super-Kamiokande’s analysis of atmospheric neutrinos, predominantly ν_μ , revealed a disappearance in ν_μ and an appearance in ν_τ with these rates depending on both event energy and direction [16]. This dataset revealed an disappearance pattern in which, for a given energy, the probability of detecting a ν_μ is dependent on the distance it traveled through the Earth between its production point in the atmosphere and the detector. SNO, with its capability of detecting the *total* rate of all three flavors of solar neutrinos through NC interactions with deuterium in its heavy water medium, directly solved the problem in 2001 by showing that the total solar neutrino flux agreed with the SSM [17]. These flavor conversions are only possible through the following theoretical framework if neutrinos have nonzero mass.

If neutrinos have zero mass, the flavor ν_e , ν_μ , ν_τ eigenstates, which dictate the states through which they interact via the weak force, cannot change with time. However, if neutrinos have nonzero mass, the addition of their mass terms to the Standard Model Lagrangian

may create different eigenstates under which they propagate through free space, the “mass eigenstates” ν_1, ν_2, ν_3 . The mass eigenstates can be written as a linear combination of flavor eigenstates. In the case that neutrinos only have Dirac mass, meaning that their mass originates through interaction with the Higgs field like other Standard Model particles, this linear combination can be parametrized with 4 parameters through the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix.

The parameters $\theta_{12}, \theta_{13}, \theta_{23}$ are the three mixing angles, for which $\theta_{12} \approx 33^\circ$ and $\theta_{23} \approx 45^\circ$. The angle δ_{CP} allows for the possibility of different mixings between the neutrinos and anti-neutrinos. If $\delta_{\text{CP}} = 0$ or $\delta_{\text{CP}} = \pi$, charge-parity (CP) is conserved for neutrinos.

For example, if a $\bar{\nu}_\mu$ is produced, it will propagate as a superposition of

$$\bar{\nu}_1, \bar{\nu}_2, \bar{\nu}_3.$$

Each of the mass eigenstates will then evolve as $e^{-iE_i t}$. For neutrino energy E ,

————— ———— ———— if the masses of the mass states are small. Since the masses

of the three mass states are different, the oscillation frequencies are different for the three states with the same neutrino energy. Assuming that the neutrinos travel near the speed of light, the probability of detecting a neutrino with flavor ν_α that originated as flavor ν_β after traveling a distance L is

where θ_{12} if $\theta_{12} \neq 0$, 0 if $\theta_{12} = 0$ (no relation to θ_{CP}). These *differences* in the neutrino masses, m_1, m_2, m_3 then enter in as oscillation parameters by determining the frequency of the oscillations. Therefore, neutrinos must not only have mass, but the mass eigenstates must also have different masses.

The goal of modern-day neutrino experiments is to precisely measure these 6 fundamental oscillation parameters $\theta_{12}, \theta_{13}, \theta_{23}, \delta_{CP}, \theta_{12}, \theta_{13}$, and θ_{CP} . (Note that $\theta_{12}, \theta_{13}, \theta_{23}$ are in degrees).

High energy accelerator or atmospheric analyses can measure θ_{12} and θ_{23} through their disappearance at small L/E and have some sensitivity to θ_{13} through $\nu_\mu \rightarrow \nu_\tau$ appearance. Measurements of θ_{12} and θ_{23} primarily come from low energy solar and reactor experiments through ν_e or $\bar{\nu}_e$ disappearance at large L/E . Reactor experiments provide greater sensitivity for all of $\theta_{12}, \theta_{13}, \theta_{23}$, and soon δ_{CP} due to their high event rates and ease of detecting IBD. The comparison of ν_μ appearance vs $\bar{\nu}_\mu$ appearance in accelerator or atmospheric analyses provides information on the neutrino mass ordering, θ_{CP} , and the octant of θ_{23} .

According to the Particle Data Group published in May 2025 [18], $\Delta m_{21}^2 = 7.5 \times 10^{-5} \text{ eV}^2$ and is confirmed to be positive as explained in Section 1.4. The sign of Δm_{31}^2 is not yet known. This unresolved question is known as the “neutrino mass ordering” problem or the “neutrino mass hierarchy” problem. A positive Δm_{31}^2 indicates that $m_1 < m_2 < m_3$ known as the “normal mass ordering” (NO), while a negative value indicates $m_1 > m_2 > m_3$ known as the “inverted mass ordering” (IO). Δm_{31}^2 is measured to be $2.5 \times 10^{-3} \text{ eV}^2$

assuming NO or 10^{-3} eV^2 assuming IO.

For the mixing angles, θ_{12} implies θ_{13} 33° is large. θ_{13} implies θ_{12} 8.6° is small. (NO) or (IO) implies θ_{13} 47° is close to the 45° value corresponding to maximal mixing. However, due to the uncertainties, it is not yet known whether θ_{13} 45° (the “upper octant”, UO), whether θ_{13} 45° (the “lower octant”, LO), or whether it is exactly 45° .

δ_{CP} is the least precisely measured oscillation parameter, despite its large implications for CP conservation or violation for neutrinos. Leading explanations regarding why normal matter far outnumbers anti-matter in the universe rely on CP violation in neutrinos [19, 20]. However, the *magnitude* of CP violation is dependent on all the neutrino mixing angles, rather than δ_{CP} alone, and is dependent on the parametrization of the PMNS matrix. The Jarlskog invariant J_{CP} δ_{CP} is independent of this parameterization and therefore more useful in determining non-conservation of the CP symmetry, with only J_{CP} representing CP conservation.

1.4 The MSW Effect for Solar Neutrinos

In addition to oscillations in vacuum, neutrino flavor conversion can be influenced by their passage through matter, somewhat analogous to the dependence of a photon’s frequency on its propagation medium’s density. This introduces a difference in the oscillations of neutrinos vs anti-neutrinos in addition to the potential effect of δ_{CP} due to the dominance of ordinary matter vs anti-matter. In fact, the driver of the solar neutrino problem is not vacuum oscillations during their travel to Earth but rather the Mikheyev–Smirnov–Wolfenstein (MSW) effect taking place within the sun [21, 22]. This effect is possible because, at these energies, only ν_e may interact with electrons in the sun through CC while all flavors can interact through NC. These CC interactions with electrons alter the ν_e state and create effective

mass states ν_1 and ν_2 under the two-flavor approximation. The masses of these states are

for ν_1 where G_F is the Fermi constant, E is the neutrino energy, and N_e is the electron density. At high enough densities such as at the core of the sun, ν_1 and ν_2 for small E . While ν_1 has the largest ν_e content and ν_2 has the largest ν_μ content in vacuum, ν_1 and ν_2 . As ν_1 travel out of the sun toward lower densities, a resonant mass density region with $N_e = N_{res}$ is reached, and ν_1 is at its minimum. As they traverse this resonant region and continue outward through decreasing density, all of the ν_1 undergo a gradual adiabatic conversion to the ν_2 mass eigenstate and remain in that eigenstate during their propagation to Earth. The location of this resonant layer is energy-dependent and absent for the lowest energy solar neutrinos. This effect is only possible if $\nu_1 \neq \nu_2$ in vacuum. If $\nu_1 = \nu_2$, ν_1 and ν_2 would always have the largest ν_e content at all densities, so this level crossing and adiabatic conversion would not occur.

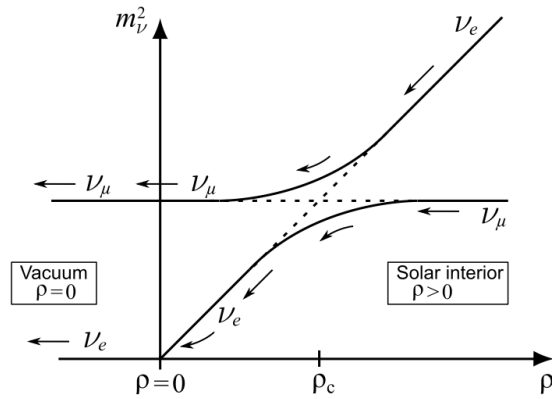


Figure 1.4: Graph illustrating the effective mass of the ν_1 and ν_2 states as a function of electron density. The dashed lines represent the unmixed flavor states (ν_μ and ν_e), which cross at the resonant density N_{res} while the solid lines represent the true effective mass eigenstates, which avoid crossing [23].

Measurements from Super-Kamiokande, SNO, and KamLAND of the ν_e solar neutrino inter-

action rate and corresponding survival probability (P_{ee}) are consistent with the predictions of the MSW effect, while measurements from Borexino and KamLAND of lower energy pp , ${}^7\text{Be}$, and pep neutrinos are consistent with the usual vacuum oscillations [24, 25]. This observation of the MSW effect confirms that θ_{12} is large. An “upturn” in P_{ee} is expected to occur between 1 MeV and 7 MeV in neutrino energy as flavor conversions transition from the MSW-dominant region with θ_{12} at higher energies to the vacuum oscillation-dominant region with θ_{12} at lower energies [26]. A measurement of P_{ee} in the transition region would be an important test of the MSW effect and electroweak physics in general. However, high rates of radioactive background radiation in this transition region have prevented observation of the solar upturn. Super-Kamiokande’s latest global solar analysis provides a 2σ significance for the presence of an upturn [27].

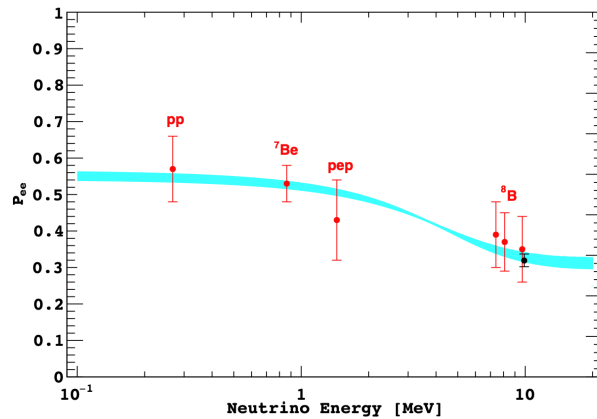


Figure 1.5: P_{ee} survival probability as a function of neutrino energy. The points represent, from left to right, the Borexino pp , ${}^7\text{Be}$, pep , and ${}^8\text{B}$ data (red points) and the SNO+SK data (black point). The blue band represents the MSW solution [18].

Matter effects also play a smaller but measurable effect as neutrinos propagate through the Earth and are an important component for the probability of ν_e or $\bar{\nu}_e$ appearance in accelerator and atmospheric analyses. Such Earth matter effects also drive a day/night asymmetry for solar neutrinos in that neutrino events arriving in a detector at night must first pass through the Earth. These matter effects increase the ν_e content in solar neutrinos at night,

but at higher energies starting from energies greater than 10 MeV . Super-Kamiokande currently finds the significance of the day/night effect assuming $7.50 \times 10^{-5}\text{ eV}^2$ [27].

Chapter 2

The NOvA Experiment

NOvA (NuMI off-Axis $\nu_\mu \rightarrow \nu_e$ appearance) is a long baseline neutrino oscillation experiment based at Fermi National Accelerator Laboratory (Fermilab) in Batavia, Illinois and uses Fermilab’s “Neutrinos from the Main Injector” (NuMI) neutrino beam [28]. A small near detector (ND) is placed 1 km away from the beam target in Fermilab, and a large far detector (FD) is placed 810 km from the target in Ash River, Minnesota 14.6 mrad off axis. As a long baseline experiment, NOvA aims to measure the oscillation parameters through the $\nu_\mu \rightarrow \nu_\tau$ disappearance and $\nu_\mu \rightarrow \nu_e$ appearance channels. NOvA began collecting data in 2014 and is scheduled to complete operations in 2027. It is a successor experiment to MINOS (main injector neutrino oscillation search), which collected data between 2005 and 2016 using the same NuMI beam but with on-axis detectors (Figure 2.1).

2.1 Neutrino Beam

The NuMI beam makes use of the main injector portion of the Fermilab accelerator complex, which used to feed protons to the Tevatron proton collider during its operation from 1983-2011 (Figure 2.2). Protons are accelerated in several stages starting with the Linac at

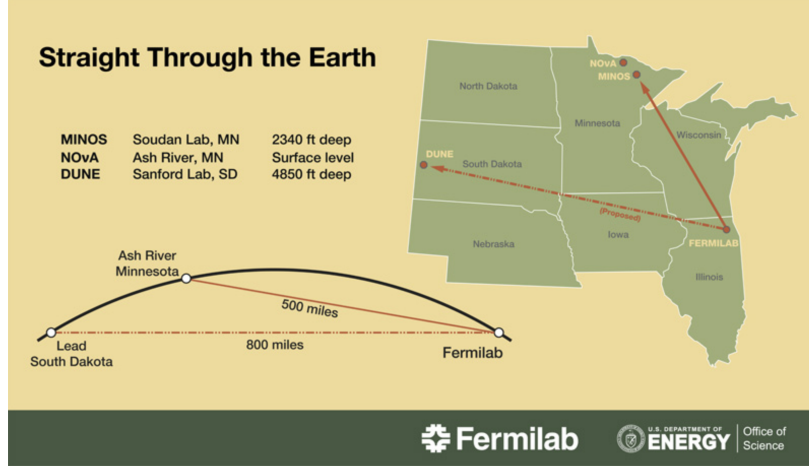


Figure 2.1: Baselines of the Fermilab long-baseline neutrino oscillation experiments including MINOS (on-axis) and NOvA (off-axis) using the NuMI beam and DUNE (on-axis) using the future PIP-II beam.

400 MeV and the Booster at 8 GeV. Some protons are delivered to both the Booster Neutrino Beam (BNB) for various short baseline neutrino experiments including MiniBooNE, MicroBooNE, SBND, ICARUS, and ANNIE. The protons are accelerated to 120 GeV at the main injector before being directed at a downward slope toward the NuMI target hall in batches of 4.0×10^{13} POT (protons on target) every 1.87 s during MINOS operation to peak 6.3×10^{13} POT every 1.2 s during 1 MW runs with NOvA [29]. The duration of each batch is 10 μ s, known as the “beam spill window”.

As generally described in Section 1.2.1, the protons collide with a graphite target, and focusing horns steer either positive or negative hadrons toward the decay pipe where they decay into neutrinos (Figure 2.3). Hadron and muon monitors determine the beam profile for surviving particles, and only neutrinos remain following several hundred meters of rock.

To maximize the data volume of appearance and disappearance events, the far detector must be placed near the oscillation maximum, which depends on baseline length and neutrino energy, which in turn depends on off-axis angle, all while taking into consideration the physically allowed geography of the experiment. At the time of the design phase for the NuMI beam and MINOS in the early 1990s, neutrino oscillation had not yet been confirmed.

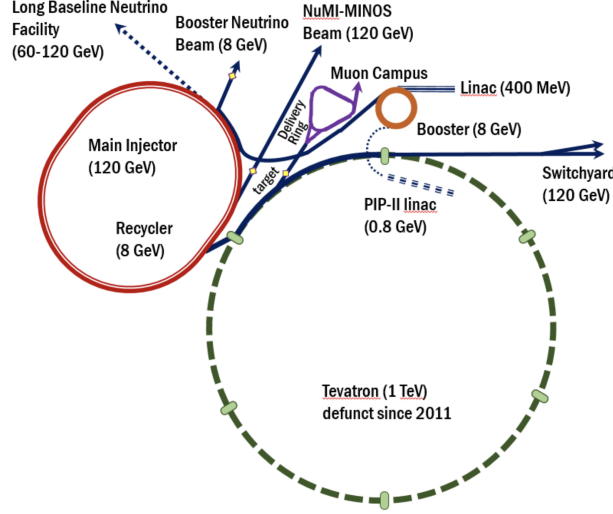


Figure 2.2: The Fermilab accelerator complex including the Linac (triple blue line), Booster (orange line), main injector (red line), neutrino beams (dark blue lines) and their targets (yellow diamonds), and the retired Tevatron (green dashed line) [30].

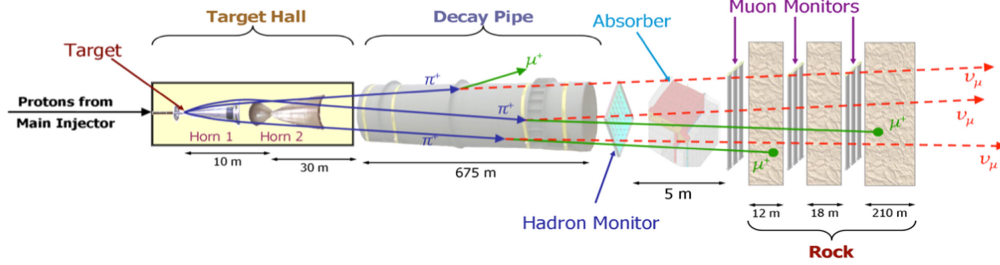


Figure 2.3: The NuMI beam shown in FHC (mode) configuration [31].

Kamiokande’s atmospheric neutrino measurements preferred much larger θ than measured today [32], and θ_{atm} was not yet shown to be nonzero. MINOS was therefore designed to be on-axis to maximize neutrino flux and demonstrate neutrino oscillations through the more common $\nu_\mu \rightarrow \nu_\tau$ disappearance channel. Since NOvA was designed as a successor to MINOS using the existing NuMI beam with minimal capability to decrease beam energy, the only adjustable parameter is θ . The NOvA far detector was placed 14.6 mrad off-axis with a corresponding peak at 2.0 GeV as a trade-off between larger flux at the 1.5 GeV appearance first oscillation maximum and the overall decrease in flux at wider angles (Figure 2.4).

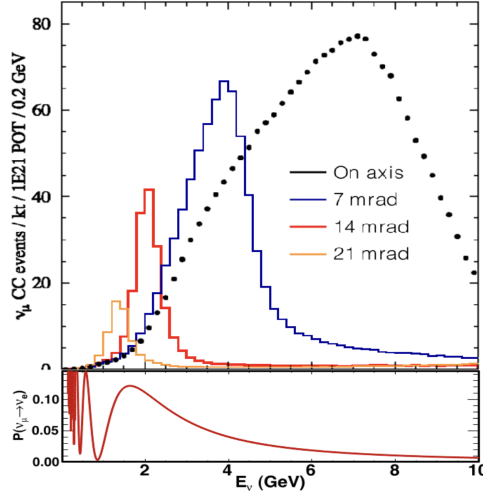


Figure 2.4: The neutrino energy spectrum for several potential off-axis angles for the NOvA FD (top) with the probability of $\nu_\mu \rightarrow \nu_\tau$ appearance (bottom).

2.2 Detector Design

Due to the lack of FD underground location at the off-axis angles, the NOvA detectors were designed to minimize detector latency in order to separate neutrino events within the beam spill time window from the large cosmic ray background. The FD and ND are functionally identical and are formed from plastic extrusions or “cells” filled with a liquid scintillator. Each cell has a wavelength shifting fiber running through it that collects scintillation light and terminates at an avalanche photodiode (APD) at the end of the cell (Figure 2.5).

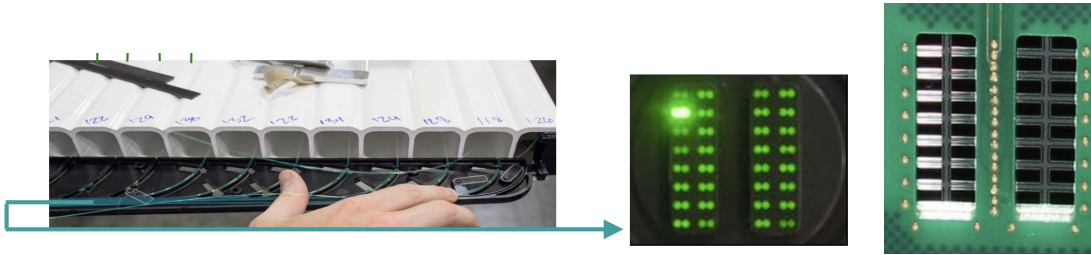


Figure 2.5: NOvA plastic extrusions each with one wavelength shifting fiber (left). Fibers from 32 cells terminate at one avalanche photodiode (right).

For both detectors, these “cells” have cross-sectional dimensions of 4 cm $\times$ 6 cm and are arranged as alternating horizontal and vertical planes. This construction generates two views, XZ and YZ, of the energy depositions in the cells (“hits”), enabling 3D reconstruction of par-

ticle tracks (Figure 2.6). The segmented design, 1 ns timing resolution of the scintillation signal, and capability for 3D reconstruction allow for powerful cosmic discrimination. The relatively low density of the oil-based medium (5% pseudocumene in mineral oil) also leads to longer radiation length, facilitating particle classification based on topologies.

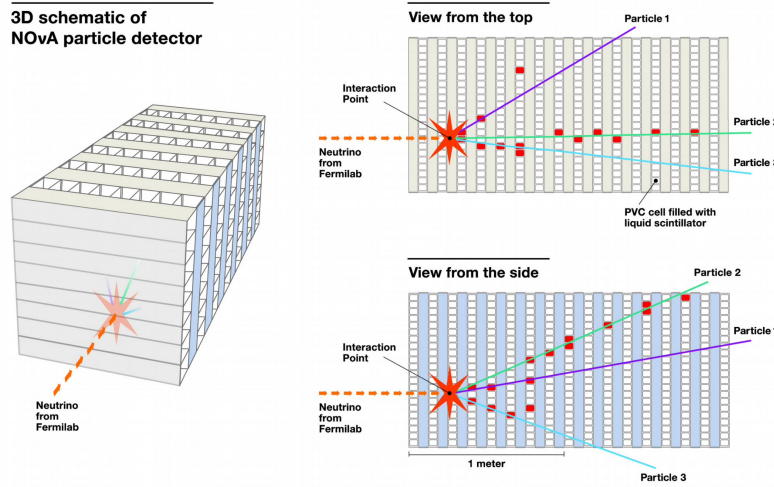


Figure 2.6: Schematic of NOvA detector and generation of top and side views from vertical and horizontal planes respectively.

The 0.3 kton near detector is located in the “MINOS underground” at Fermilab at a depth of 105 m and 1 km to the target. The “active region” of the ND volume is 3.8 m $\times$ 3.8 m $\times$ 12.7 m, consisting of 192 planes with 96 cells each for a total of 18,432 cells. In addition to the active region, the ND features a “muon catcher” placed at the back of the detector. The muon catcher was repurposed from an earlier prototype of the ND and contains 10 cm thick steel planes alternating between the liquid scintillator planes for higher muon stopping power. The muon catcher is a narrower 3.8 m $\times$ 2.6 m $\times$ 3.2 m and is placed on the floor behind the detector with the longer axis spanning the y direction in order to contain slightly downward-going muons from neutrino interactions in the active region (Figure 2.7). It consists of 22 liquid scintillator planes with 96 cells in vertical planes and 64 cells in horizontal planes for a total of 1,760 cells. The detector is located in a separate alcove of the MINOS underground compared to where the MINOS ND was previously located in order to have it centered at

the 14.6 mrad off-axis angle. Due to its proximity to the beam, however, the ND captures a wider range of angles, resulting in a wider neutrino energy spectrum.

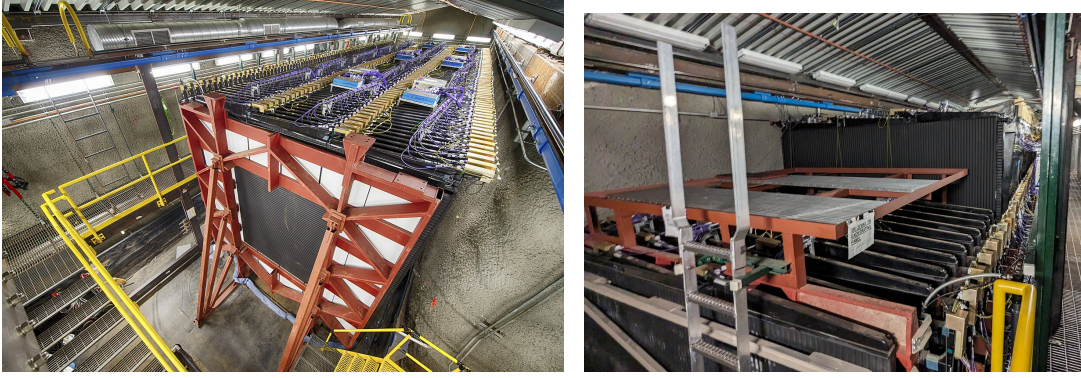


Figure 2.7: NOvA ND front view (left) and back view showing muon catcher (right).

The 14 kton far detector is located in Ash River, Minnesota 810 km from the target. It is 15.2 m $\times$ 15.1 m $\times$ 59.6 m, consisting of 896 planes with 384 cells each for a total of 344,064 cells.

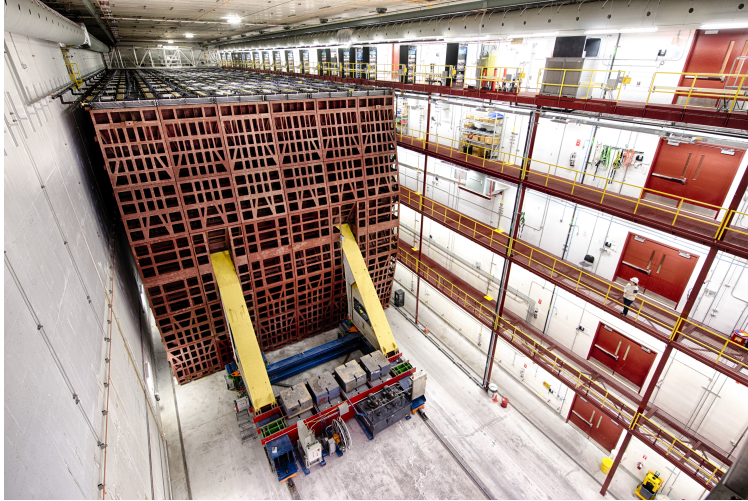


Figure 2.8: NOvA FD front view.

2.3 Event Topologies

Neutrinos interacting in the detector, usually with a nucleon from a carbon atom, may interact via either CC or NC (Figure 2.9). In a CC interaction, the interaction produces

a lepton and a hadronic component in which the lepton flavor corresponds to the neutrino flavor by lepton number conservation. A $\bar{\nu}_\mu$ CC interaction will produce a muon and a $\bar{\nu}_\mu$ CC interaction will produce an electron. Neutrinos interacting via NC only produce a hadronic component, so the neutrino flavor cannot be determined. Since the W production threshold of 3.4 GeV is above typical NOvA neutrino energies, $\bar{\nu}_\mu$ CC events are not observed, and ν_μ can only interact via NC.

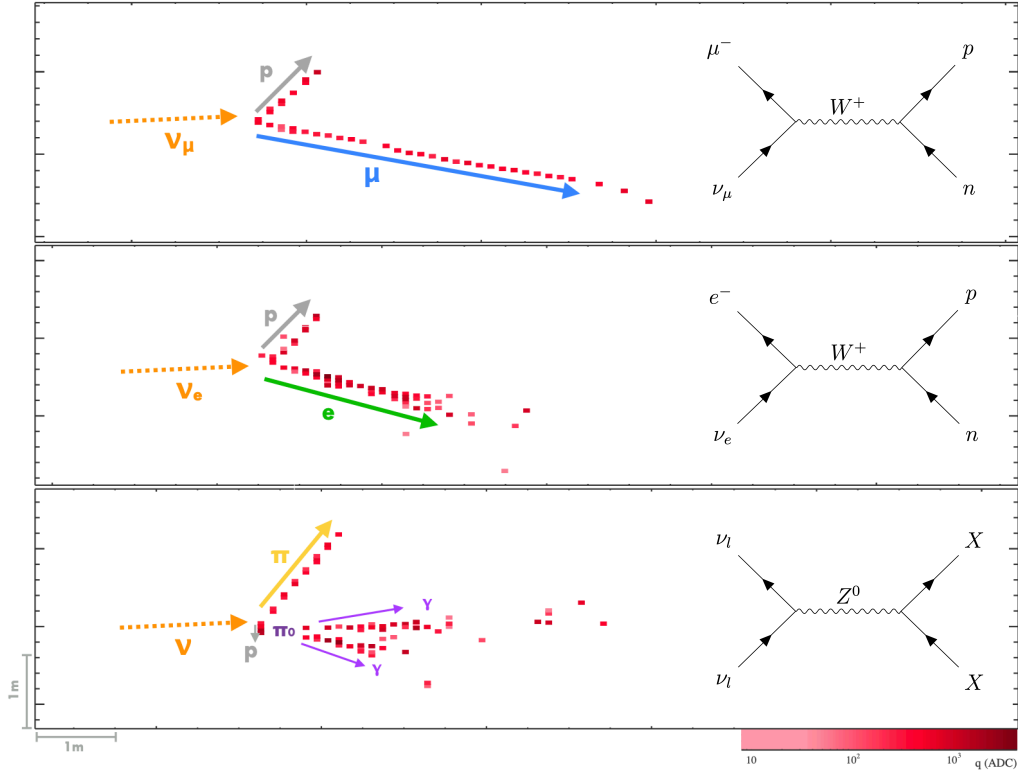


Figure 2.9: Example NOvA event displays of neutrinos interacting via $\bar{\nu}_\mu$ CC (top), $\bar{\nu}_e$ CC (middle), and ν_l NC (bottom) with corresponding Feynman diagrams. Dashed arrows illustrate the interacting neutrinos, which do not deposit energy, and solid arrows highlight the resulting visible particles.

As minimum ionizing particles, muons deposit their energy as long, thin tracks. Charged pions and protons also appear as tracks, depositing more energy per radiation length, with protons depositing the most energy over a short length. Electrons and photons appear as wider and shorter electromagnetic showers. However, unlike electron showers, photon

showers also exhibit a short gap between their production point and the start of the shower. Neutral pions do not deposit energy until decaying into a pair of photon showers. Neutrons rarely deposit energy in the scintillator except for occasional, small proton recoils and are therefore mostly invisible. Kaons are very rarely produced in NOvA neutrino interactions as they would require greater neutrino energies.

Within CC interactions, neutrinos may interact via various interaction modes including quasielastic scattering (QE), resonant pion production (Res), and deep inelastic scattering (DIS). In QE, the neutrino scatters off a down quark in a neutron, converting it into an up quark via a W and producing a proton. The reverse is true for anti-neutrinos. QE is the cleanest of the interaction modes in that it typically only produces a proton in the hadronic portion of the event. Res involves exciting a nucleon that later emits a charged pion as it deexcites. DIS is the messiest of the three since it involves breaking apart the nucleon, producing various hadrons in the process. A less common interaction mode is meson exchange current (MEC) in which the neutrino interacts with two nucleons previously coupled via a virtual pion.

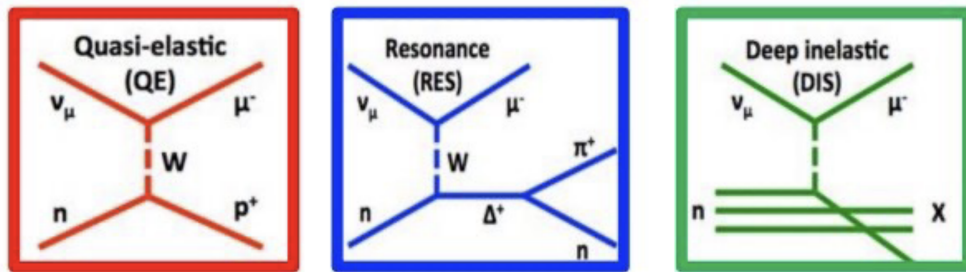


Figure 2.10: Feynman diagrams illustrating QE (left), Res (middle), and DIS (right) [33].

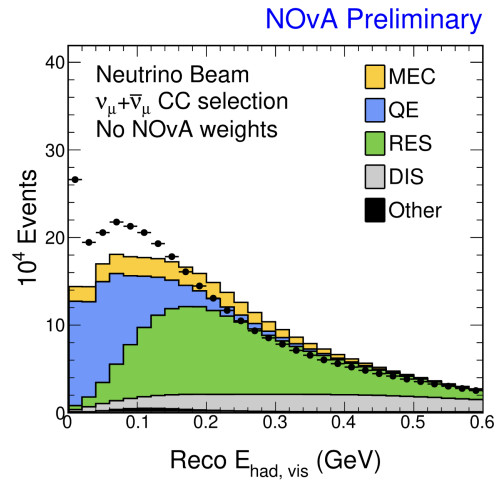


Figure 2.11: Content of interaction modes as a function of hadronic energy before NOvA cross-section tuning [34].

Chapter 3

NOvA Event Reconstruction and Simulation

An overview of the NOvA reconstruction chain is given in Figure 3.1 and in [35]. The beam spill timing information is used as the trigger. A $550\text{ }\mu\text{s}$ window is saved around each beam spill. This window is separated into the actual $10\text{ }\mu\text{s}$ beam spill window within which neutrino interactions may be found and the sideband window. A periodic trigger (10 Hz for FD and 1 Hz for ND) is also collected to study the cosmic ray background in addition to various other data-driven triggers for other analyses besides the neutrino oscillation analysis. The scintillation light waveforms from the APDs are digitized by analog-to-digital converters (ADCs) in the front-end electronics. The “CalHit” algorithm first fits a single hit time and ADC value corresponding to the collected charge. Since each $550\text{ }\mu\text{s}$ contains several cosmic events as well as a potential neutrino event, a “slicer” algorithm then clusters hits near each other in both space and time and assigns them to a single “slice”. Each slice is considered one event. The “ElasticArms” vertex algorithm then determines the interaction vertex location. This vertex position serves as a seed to the “prong” segmentation, which uses the fuzzy k-means clustering algorithm to cluster hits resulting from the same particle [36]. Each such

cluster is referred to as a prong. The Kalman track fitting algorithm is also applied to be used in the selection [37]. Kinematic quantities such as prong length and direction are then calculated for each of the prongs. Figures 3.2-3.3 demonstrate this process with an example event. In parallel, each prong is assigned to its corresponding particle type, and the overall event is classified (see Section 3.2). Finally, all of this information is used to determine the energy for each prong as well as the overall event. This final energy estimation step is currently handled during the analysis step (Section 4.3).

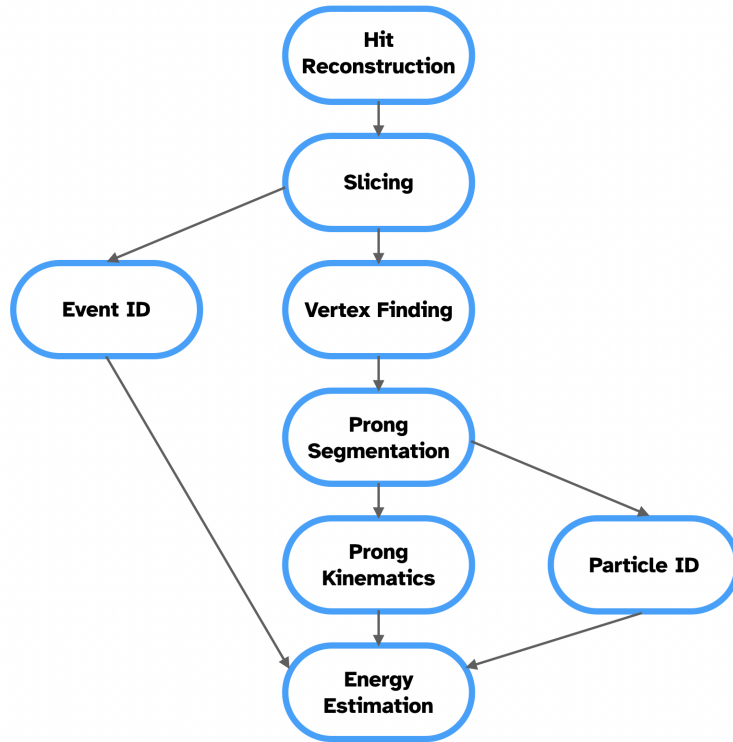


Figure 3.1: Overview of the NOvA reconstruction chain.

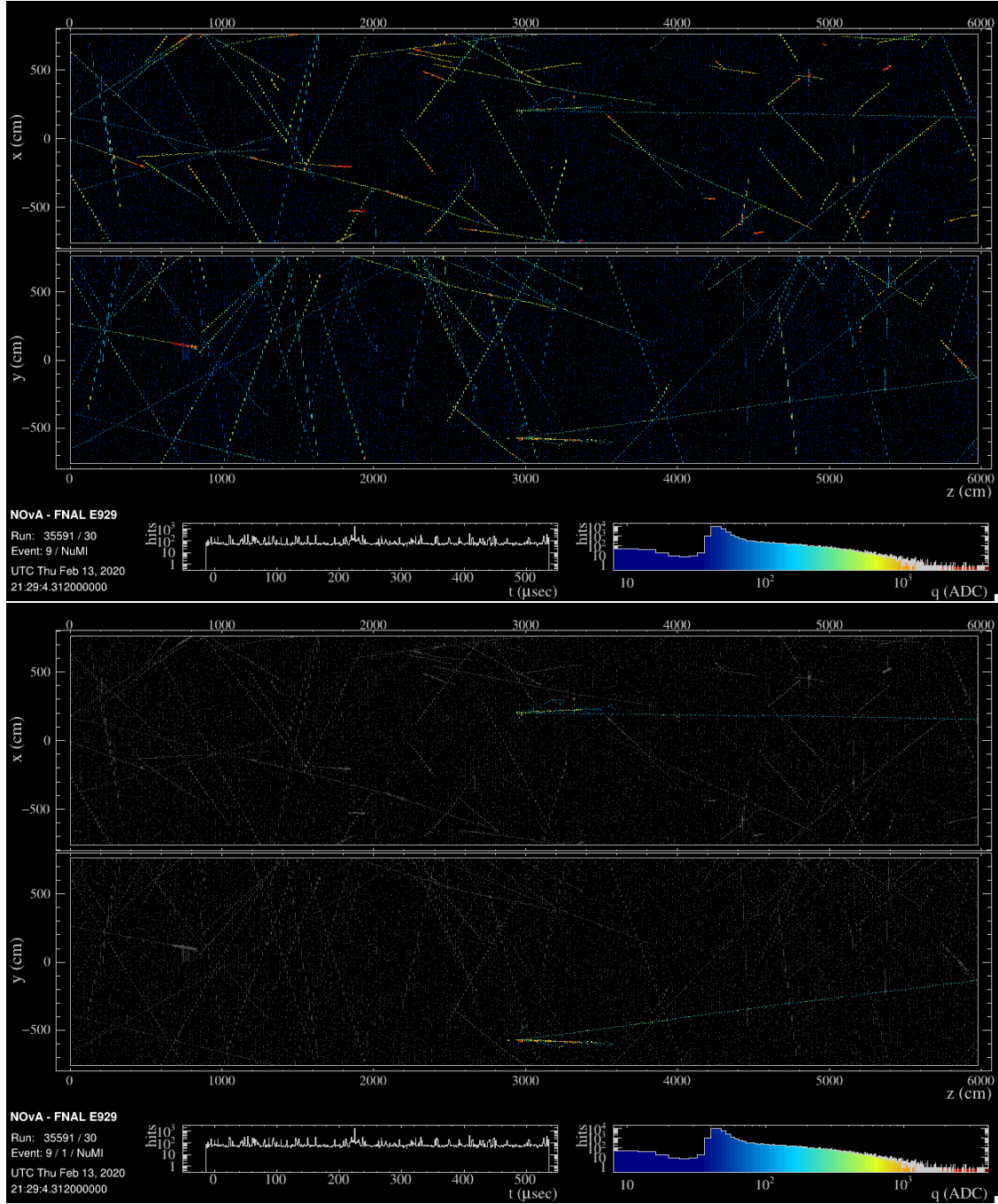


Figure 3.2: An example event display of a full 550 μ s window including a simulated event (top). After the slicer is applied, hits not in the slice are greyed out (bottom). CC

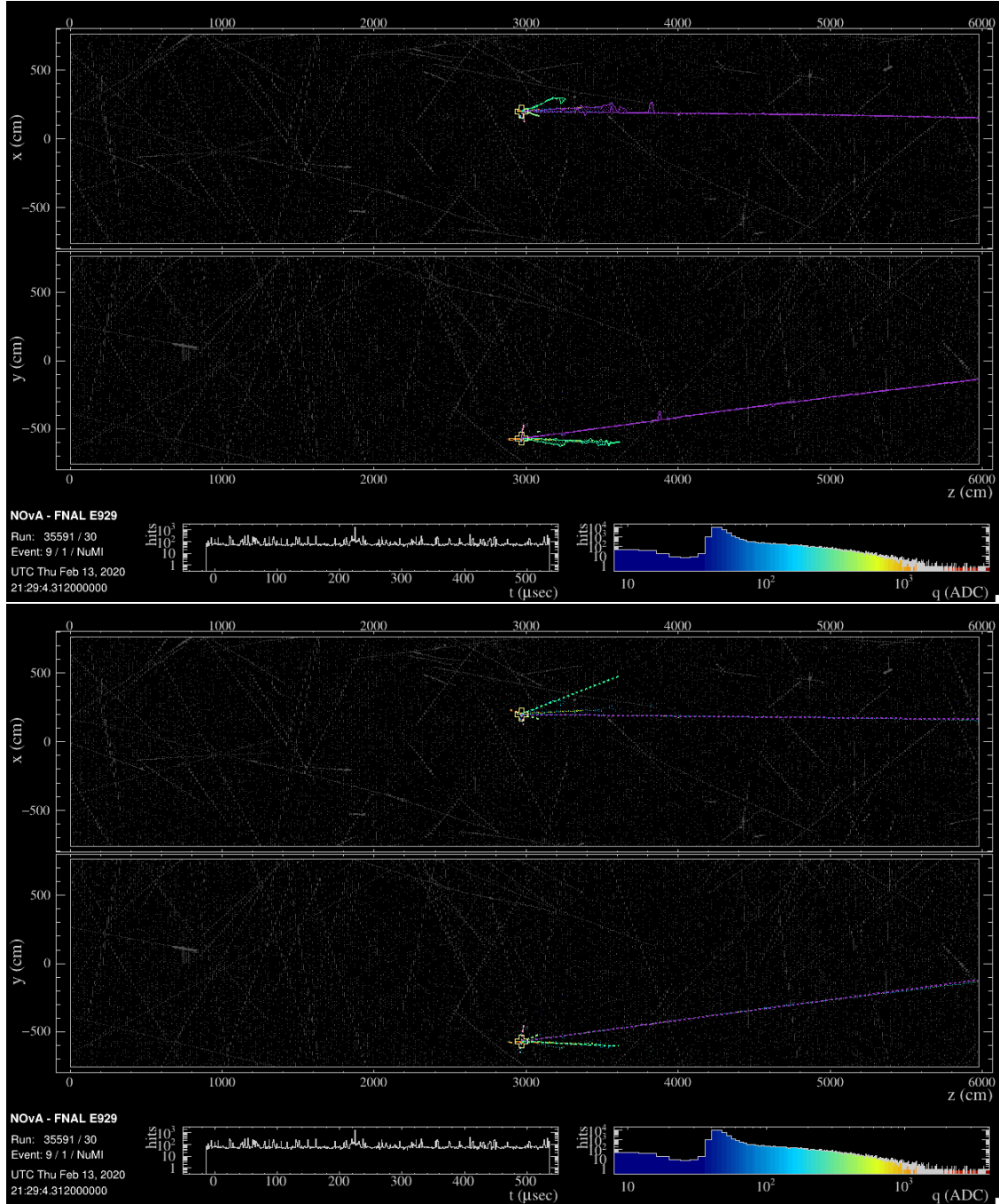


Figure 3.3: The same event as Figure 3.2 with reconstruction products shown. The vertex is shown as a yellow plus, and 6 prongs along with their direction and length are shown as purple, turquoise, orange, blue, pink, and light green shaded regions (top). The same prongs are alternatively shown as dashed lines to indicate their calculated length and direction (bottom).

3.1 Simulation

The beam is first modeled using the PPFX package used by all Fermilab experiments using the NuMI beam [38]. Monte-Carlo (MC) simulations of neutrino beam interactions in both detectors are then generated with GENIE v3.0.6 [39]. The resulting energy depositions in the detector material are simulated with GEANT4 v10.4 [40]. Cosmic ray interactions are not simulated and are entirely sourced from periodic triggers. A small portion of the FD 10 Hz trigger dataset is reserved to be used for the FD beam MC. Smaller 12 μ s windows are taken from this 10 Hz trigger dataset, centered at the beam spill time, and one beam MC event is overlaid within each window. The event reconstruction chain is then run on this overlaid dataset. This approach ensures that the cosmic background and noise conditions of the MC dataset are realistic.

3.2 Event and Particle Classification

Since the NOvA oscillation analysis published in 2017, event classification has been conducted with a convolutional neural network (CNN) known as “EventCVN” [41]. “CVN” means convolutional visual network and is synonymous to CNN. This network was the first CNN to be used in a major high energy physics analysis [42]. This network takes as input two “pixel map” images corresponding to the XZ view and YZ view (Figure 3.4). These images are generated by scaling the energy depositions in each hit to an 8-bit integer, with the maximum 255 value corresponding to 278 MeV. The images are then cropped to around the event such that the first hit plane is placed at in the image, and the median hit along or is placed at in the middle. These dimensions correspond to roughly 14 m 4 m in . The dimensions and cropping procedure ensure that the majority of neutrino interaction events are contained, excluding only the highest energy muon tracks (> 3 GeV).

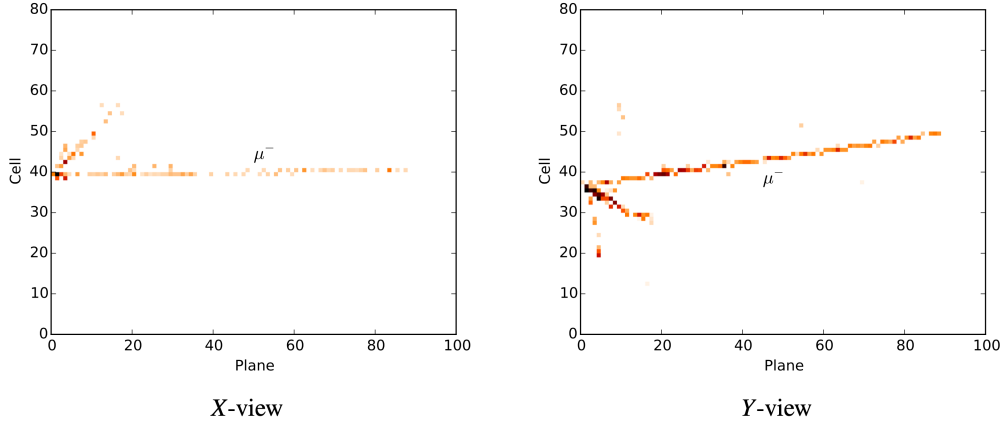


Figure 3.4: An example pixel map for a $\bar{\nu}\nu$ CC event with the XZ view (left) and YZ view (right).

The two views are fed separately in a Siamese MobileNet CNN architecture [43] in which convolution layers are applied separately to each view before they are combined in later layers (Figure 3.5). The network is trained to output a length 4 softmax vector that can be interpreted as the probabilities that the event is a $\{\bar{\nu}\nu$ CC, $\nu\nu$ CC, NC, Unknown (cosmic) $\}$. (The softmax operation is $\frac{e^{x_i}}{\sum_j e^{x_j}}$ for the i -th element of a vector with n elements.) Since there is no truth information associated with cosmic ray events sourced from real data, these are given the “Unknown” designation.

Since the following analysis published in 2018, prong classification has been conducted with the “ProngCVN” network. ProngCVN is a 4-view Siamese GoogLeNet CNN [44] (Figure 3.6). In addition to the 2 full event pixel maps for contextual information, it takes as input prong pixel maps, which only contain hits assigned to the single prong of interest according to the upstream prong segmentation algorithm for a total of 4 views. ProngCVN is only trained on the 5 most relevant particle types, $\{\bar{\nu}, \nu, \mu^+, \mu^-, e^+\}$, removing others and cosmics from the training set.

EventCVN and ProngCVN were also used in the 2020 NOvA oscillation analysis and the 2024 analysis presented in Chapters 4-6. The TransformerCVN network described in Section 3.3 is planned to be used for NOvA’s final oscillation analysis in Chapter 7.

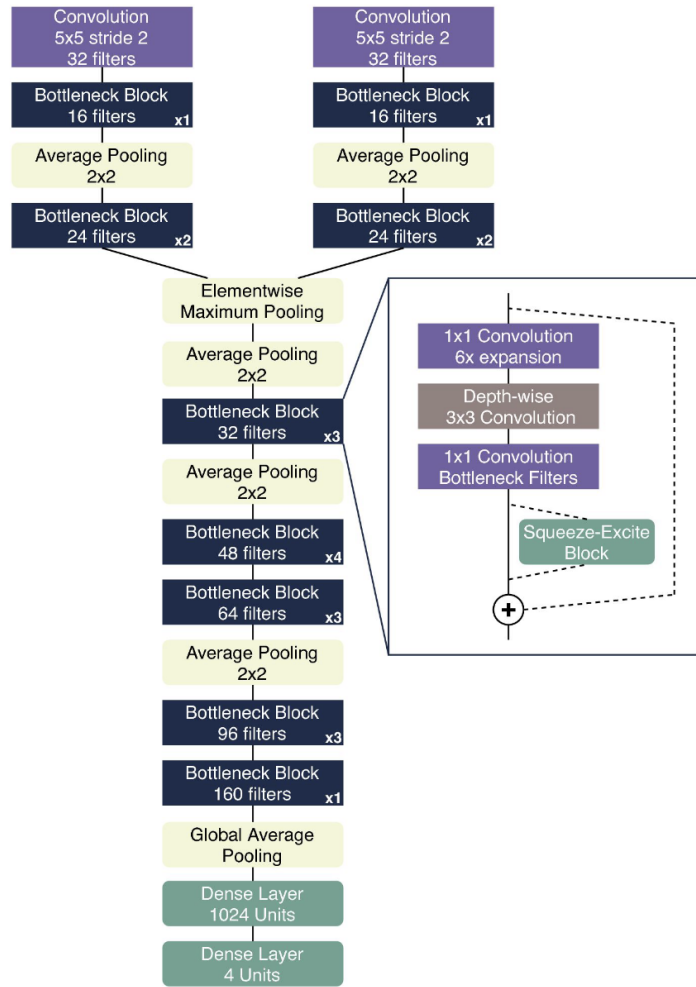


Figure 3.5: The 2-view Siamese MobileNet architecture of EventCVN.

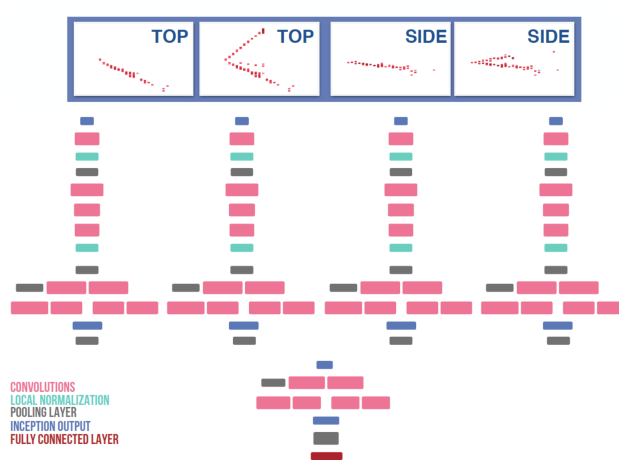


Figure 3.6: The 4-view Siamese GoogLeNet architecture of ProngCVN.

3.3 TransformerCVN

TransformerCVN is a network for simultaneous event and particle classification with a combined Transformer-CNN architecture. The model was designed by Alexander Shmakov, a UC Irvine computer science PhD student, under the supervision of Professor Pierre Baldi. The network architecture is shown in Figure 3.7. For each event, the network takes as input the pixel map image for the overall event as well as the prong pixel map for each reconstructed prong found in the event.

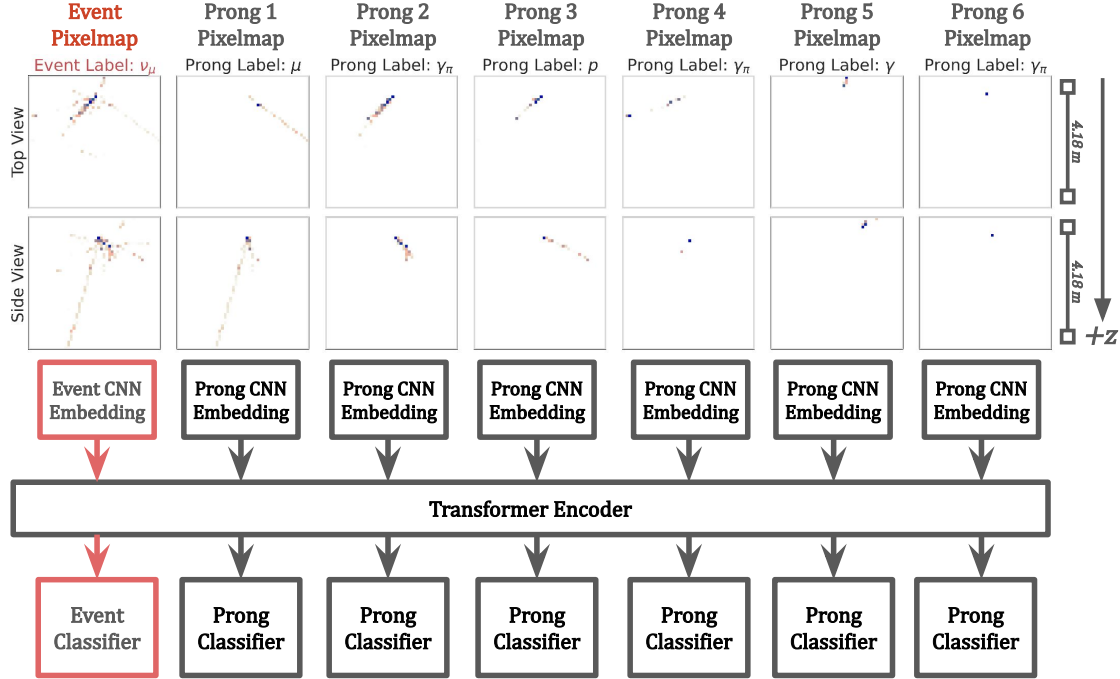


Figure 3.7: TransformerCVN network architecture featuring DenseNet CNNs to embed event and prong pixel maps into a transformer encoder.

We use a densely-connected CNN (DenseNet) [45] to reduce each of the pixel maps to 1 x 256 abstract feature representation vectors. We also have an alternative CNN architecture with a modified DenseNet that uses sparse convolutions, only performing convolutions for which the center pixel is non-zero. (The name “DenseNet” here refers to the presence of skip connections that connect each convolutional layer with every preceding layer, not that the convolutions themselves are performed over all pixels.) However, at the time of development,

it is not possible to trace this sparse CNN in order to compile it into TorchScript format for use in the NOvA C++ ART framework used in large-scale data processing and simulation production campaigns. We therefore use the dense CNN in production. We set a conservative limit of 10 prongs per event during training keeping only the highest energy prongs for events with greater than 10 prongs, but 90% of events have fewer than 7 prongs. This limit does not apply to inference. Each dense CNN embedding has 960k parameters, comparable to the total 1.0M parameters in the EventCVN MobileNet architecture. The number of floating-point operations and memory required scales with the number of input prongs.

The feature representation vectors are input into a transformer encoder [46]. Transformer networks were originally developed for natural language processing and operate sets of objects such as words in a sentence. At the heart of the transformer is the attention mechanism, which determines the relevance of each input object to all others represented as attention weights. This allows for effective learning of relationships and contextual sharing of information between input prongs. This transformer architecture follows the canonical formulation described in [46], with one major exception: we elect not to add the position embeddings to our feature representation vectors, which are typically used to encode the order of words in natural language. This avoids imposing an inherent order to the pixel map inputs, maintaining permutation invariance. We do however add a *type* embedding to differentiate prong and event pixel maps by concatenating the event vector separately from the prong vectors. This ensures the transformer is aware of the different objectives associated with these two inputs types while still allowing information to be shared between all components of an event. The transformer encoder has 2.4M parameters.

Separate decoders (fully-connected neural networks) then perform the event and prong classification tasks. Each has 44k parameters. Initially, the event classifier gave softmax scores corresponding to the probability that the event is a $\{ \quad \text{CC QE}, \quad \text{CC Res}, \quad \text{CC DIS},$
 $\text{CC Other}, \quad \text{CC QE}, \quad \text{CC Res}, \quad \text{CC DIS}, \quad \text{CC Other}, \text{NC}, \text{Unknown (cosmic)} \}$.

However, we later found that attempting to separate the interaction modes is slightly detrimental to the main flavor-tagging objective, so we reduced the target event classes to $\{\text{CC}, \text{NC}, \text{NC}, \text{NC}, \text{Unknown (cosmic)}\}$, matching EventCVN’s task. The prong targets are $\{\text{e}, \mu, \tau, \text{photon}, \text{neutron}, \text{pion}, \text{kaon}, \text{Other particle}, \text{Unknown (cosmic)}\}$. In an attempt to identify neutrons and neutral pions, which do not deposit energy in the scintillator detector, the photon class was split into $\text{photon}^{\text{mother}}$ where $\text{photon}^{\text{mother}}$ refers to a photon with a mother e . Simulated prongs corresponding to particles other than the above classes were given the “Other particle” designation, a minority class primarily consisting of kaons.

Training was done on UCI group GPU servers using the PyTorch Lightning framework. The loss function representing the classification task is a sum of two focal loss terms [47] for the event and prong classification tasks. The event loss can be weighted relative to the prong loss, and we ultimately upweighted the event loss 9:1. We use a cosine learning rate schedule, which uses a decaying learning rate with periodic hard resets to ensure the loss does not get stuck in a local minimum.

3.3.1 Network Performance

In order to directly compare to the existing EventCVN and ProngCVN networks, the first version of the network was trained on NOvA’s 5th production campaign used in the 2020 and 2024 analyses. We used 6.32M, 0.17M, and 0.33M events in the training, validation, and testing set respectively. The validation set was used for hyperparameter optimization, reserving the testing set for the final model only. Separate networks were trained for FHC and RHC, and the FHC network is shown here.

Figures 3.8 and 3.9 show the raw network softmax scores for CC , NC , and Unknown separated by true event or prong class. Tables 3.1 and 3.2 show the corresponding performance metrics including one-vs-rest ROC AUC for the event and prong classification tasks. These scores show that the network achieves very similar event classification performance and increased

prong classification performance compared to EventCVN and ProngCVN.

Metric	Event		Prong	
	TransformerCVN	EventCVN	TransformerCVN	ProngCVN
Accuracy	0.894	0.897	0.754	0.726
Precision	0.894	0.908	0.761	0.760
ROC AUC	0.978	0.982	0.940	0.929

Table 3.1: Aggregated metrics accuracy, precision, and area under a standard ROC curve (ROC AUC) for event classification and prong classification. All metrics were aggregated in a one-vs-rest manner across each task’s labels.

Event class	CC	CC	NC	Cosmic
TransformerCVN	0.981	0.987	0.965	0.998
EventCVN	0.983	0.987	0.965	0.999

Prong class					
TransformerCVN	0.972	0.987	0.934	0.896	0.925
ProngCVN	0.956	0.981	0.926	0.884	0.911

Table 3.2: ROC AUC values for individual labels in a one-vs-rest manner for either event classification (top) or prong classification (bottom).

Figures 3.10-3.14 show confusion matrices assigning the event/prong class with the highest score as the predicted class. These are particularly useful for prong classification given the way prong scores are used in analyses to determine the appropriate energy estimation algorithm. Figure 3.12 shows a confusion matrix for the full set of prong classes TransformerCVN was trained on, which indicates that TransformerCVN has some ability to identify photons from neutral particles. However, this is a particularly difficult task, and the network struggles with this as well as the minority “Other” class (mostly kaons). In Figure 3.13, only the typical five prong classes are used, summing the photon score across our split classes and removing “Other” prongs and cosmics. This allows us to compare to the 5-class ProngCVN in Figure 3.14.

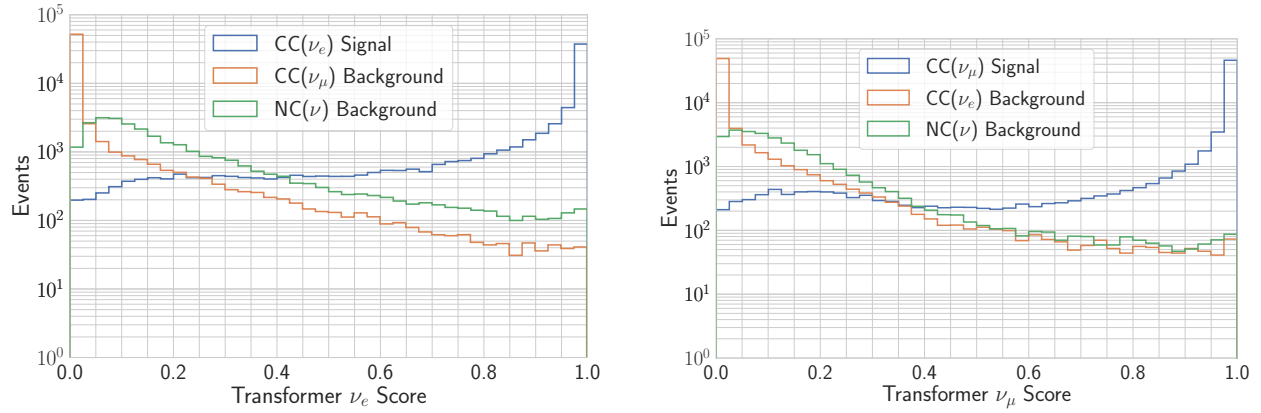


Figure 3.8: Event network output score distributions for ν_e and ν_μ .

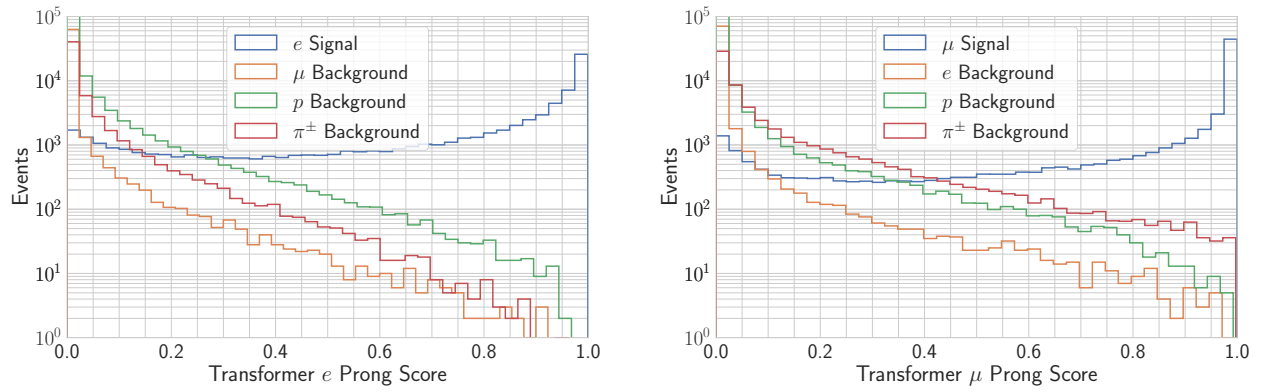
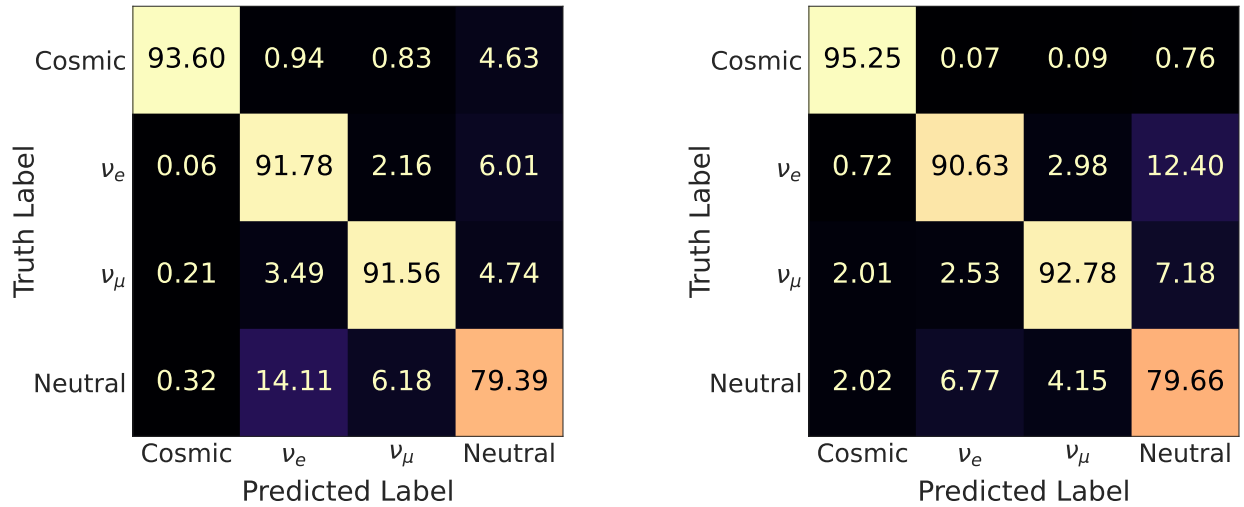


Figure 3.9: Prong network output score distributions for e and μ prongs.



(a) “Efficiency” matrix, normalized along truth labels.

(b) “Purity” matrix, normalized along predictions.

Figure 3.10: TransformerCVN event confusion matrices.

Truth Label	Cosmic	93.13	0.55	0.67	5.65
	ν_e	0.02	88.48	1.41	10.09
	ν_μ	0.16	2.32	90.17	7.35
	Neutral	0.09	6.17	2.65	91.10
		Cosmic	ν_e	ν_μ	Neutral
		Predicted Label			

(a) “Efficiency” matrix, normalized along truth labels.

Truth Label	Cosmic	97.61	0.05	0.08	0.74
	ν_e	0.27	94.91	2.04	16.76
	ν_μ	1.56	1.82	96.01	8.95
	Neutral	0.56	3.22	1.87	73.55
		Cosmic	ν_e	ν_μ	Neutral
		Predicted Label			

(b) “Purity” matrix, normalized along predictions.

Figure 3.11: EventCVN event confusion matrices.

Truth Label	e	83.73	0.59	7.58	0.89	0.01	5.08	2.03	0.00	0.08
	μ	0.77	88.30	5.67	3.71	0.01	0.70	0.56	0.00	0.28
	p	1.42	0.59	91.01	3.15	0.04	1.65	2.03	0.00	0.12
	$\pi^\pm$	1.98	4.44	45.95	37.87	0.06	4.29	5.18	0.00	0.23
	γ_n	3.86	0.35	50.84	3.02	5.79	15.08	20.42	0.00	0.64
	γ_π	9.66	0.77	24.81	3.56	0.16	50.66	10.12	0.00	0.26
	γ_{other}	5.26	0.50	32.31	4.72	0.25	17.86	38.89	0.00	0.21
	other	3.27	1.26	75.83	11.90	0.05	2.87	4.68	0.00	0.14
	cosmic	0.29	0.26	3.29	0.14	0.01	0.45	0.09	0.00	95.48
		e	μ	p	$\pi^\pm$	γ_n	γ_π	γ_{other}	other	cosmic
Predicted Label										

(a) “Efficiency” matrix, normalized along truth labels.

Truth Label	e	83.88	0.92	3.08	2.22	1.55	7.76	4.32	0.00	0.87
	μ	0.51	91.41	1.54	6.16	0.61	0.72	0.80	0.00	1.98
	ρ	2.41	1.56	62.83	13.31	7.53	4.27	7.33	0.00	2.06
	$\pi^\pm$	1.29	4.46	12.12	61.05	4.64	4.25	7.14	0.00	1.54
	γ_n	0.30	0.04	1.62	0.59	50.64	1.81	3.41	0.00	0.52
	γ_π	7.66	0.94	7.98	7.00	14.53	61.28	17.03	0.00	2.12
	γ_{other}	3.79	0.56	9.46	8.45	20.24	19.67	59.56	0.00	1.60
	other	0.13	0.08	1.24	1.19	0.20	0.18	0.40	0.00	0.06
	cosmic	0.03	0.04	0.12	0.03	0.07	0.06	0.02	0.00	89.26
			e	μ	ρ	$\pi^\pm$	γ_n	γ_π	γ_{other}	other
Predicted Label										

(b) “Purity” matrix, normalized along predictions.

Figure 3.12: TransformerCVN full prong confusion matrices.

Truth Label	e	μ	p	$\pi^\pm$	γ
	82.29	0.58	6.84	0.72	9.56
	0.68	88.42	5.38	3.42	2.10
	1.28	0.59	89.24	2.93	5.96
	1.68	4.43	43.47	35.77	14.65
	5.87	0.54	23.70	2.81	67.08
	e	μ	p	$\pi^\pm$	γ
Predicted Label					

(a) “Efficiency” matrix, normalized along truth labels.

Truth Label	e	86.31	0.91	3.00	2.03	6.97
	μ	0.48	91.74	1.58	6.42	1.02
	p	2.29	1.56	66.50	13.93	7.37
	$\pi^\pm$	1.14	4.46	12.38	65.08	6.93
	γ	9.79	1.34	16.54	12.54	77.71
		e	μ	p	$\pi^\pm$	γ
Predicted Label						

(b) “Purity” matrix, normalized along predictions.

Figure 3.13: TransformerCVN 5-class prong confusion matrices. TransformerCVN’s prong outputs were reduced to the 5 common prong classes by removing other particles and cosmoics and adding scores among the split photon class.

Truth Label	e	μ	p	$\pi^\pm$	γ
	77.94	0.29	3.02	4.39	14.35
	0.65	82.36	2.03	12.54	2.42
	1.87	0.25	64.73	17.72	15.43
	2.13	1.73	12.64	68.22	15.27
	7.35	0.19	6.32	10.68	75.46
	e	μ	p	$\pi^\pm$	γ
Predicted Label					

(a) “Efficiency” matrix, normalized along truth labels.

Truth Label	e	82.37	0.51	2.28	4.22	8.34
	μ	0.46	96.25	1.02	8.06	0.94
	p	3.36	0.74	82.93	28.90	15.23
	$\pi^\pm$	1.46	1.97	6.19	42.51	5.76
	γ	12.35	0.54	7.58	16.32	69.74
		e	μ	p	$\pi^\pm$	γ
Predicted Label						

(b) “Purity” matrix, normalized along predictions.

Figure 3.14: ProngCVN 5-class prong confusion matrices.

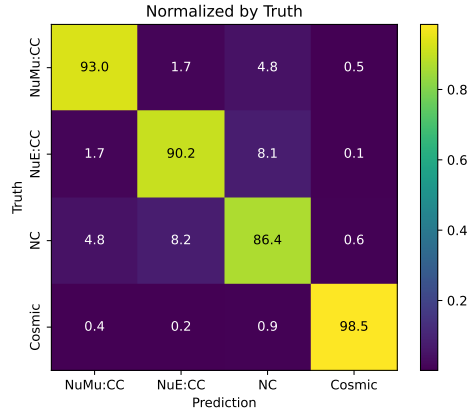
The final version of the network was trained on NOvA’s 6th production campaign in preparation for the final analysis. Several changes to the reconstruction chain were made in this campaign that affect the prong reconstruction and therefore the pixel maps input to TransformerCVN. These changes include moving from the ElasticArms vertex reconstruction algorithm in favor of “VertexCVN”, a CNN for vertex regression, and implementing “OPALProng”, an improved view-matching algorithm for the creation of 3D prongs [48]. Previous studies for all of EventCVN, ProngCVN, and TransformerCVN showed that there was minimal drop in classification performance when using the network trained on the wrong horn current. Network performance therefore benefits from combining the FHC and RHC datasets during training simply due to the larger dataset size. Since production 6 is also larger than production 5, the dataset for the TransformerCVN production 6 network is 29M, 1M, and 3M events in the training, validation, and testing set respectively.

Table 3.3 shows the one-vs-rest classification performance on the production 6 test set. We see that after this retraining, the new TransformerCVN model achieves marginally better performance than EventCVN, although the EventCVN was not retrained. ProngCVN was further optimized for production 6 [49].

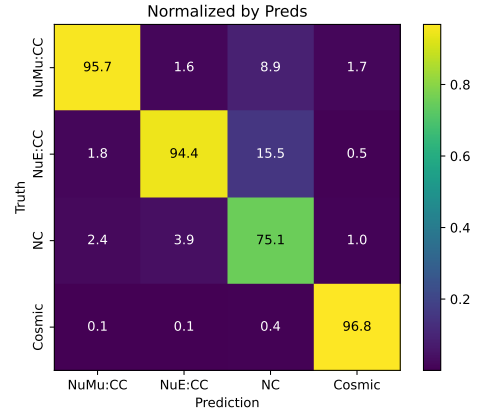
Event class	CC	CC	NC	Cosmic
TransformerCVN	0.988	0.991	0.973	0.999
EventCVN	0.985	0.990	0.968	0.999

Prong class					
TransformerCVN	0.976	0.987	0.936	0.899	0.935
ProngCVN	0.952	0.973	0.921	0.875	0.913

Table 3.3: ROC AUC values for individual labels in a one-vs-rest manner for either event classification (top) or prong classification (bottom) on the production 6 dataset. Note that EventCVN was not retrained for this dataset.

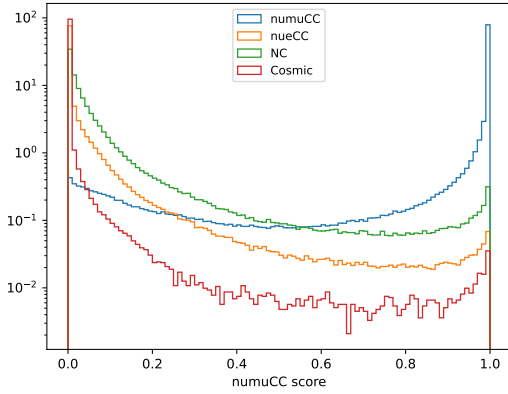


(a) “Efficiency” confusion matrix.

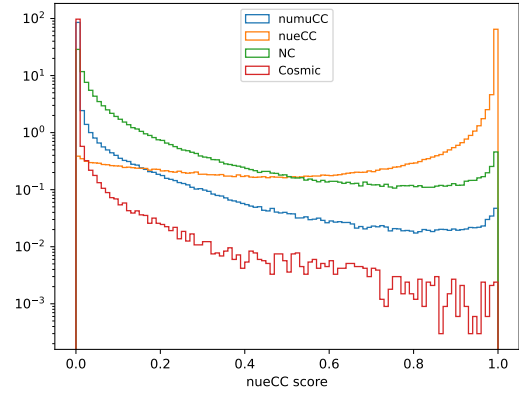


(b) “Purity” confusion matrix.

Figure 3.15: Event classification confusion matrices for retrained TransformerCVN.

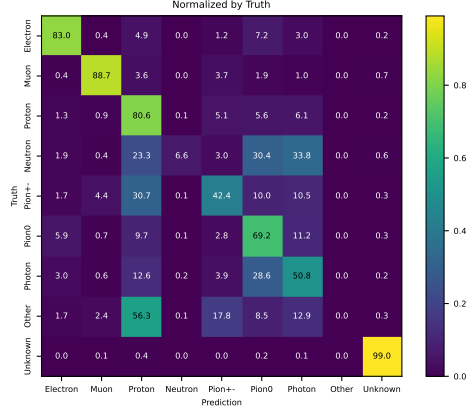


(a) score distributions.

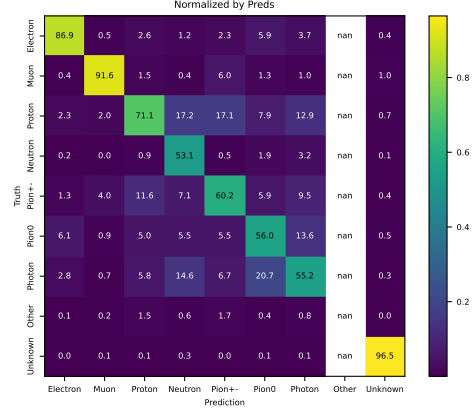


(b) score distributions.

Figure 3.16: Event output score distributions separated by true class for retrained TransformerCVN.

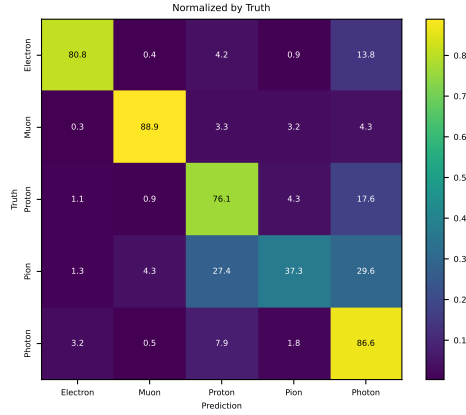


(a) “Efficiency” confusion matrix.

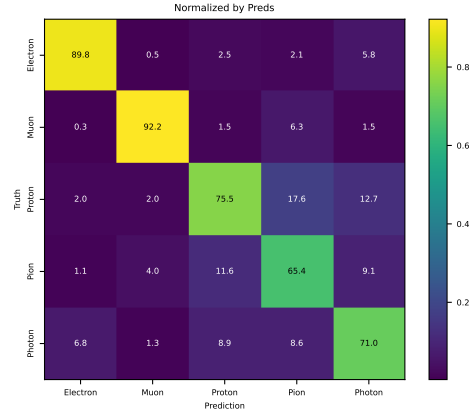


(b) “Purity” confusion matrix.

Figure 3.17: Prong classification confusion matrices for retrained TransformerCVN.



(a) “Efficiency” confusion matrix.



(b) “Purity” confusion matrix.

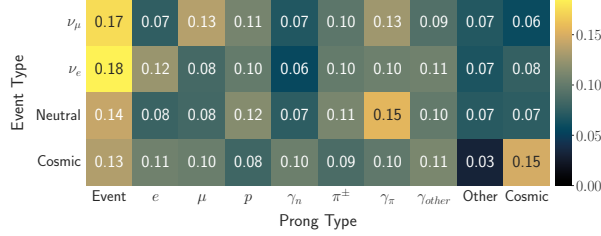
Figure 3.18: 5-class prong classification confusion matrices for retrained TransformerCVN.

3.3.2 Interpretability

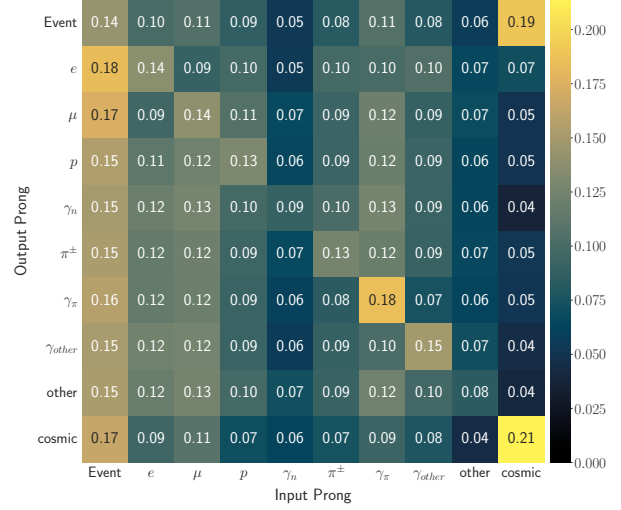
The unified event/prong classification approach and attention mechanism of Transformer-CVN make it possible to study how relationships between different particle and event types affect the classification tasks. The attention weights provide a numerical importance of each input to each output, and we can compute the total attention across the entire encoder stack by simply taking the product of individual attention maps. We can also aggregate attention scores across all events in the dataset while accounting for the different input types in each event: each prong’s truth labels define its *type*, and we incorporate the additional event pixel map as an *event-type* pseudo-prong. For every event type, we average attention scores for each true prong type in the event, summing the attention score for all prongs of the same type in each event. Put another way, we find the importance of every prong *type* for a given event, and then average this importance across events, grouping by event class. This provides the importance of different types of prongs for providing the context necessary to make an event-level prediction.

Figure 3.19 summarizes this analysis. For the impact of prong types on event classification, we see the large impact of the presence of particles largely unique to certain interaction types, namely electrons for $\bar{\nu}\nu$ CC and muons for $\bar{\nu}\nu$ CC. We also notice that the NC events, which are generally more varied as they consist primarily of hadronic components, depend less on leptons and highly on the (photon from a) π^0 for classification. It is interesting to note that this π^0 is important for both $\bar{\nu}\nu$ CC and NC neutrinos, indicating it may be most important for separating between $\bar{\nu}\nu$ CC and NC. For the influence of prong types on prong classification, we see from Figure 3.19b that the diagonal elements are generally high as expected, but it is clear that the network relies heavily on the full event pixel map even for prong classification while getting a smaller amount of additional context from the other prong images.

One method for interpreting the learned behavior of convolution networks is to use saliency



(a) Input prong attention scores for event classification.



(b) Input prong attention scores for prong classification.

Figure 3.19: The average attention score for each prong type for a given event or prong type. Values are normalized so that each row sums to one. Higher values indicate the attention layer found the overall contribution of the given prong type inputs more important for classification compared to other prong types. The “Event” column refers to the full event pixel map input.

maps [50]. Saliency refers to the derivative of the network output with respect to each input pixel, —. This produces pixel maps representing how strongly the output probability will change with respect to an infinitesimal change in intensity for every input pixel. We produce these saliency maps for every prong’s output and event output with respect to all input pixel maps, both prong and event (Figure 3.20).

Similarly to the attention maps, pixel saliency values calculated for individual events is somewhat noisy and less informative, especially due to our input sparsity. We therefore aggregated saliency maps across multiple events to extract an average, interpretable result for understanding the underlying physics. We rotate all of the pixel maps so that the tracks within them all face towards the + axis and average each output head’s saliency over these rotated pixel maps. We also remove prongs with length greater than 50 pixels (488 cm) and reconstructed energy greater than 4 GeV to focus this analysis on topological differences between prong types rather than differences in obvious quantities such as prong lengths and

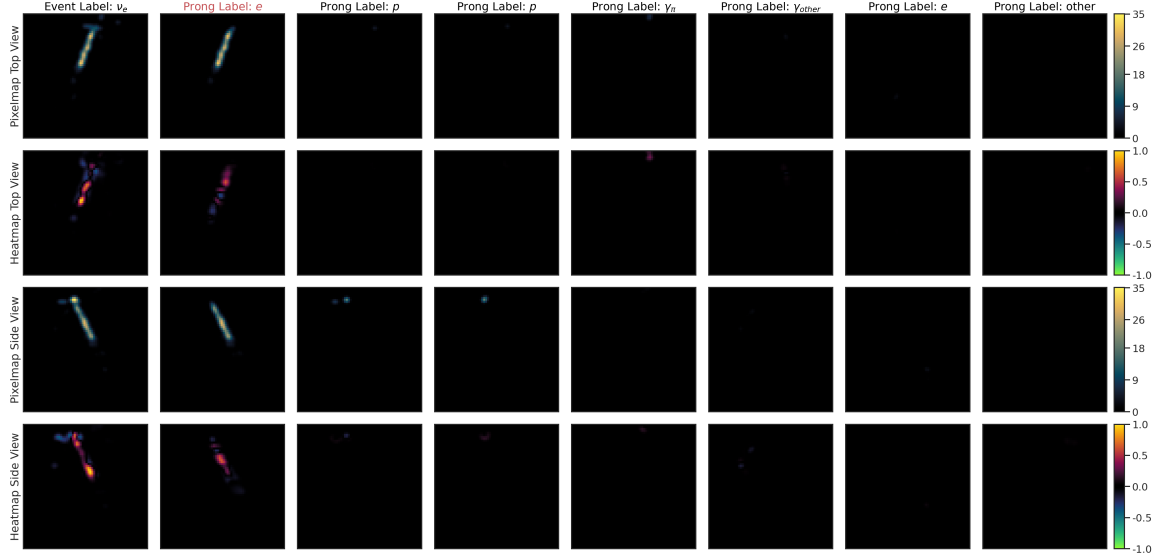
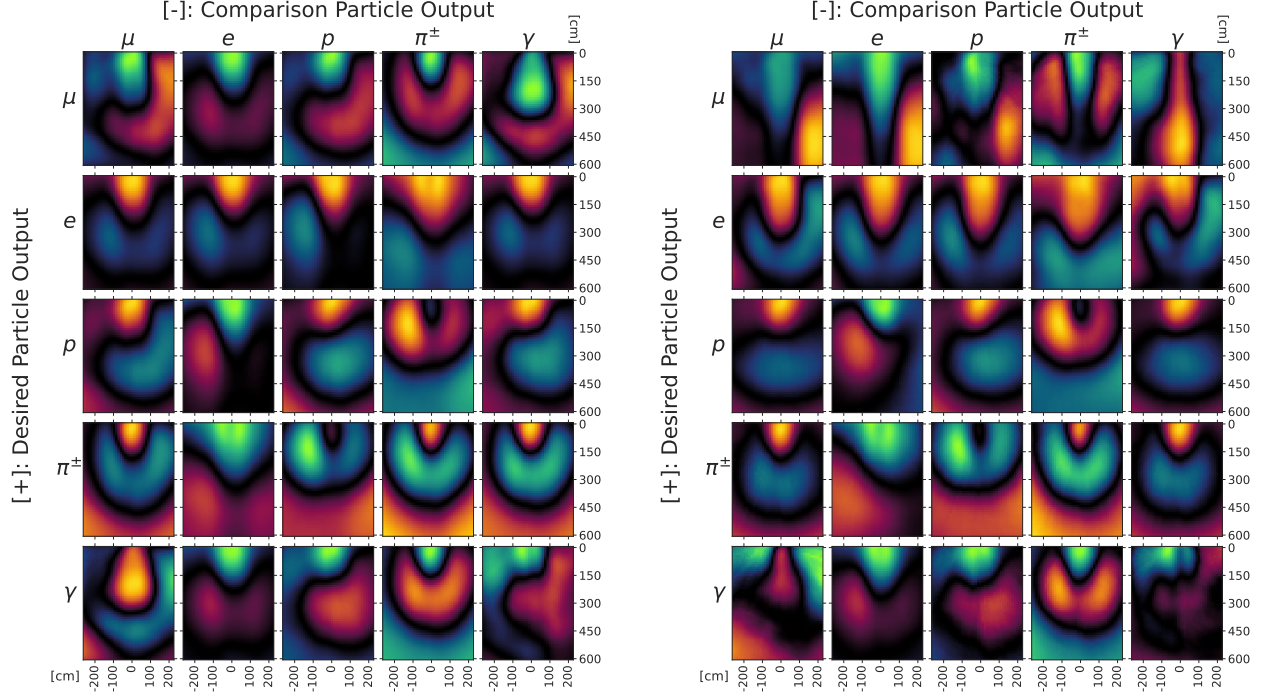


Figure 3.20: Saliency map for one example event. The 1st and 3rd rows represent the input pixel maps for the top and side views respectively. The 2nd and 4th rows represent the corresponding saliency values that led to the classification of the leading prong as an electron with hits in red counting in favor of classification as electron and hits in blue counting against.

energies. We note that this aggregation is imperfect and introduces some artifacts, which may explain certain phenomenon such as the asymmetry along the vertical axis in some of the maps arising from direction reconstruction.

Figure 3.21 shows a grid of saliency maps for five major prong particle labels. These maps indicate which regions correlate the most with the network predicting the *Desired Particle* label rather than the *Comparison Particle* label. We may choose to either aggregate over all prong pixel maps (Figure 3.21a) or only those for which the truth label matches the Desired Particle (Figure 3.21b). This true prong label filtered figure is particularly useful as it allows us to use the off-diagonal elements as a way to diagnose potential reasons for misclassification by studying the blue regions.

The saliency map (Figure 3.21b, top left) shows that heavy activity close to the vertex makes the network less likely to classify the prong as a muon. An outline of a typical muon track can be seen in this plot along which the saliency increases farther from the vertex.



(a) Saliency map grid in which all events are included in every map. This grid must be anti-symmetric.

(b) Saliency map grid in which only events with true particle type matching the Desired Particle row are included in the aggregation. This filtering highlights differences between different prong types.

Figure 3.21: Grid of aggregated saliency maps and difference maps for every pair of prong types. The individual event pixel maps are rotated such that the prong direction is downward oriented and translated so that the vertex is at the center-top. The diagonal pixel maps present saliency maps for the output associated with each particle type. The off-diagonal maps presents pair-wise difference saliency maps: computed as simply $\frac{S_{ij} - S_{ji}}{S_{ii} + S_{jj}}$, each renormalized to highlight the differences. Red indicates a positive correlation with the Desired Particle's classification probability, while blue represents a negative correlation.

One artifact of these saliency maps is that regions of the map where there is little activity across all events in the given particle type can have more volatile saliency values as can be seen by the red regions on either side of the muon track and unpredictable behavior near the edges of all maps. Conversely, the μ saliency map (as well as all off-diagonal elements in the μ row) shows that hits close to the vertex makes the network more likely to identify the typical shorter, more diffuse pattern of electromagnetic showers, while activity farther from the vertex may lead to misclassifying the electron. Track classification usually relies on the length of the track since muons deposit energy over much longer tracks while protons

leave the shortest, most intense tracks. This can be seen from any of the off-diagonal plots comparing any pair among π^+ , π^- , or e^- since hits closer to the vertex contribute more toward classification of the particle producing shorter tracks. π^+ vs π^- separation is a difficult task since both particles produce showers, but photons can be identified from the small gap they leave between the vertex and the start of the photon’s shower. This can be seen from both the π^+ and π^- maps since hits at the vertex make electron classification more likely while hits a short distance away in any direction contribute toward photon classification.

We may also compare the track profile between the different classes by integrating the saliency maps presented above across the width of the detector to produce one-dimensional saliency with respect to vertex distance along the track (Figure 3.22). This provides the typical pattern we expect every prong type to form and indicates the most important regions for each prong class. We again notice that the saliency for muons initially increases as the distance from the vertex increases, and it remains flat for a long portion of the track since muon prongs extend farther than the other prong types. Conversely, electron saliency drops sharply as distance increases. We also notice the delayed shower signature expected from photon tracks since saliency starts low and then increases with greater distance, differentiating them from electrons that do not leave a gap. The photon saliency then decreases as the network differentiates them from the long track-producing particles. We also highlight that, unlike the saliency maps, these plots are un-normalized. This lets us also see that the network is generally biased towards predicting easy muon targets since they have positive salience across the distance, while the network is biased against predicting pions since their saliency is negative throughout.

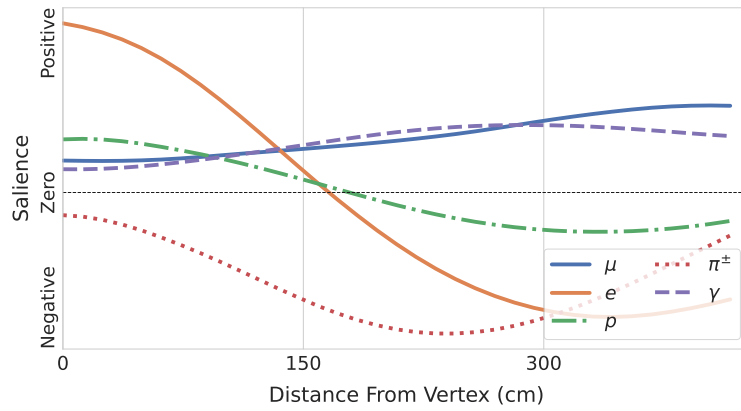


Figure 3.22: Integrated salience map demonstrating the profile of different prong types. Presents the average track-length profile of different particles with hits a certain distance away from the vertex either contributing or detracting from the classification likelihood.

Chapter 4

The 3-Flavor Oscillation Analysis

Chapters 4-6 focus on the 2024 analysis, named as such because preliminary results were presented at the “Neutrino 2024” conference in June 2024 in Milan, Italy [51]. Final results were published in January 2026 [52]. This analysis uses NuMI neutrino beam data collected over the course of 10 years during 2014-2023. This dataset includes POT (the number of protons collided with the graphite target) in FHC mode and POT in RHC mode. This doubles the POT FHC dataset available in the 2020 analysis (Figure 4.1) [53].

4.1 Sensitivity to Oscillation Parameters

As discussed in Section 1.3, as a long baseline oscillation analysis relying on appearance and disappearance with an 810 km 1.5 GeV, NOvA is sensitive to a certain subset of the oscillation parameters. With the high statistics (Section 4.2) and energy resolution (Section 4.3) of events, the energy spectrum of the disappearance channel provides a measurement of and (Figure 4.2). The height of the “dip” approximately gives the probability of disappearance at the first oscillation maximum through .

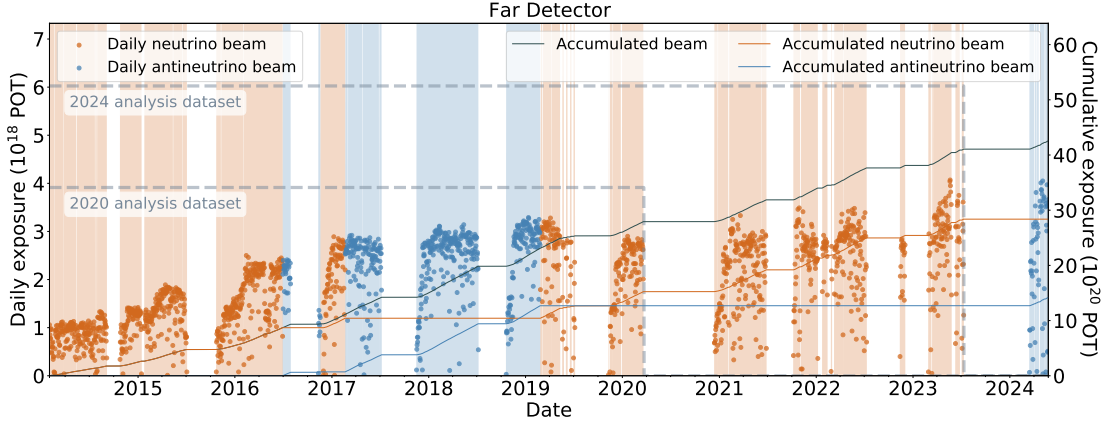


Figure 4.1: Time series showing the daily FHC neutrino beam (orange) or RHC antineutrino beam (blue) POT recorded by NOvA, from the start of Period 1 to May 31, 2024. Also plotted are lines for the cumulative neutrino beam POT (dark orange), cumulative antineutrino beam POT (dark blue) and total accumulated POT (grey). The boxed regions show the records included in the 2020 analysis dataset and the 2024 analysis dataset.

The position of this oscillation is specified by $\sin^2(\theta_{23})$ as it determines the oscillation frequency in L/E with fixed L .

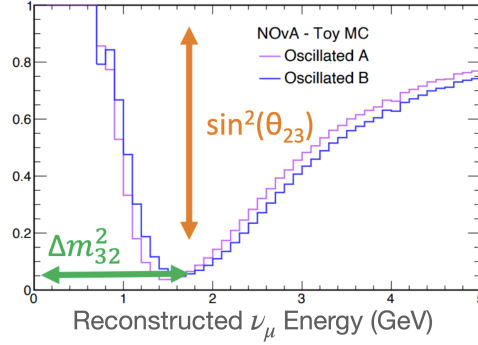


Figure 4.2: Illustration of NOvA’s sensitivities to Δm^2_{32} and $\sin^2(\theta_{23})$ through the energy spectrum of the ν_μ disappearance channel using toy MC.

Similarly, the rate of $\bar{\nu}_\mu$ appearance at the oscillation maximum provides $\sin^2(\theta_{23})$. Several effects can change the rate of $\bar{\nu}_\mu$ appearance between ν_μ and $\bar{\nu}_\mu$, and these are often illustrated through “bi-probability” or “bi-event” plots (Figure 4.3). One effect is the inherent asymmetry between neutrinos and anti-neutrinos determined by θ_{CP} (panel 2). As the value of θ_{CP} increases, ellipses are traced out on these plots, and their directions are opposite between

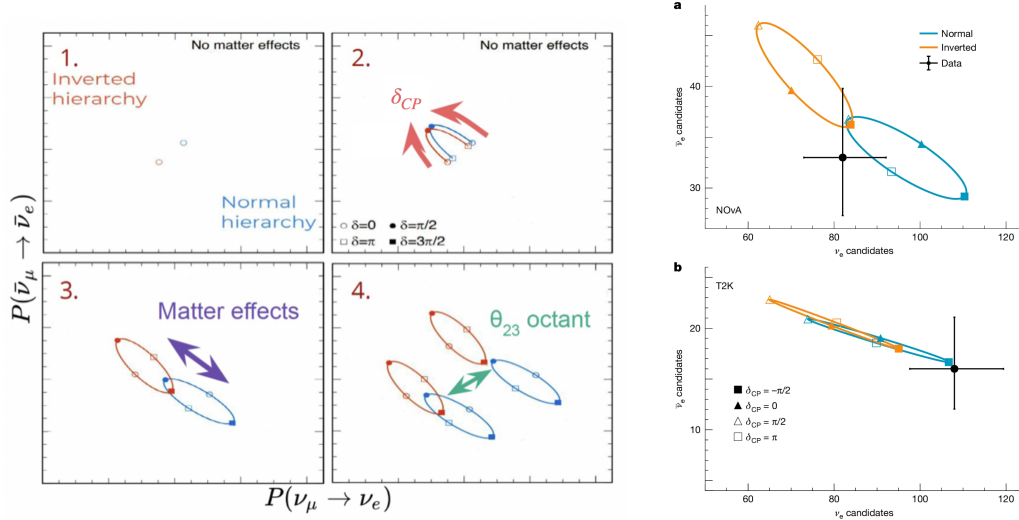


Figure 4.3: Illustration of NOvA’s sensitivities to δ_{CP} , the sign of Δm^2_{31} (the mass ordering), and the octant of θ_{23} by comparing the rate of $\nu_\mu \rightarrow \bar{\nu}_e$ vs $\nu_\mu \rightarrow \nu_e$ appearance between FHC and RHC modes through the $\nu_\mu \rightarrow \bar{\nu}_e$ appearance channel (left). The right plot, taken from [54], shows how this bi-probability plot compares with the T2K experiment, which has a shorter baseline and is therefore less susceptible to matter effects. The data point reflects the 2020 NOvA results (right, top) and T2K Run 1–10 (right, bottom). The effect of the octant is not shown on the right plot.

the mass orderings (clockwise for IO and counter-clockwise for NO). δ_{CP} causes an increase in $\nu_\mu \rightarrow \bar{\nu}_e$ appearance and decrease in $\nu_\mu \rightarrow \nu_e$ appearance, and the opposite is true for $\delta_{CP} = \pi$. The second is matter effects arising from the fact that the neutrinos travel through ordinary matter (Section 1.4). This generally leads to an increase in $\nu_\mu \rightarrow \bar{\nu}_e$ appearance probability and decrease in $\nu_\mu \rightarrow \nu_e$ appearance probability under NO, but the opposite is true under IO (panel 3.). Finally, if $\theta_{23} = 45^\circ$, these ellipses split depending on whether δ_{CP} is in the UO ($\delta_{CP} = 0$, greater $\nu_\mu \rightarrow \bar{\nu}_e$ appearance) or LO ($\delta_{CP} = \pi$, less $\nu_\mu \rightarrow \bar{\nu}_e$ appearance) (panel 4). Due to the conglomeration of all these effects and low energy resolution of events, the $\nu_\mu \rightarrow \bar{\nu}_e$ appearance channel does not provide as much sensitivity to δ_{CP} , relying instead on the value from $\nu_\mu \rightarrow \nu_e$ disappearance.

T2K (Tokai to Kamioka) is another long baseline neutrino oscillation experiment that uses the J-PARC Neutrino Beamline in Tokai, Japan and Super-K as the far detector with 295 km 0.6 GeV at a 44 mrad (2.5°) off-axis angle. Due to the shorter baseline, the majority

of the effect resulting in asymmetry between μ^+ appearance and μ^- appearance would arise from δ_{CP} . T2K is therefore more sensitive to this parameter. For NOvA, however, the higher degree of matter effects due to its longer baseline creates a greater separation between NO and IO. NOvA is therefore more sensitive to the mass ordering except in the case where the effect from δ_{CP} is equal and opposite, i.e. normal ordering with $\delta_{CP} = 0$ and inverted ordering with $\delta_{CP} = \pi$ are degenerate.

4.2 Event Selection

The sources of background for the NOvA μ^+ and μ^- selections can be broadly separated into the cosmic ray background and beam-related backgrounds. Geometric information can be used to separate cosmics from beam events. Since neutrinos do not deposit energy in the detector until their interaction, an event with a reconstructed vertex near the detector walls is more likely to be a cosmic. This is known as a “fiducial volume” cut. Alternatively, a “containment” cut removes events with any reconstructed prong activity near the walls. These cuts are typically tighter for the top wall compared to the others. Beam events must also have a reconstructed neutrino direction that aligns with the beam axis.

For the μ^+ appearance selection ($\mu^+ \rightarrow \mu^+ \nu_\mu$ CC), the beam-related backgrounds include $\mu^+ \rightarrow \mu^+ \nu_\mu$ CC, NC, “wrong-sign” $\mu^- \rightarrow \mu^- \nu_\mu$ CC, and the intrinsic beam $\mu^+ \rightarrow \mu^+ \nu_\mu$ CC. $\mu^+ \rightarrow \mu^+ \nu_\mu$ CC can be distinguished from $\mu^+ \rightarrow \mu^+ \nu_\mu$ CC and NC based on the topological patterns of the resulting particles and has been done with CNNs since the 2017 analysis (Section 3.2). NC is the most significant background for the μ^+ selection due to the difficulty of separating electromagnetic showers from electrons vs photons. Although theoretically possible and important for atmospheric neutrino oscillation analyses, NOvA does not attempt to separate the wrong-sign background due to the difficulty and low contamination in the beam profile. It is not possible to separate an oscillated μ^+ from a unoscillated intrinsic beam μ^+ based on event topology, but neutrino energy can be used since oscillated μ^+ must be near the 1.5 GeV oscillation maximum.

The disappearance selection (CC) is comparatively a purer sample due to the ease of identifying muon tracks. “wrong-sign” events are considered signal because the disappearance probability is not strongly affected by any of the ϵ_{CP} or matter effects. The beam-related backgrounds are therefore only CC and NC. As cosmics are predominantly muons, their tracks may be confused as ν_{μ} events.

The selection procedure is unchanged compared to the 2020 analysis (Figure 4.4) [53]. The basic quality cuts ensure that at least one Kalman track exists in the event and that there are at least 20 hits across at least 4 planes. A “veto” cut is also applied along with the quality cuts to remove events that are clearly cosmics, which removes events with a reconstructed track passing through two walls (“through-going” muon), or passing through one wall and with an angle θ_{Beam} . The containment cut requires all reconstructed prongs to be at least 60 cm from the top wall and 12 cm-18 cm from the other five walls. For ν_{μ} only, Kalman tracks are also required to be at least 6 planes from the front and back walls. Three separate PID cuts are then applied. The first is a cut on the EventCVN CC score. The second is a cut on a BDT trained to separate signal ν_{μ} events from background cosmic events based on event-level and Kalman track-related variables. The third is a cut on “RemID”, a k-nearest neighbors algorithm trained on similar variables. These three cuts are simultaneously tuned to maximize the ϵ_{signal} figure of merit.

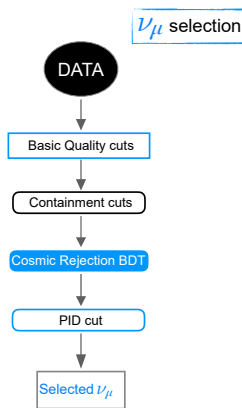


Figure 4.4: Cutflow chart illustrating the ν_{μ} selection.

The selection consists of a “Core” sample providing the bulk of the sensitivity for the analysis, a “Peripheral” sample, and a “LowE” sample new to the 2024 analysis (Figure 4.5). The Core sample follows a similar selection procedure as the selection. The basic quality cuts require that a reconstructed vertex and at least one prong exists. The veto cut is also applied at this step. The “Core Preselection” includes the same prong containment cut, a cut requiring the event to have between 30 and 150 hits, and a cut requiring the longest prong to be between 100 cm and 500 cm. The final selection is similarly performed by several machine-learning based classifiers. The first is a cut on the EventCVN CC score. The second is a cut on a BDT trained to separate signal events from *all* backgrounds based on event-level and prong-related variables.

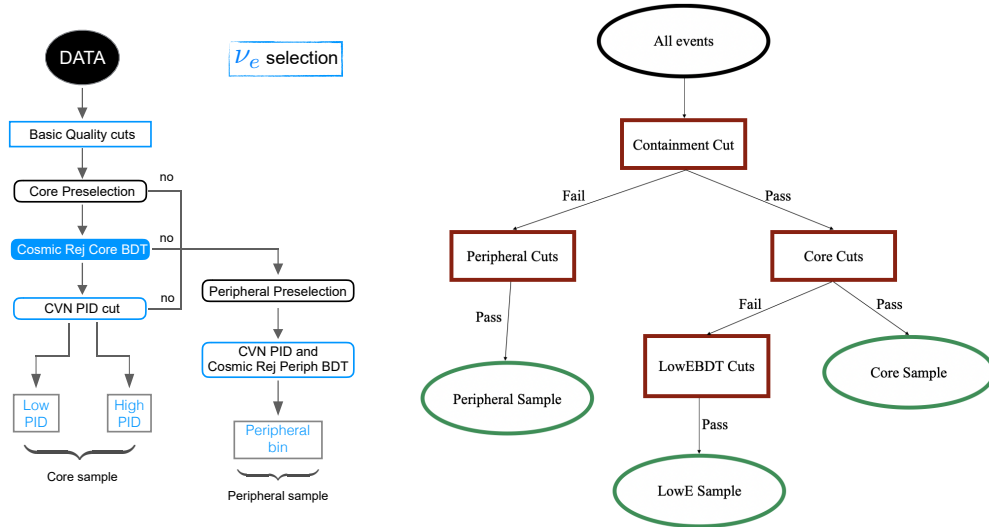


Figure 4.5: Cutflow chart illustrating the 2020 (left) and 2024 (right) selections.

Since appearance events are comparatively rare and valuable for the oscillation analysis, additional effort is made to keep as many of these events in the selection. For the 2020 analysis, an event that fails any of the cuts following the basic quality cuts is subsequently considered for the Peripheral sample as the “Peripheral Preselection”. For this 2024 analysis, however, the separation between the Core and Peripheral samples was simplified by only considering events that fail the containment cut as eligible for the Peripheral sample, since 90% of these previous Peripheral Preselection events failed the containment cut. This

redefinition based on geometry alone results in completely independent Core and Peripheral samples. Cuts on the EventCVN π^0 CC score and a separate BDT model trained for the Peripheral sample are then applied.

The 2024 analysis also features an additional cut on the EventCVN Cosmic score for all samples. This cut is particularly effective in the Peripheral sample with greater cosmic background, increasing the figure of merit $FOM_{\text{Peripheral}}$ by 1.6% in the Core and 26% in the Peripheral sample compared to a 2020-like selection.

To tune the cuts on the machine-learning based classifiers for the Peripheral sample, the cuts on the EventCVN π^0 CC score, EventCVN Cosmic score, and BDT are simultaneously tuned to optimize the figure of merit (Figure 4.8). The Core sample is additionally split into two separate bins based on the EventCVN π^0 CC score. The EventCVN Cosmic score cut, the BDT cut, the EventCVN π^0 CC score cut for the “low PID” sample, and the EventCVN π^0 CC score cut for the “high PID” sample are simultaneously tuned to optimize $FOM_{\text{Low}} + FOM_{\text{High}}$ (Figures 4.6 - 4.7).

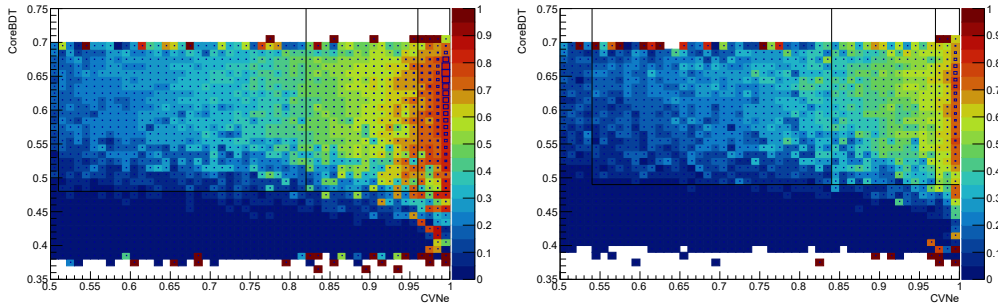


Figure 4.6: The Core PID cut optimization projected onto 2D Core BDT vs EventCVN π^0 CC score for FHC (left) and RHC (right). The color (z-axis) represents signal purity and the box size is proportional to the amount of signal content in each bin. Black lines show optimal cut values.¹

The “LowE” sample is new to the 2024 analysis. As previously discussed in Section 2.1, the

¹Note that there are three vertical lines indicating π^0 CC score cut values. The cuts are originally designed to have three Core PID bins: a “Low”, “Mid”, and “High” PID bin. However, due to the very poor signal to NC background ratio in the low bin, this bin was excluded, and the original Mid bin is referred to as the Low bin in Figure 4.5 and all publications.

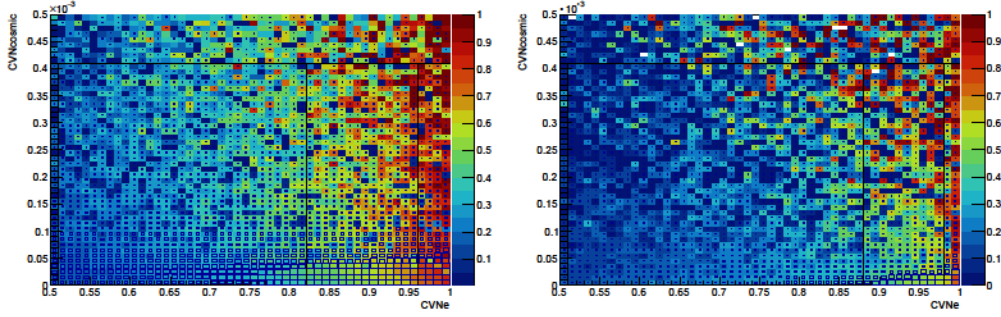


Figure 4.7: The Core PID cut optimization projected onto 2D EventCVN Cosmic score vs EventCVN ν_e CC score for FHC (left) and RHC (right).¹

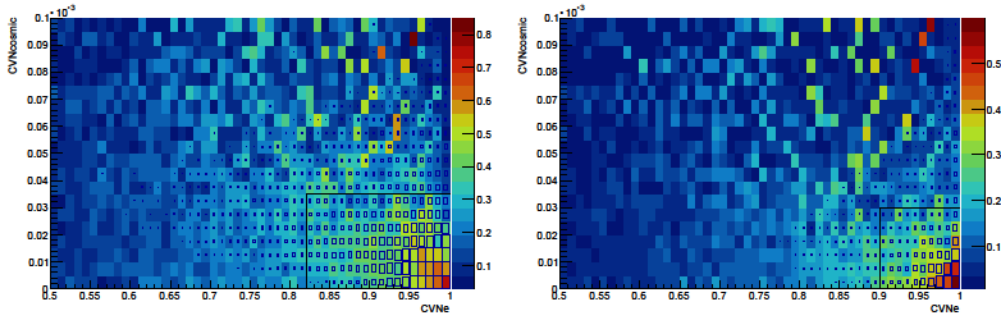


Figure 4.8: The Peripheral PID cut optimization projected onto 2D EventCVN Cosmic score vs EventCVN ν_e CC score for FHC (left) and RHC (right).

neutrino beam energy peak slightly above the energy region with the greatest sensitivity to the oscillation parameters, especially the mass ordering. The LowE sample aims to recapture Core low energy events rejected by the PID cuts. For these contained rejected events, very loose cuts are applied to the PID variables to create the “LowE Preselection”. Another BDT is then trained on slice and prong variables as well as the EventCVN scores themselves as input.² A single cut on this “LowEBDT” output score is then chosen to optimize $\text{FOM}_{\text{LowE}}^2$. This sample is only added to the FHC ν_e selection due to its greater POT as a RHC $\bar{\nu}_e$ LowE sample would have very few events.

Table 4.1 shows the final event counts for all samples for the 2024 analysis.

²Although the procedure for this LowE sample may appear similar to the procedure and motivation for the Super-K work in Chapters 8–13, I was not involved in this work, and the credit for the idea and implementation for the LowE sample goes to Cullen Sullivan.

Sample	FHC			RHC	
			LowE	μ^+	μ^-
μ^+	372.3	4.3	0.3	24.4	0.2
μ^-	24.5	0.1	0.0	71.5	0.2
LowE	0.4	125.3	3.4	0.0	2.1
Beam	0.0	1.8	0.1	0.0	18.9
NC	0.1	26.1	0.8	0.0	6.5
Cosmic	5.5	16.8	5.3	0.8	2.0
Others	4.4	5.5	0.5	0.7	1.1
Signal	1.5	0.8	0.1	0.2	0.1
Background					
Predicted	408.6	180.7	10.5	97.7	31.1
Observed	384	169	12	106	32

Table 4.1: Observed and predicted numbers of μ^+ , μ^- and LowE events in the neutrino beam, and μ^+ and μ^- events in the antineutrino beam [52]. The low PID, high PID, and Peripheral samples have been combined in the μ^+ columns, while the LowE sample is shown separately due to its novel status in this analysis. Wrong-sign is considered signal in the μ^+ and μ^- samples, but background in the μ^+ and μ^- samples. Others refers to the sum of all remaining oscillation channels that are not listed explicitly in this table. The best-fit point for the prediction is extracted from a frequentist fit to the data with the Daya Bay constraint on θ_{13} [55].

4.3 Energy Estimation

The energy estimation methods for the 2024 analysis are all based on traditional fits to reconstructed variables. For the μ^+ -energy related variables, the ProngCVN scores are used to determine the correct energy estimator to separate the electromagnetic (EM) and hadronic (Had) portions of the event. The particle type scores are first condensed into EM and Had scores by adding the softmax probabilities of the applicable particle types.

$$\text{EM} = \sum_{i \in \text{EM}} \text{ProngCVN}_i, \quad \text{Had} = \sum_{i \in \text{Had}} \text{ProngCVN}_i, \quad \pm$$

Each prong is then classified according to the category with the highest score. The deposited calorimetric energy across all prongs classified as EM is first totaled to give E_{EM} . All remain-

ing calorimetric energy in the event (including that not assigned to any prong) is assigned to E_{Had} . A quadratic fit in these two variables based on true neutrino energy from MC is then performed to give reconstructed E_{reco} (Figure 4.9).

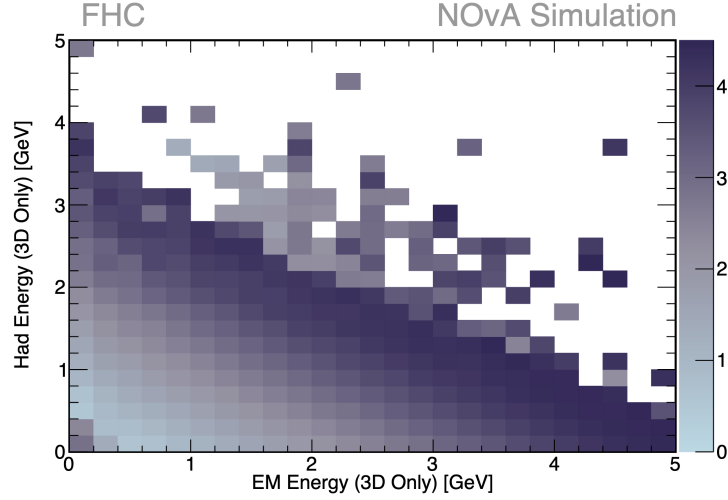


Figure 4.9: Histogram of the weighted average true E_{reco} energy (color scale in GeV) as a function of E_{EM} and E_{Had} [56].

For the E_{reco} -energy related variables, the longest reconstructed Kalman track is considered the muon. A linear fit to true muon energy vs track length is then conducted to give the reconstructed E_{muon} . All other visible energy in the event is attributed to the hadronic portion E_{Had} . A second linear fit is then conducted to true neutrino energy E_{true} vs E_{Had} to give reconstructed E_{reco} (Figure 4.10).

For the Core sample, events are required to have $1 \text{ GeV} < E_{\text{EM}} < 5 \text{ GeV}$ (Figure 4.11). Both the Low and High PID bins are then binned into 0.5 GeV bins. The LowE sample events are required to have $0.5 \text{ GeV} < E_{\text{EM}} < 1.5 \text{ GeV}$ and have two 0.5 GeV bins. The Peripheral sample is required to have $0 \text{ GeV} < E_{\text{EM}} < 4.5 \text{ GeV}$. However, events failing the containment cut may deposit significant energy outside of the detector, making energy estimation difficult. The Peripheral sample is therefore not binned by energy. Instead, a new $E_{\text{EM}} > 2.5 \text{ GeV}$ cut for the Peripheral sample is applied in this analysis to remove background unoscillated events that dominate above the oscillation maximum [57].

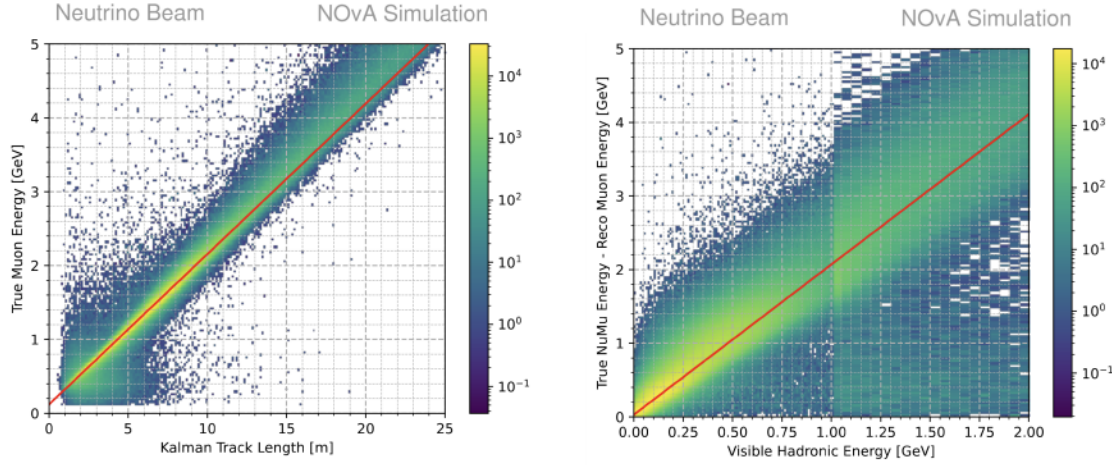


Figure 4.10: Histograms of the true quantity vs the reconstructed quantity for the primary muon (left) and hadronic part (right) in the ν_μ CC events. The red line shows a linear fit to the histogram [56].

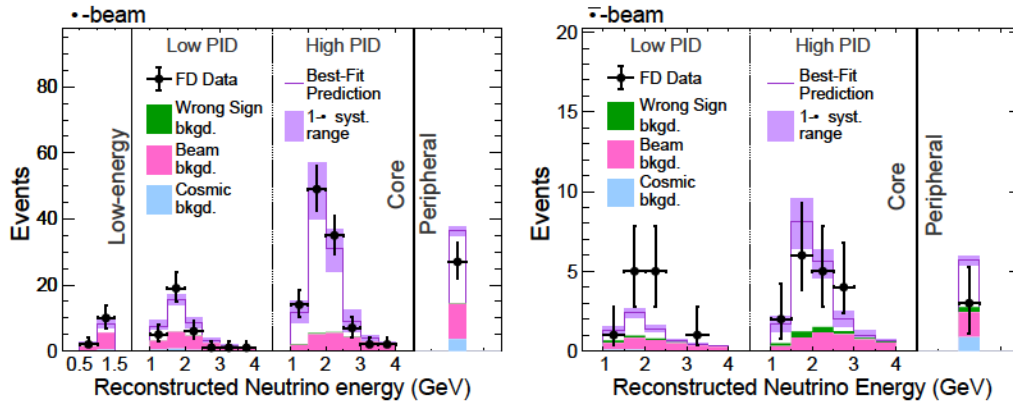


Figure 4.11: The ν_e selection E_{ν_e} spectra for FHC (left) and RHC (right).

For the ν_μ selection, events are required to have $0 \text{ GeV} < E_{\nu_\mu} < 5 \text{ GeV}$ (Figure 4.12). The ν_μ selection is further divided into four quartiles by hadronic energy fraction $E_{\text{Had}}/E_{\nu_\mu}$ [58]. Variable bin sizes are used for E_{ν_μ} , as low as 0.1 GeV in the region of the oscillation dip.

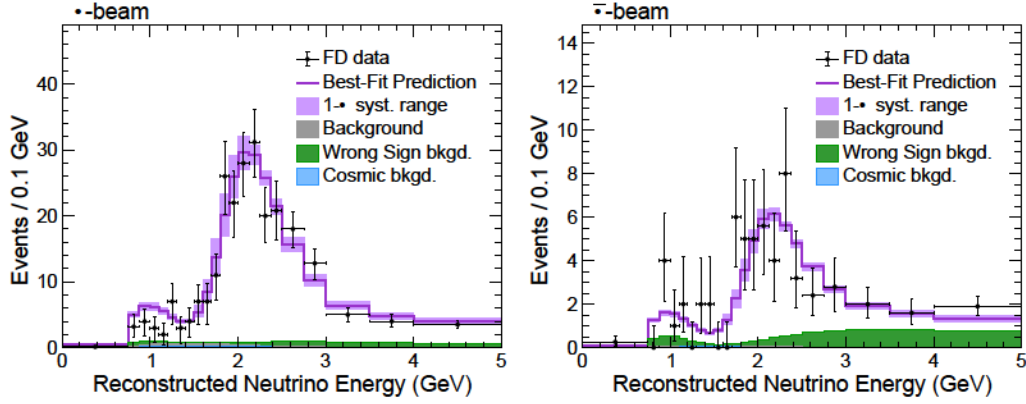


Figure 4.12: The ν_μ selection E_{ν_μ} spectra for FHC (left) and RHC (right).

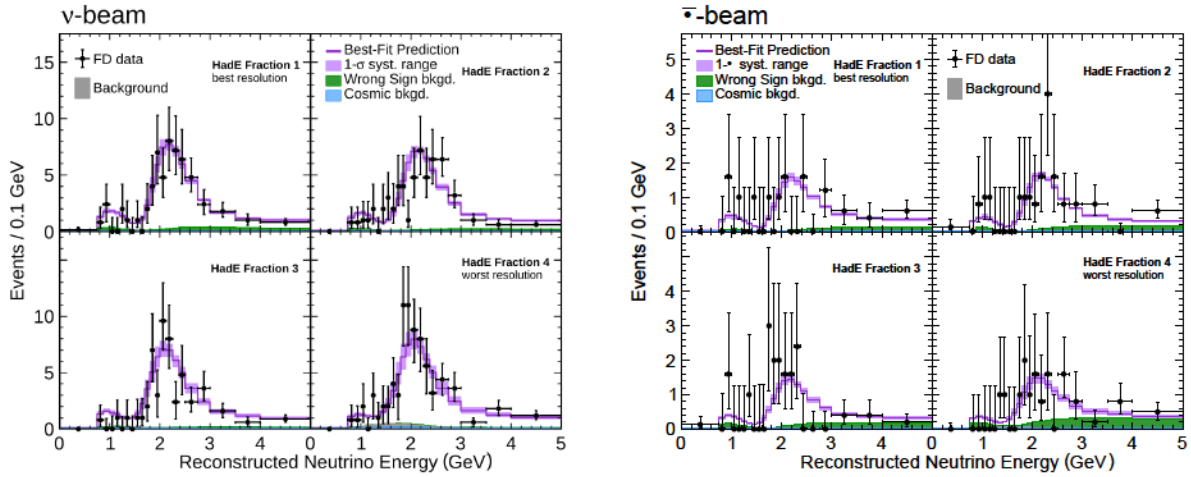


Figure 4.13: The ν_μ selection E_{ν_μ} spectra separated by hadronic energy quartile for FHC (left) and RHC (right).

4.4 Systematic Uncertainties

The impacts of systematic uncertainties are evaluated by varying the simulation via event reweighting or simulating alternative event samples. The event selection and energy estimation procedures are then repeated to produce alternative FD predictions.

4.4.1 Calibration

A major source of uncertainty is the detector calibration. The calibration uncertainty quantifies the potential error in the calibration of the APD’s response to scintillation light measured in ADC values to deposited calorimetric energy. The primary calibration source is cosmic ray stopping muons. The calibration is then applied to MC, and the error in reconstructed vs true calorimetric energy is calculated. A conservative 5% calibration error is assumed corresponding to the maximum observed error [59]. Both a 5% absolute calibration error affecting both FD and ND and a relative 5% error in which FD and ND are uncorrelated are included. A “calibration shape” uncertainty is also included to quantify the difference in a cell’s response between the ends and the middle of the cell.

4.4.2 Detector Response

The light level model refers to the model that determines the actual amount of scintillation light produced for various particles. A general “light level” error is set at 5% by comparing various light models. A separate 6.2% “Cherenkov” light level error is included to account for uncertainties in the amount of Cherenkov light protons produce in addition to scintillation light, and a “detector aging” systematic accounts for the steady 0.3%/year decrease in light level over time.

4.4.3 Lepton Reconstruction

Since the energy of muons are calculated by track length instead of deposited energy, a separate “muon energy scale” error is included for the uncertainty in this calibration, at most 0.15% for uncorrelated effects in the FD and ND.

4.4.4 Cross Sections

The cross sections of neutrino interactions on carbon are extensively studied with the ND [60, 61, 62]. These studies help inform the nominal NOvA MC as well as alternative samples. Alternative samples are produced by varying parameters known as “knobs” in the GENIE simulation. 30 knobs are identified to have large effects on the cross sections, and alternative MC is made for each of these. These include modifications to each of the QE, DIS, Res, and MEC interaction modes, the final state particles, and radiative corrections. For an additional 48 knobs with smaller effects, principle component analysis (PCA) is used to reduce the effects of these 48 to the first 8 principle components.

In addition to neutrino cross sections, a new “GEANT4Reweight” systematic was added to the 2024 analysis to include uncertainties in secondary interactions of hadrons. Knobs for the cross sections of charged pions (5 for different interactions for each of π^+ and π^-) and protons (1 knob) are varied during GEANT4. Neutrons are treated separately in 4.4.6. PCA is performed to reduce these 11 knobs to 3 principle components.

4.4.5 Beam Flux

The beam flux systematic refers to uncertainties in the modeling of the NuMI beam. 100 different PPFX models are used to vary the flux of hadrons produced at the target. Another 20 PPFX models describe adjustments to the parameters of the beam transport including the dimensions of the target, positions of the focusing horns, and their current. Together this makes for 2000 possible beam models. PCA is again used to reduce these to only 5

principle components.

4.4.6 Neutron Model

As neutrons are effectively invisible in scintillator detectors such as NOvA, the neutron model can be a large source of error. MC with an alternative MENATE neutron model with better modeling of neutron-carbon interactions [63] instead of the standard GEANT4 model is used for the neutron model error.

4.4.7 ND/FD Uncorrelated Errors

This error category includes several others effects for mostly uncorrelated sources of error between ND and FD. “Rock” events, referring to neutrinos interacting in the space surrounding the detector with produced particles subsequently interacting in the detector, are not accounted for in FD. A 100% uncertainty in the predicted FD rock rate is therefore assigned as the “rock scale” uncertainty. Cosmic ray events are included in FD predictions as described in Chapter 3. These cosmics are taken from real data, avoiding systematic errors due to cosmic background simulation. Although the cosmic dataset is large after 10 years of NOvA running, we still include a bin-by-bin “cosmic scale” uncertainty based on the statistical error of the cosmic event counts.

Finally, the “normalization” error contains several errors assigned for other small effects. These include a POT accounting uncertainty (0.55%), detector mass accounting (0.19%), air bubbles present in the FD scintillator but not in the ND (0.4%), and pileup of neutrino events in ND (0.5%).

4.4.8 Near-to-Far Extrapolation

The main purpose of the ND in the context of a oscillation analysis is to constrain systematic errors using high statistics measurements near the beam target. Due to the functionally

identical construction of NOvA’s ND and FD, this can be done through “extrapolation”.

The ND MC is first split into its various signal and background components through “decomposition”. The ND MC is then compared to data in each reconstructed energy bin, and this data/MC difference is split proportionally among the components. For each component, a histogram of true energy vs reconstructed energy is then made for both ND and FD MC, which can be used to create a matrix mapping ND to FD reconstructed energy without oscillations. The component’s proportion of ND data/MC difference is then propagated to FD using these matrices. Finally, oscillation weights are applied to create the FD predictions.

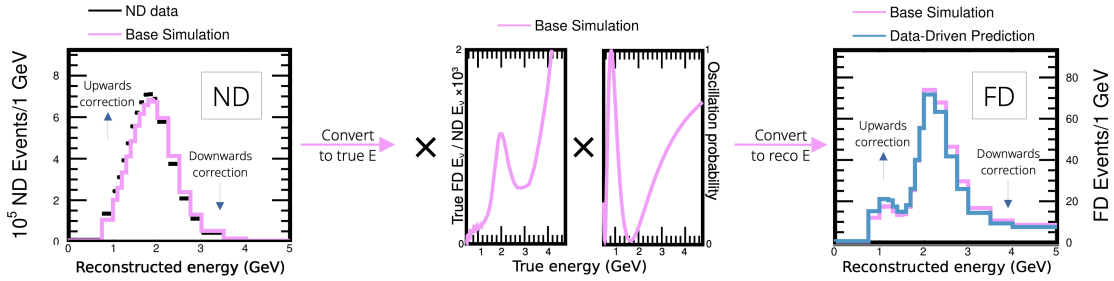


Figure 4.14: Illustration of the extrapolation procedure [51].

This approach is useful for significantly reducing systematic errors that are correlated between FD and ND. Figure 4.15 how systematic uncertainties are affected by using extrapolated FD predictions instead of pure FD predictions. Strongly correlated errors such as cross section uncertainties and beam flux are significantly reduced while the uncorrelated portions of the calibration-related errors are slightly increased.

4.4.9 Error Summary

Table 4.2 shows a summary of systematic uncertainties on the oscillation parameters. The calibration and lepton reconstruction errors are the largest errors for θ_{13} due to its dependence on the position of the oscillation dip in the ν_μ energy spectrum. The cross section uncertainty is considerable for θ_{13} and δ_{CP} . However, all parameters are statistics limited.

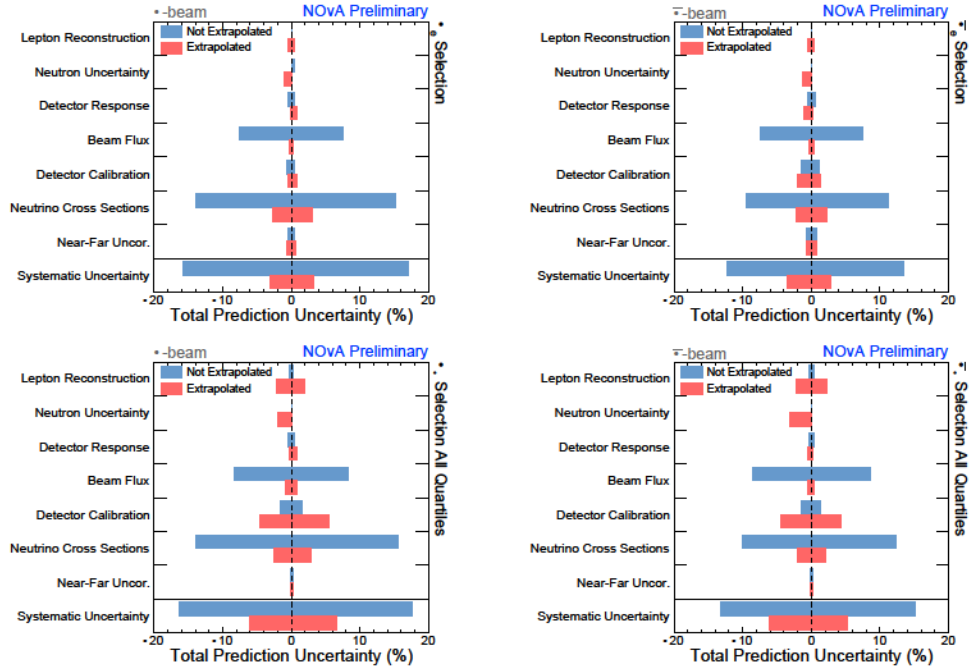


Figure 4.15: Summary of all errors on the integral number of ν_e (top) and ν_μ (bottom) signal and background events for FHC (left) and RHC (right). The two bars compare extrapolation (red) to no extrapolation prediction (blue).

Source	$ \Delta m_{32}^2 $ ($\times 10^{-5} \text{ eV}^2$)	$\sin^2 \theta_{23}$ ($\times 10^{-3}$)	δ_{CP} ($\times 10^{-2} \pi$)
Beam flux	+0.06/-0.12	+0.40/-0.75	+0.16/-1.00
Calibration	+1.40/-1.80	+3.20/-16.0	+1.70/-17.0
Detector response	+0.17/-0.27	+0.36/-3.90	+0.37/-3.30
Lepton Reco.	+0.90/-1.40	+2.70/-4.50	+0.61/-2.90
ND – FD Uncor.	+0.25/-0.34	+2.40/-2.50	+0.79/-4.00
Cross Sections	+0.65/-1.10	+3.10/-5.10	+2.00/-9.70
Neutron model	+0.25/-0.39	+2.70/-0.49	+0.43/-1.30
Systematic Unc.	+1.70/-2.30	+6.60/-18.0	+2.90/-21.0
Statistical Unc.	+3.20/-4.40	+22.0/-85.0	+6.60/-74.0

Table 4.2: Systematic and statistical uncertainties on the oscillation parameters Δm_{32}^2 , $\sin^2 \theta_{23}$, and δ_{CP} evaluated at the best-fit point extracted from a frequentist fit with the Daya Bay 1D constraint on $\sin^2 2\theta_{13}$ [55]. These predicted $\pm 1\sigma$ ranges are given as absolute error on the best-fit values. Asymmetric errors on δ_{CP} and $\sin^2 \theta_{23}$ arise from limited sensitivity and degeneracy in parameters, leading to several highly probable regions that vary with the best-fit values.

Chapter 5

Bayesian Analysis Framework

Prior to the 2024 analysis, all NOvA oscillation analyses were conducted primarily under a frequentist framework. The 2020 analysis was the first to be later reanalyzed under a Bayesian framework as a cross-check [64]. Bayesian sampling does not require the computationally-expensive Feldman-Cousins corrections necessary for a frequentist fit on low statistics (e.g. ν_{μ} appearance) and bounded parameters (e.g. θ_{CP}) [65, 66], so it is considerably faster to accurately complete and does not require extensive computing time on an off-site computing grid such as NERSC. A single fit³ can be used to generate all 1D credible intervals as well as any 2D credible contours between any pair of parameters. Separate fits are also not necessary to implement different external constraints, as these can be implemented through a straightforward reweighting of the posterior. For these reasons, the Bayesian results were considered the primary results for the 2024 analysis with the frequentist results serving as the cross-check for a selection of contours.

³I acknowledge that the word “fit” may be inappropriate in the context of a Bayesian framework but will occasionally use it here to indicate a set of MCMC chains performing Bayesian sampling.

5.1 Markov Chain Monte Carlo

The Bayesian sampling procedure is grounded in Bayes’s theorem and is performed through Markov Chain Monte Carlo (MCMC) [67, 68]. MCMC makes it possible to efficiently sample a high-dimensional space, which is necessary for this analysis in order to include the many systematic uncertainties as nuisance parameters to sample. Each MCMC chain begins with a “prior” distribution for each parameter being measured. We use a uniform prior for the 4 oscillation parameters as measured: θ_{12} , θ_{23} , θ_{13} , ϕ_{CP} . [64] determined that the choice of prior did not have a large effect on the results except in the case of ϕ_{CP} . For ϕ_{CP} , we consider flat priors in $[-\pi, \pi]$ in addition to $[0, 2\pi]$. This prior is also set to 0 outside of $[-\pi, \pi]$ to address the cyclical nature of the parameter. 1 additional binary parameter representing the mass ordering is also included in order to address the discontinuity in θ_{13} .

In addition to the oscillation parameters, systematic uncertainties described in Section 4.4 in a total of 68 implementations are also included as nuisance parameters for the MCMC chain to explore. These uncertainties post-extrapolation affect the FD predictions as the oscillation parameters would. The priors for these parameters are Gaussian with width equal to their corresponding uncertainty. All systematics are considered uncorrelated a priori.

A random first point is chosen in the available sample space. At each step in the chain, a new point in the sample space is selected, and this point may be either accepted or rejected. A target distribution based on Bayes’s theorem, $P(\theta | D) \propto P(D | \theta) P(\theta)$, determines the posterior probability of θ given the data. If the new point is more likely, it is always accepted. If the point is less likely, it may be rejected with some determined probability, and the chain returns to the previous accepted point. Various methods and criteria exist for the selection of subsequent points. At the termination of the chain, the histogram of a parameter’s value across all steps forms the 1D posterior distribution for the parameter.

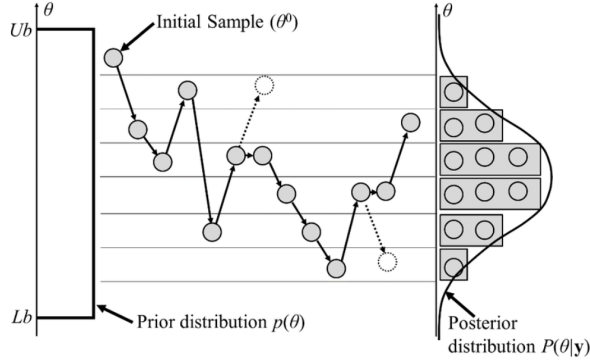


Figure 5.1: Illustration of MCMC, with the ARIA approach in particular [69].

NOvA uses two different MCMC methods for the oscillation analysis, Metropolis-Hasting [70, 71], more commonly referred to now as MR T to credit all original inventors, and Hamiltonian Monte Carlo [72]. At NOvA, we nickname these methods according to the inventors: “ARIA” for MR T and “Stan” for Hamiltonian Monte Carlo.

ARIA is a more straightforward, brute-force approach to point selection. The points are chosen randomly, in this case drawing from a multivariate Gaussian centered at the current point. Since this approach results in high correlation between subsequent steps, the chain is thinned out at the end, removing the majority of steps.

Stan uses a more sophisticated approach to pick subsequent steps. Under Hamiltonian Monte Carlo, the negative log-posterior distribution is seen as a energy potential. To generate a proposal, a supposed particle at the current point is given a random momentum, and its Hamiltonian dynamics within that potential are simulated. After a specified integration time, the resulting location of the particle in the potential is chosen as the next point. This approach allows the chain to efficiently traverse the target distribution. Rejecting points is much less frequent under Stan. Since subsequent steps are significantly less correlated, chain thinning is also not necessary for Stan. Although each step takes considerably longer to select, it is possible to generate an accurate posterior distribution with much fewer steps compared to ARIA.

In the oscillation analysis, separate fits are run using both ARIA and Stan. For the special binary mass ordering parameter, we always use an ARIA-like approach in which we introduce some probability to force the chain to jump to the other mass ordering. Comparing ARIA and Stan distributions is useful for validation of the fitting framework since the posterior distributions should agree given enough steps (Figure 5.2).

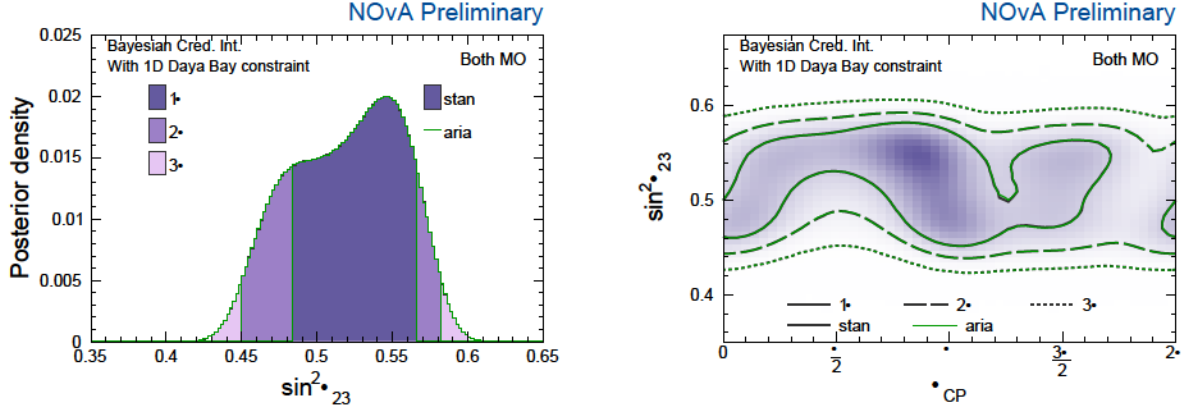


Figure 5.2: Comparison of posterior distributions of $\sin^2(\theta_{23})$ (left) and δ_{CP} (right) between Stan (black) and ARIA (green) MCMC methods. These fits were performed with real FD data, but this validation was first performed with fake FD data before unblinding. The distributions represent 740M ARIA steps after keeping only every 8th step and 125M Stan steps, both collected over 5000 chains.

Following the fits, a “Bayes factor” can be calculated to indicate the preference for one model over the other. This is useful for answering binary questions such as the mass ordering, the θ_{23} octant, or existence or non-existence of CP violation. The Bayes factor is the ratio of posterior probabilities between two hypotheses. However, this approach is not compatible for the case of the CP violation question since CP is conserved only under discrete values of δ_{CP} or J_{CP} rather than a continuous range of values. We therefore use the Savage-Dickey method [73, 74] to calculate Bayes factors. In this approach, the Bayes factor represents the ratio of the posterior probability density to the prior probability density at the bin containing the null hypothesis, i.e. CP conservation.

Chapter 6

Results with 10-Year Dataset

The results shown here are the primary results of the NOvA 2024 analysis [52]: the Bayesian results using Stan MCMC with 125M steps over 5000 chains. A summary of the measured neutrino oscillation parameters is shown in Table 6.1, and a summary of the 1D and 2D posterior distributions is shown as triangle plots in Figures 6.2 - 6.4. 1σ , 2σ , and 3σ credible intervals indicate regions of the posterior distribution that contain 68.27%, 95.45%, and 99.73% of the MCMC steps, respectively. These credible intervals can be calculated by marginalizing either simultaneously over both mass orderings or conditional for each ordering. Because the sensitivity to θ_{13} is limited, external input from the Daya Bay experiment [55] is incorporated to improve precision. Three treatments of θ_{13} are considered: no constraint (using only NOvA data), the Daya Bay θ_{13} constraint (1D Daya Bay), and a two-dimensional Daya Bay constraint on θ_{13} and δ (2D Daya Bay). Figure 6.1 demonstrates the effect of introducing these constraints with further discussion in Section 6.1 on the mass ordering preference. All other oscillation parameters are fixed to their values from the 2023 PDG [75].

Reactor Constraint	Fit Constraint	Parameter	Normal MO			Inverted MO		
			HPD	1	range	HPD	1	range
None	Conditional	eV	0.89	[0.00, 0.24]	[0.50, 1.15]	1.57	[1.32, 1.78]	
			0.547	[0.465, 0.509]	[0.517, 0.562]	0.474	[0.453, 0.537]	
			2.431	[2.392, 2.467]		-2.476	[-2.515, -2.440]	
			8.5	[7.1, 10.0]		9.6	[8.3, 11.2]	
	Non-Conditional	eV	0.89	[0.00, 0.24]	[0.50, 1.15]	N/A	[1.33, 1.78]	
			0.547	[0.450, 0.576]		N/A	[0.466, 0.482]	
			2.431	[2.392, 2.480]		N/A	[-2.491, -2.456]	
			8.5	[6.6, 10.6]		N/A	[9.0, 10.3]	
1D	Conditional	eV	0.93	[0.04, 0.29]	[0.62, 1.14]	1.50	[1.30, 1.69]	
			0.548	[0.484, 0.566]		0.545	[0.489, 0.565]	
			2.431	[2.397, 2.467]		-2.479	[-2.515, -2.443]	
	Non-Conditional	eV	0.93	[0.02, 0.34]	[0.58, 1.15]	N/A	[1.37, 1.63]	
			0.548	[0.459, 0.575]		N/A		
			2.431	[2.376, 2.485]		N/A	[-2.480, -2.472]	
2D	Conditional	eV	0.92	[0.09, 0.28]	[0.59, 1.14]	1.55	[1.33, 1.72]	
			0.550	[0.473, 0.498]	[0.508, 0.569]	0.554	[0.472, 0.492]	[0.510, 0.574]
			2.439	[2.408, 2.472]		-2.497	[-2.531, -2.467]	
	Non-Conditional	eV	0.92	[0.05, 0.36]	[0.53, 1.16]	N/A		
			0.550	[0.470, 0.571]		N/A		
			2.439	[2.400, 2.477]		N/A		

Table 6.1: Summary of the oscillation parameters measured by NOvA for various reactor constraint options. The “HPD” represents the point with the highest posterior density. The indicated regions correspond to 1-creditable intervals. The results for marginalizing over both orderings and for marginalizing separately over each ordering are shown, though the HPD is always in the normal ordering under non-conditional marginalization.

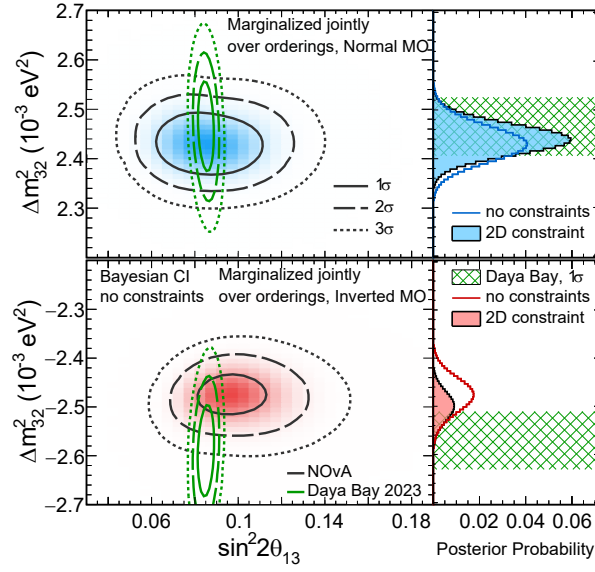


Figure 6.1: Illustration of the 1D and 2D Daya Bay reactor constraint under the assumption of NO (top) or IO (bottom). The binned NOvA posterior shown in color in the left panels is without any external constraints on $\sin^2 2\theta_{13}$, and the green contours show the Daya Bay measurements. The panels on the right illustrate the effect of applying the 2D constraint on the NOvA posterior.

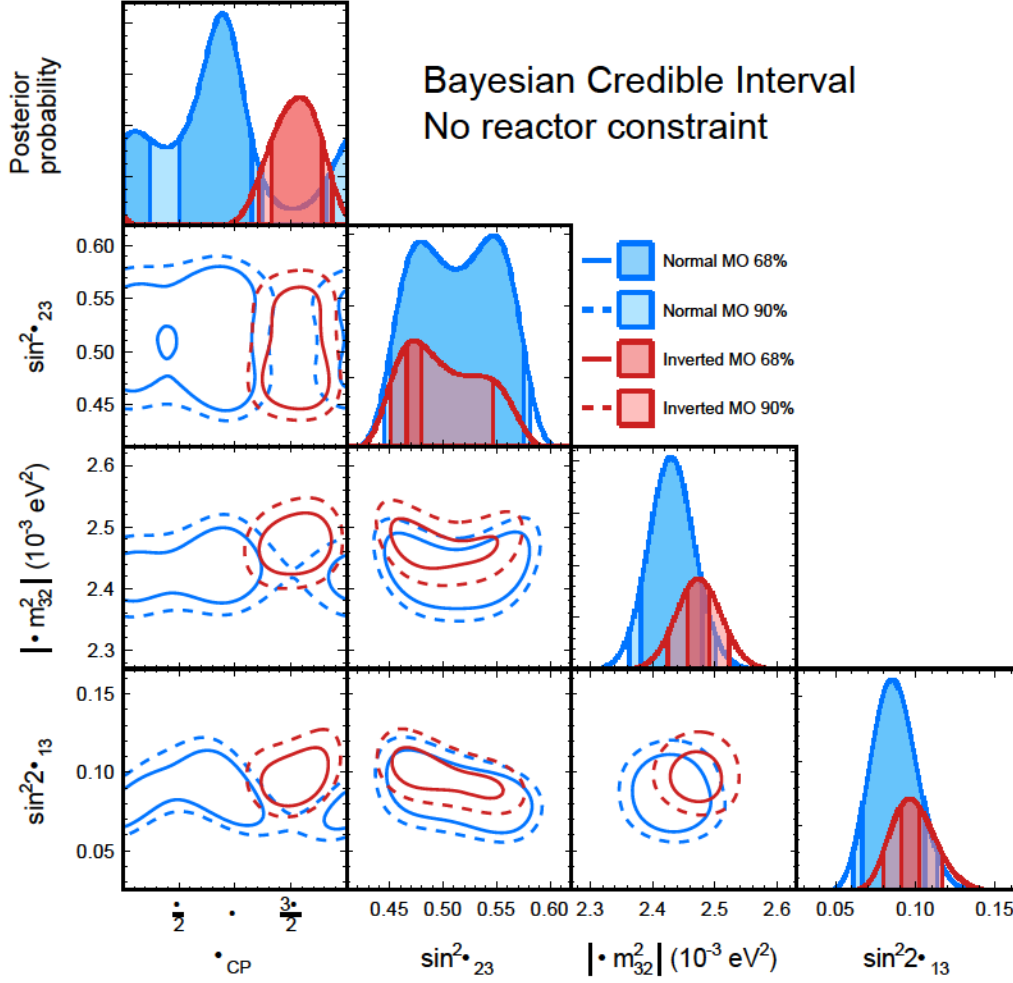


Figure 6.2: Triangle plot of marginalized posterior probability densities. The two-dimensional posteriors have 68% and 90% credible levels indicated. The one-dimensional posterior probability densities are normalized to 1, and have 68% and 90% credible intervals indicated in different shades. These plots are calculated by marginalizing over both mass orderings (MO).

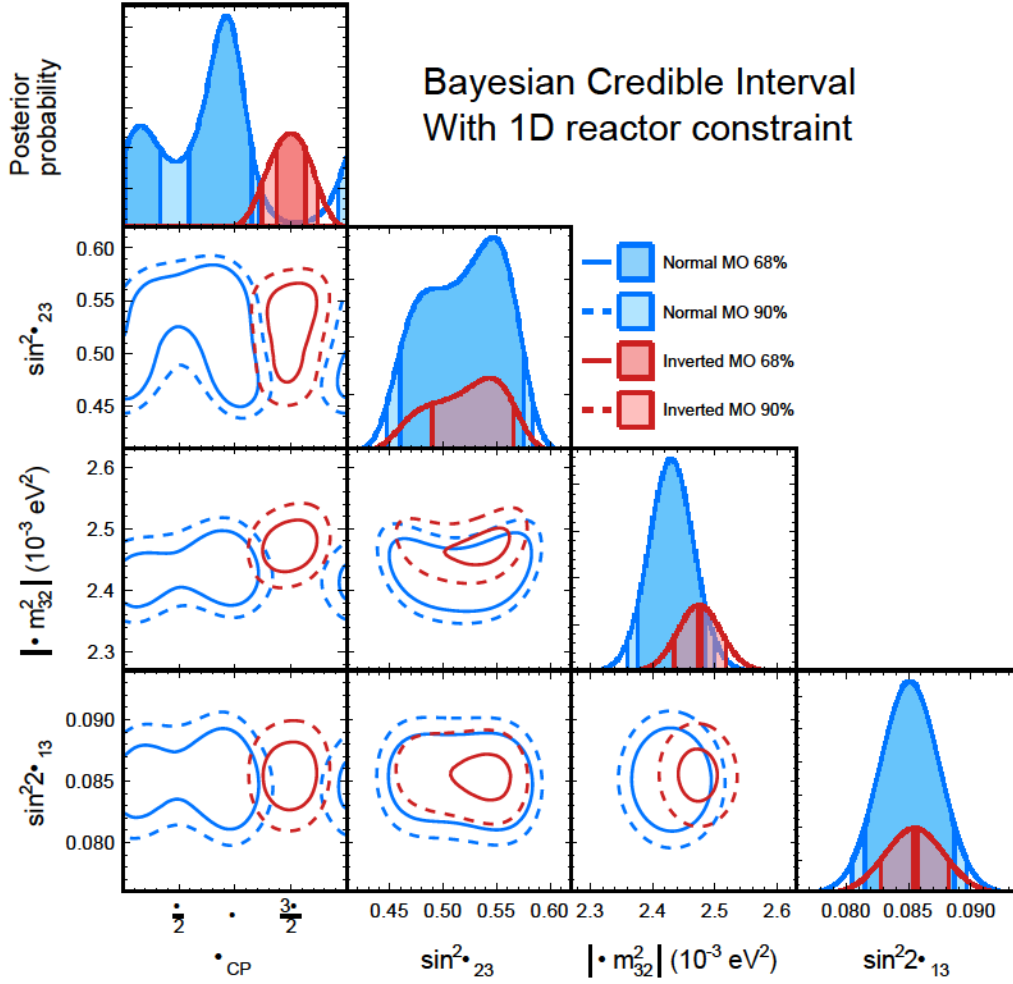


Figure 6.3: Triangle plot of marginalized posterior probability densities with the 1D Daya Bay constraint on $\sin^2 2\theta_{13}$ applied. The $\sin^2 2\theta_{13}$ axis is rescaled accordingly.

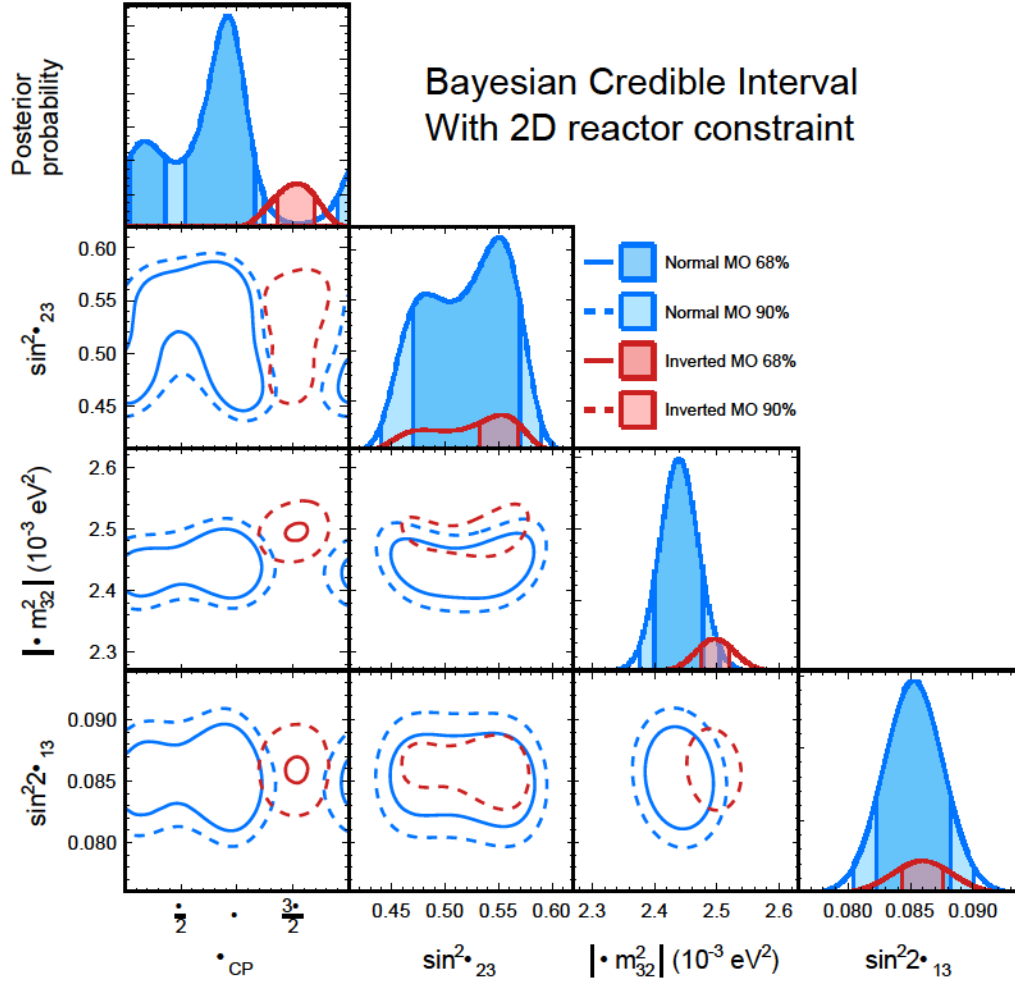


Figure 6.4: Triangle plot of marginalized posterior probability densities with the 2D Daya Bay constraint on $\sin^2 2\theta_{13}$ and $\sin^2 \theta_{23}$ applied.

6.1 Mass Ordering and Octant

Table 6.2 shows NOvA’s preferences for the normal ordering and the upper octant of θ_{12} , quantified using Bayes factors and the integrated posterior probabilities associated with these hypotheses. Results are presented for three treatments of external reactor constraint: no constraint, and the 1D and 2D Daya Bay constraints, with no strong preferences are found for either the mass ordering or octant of θ_{12} .

Bayes Factors	Unconstrained	1D Daya Bay	2D Daya Bay
Normal Ordering Preference	2.4 (70%)	3.3 (77%)	6.6 (87%)
Upper Octant Preference	1.3 (56%)	2.1 (68%)	2.0 (66%)

Table 6.2: Bayes factors, with percentage posterior probability preferences reported in parentheses, illustrating the preference for the normal ordering over the inverted ordering (top row) and for the upper octant over the lower octant (bottom row). Values are derived from fits using no external constraint on θ_{12} , 1D and 2D constraints from Daya Bay.

Figure 6.5 further illustrates how different constraints on θ_{12} and θ_{13} from Daya Bay affect the MO preference, presenting the 1D posterior probabilities for θ_{12} in both normal and inverted orderings across the three choices on external constraint. As can be seen by comparing the θ_{12} posteriors in Figure 6.5 and the conditional HPDs in Table 6.1, the preferred θ_{12} value is consistent under the various reactor constraint options for the normal ordering but shifts in the inverted ordering after applying Daya Bay’s constraint $\theta_{12} > 45^\circ$. Therefore, the preference for the normal ordering is the greatest under the 2D constraint at 87% since the normal ordering implies greater consistency between NOvA’s and Daya Bay’s measurements.

Under the 1D constraint, NOvA has a slight 68% preference for the upper octant where $\theta_{12} > 45^\circ$, with a measurement of $\theta_{12} = 50^\circ$, although it decreases to 66% under the 2D constraint.

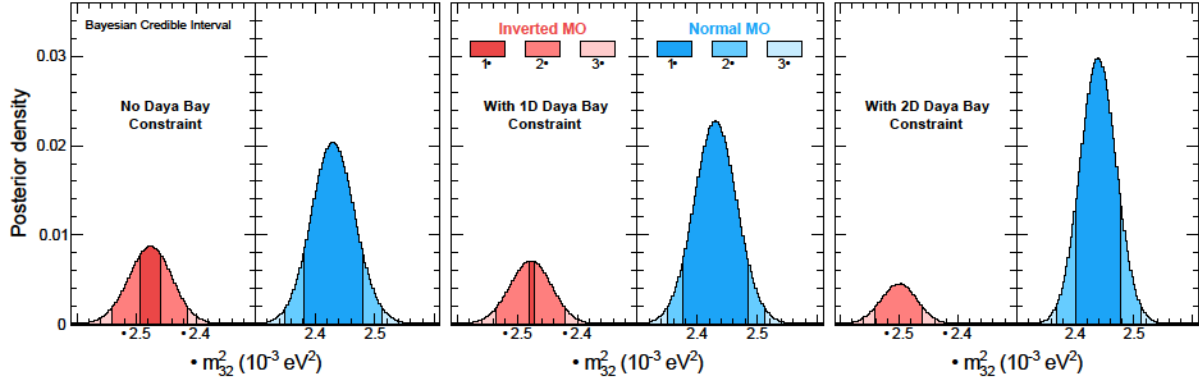


Figure 6.5: Marginalized posterior probability densities for Δm_{32}^2 from a fit to data with different options of the Daya Bay constraint applied (left panels without any constraint, middle panels with θ_{13} constraint, right panels with $\theta_{13}-\Delta m_{32}^2$ constraint).

6.2 CP Violation

Credible regions on the CP violation phase δ_{CP} are presented in Figure 6.6. The two orderings exhibit HPDs in different regions of parameter space. Under the inverted ordering, the HPD is at the CP-violating value 1.5π . Under the normal ordering, the HPD is 0.93π , closer to the CP-conserving value π .

Credible intervals for the Jarlskog invariant, with priors uniform in δ_{CP} and in $\sin \delta_{\text{CP}}$, are shown in Figure 6.7. In the inverted ordering, which is less favored by the NOvA analysis, the majority of the probability density is away from the CP-conserving point at $J_{\text{CP}} = 0$. Bayes factors for the preference of CP violation over CP conservation were extracted using the Savage-Dickey method based on the Jarlskog invariant. The Bayes factors for CP violation are 1.1 (51%) in the normal ordering (nearly equally compatible with both CP conservation and CP violation) and 4.3 (81%) in the inverted ordering.

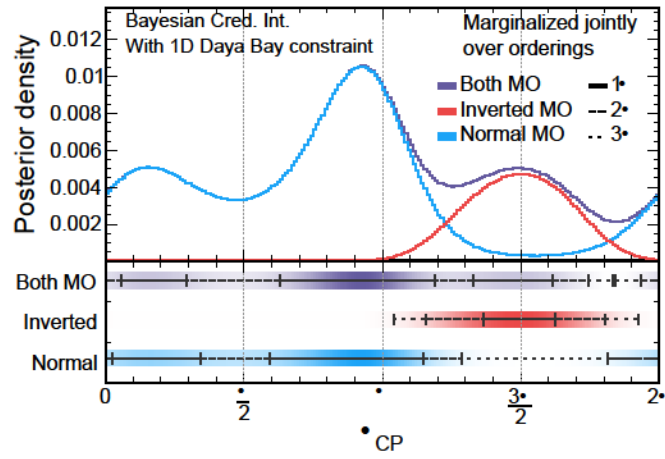


Figure 6.6: Posterior probability densities for δ_{CP} marginalized jointly over the MOs, shown for both MO (purple), inverted ordering (red), and normal ordering (blue), with the 1, 2, and 3 σ intervals shown below. Posterior extracted from a fit with an external 1D constraint on θ_{13} from the Daya Bay experiment.

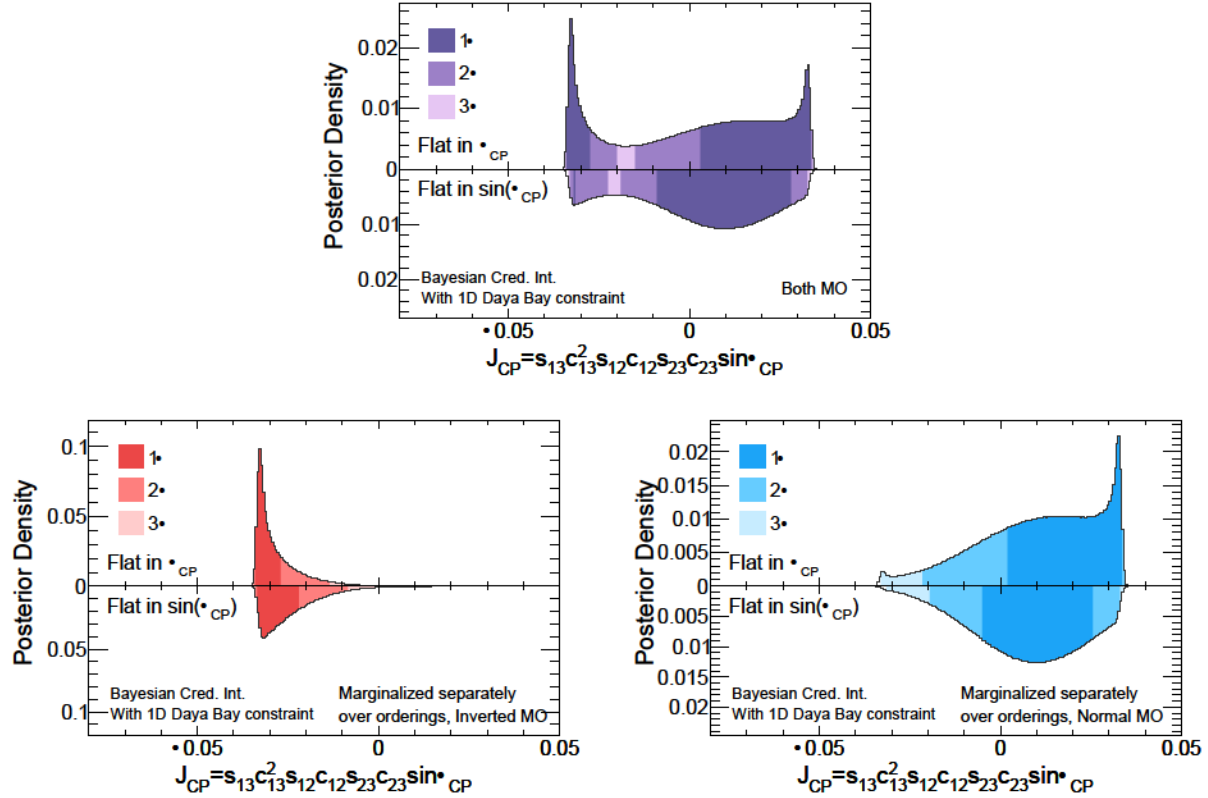


Figure 6.7: Posterior probability densities for the Jarlskog invariant. On each plot, the top panel shows the posterior with a prior uniform in δ_{CP} , and the prior uniform in $\sin(\delta_{CP})$ in the upside-down canvas. $J_{CP} = 0$ represents CP conservation, and non-zero values represent CP violation. The 1D Daya Bay constraint is applied. Plots are shown marginalized over both mass orderings (purple, top) and separately over inverted ordering (red, left) and normal ordering (blue, right).

6.3 Comparisons to Other Experiments

Figure 6.8 compares the measurements of $\sin^2 \theta_{23}$ versus δ_{CP} from NOvA with corresponding results from other accelerator or atmospheric neutrino experiments sensitive to these parameters. In the less favored inverted ordering, the two experiments' credible intervals show strong overlap and are centered around δ_{CP} near $3\pi/2$. In the normal ordering, T2K data tend to favor values of near $3\pi/2$ but NOvA data prefer the $[0, \pi]$ region. However, the 68% credible intervals of NOvA and T2K do overlap, indicating that current data are insufficient to draw strong conclusions about the value of δ_{CP} , even when combining results across experiments.

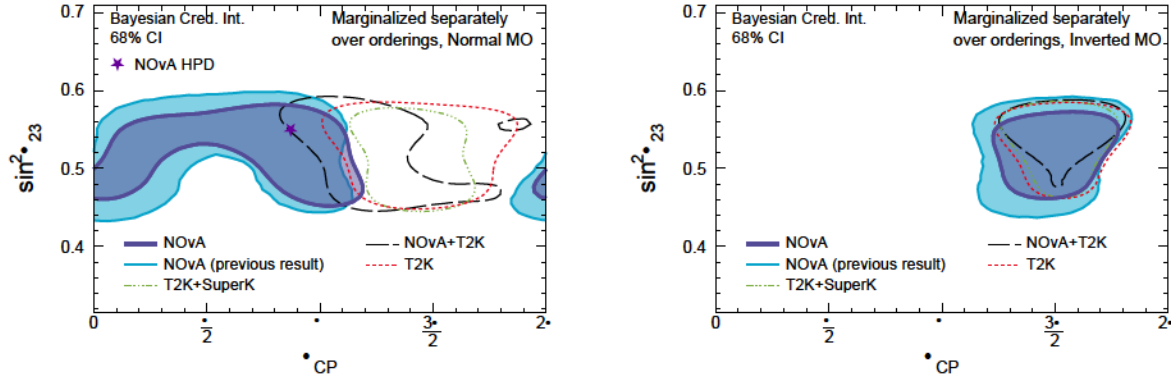


Figure 6.8: NOvA 2024 results for $\sin^2 \theta_{23}$ vs δ_{CP} (purple) in the normal (left) and inverted (right) orderings, compared with NOvA 2020 results [64] (blue) and results from other experiments [76, 77, 78], including the joint NOvA-T2K analysis [54] (black).

A comparison of the NOvA Δm_{32}^2 vs $\sin^2 \theta_{23}$ measurement with similar measurements from other experiments in the normal ordering is shown in Figure 6.9.

Figure 6.10 shows the central values and 1σ credible intervals of Δm_{32}^2 from NOvA alongside recent results from accelerator (green), atmospheric (blue), and reactor (red) experiments, as well as combined fits (purple). Agreement among experiments is stronger in the normal ordering, contributing to the increased preference for this ordering in the NOvA analysis when the 2D reactor constraint from Daya Bay is included. At the moment, the measurement of Δm_{32}^2 with the 1D constraint represents the most precise single experiment measurement

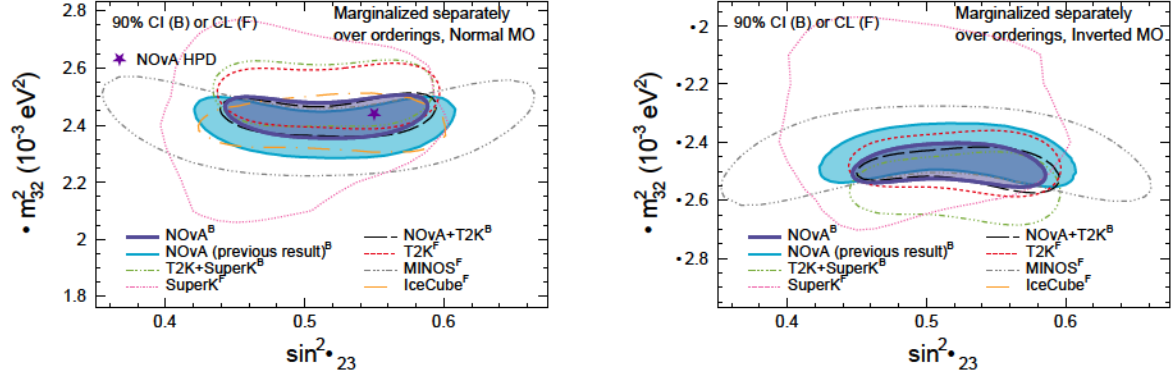


Figure 6.9: NOvA 2024 results for $\Delta m_{32}^2 - \sin^2 \theta_{23}$ (purple) in the normal (left) and inverted (right) orderings, compared with NOvA 2020 results [64] (blue) and results from other experiments [76, 79, 78, 80, 77], including the 2024 joint NOvA-T2K analysis [54] (black). Contours labeled B are from Bayesian analyses, while those labeled F are from frequentist analyses, used when Bayesian results were not available.

at 1.4% (1.5%) precision in the normal (inverted) ordering.

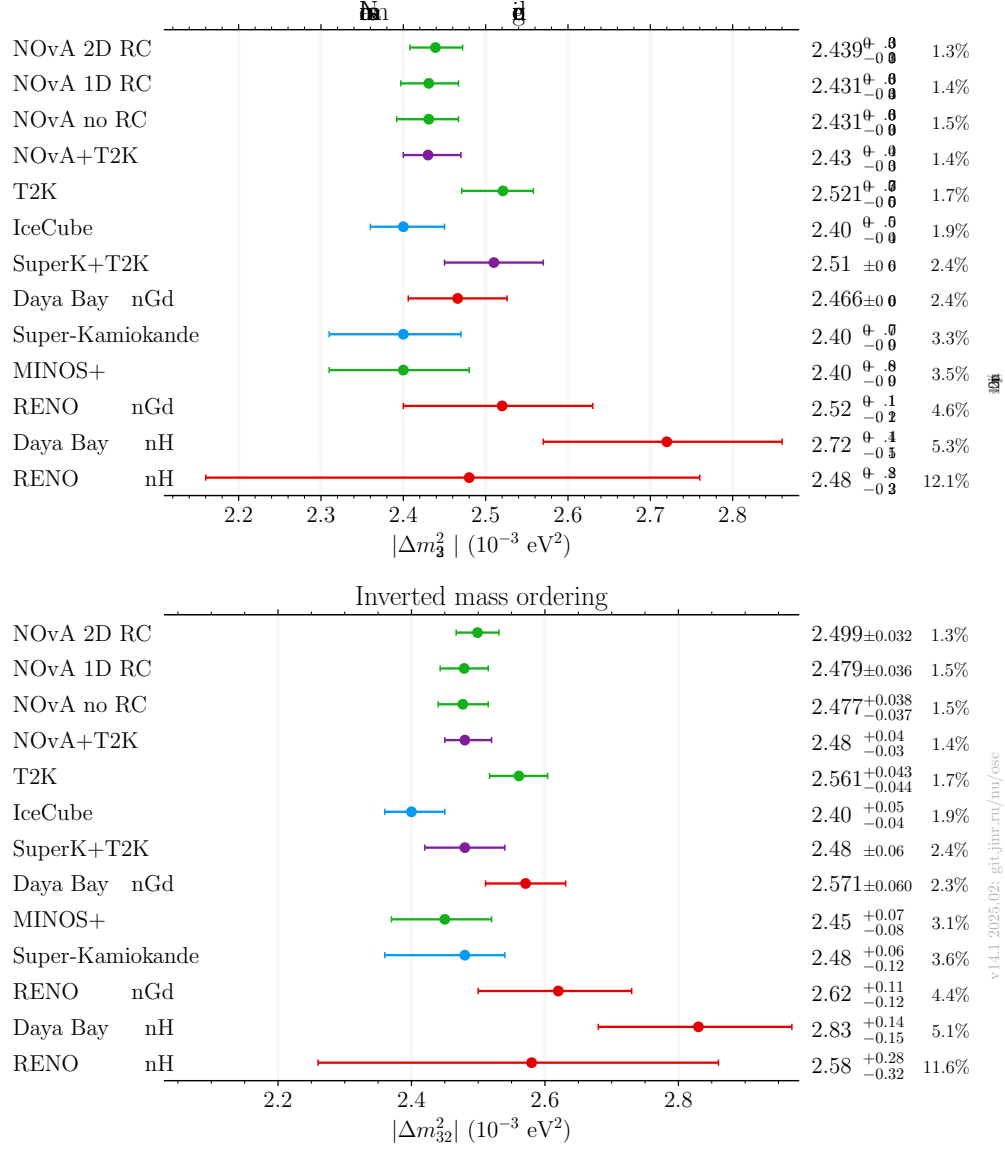


Figure 6.10: Comparison of central values and 1-sigma intervals for absolute value of $|\Delta m^2_{32}|$ in the normal ordering (top) and inverted ordering (bottom) from various accelerator (green), reactor (red), atmospheric (blue) experiments, and the two joint fits (purple). Sources for the results from top to bottom starting with the fourth line are as follows: [81, 79, 76, 55, 77, 80, 82, 83, 84]. Comparisons for other oscillation parameters, can be found in [85].

Chapter 7

Summary and Future Measurements

This analysis of FHC POT and RHC POT neutrino beam data collected over 10 years provides improved constraints on θ_{12} , θ_{13} , and δ_{CP} . Under a constraint from Daya Bay, the preference for the upper octant is 68%, and the preference for the normal ordering is 77%. While there is a preference for CP-violation under the inverted ordering, there is no such preference in the normal ordering with an HPD at $\delta_{CP} = 0$. NOvA currently provides the most precise measurement of θ_{13} , with a precision of 1.5%. This enhanced precision in measuring θ_{13} enables the complementarity between accelerator and reactor experiments to be leveraged by incorporating a 2D θ_{12} – θ_{13} constraint from Daya Bay, which was easily implemented under the Bayesian analysis framework. This constraint strengthens the preference for the Normal MO to 87%.

All of NOvA’s oscillation parameter measurements are very much statistically limited. Preparations are currently underway for one more oscillation analysis planned after the experiment’s conclusion in 2027. This date is approximately determined by the time that portions of the NuMI beam must be decommissioned to prepare for the upgraded beamline for the future DUNE experiment (Deep Underground Neutrino Experiment).

Following the cutoff for data used in the 2024 analysis with 1.5×10^{12} FHC POT and 1.5×10^{12} RHC POT, an additional 1.5×10^{12} RHC POT was collected in 2024 (Figure 4.1), aiming to match the total FHC dataset. In August 2024, a transformer caught fire at a Fermilab substation that supplies power to the main injector. Following additional inspections, other transformers were deemed in need of repair, leaving 3 of 8 operable transformers. Although the BNB was able to resume operations soon after, the remaining infrastructure could not support NuMI beam running. Repairs to the accelerator and power infrastructure are scheduled to complete in October 2026. During the end of 2026 into early 2027 before the end of NuMI operation, NOvA expects to collect an additional 1.5×10^{12} POT, in FHC mode at the request of other experiments.

This analysis (referred to here as the “2027 analysis”) will therefore have approximately 1.5×10^{12} FHC POT and 1.5×10^{12} RHC POT. As discussed in Chapter 3, this analysis will feature several changes to the reconstruction chain. These include the TransformerCVN network for event and particle classification, the VertexCVN network for vertex regression, and the OPALProng view-matching algorithm. Several machine learning algorithms for energy regression have also been included in the 6th production campaign for their potential use in the 2027 analysis. One of these is “RegCVN”, a term for several CNNs for π^0 , η , electron, and muon energy regression that use the same pixel map images as input [86, 87]. Another is “TransformerEE”, a transformer model that takes reconstructed event and prong variables as input for π^0 and muon energy regression [88, 89]. Other improvements to this production include updated event simulation with GENIE v3.6.2 with more accurate QE simulation and cross-section weights, updated detector simulation with GEANT4 v11.0, particle-specific calibration uncertainties using data from the NOvA test beam detector [90], and using MENATE (Section 4.4.6) as the default neutron model.

On a more immediate timescale, the JUNO experiment (Jiangmen Underground Neutrino Observatory) began operations in August 2025. Through reactor neutrino oscillations, it ex-

pects to provide world-leading measurements of oscillation parameters θ_{12} , θ_{13} , θ_{23} , and δ_{CP} [91]. The most important in the context of the NOvA oscillation analysis is θ_{13} , which is expected to be at 0.8% precision with 100 days of data. This will considerably improve on the current 2D reactor constraint from Daya Bay and, depending on its value, may provide considerably stronger constraints on the mass ordering in conjunction with NOvA and T2K within months.

For the θ_{12} and δ_{CP} parameters, next-generation beam neutrino experiments are needed to fully resolve the CP violation and octant questions. DUNE will be the successor long baseline neutrino oscillation experiment to NOvA [92]. It will have a longer 1300 km baseline between the near detector at Fermilab and the far detector at Sanford Underground Research Facility in Lead, South Dakota (Figure 2.1), providing increased sensitivity to the octant and mass ordering. Both detectors will use liquid argon time projection chamber (LArTPC) technology for increased energy resolution and mm-scale spatial resolution. The detectors will be on-axis to the upgraded Proton Improvement Plan II (PIP-II) 1.2 MW neutrino beam, upgradable to 2.4 MW. Many of the NOvA reconstruction methods discussed in Chapter 3 including EventCVN, RegCVN, TranformerEE, and TransformerCVN have been adapted for DUNE LArTPC simulation.

Hyper-Kamiokande (Hyper-K) will be the successor collaboration to both the Super-K detector and T2K long baseline experiment [93]. Since Hyper-K is being built near Super-K, the baseline will be the same at 295 km, 44 mrad off-axis, providing optimal sensitivity for δ_{CP} . The J-PARC neutrino beam will be upgraded to 1.3 MW. The DUNE far detector will begin data collection at the end of 2029 with first neutrino beam data in 2032. Hyper-K and the associated neutrino beam will begin operations in 2028. These experiments are expected to eventually resolve the remaining questions under the standard model 3-flavor neutrino oscillation framework.

Chapter 8

The Super-Kamiokande Experiment

Super-Kamiokande (Super-K or SK) is a cylindrical water Cherenkov detector in Gifu, Japan with 50 kton of ultrapure water [94]. It is located inside the Kamioka mine 1 km under Mt. Ikeno, with a water-equivalent overburden of 2.7 km. Super-K started operations in April 1996 as a successor to the previous Kamiokande experiment. The “NDE” in the name stands for “nucleon decay experiment” in reference to Kamiokande’s original goal of observing nucleon decay, though it may refer to “neutrino detection experiment” in later contexts. Although searches for nucleon decay continue today in Super-K, the primary goal of Super-K is neutrino oscillation measurements using various sources. Super-K studies neutrinos from various sources, including atmospheric neutrinos and solar neutrinos, and contributes major analyses in accelerator neutrinos through the T2K experiment. It also leads the search for the DSNB.

8.1 Detector Design

The Super-K cavern consists of a cylindrical tank portion with a roughly hemispherical accessible dome above it containing detector systems. The detector is divided into a 32.5 kton

inner detector (ID) and a 17.5kton outer detector (OD) by a steel structure 2 m from the tank walls that holds photomultiplier tubes (PMT) for both the ID and OD. Up to 11,146 20-inch PMTs can be installed in the ID, and there are 1,885 8-inch PMTs in the OD. The ID has a diameter of 33.8 m and height of 36.2 m. Super-K analyses use cylindrical coordinates with ϕ oriented along the dome's access tunnels and the origin being at the center of the ID.

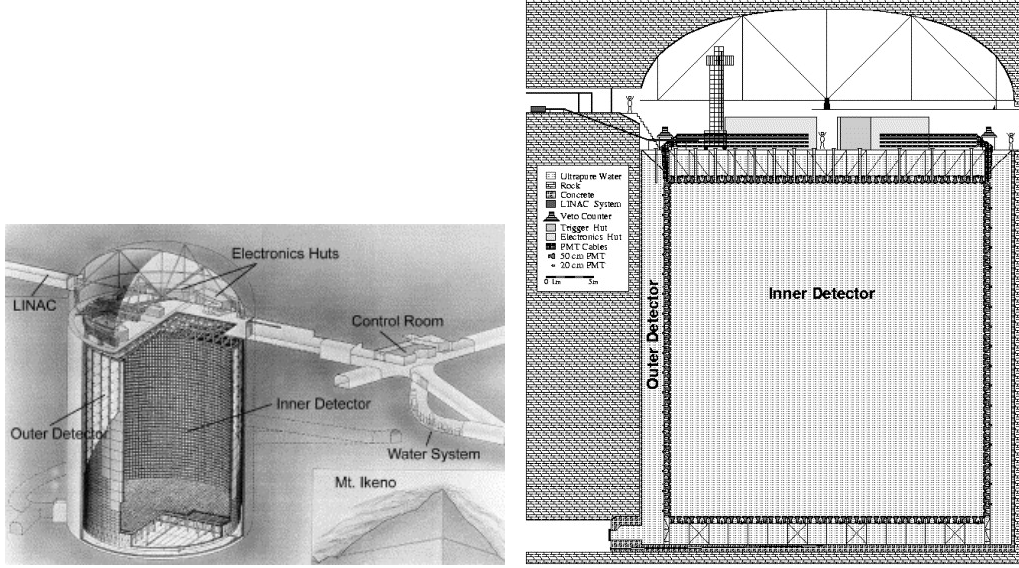


Figure 8.1: Diagram of Super-K [94].

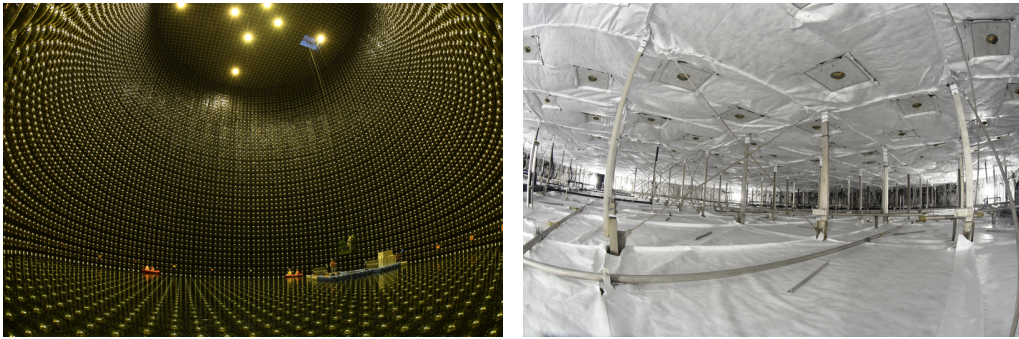


Figure 8.2: Pictures of Super-K ID (left) and bottom of OD (right).

As particles travel through the detector volume with speeds above the speed of light in water, Cherenkov light is emitted as cones with a maximum angle of 42° corresponding to particles with energies well above Cherenkov threshold. This light is then detected by the PMTs. The purpose of the optically-separated OD, acting as an anti-coincidence, is to exclude events

that do not originate in the ID, which is important for rejection of the 2 Hz cosmic ray muon background.

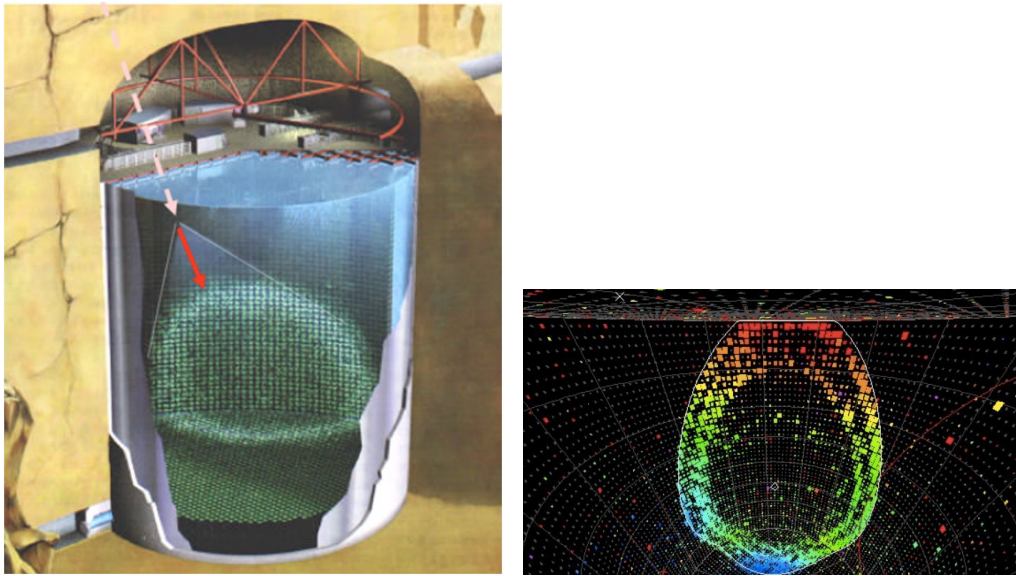


Figure 8.3: Illustrations of Cherenkov cones in Super-K.

External magnetic fields can influence the detection efficiency of PMTs by distorting the path of electrons between the photocathode and first dynode. Super-K therefore has multiple current-carrying geomagnetic compensation coils attached to the edges of the detector wrapped around the inside of the tank. These coils reduce the 450 mG Earth's magnetic field to 32 mG.

The water must be kept at ultrahigh purity both to prevent absorption of Cherenkov light before it can be detected at the PMTs and to minimize radioactive impurities such as that pose a significant background for low energy analyses (Section 10.1). These impurities are found in greater concentrations near the detector walls. Supply water is first drawn from naturally occurring sources in the mine. The water is passed through a 1 μm filter, a heat exchanger to cool the water to 13 $^{\circ}\text{C}$, ion exchange resins, cartridge polishers, a system to dissolve radon-free air, a degasifier to remove dissolved gasses in the water, an ultraviolet sterilizer, and a reverse osmosis filter (Figure 8.4).

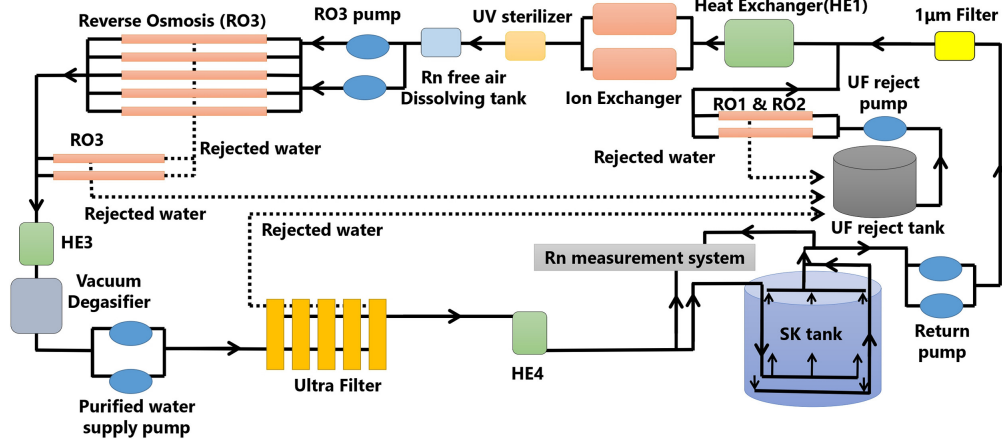


Figure 8.4: Diagram of the Super-K water system [95].

The water in the tank is continuously recirculated at a rate of 60 tons/h, maintaining a typical attenuation length of ~ 100 m. While a greater recirculation rate helps prevent the buildup of impurities, it also increases the mixing between the low-purity regions near the ID walls and relatively high-purity regions near the center. Water currents are therefore carefully designed to increase recirculation near the walls and push impurities toward the bottom half of the detector, keeping the center-top half relatively still. Note the lack of inlets/outlets in the center of the ID top (Figure 8.5). The water temperature must also be controlled to within 0.01°C and kept highly uniform to prevent convection that may disturb the still water region.

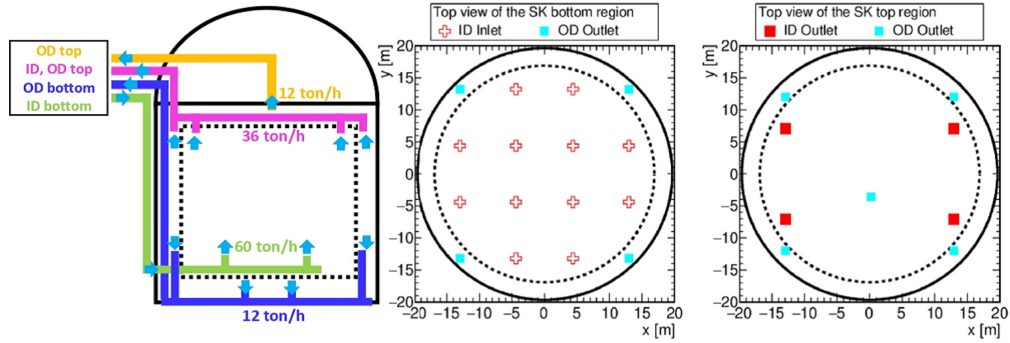


Figure 8.5: Schematic view of the water circulation paths in the Super-K tank (left). The colored arrows show the direction of water flows. The locations of the water inlets and outlets at the bottom (middle) and top (right) of the tank are shown. The solid lines indicate the tank boundary, and the lines indicate the ID boundary [95].

To further reduce the concentration of radon, a 60 cm “buffer gas” layer is kept between the OD water level and the top of the tank. This gas is continuously filtered to remove radon, and the low-radon air degasses the water in the tank.

Water Cherenkov detectors also have the ability to detect neutrons through neutron capture on hydrogen. The neutron is captured after approximately 0.20 ms and produces a characteristic 2.2 MeV photon. In order to further improve neutron detection, gadolinium sulfate was added to the water in July 2020. Gadolinium has a higher neutron capture cross section, resulting in faster 0.11 ms capture times, and with a cascade of several photons with total 7.9 MeV energy.

Table 8.1 shows a list of Super-K phases divided by key events in detector conditions. SK-I is the first phase of the experiment lasting five years until July 2001. However, a PMT implosion incident occurred during the following detector refurbishment. One of the bottom PMTs shattered, and the shockwave in the water from its implosion caused a chain reaction shattering all PMTs in the bottom half of the detector. The remaining PMTs were then evenly redistributed, Super-K resumed limited operations during SK-II until replacement PMTs could be acquired. SK-III started upon installation of the replacement PMTs. The next SK-IV phase is among the most important phases for Super-K analyses lasting from October 2008 to May 2018 with 2970 days livetime. Aside from being the longest of all SK phases at nearly ten years, SK-IV is characterized by improved front-end electronics [96] and low levels of radioactive impurities compared to previous phases. SK-V began with the end of the open tank work in February 2019, and SK-VI marks the start of the SK-Gd era with the dissolution of 13 tons of gadolinium sulfate [97] into the water. Another 26 tons were dissolved before the start of SK-VII in June 2022. Following loss of electrical contact with some of the geomagnetic compensation coils beginning in October 2023, SK-VIII was started in September 2024 after their repair.

Phase	Duration	ID PMTs	Photocoverage	Gd (% w/w)
SK-I	Apr 1996 – Jul 2001 (1496 days)	11146	40%	0
SK-II	Oct 2002 – Oct 2005 (791 days)	5182	19%	0
SK-III	Jul 2006 – Aug 2008 (548 days)	11129	40%	0
SK-IV	Oct 2008 – May 2018 (2970 days)	11129	40%	0
SK-V	Feb 2019 – Jul 2020 (379 days)	11129	40%	0
SK-VI	Jul 2020 – Jun 2022 (552 days)	11129	40%	0.01
SK-VII	Jun 2022 – Sep 2024 (596 days)	11129	40%	0.03
SK-VIII	Sep 2024 – (240+ days)	11129	40%	0.03

Table 8.1: List of Super-K operations phases, their duration, number of ID PMTs, photocoverage of those PMTs, and the concentration of gadolinium dissolved in % by weight.

8.2 Data Acquisition

The PMT waveforms are first processed by front-end electronics, referred to as Analog-Timing-Modules (ATM) before SK-IV [94] and QBEE since SK-IV [96]. These electronics integrate and digitize the waveforms, providing a single raw time and integral raw charge for each hit. Since SK-IV, all subsequent data processing is software-based. There are four “electronics huts” each with five Front-end PCs responsible for one quadrant of the detector. Raw data is transferred to a “central hut” where Merger PCs run a software trigger algorithm. This software trigger is a simple hit coincidence trigger that records an event when the number of collected hits in a sliding 200 ns window exceeds a certain threshold. Events are categorized into “high energy” (HE), “low energy” (LE), and “super low energy” (SLE) triggers depending on the number of hits. This data is then sorted by an organizer PC before being saved to long-term storage.

Toward the end of SK-IV, several efforts were made to increase the trigger efficiency at low energies. Figure 8.7 shows the trigger efficiency over time of SLE triggers for the lowest energy bins. The gradual increase in trigger efficiency is driven by the improvement in PMT gain and water transparency throughout SK-IV [27]. Following a reduction in the radon content in the purified water supply [95], the SLE trigger threshold was reduced from 34

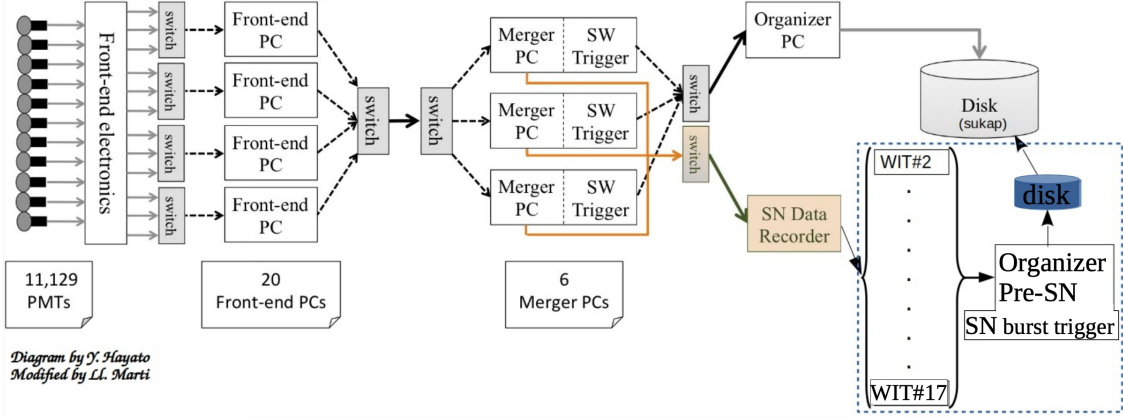


Figure 8.6: Illustration of the Super-K DAQ since SK-IV from the PMT readout to processing for long-term storage through both the software trigger and WIT system.

hits in 200 ns to 31 in May 2015. This change in threshold brought the trigger efficiency at the 3.49 MeV electron kinetic energy³ threshold to nearly within a 22.5 kton fiducial volume defined as >2 m from the ID walls.

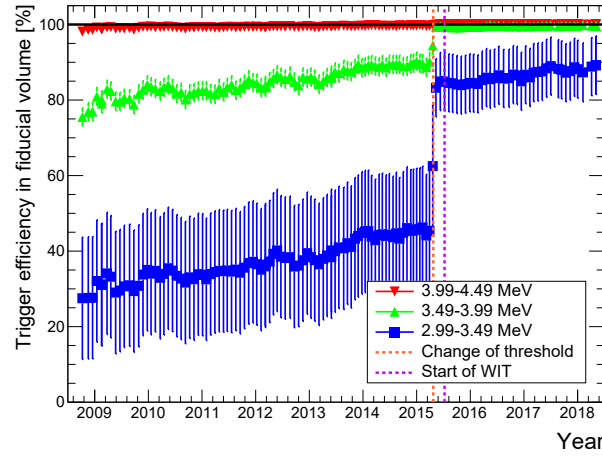


Figure 8.7: The time variation of the SLE trigger efficiency estimated by the MC simulation of solar neutrinos shown as the average of 30-day intervals. The green and red plots represent the lowest two energy bins of [27], and the 2.99-3.49 MeV energy bin (blue) is discussed in this article. After changing the trigger threshold (orange dashed line) from 34 hits to 31 hits, trigger efficiency improved significantly. The start of data-taking with the separate WIT system (violet dashed line) is shown for reference.

³Energy values in the Super-K solar analysis generally refer to the kinetic energy of the recoil electron produced in neutrino elastic scattering, excluding the electron rest mass. While calculations are performed using total energy (0.511 MeV), all values reported here are in kinetic energy unless otherwise noted. For increased clarity, these values are reported with two decimal places, ending in .x9.

In parallel, raw data before the software trigger is sent to a supernova data recorder machine, whose original role was to run a supernova burst trigger algorithm and save raw data in the event of a supernova. However, this functionality was later disabled and is now handled by the downstream WIT system.

8.3 Wideband Intelligent Trigger

The independent WIT system was also installed toward the end of SK-IV in July 2015 [98, 99]. In an alternative approach to improving trigger efficiency, WIT acts as a separate DAQ running in parallel to the hit coincidence trigger. The WIT system consists of a collection of high performance computers and receives the same data as the software trigger as unsorted 23 ms blocks. WIT had 6 computers with 16 CPU cores each during SK-IV and currently has 4 16-core, 3 28-core, and 10 32-core computers. With hyperthreading enabled, each physical CPU core can process two 23 ms blocks in parallel. The WIT trigger algorithm applies a variable and much looser threshold of 11 hits above the current dark noise rate in 230 ns (the “pre-trigger”). The efficiency of the pre-trigger is near 100% down to 2.49 MeV kinetic energy (Figure 8.8) [99]. A second filter known as “STORE” then creates a list of potential vertex positions from the PMT four-hit combinations and requires that at least effectively 6.7 hits are coincident within 5 ns after time-of-flight subtraction for at least one of the potential vertices [100]. The process then conducts full vertex reconstruction with the “BONSAI” algorithm (Chapter 9). Events that have between 6.7 and 11 coincident hits from the STORE filter above are required to pass the 2 m fiducial volume cut, and events with 11 or more coincident hits must simply be within the inner detector volume. Each block is then sent to a separate organizer machine that sorts the blocks in time. The WIT dataset had not previously been used as a source of signal data for the solar event selection, but it has been used in [27] to reject low energy neutron cloud events resulting from background cosmic ray muon spallation [100]. In later phases, the WIT system has been used to serve

as a trigger and raw data buffer for potential pre-supernova and supernova burst neutrinos [101].

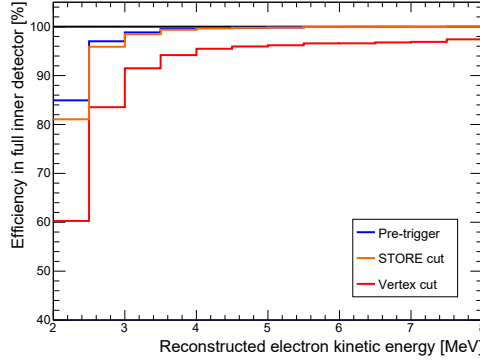


Figure 8.8: The trigger efficiency of the WIT system as a function of reconstructed electron kinetic energy based on solar MC. The blue, orange, and red plots show the cumulative efficiency following the pre-trigger, the cut on the number of hits coincident with potential vertices, and the cut on the final vertex position, respectively.

8.4 Calibration Sources

Super-Kamiokande maintains several calibration sources for determining the energy scale, tuning MC, and calculating systematic errors. An electron linear accelerator (LINAC) was thoroughly integrated into the design of the Super-K cavern (Figure 8.1) and supplies low energy electrons at several energies [102]. The LINAC is housed in a tunnel above the dome floor starting from the edge of the cavern. Electrons are generated via thermionic emission by heating a tungsten filament. The electrons are accelerated along the LINAC tunnel, eventually bending down toward the cavern floor, along the $-z$ -axis of the cavern, then down into the detector (Figure 8.9). Electrons exit through a 0.1 mm titanium window into the water. Energy modes of 3.8 MeV, 4.2 MeV, 5.0 MeV, 6.2 MeV, 8.1 MeV, 13.0 MeV, and 18.4 MeV in electron kinetic energy upon entering the water are available, but the 5.0 MeV and lower modes have been nonfunctional since the end of SK-IV. A germanium detector is used to precisely measure the energy to 0.002 MeV resolution.

A deuterium-tritium (DT) generator is used to perform calibration with neutrons [103]. The

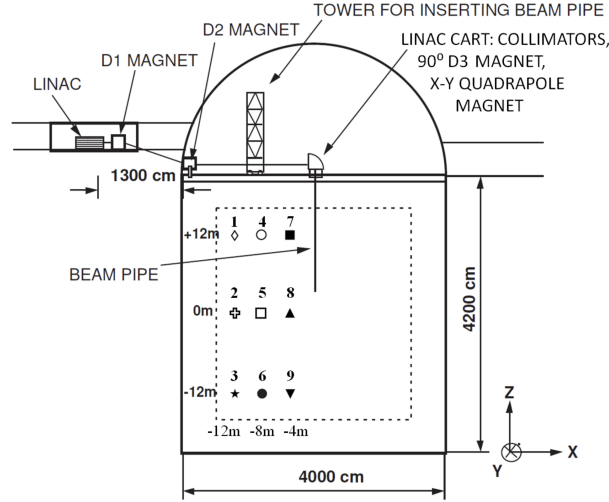


Figure 8.9: Schematic of the LINAC geometry feeding into the detector. Data taking positions are indicated with numbered shapes.

DT is lowered into the tank and then generates 14.2 MeV neutrons via ${}^{16}\text{O}(\alpha, n){}^{19}\text{F}$. The neutrons then produce ${}^{16}\text{N}$ through ${}^{16}\text{O}(n, p){}^{16}\text{N}$. The DT is then raised 1–2 m to reduce shadowing while the ${}^{16}\text{N}$ decays, producing electrons over 40 s: ${}^{16}\text{N} \rightarrow {}^{16}\text{O} + e^- + \bar{\nu}_e$. Although energy calibration is possible with these electrons because the ${}^{16}\text{N}$ spectrum is well known, it is less suitable than the LINAC due to its broad spectrum. The main advantage of the DT is its compact size, allowing for calibration at all calibration ports including along ${}^{16}\text{O}$, ${}^{16}\text{N}$, and ${}^{16}\text{F}$.

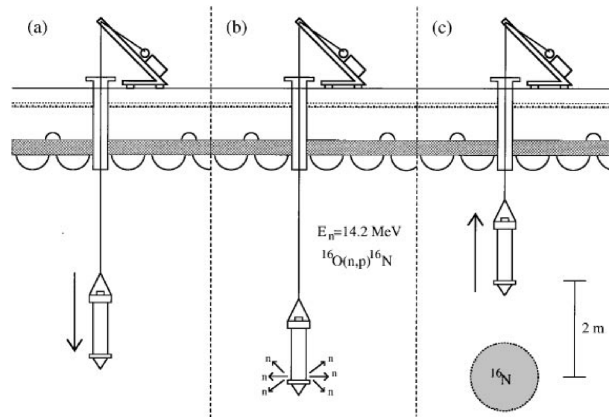


Figure 8.10: An overview of DT data taking. The DT is lowered to the position where data is to be taken (a), fired in at that location (b), and withdrawn 2 m and data collected (c).

Other calibration sources are also maintained for purposes other than calibrating the energy scale. One is a nickel-californium radioactive source [94]. ^{252}Cf undergoes fission, emitting neutrons that are captured by nickel, which then emits 9 MeV gammas. The compact 1 kg ball can easily access any part of the detector, and is therefore used for systematic error calculations such as for the vertex reconstruction.

Some calibration sources are also permanently installed in the detector including an LED and a scintillator ball connected via optical fiber to an external xenon flash lamp. The xenon light is used for low-level calibrations to periodically test the PMT response and calibrate their collected charge values to individual photoelectrons.

Chapter 9

Low Energy Event Reconstruction

40 μs ($-5 \mu\text{s}$ to $35 \mu\text{s}$) of PMT data is saved around each HE and LE trigger, while 1.5 μs ($-0.5 \mu\text{s}$ to $1.0 \mu\text{s}$) is saved for SLE and WIT triggers due to the high event rates. Each of these triggers corresponds to one event record. The collection of PMT hit times in each event is then used for vertex reconstruction. (The charge collected by PMTs is not used during low energy event reconstruction, and each hit is assumed to be one photoelectron.)

The following event reconstruction chain assumes that these events produce only a single electron or positron, as is expected of electron elastic scattering or prompt IBD events. The vertex reconstruction algorithm used since SK-II is BONSAI (branch optimization navigating successive annealing iterations) [104, 105]. Since low energy electrons travel less than 10 cm in the detector, low energy events are treated as point sources at the interaction vertex. BONSAI begins by considering all possible combinations of four hits, each of which triangulates a “test vertex” in space and time. Unphysical test vertices such as those outside of the detector are discarded, and groups of test vertices closer than 0.6 m are combined into single test vertices. For a given test vertex, both the time of flight () and the assumed vertex emission time () are subtracted from each measured PMT hit time () to calculate

the timing residual. A likelihood is then evaluated for the vertex using a probability density function (PDF) (), which is determined based on the PMT timing residual distributions of LINAC calibration events (Figure 9.1).

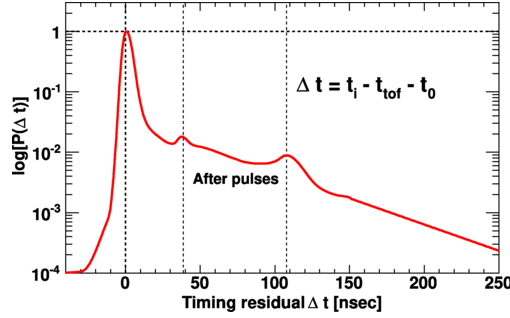


Figure 9.1: The PDF of the timing residuals based on LINAC events used for the vertex reconstruction maximum likelihood fit. The distribution is skewed toward positive due to scattering. The second and third peaks around 40 ns and 110 ns are due to PMT after pulses [27].

A finer grid search then begins in a 2 m region surrounding the test vertices with the greatest likelihoods with a final precision of 10 cm. Figure 9.3 shows the vertex resolution as a function of energy.

The direction of the recoil electron can be fit by comparing the hit patterns to 42° Cherenkov cones corresponding to various potential directions. The likelihood of a candidate direction is calculated only with hits within 30 ns after time-of-flight subtraction. This likelihood is

$$L(\theta) = \frac{1}{N} \sum_{i=1}^N \frac{1}{\Delta t_i} \exp\left(-\frac{|\theta - \theta_i|}{\Delta t_i}\right)$$

in which θ is the angle between the candidate direction and the direction between the vertex and i th PMT, Δt_i is the typical acceptance of a PMT as a function of E , and $P(\theta)$ is a PDF determined from the true electron direction of solar MC. An energy-independent PDF is used for the first calculated direction. However, lower energy electrons exhibit more

scattering, resulting in fuzzier Cherenkov cones and reduced direction resolution. After energy calculation, the direction is therefore recalculated using an energy-dependent PDF as in Figure 9.2.

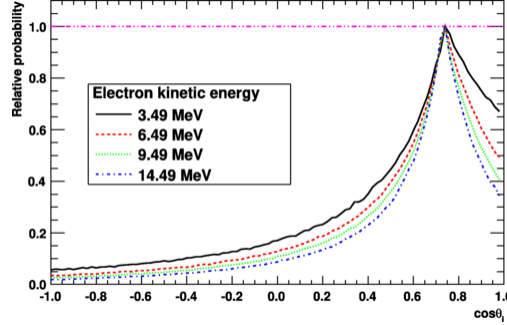


Figure 9.2: The distribution of the opening angle between the direction of the MC electron and the vector from the reconstructed vertex to each hit-PMT’s position. The black solid, red dashed, green dotted, and blue dash-dotted lines show the distribution of 3.49 MeV, 6.49 MeV, 9.49 MeV, and 14.49 MeV in electron kinetic energy, respectively [27].

For the calculation of electron energy, hits within a 50 ns window after time-of-flight subtraction are used. This number of hits is first scaled to an “effective” number of hits by applying several corrections to account for relative differences of PMT performance, the dark noise rate, late arrival hits from scattered light (hits outside the 50 ns window but within a 100 ns window), multiple photoelectrons, the effective surface area of PMTs, and light attenuation in the water. is then converted to the electron total energy through a quadratic fit using single electron MC. As a rough approximation, 6 hits represents 1 MeV of kinetic energy above the 0.26 MeV Cherenkov threshold.

The reconstruction procedure can be summarized as:

1. Vertex reconstruction with BONSAI
2. Direction fit with energy-independent PDF
3. Energy calculation with direction from 2.
4. Direction fit with energy-dependent PDF
5. Energy calculation with direction from 4.

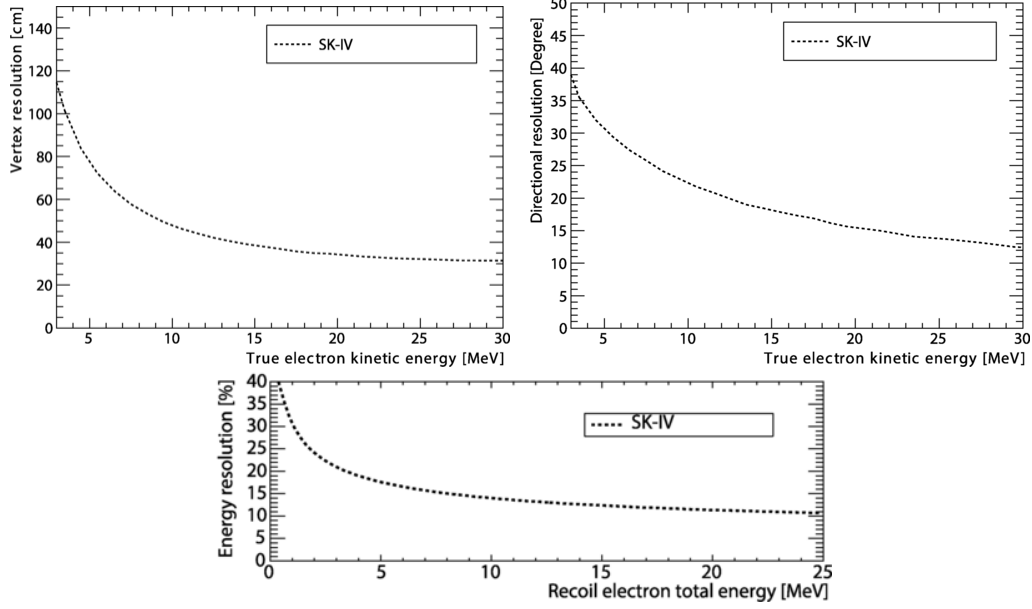


Figure 9.3: The vertex (left), direction (right), and energy (bottom) resolution in SK-IV as a function of the true electron kinetic energy based on solar MC [27].

Following these basic reconstructed quantities, additional variables can be calculated that may be useful in event selection. A simple but very effective quantity is the closest distance between the vertex and ID wall, $d_{\text{v-w}}$. A “fiducial volume” cut is often applied to $d_{\text{v-w}}$, typically 2 m or 1 m. Another variable is the “effective wall” distance d_{eff} , the distance to the wall along the reverse of the direction vector. A cut on this variable is often referred to as an “external event cut” as it more effectively removes events originating from the ID wall.

More complex and powerful variables are related to the goodness of the vertex and hit patterns. These goodness variables help reject poorly reconstructed events, such as radioactive background events originating near the ID walls (Section 10.1) coincident with noise hits that pull the reconstructed vertex inside the fiducial volume. χ^2_{vtx} represents the goodness of the vertex fit and is higher for smaller PMT hit timing residuals [104]. It is calculated by comparing the timing residuals to two Gaussians: one with a width of $\sigma_{\text{vtx}} = 5$ ns to encompass selected hits, and the other with an $\sigma_{\text{wall}} = 60$ ns width characteristic of the PMT timing

resolution for a single photoelectron.

$$\chi^2 = \frac{1}{N} \sum_{i=1}^N \frac{(h_i - e_i)^2}{h_i + e_i}$$

represents the goodness of the hit pattern and is lower for better azimuthal symmetry of the Cherenkov cone [104]. The reconstructed vertex and direction are used to determine an “expected” Cherenkov cone with an expected hit pattern in which the event’s hits are evenly spaced around the cone with azimuthal angle $\Delta\phi = 2\pi/N$. The actual hits are ordered by azimuthal angle to make the ordered set $\{h_i\}$. The actual and expected distributions are then compared as

$$\chi^2 = \frac{1}{N} \sum_{i=1}^N \frac{(h_i - e_i)^2}{h_i + e_i}$$

A cut on the combined χ^2 goodness variable is often applied, as it is one of the most powerful event separation variables (Section 11.1, Figure 11.4).

While χ^2 checks the symmetry of the cone, another likelihood is calculated for the consistency of the pattern with a *single* cone rather than multiple, which would be indicative of events other than neutrino elastic scattering [106]. The hit pattern is compared to the likelihood function

in which P is a PDF generated from solar MC and the other variables are defined as above.

Another variable identifies small clusters of hits, typical for events originating from the PMT glass [107]. This variable is $C = \frac{N_c}{N}$, where N_c is the radius of the smallest sphere containing one fifth of the hit PMTs in the event, and N is the maximum number of hits in any 20 ns timing window without time-of-flight subtraction.

Finally, a “multiple scattering goodness” (MSG) is calculated, which represents the goodness

of a direction fit performed by an alternative direction fitter [107]. Rather than using a likelihood function, this fitter reconstructs direction by drawing 42° cones corresponding to the vectors connecting the vertex and each PMT hit within 20 ns with time-of-flight subtraction. This therefore forms cones that may intersect along lines. The number of such intersections is . Clusters of points of intersections will form around likely direction vectors, defining candidate directions (Figure 9.4). The intersection unit vectors assigned to the candidate cluster are added together. The magnitude of this vector sum is compared to the largest possible: . MSG represents the ratio of the two and is intended to quantify the amount of elastic scattering the electron exhibits. Poor direction fits are also associated with background ^{212}Pb beta-gamma decays (Section 10.1). Since multiple scattering is purely dependent on the electron energy, MSG may also be interpreted as an alternative electron energy estimation method.

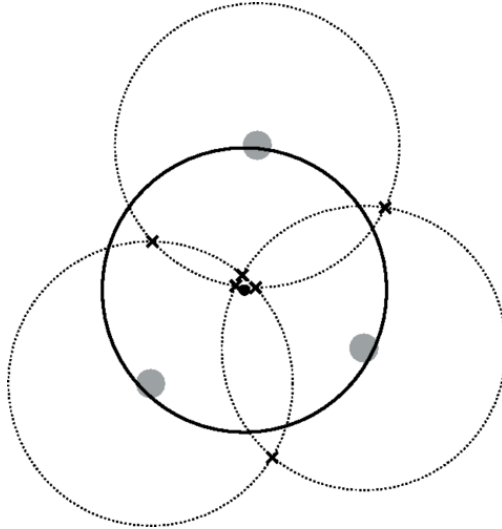


Figure 9.4: Illustration of the MSG calculation. The gray points represent PMT hits and the dotted circles surrounding them are the projections of the 42° cones centered around each hit. The black crosses give the intersection points of the cones. The vectors from the event vertex position to these intersection points are taken as event direction candidates. The black dot shows the event best fit direction and the black solid circle is the projection of its Cherenkov cone onto the inner detector wall. The intersections will cluster around the event direction [107].

Chapter 10

The Solar Analysis

The most recent solar analysis was published in May 2024 and uses data through the end of SK-IV with a total livetime of 5805 days [27]. These results include measurements of the total solar neutrino flux, seasonal variation of the flux, the day/night asymmetry affecting the $\bar{\nu}_e$ content due to Earth matter effects, the θ_{12} and θ_{13} parameters, and a neutrino energy spectrum to study the solar upturn. Chapters 11-12 are an extension of this analysis to consider energies below the 3.49 MeV threshold in the 2970 days livetime of SK-IV.

10.1 Sources of Background

At low energies, a leading source of radioactive background is beta-gamma decay of ^{232}Tl present in the PMT glass. This decay has a Q-value of 5.0 MeV and most often includes a 2.6 MeV gamma. These events therefore occur near the detector walls, but occasional coincidental triggers with dark noise hits can result in these events being reconstructed within the 2m fiducial volume and inward-going. Another significant source is ^{210}Bi beta decay with a Q-value of 3.3 MeV resulting from the decay of ^{210}Po gas [95, 108]. ^{210}Pb is

produced through the uranium-thorium chain prevalent underground and, as a noble gas with a 3.82 day half-life, can easily spread through mine tunnels and the detector regions. This ^{222}Rn background is the most significant due to its ability to enter the fiducial volume and the fact that the single ^{214}Pb beta is indistinguishable from a neutrino elastic scattering recoil electron. As explained in Section 8.1, the water currents within the inner detector are designed in such a way as to collect as much of this radon as possible near the bottom of the detector and are suppressed through temperature control within the clean center-top region [95, 109] (Figure 10.1). Although the energies of these events resulting from radioactive impurities are below even those considered in this analysis, poor energy resolution and coincidental triggers cause these decays to be the dominant background source as high as 6 MeV.

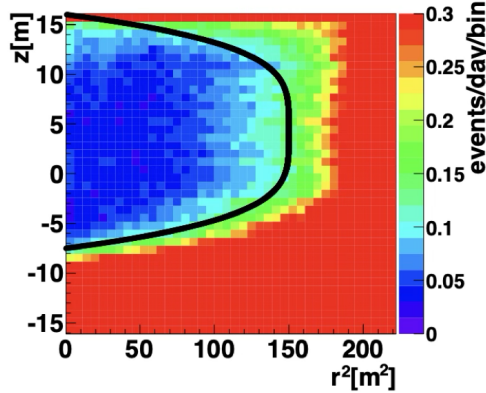


Figure 10.1: SK-IV vertex distribution in 3.49 MeV - 3.99 MeV. The black line is a tight fiducial volume cut boundary [107].

At higher energies, the leading source of background is again radioactive decay but caused by isotopes created during cosmic ray muon spallation [110]. The dominant spallation products at these low energies are radioactive isotopes created by the interaction of hadronic showers [100] with ^{22}Ne . The leading spallation-induced radioactive isotope is ^{26}Al , which is long-lived, lasting up to several seconds, and can be found meters away from the showering vertex before undergoing beta decay [106, 111].

10.2 Event Selection

To remove the ^{214}Bi and ^{214}Pb decay background, cuts are applied on the variables introduced in Chapter 9 with increasing selectivity at lower energies. A 2 m fiducial volume cut on z_{fid} is applied for all events, but the 8.8 kton tight fiducial volume (TFV) cut boundary (Figure 10.1) is applied for $E < 4.99$ MeV. The cut on the TFV_{cut} event quality is $\text{TFV}_{\text{cut}} > 0.5$, $\text{TFV}_{\text{cut}} > 0.6$, and $\text{TFV}_{\text{cut}} > 0.7$ for $E < 4.99$ MeV, $4.99 \text{ MeV} - 6.99 \text{ MeV}$, and $E > 6.99 \text{ MeV}$ respectively. A cut on the small cluster radius $r_{\text{small}} < 68 \text{ cm}$ is applied to certain energy-dependent detector regions starting at $E < 4.99 \text{ MeV}$. The cut on the effective wall distance is 4 m for $E > 7.49 \text{ MeV}$, 6 m for $4.99 \text{ MeV} - 7.49 \text{ MeV}$, and 10 m, 13 m, or 12 m for $E < 4.99 \text{ MeV}$ depending on if the intersection point is located in the top, bottom, or barrel wall respectively.

Three cuts are applied to remove the spallation-based background: a spallation log-likelihood cut that compares the proximity in location and time of events to cosmic ray muon tracks, a multiple spallation cut removing small clusters of low energy events that often result from a spallation-induced hadronic shower, and the neutron cloud cut that compares the proximity of events to spallation neutron event candidates in WIT data before first reduction [100].

More minor cuts remove flashing PMTs and detector elements far from the walls including temperature sensors and the permanent calibration sources. Bad runs are excluded including periods of high background resulting from high convective currents⁴.

Following the final selection, events are split into bins of 0.5 MeV up to 13.49 MeV, 1.0 MeV up to 15.49 MeV, and 15.49 MeV – 19.49 MeV. The energy bins below 7.49 MeV are each further split into three bins of MSG: 0.00 – 0.35, 0.35 – 0.45, 0.45 – 1.00.

⁴This is traditionally called a “cut” rather than a reduction in livetime because only $E < 5.49 \text{ MeV}$ events are excluded during convection periods.

10.3 Signal Extraction

Event selection for the solar analysis is conducted with reconstructed quantities and related goodness variables (Chapter 9), but no cut using the solar direction is applied. Instead, the number of signal and background events present in the selection is determined with an extended maximum likelihood function fit on the solar angle distribution, $\frac{dN}{d\theta_{\text{sun}}}$ for electron direction θ_e and solar direction θ_{sun} at the time of the event. The distribution of data is compared to expected distributions (“shapes”) for signal and background.

10.3.1 Solar Distribution Shapes

The signal shape is generated using polynomial fits to MC. For background, two different data-driven methods are used in [27] to generate background shapes. In the “scramble” method, the background shape for $\frac{dN}{d\theta_{\text{sun}}}$ is calculated by filling this solar angle distribution with $\frac{dN}{d\theta_{\text{sun}}}$ for all possible pairs of events θ_e and θ_{sun} in the selection to give the background shape $\frac{dN}{d\theta_{\text{sun}}}$. If the signal-to-background ratio is poor, there will be many more pairs of background events than pairs of signal events, so the scramble method is more appropriate for low signal-to-background ratio. With better signal-to-background ratio, the presence of signal events will significantly warp the background shape. To account for the presence of signal, bins are downweighted according to the expected signal content from data and an assumed background shape. For the first fit, the assumed shape is flat in $\frac{dN}{d\theta_{\text{sun}}}$. The fitting procedure is then repeated after reweighting the bins using the first background shape as the assumed shape. This iterative process is stopped after two fits since the first and second background shapes are generally in good agreement. The scramble method is somewhat susceptible to the presence of signal as statistical fluctuations in the signal will still warp the shape, so the scramble method becomes less appropriate at higher energies.

In the “standard” method, polynomial fits are applied to the event direction distributions in detector coordinates of the zenith (θ_z) and azimuthal (ϕ) angles: $\frac{dN}{d\theta_z d\phi}$ and $\frac{dN}{d\theta_z}$. Since

is cyclical, the boundary conditions $\phi = 0$ and $\phi = 2\pi$ are imposed. A toy MC is then generated based on these shapes, and these detector angle distributions are refit before projecting to the solar angle θ_{sun} [106]. A similar downweighting approach to account for the presence of signal is used for these first and second fits. To assign an uncertainty to the standard background shape method, a covariance matrix is calculated between the polynomial fit coefficients. Each coefficient is then independently varied by $\pm 1\sigma$, and these shifts along with the covariance matrix are carried throughout the toy MC and second fit in order to calculate an uncertainty in the subsequent likelihood fit. The standard method assumes there is no correlation between ϕ and θ_{sun} , so does not give as accurate a shape compared to the scramble method's full two-dimensional fit. However, the toy MC between the two fits reduces warping of the shape due to statistical fluctuations in the signal. The SK solar analysis uses the standard method for the entire solar sample and reserves the scramble method as an alternative method for the purpose of calculating the background shape systematic error if the discrepancy between the two methods is greater than the standard method's uncertainty estimation. Comparisons between the two at low energies are shown in Section 11.3. Both signal and background shapes are converted to polynomial probability density functions prior to the maximum likelihood fit.

10.3.2 Extended Likelihood Method

The likelihood function for the number of signal events in a dataset is

Separate background predictions are made for each selection bin, whereas signal predictions are made over the entire analysis energy region based on MC and then split into the respective analysis bins. The free parameters are N_b , the number of background events in the i -th selection bin, and N_s , the total number signal events in the dataset. N is the number of

observed events in the i -th selection bin, N_{bins} is the number of selection bins, f_i is the fraction of signal MC events in the i -th selection bin, and P_{sig}^i and P_{bkg}^i are the values of the PDFs calculated from the signal and background shapes for selection bin i at the $\cos\theta_{\text{Sun}}$ bin corresponding to the i -th observed for data event j .

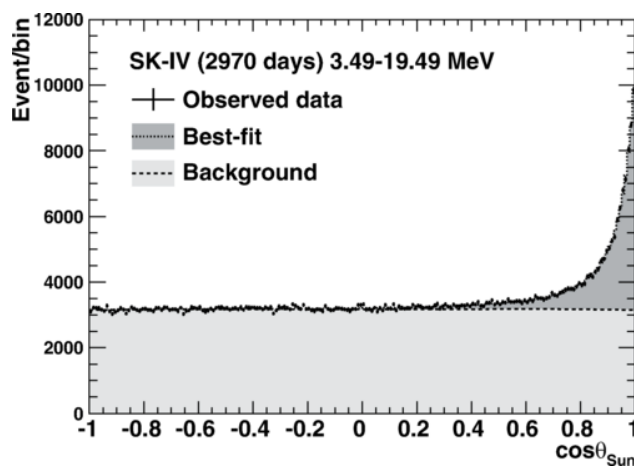


Figure 10.2: Solar angle distribution for the full energy range of 3.49–19.49 MeV using the SK-IV dataset. The black points show the observed data, the dark and light gray shades show the background and the best-fit signal, respectively [27].

Chapter 11

Extension to 3.49 MeV

By far the largest source of background below the analysis threshold in 2.49 MeV - 3.49 MeV is the radioactive beta decay. In the WIT dataset after first reduction cuts, there are radioactive background events per day compared to solar neutrino interactions. There are events per day due to the spallation isotope background.

11.1 Event Selection

The event rate of radioactive background increases exponentially with decreasing energy approaching the true energy of these betas and gammas. Because radioactive background is only found in higher reconstructed energy regions due to finite detector energy resolution, the event rate increases exponentially with decreasing reconstructed energy approaching the true energy of these underlying betas and gammas. The background rejection of the solar neutrino selection cuts used in previous SK solar analyses [27, 107, 104] is not sufficient to observe a signal below the previous 3.49 MeV threshold. Several different event selection approaches were studied to improve background rejection in this energy region including several machine learning based methods.

11.1.1 Development of Machine Learning Methods

Through cooperation with the Water Cherenkov Machine Learning Group (WatChMaL) [112], several advanced machine learning methods that take the raw PMT hit times and charges as input were studied. As all of the networks discussed here perform binary classification, the output is a 1×2 softmax vector corresponding to the probabilities that the event is signal or background.

The first studied class of networks is convolutional neural networks (CNNs) operating on 2D event display images. Typical Super-K event displays contain a rolled-out and flattened rectangle corresponding to the cylinder barrel and the top and bottom circular endcaps above and below the barrel (Figure 11.1). Each pixel corresponds to a PMT and has two channels corresponding to the hit time and collected charge. Several ResNet [113] architectures with slight modifications⁵ by WatChMaL were compared, and the best performing model was a ResNet50 model [114].

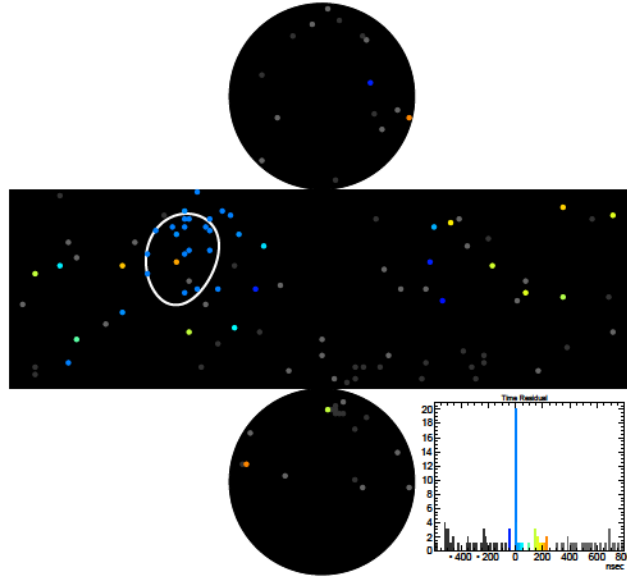


Figure 11.1: An example 3.85 MeV event display. The pixel color represents the PMT hit time as shown by the histogram on the lower right. The white ring indicates the Cherenkov cone implied from the reconstructed electron direction.

⁵The initial 7×7 convolution layer was replaced with 2 1×1 convolution layers to perform channel-mixing for detectors with multi-PMTs that have 19 charge and time channels per PMT [112].

Unlike for tracking detectors, CNNs are not as suitable for the geometry of water Cherenkov detectors. While tracking detectors constrain the trajectory of particles either in 3D space or through multiple 2D views with a regular rectangular grid, water Cherenkov data is collected on a surface in 3D space. While the placement of PMTs is typically regular, it is inherently non-Cartesian particularly on the curved endcaps where a standard grid cannot be maintained without overlap or gaps. 3D convolutions can be used, but the 3D images would be very sparse, and pixels that correspond to locations without PMTs will always be zero. 2D images must distort the geometry of the detector, although extensive studies have been done on alternative PMT-to-image maps such as duplicating and rearranging PMTs to preserve the boundary conditions of the barrel and between the barrel and endcaps [112, 115, 116].

Another approach that naturally resolves the geometry issue is graph neural networks (GNNs). GNNs operate on graphs, a set of points connected by edges potentially with features associated with each point and edge. The most natural way to apply GNNs to water Cherenkov detector data is to treat each hit PMT in an event as a point in the graph. For the dataset, the PMT hit time, charge, 3D position, and 3D orientation vector were embedded in each point as $\mathbf{p} = [t, q, x, y, z, \theta, \phi]^T$. The points were connected as a k -nearest neighbor graph according to the positions, optimizing k as a hyperparameter (Figure 11.2). No features were embedded in the edges. Several GNN architectures were investigated [114] including the original graph convolution network (GCN) [117], residual GCN (ResGCN), graph isomorphism network (GIN) [118], and dynamic edge convolutional neural network (DyEdCNN) [119]. The PointNet network [120] for point clouds was also tested. Among these networks, the network with the best performance was the DyEdCNN. As the name implies, this network was originally developed for point clouds rather than graphs with fixed edge connections. The graph edges are recomputed at each layer, and the graph structure can change throughout the network.

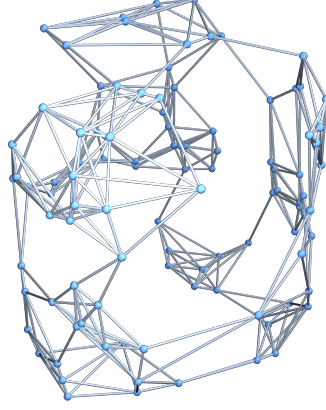


Figure 11.2: An example graph structure. The blue spheres represent PMT hits on the cylindrical Super-K ID walls. In a 5-nearest neighbor graph, each PMT is connected to the 5 closest hit PMTs through edge connections indicated by grey lines.

Boosted decision trees (BDTs) were trained on the reconstructed variables used in event selection for the existing solar analysis from Section 9. The final list of variables is shown in Table 11.1 and distributions of the 4 most important variables as determined by the XGBoost “Total gain” metric are shown in Figure 11.4. In addition to existing variables,

Index	Description
1	Number of PMT hits in the event
2-3	Reconstructed BONSAI vertex position in the detector as x and y
4-6	Reconstructed direction as a unit vector
7	Closest distance from the vertex to the inner detector wall (cm)
8-10	Point of intersection with the wall along a line starting from the vertex aligned with the reverse direction vector as x , y , and z
11	Distance from the vertex to the wall along this line (effective wall distance) (cm)
12	Distance from the vertex to the TFV boundary
13	Value of a new 2D polynomial fit to the background vertex positions in x and y
14	Reconstructed total event energy
15	Vertex fit goodness, χ^2 , higher for smaller PMT hit timing residuals
16	Hit pattern goodness, χ^2 , lower for better azimuthal symmetry of the Cherenkov cone
17	
18	Multiple scattering goodness (MSG), higher for better quality direction reconstruction associated with higher energy and less electron scattering
19	Smallest radius containing one fifth of the hit PMTs, smaller for events originating at or near PMT glass

Table 11.1: Variables input to the BDT.

two new variables are added and are intended to have the BDT induce a cut related to the TFV without explicitly applying such a cut. One of these (Index 12) is simply the closest distance of the vertex to the 8.8kton TFV cut boundary (Figure 10.1) with negative values indicating events outside the boundary. The other (Index 13) is a 2D polynomial fit to the background vertex distribution (Figure 11.3). XGBoost v1.5.0 was used to train, test, and implement the BDTs [121].

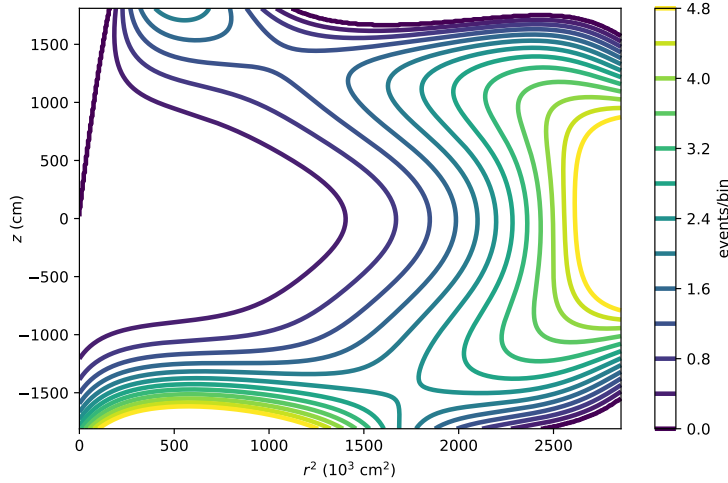


Figure 11.3: A 2D polynomial fit to the background vertex distribution. The polynomial is 5th-degree in z and 4th-degree in r^2 as this choice minimized the χ^2 . The

Several “hybrid” approaches that use both PMT hit information and reconstructed variables were also explored. One approach is simply including the signal output score of a hit-trained network as an additional variable input to the BDT. Another more sophisticated approach is to build a network that operates on both input types. For this network, the DyEdCNN, was used to serve as the input embedding for the PMT hit information. The final softmax operation was removed, and a 100 feature representation vector rather than a 10 vector is output. A fully-connected neural network was used to embed the reconstructed variable inputs and also outputs a 100 . These vectors are concatenated into a 200 vector that is input to a final fully-connected network. All portions of this hybrid network are trained simultaneously and share weights.

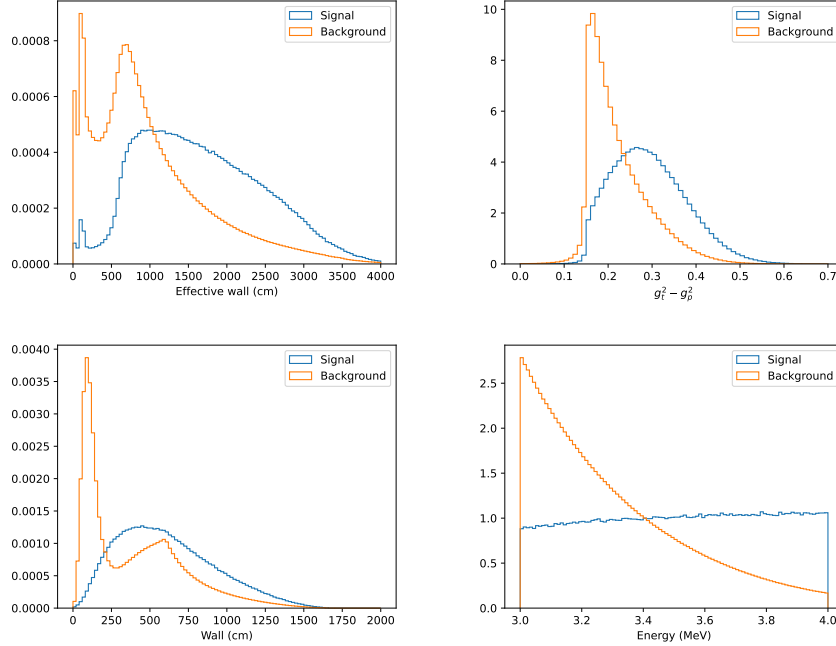


Figure 11.4: Distributions of the most important reconstructed variable inputs to the BDT in the WIT dataset for signal (blue) and background (orange). These distributions are following WIT online trigger cuts and offline preselection cuts, which accounts for hard and soft cutoffs in these variables (See Section 8.3 and 11.1.2). Histograms are normalized to an area of 1.

Figure 11.6 compares the performance of several of these methods. Although for a large portion of the ROC curve, the DyEdCNN and hybrid network perform the best, the ROC curves intersect as the BDT outperforms in the low acceptance regime necessary for this very low energy context with low signal-to-background ratio. At the same signal efficiency of the cuts used in the previous SK solar analysis [27] (the “Solar Analysis Cuts”), the BDT achieves equal or better performance than any more sophisticated machine learning method. This analysis therefore uses a BDT since it is also less computationally intensive than other methods.

Several other approaches to modify the training procedure were studied to improve the performance at very low acceptance. However, typical methods used to address “class-imbalanced” datasets including class reweighting and using alternative loss functions such as the “Recall loss” and “Focal loss” [122, 123, 47] were found to be less effective than the following method:

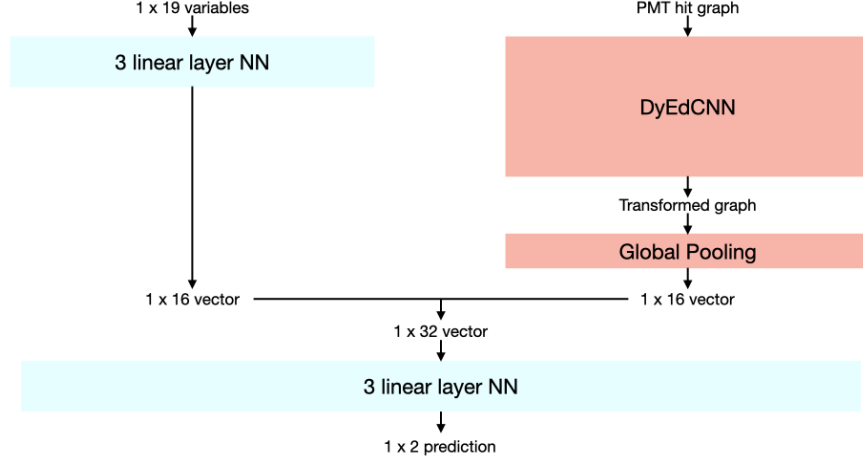


Figure 11.5: Architecture of the hybrid network.

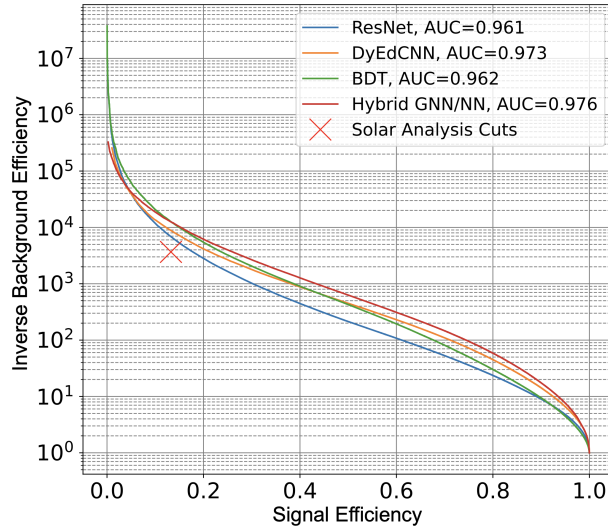


Figure 11.6: “Rejection” ROC curve showing the background rejection as the inverse of the false positive rate (FPR) vs the true positive rate (TPR, or signal efficiency) for the WIT 2.49 MeV - 3.49 MeV testing dataset. This plot includes the ResNet50 CNN, the DyEdCNN GNN [119], the BDT, and the hybrid network (Figure 11.5). The solar analysis cuts used in [27] is shown as a red $\times$. The legend shows the area under the curve (AUC) values on a standard TPR vs FPR ROC curve.

three separate BDTs are trained in succession (Figure 11.7). The first is trained on a subset of the overall training sample with equal parts signal and background. The first BDT is then used to perform inference on the rest of the training dataset, and a loose cut on the BDT output score is applied, chosen to reduce background by a factor of ten. A second BDT is then trained on another balanced subset of surviving signal and background events using

the same variables as in Table 11.1. This process is repeated to train a third model with only events that pass these cuts on the scores of the first two models. A completely separate hold-out test set is used to evaluate the model. Although each individual BDT is trained on a balanced dataset, this approach requires much more background than signal for training overall.

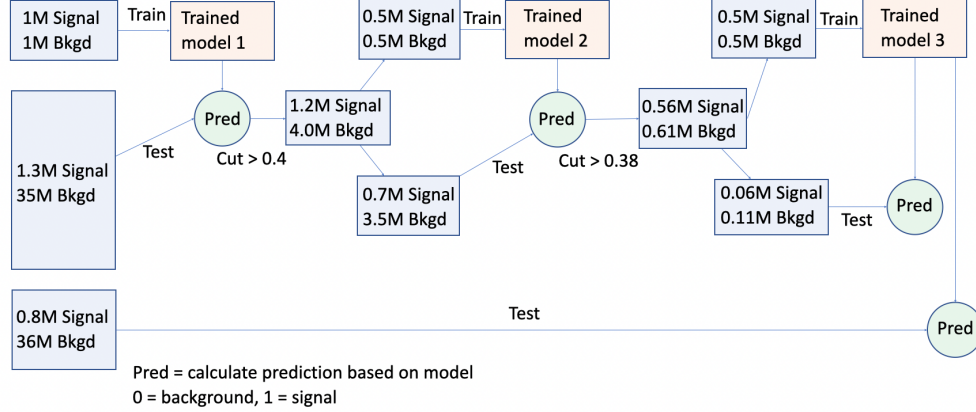


Figure 11.7: Illustration of the “triple-fit” BDT training procedure.

Figure 11.8 shows that the BDT models trained through this procedure get progressively better at rejecting the most signal-like background events at $\sim 50\%$ signal efficiency at the cost of better performance at higher signal efficiencies. For the WIT dataset, the final BDT model offers increased background rejection in 2.49 MeV - 3.49 MeV for the same signal efficiency as the existing cuts by a factor of 6.8.

11.1.2 Dataset Generation

For the signal Monte-Carlo (MC) simulation, we follow the same procedure as in [27]. Solar neutrino recoil electron events are generated within the inner detector volume. For the purpose of generating a training dataset, however, the direction of the recoil electron was randomized drawing from a spherically uniform distribution. This approach prevents the network from potentially learning about the solar direction and causing biases in the solar angle distribution so that any intentional use of the solar direction can be reserved for the

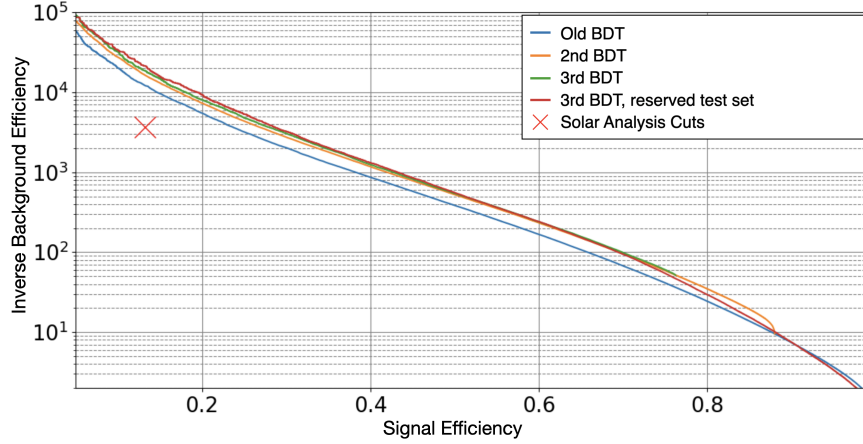


Figure 11.8: Rejection ROC curve showing the effect of the “triple-fit” procedure for training BDTs. The blue curve shows the performance of the first BDT model in Figure 11.7, the orange shows the second model, and the green curve shows the third model that is actually implemented. The red curve represents this third model on the completely separate hold-out test set.

post-selection solar angle fitting just as in the existing analysis. Detector simulation is then performed with GEANT3 [124] modified for SK-IV.

We generate two separate datasets corresponding to each of SLE and WIT data and train separate BDTs for each of these datasets. For the SLE dataset, the detector simulation is run with simulated dark noise. For the WIT dataset, dark noise simulation is turned off and events are instead overlaid onto “dummy triggers” containing only real dark noise hits. The respective trigger algorithm is then run on each of the datasets. Separate testing datasets are also generated in the same manner but without the recoil electron direction randomization so that these datasets respect the actual solar direction. These signal testing datasets are also used for the signal extraction step in Section 11.2.

Following signal dataset generation but before training, a series of loose preselection cuts based on some of the variables in Table 11.1 are applied. These cuts are chosen to exactly match those used on data as events failing these cuts have been deleted. These include loose cuts on the distance to wall, effective wall distance, $\cos\theta_{\text{wall}}$, and, notably, the energy. This cut in kinetic energy is 2.99 MeV, 2.49 MeV, and 2.24 MeV for the SLE 34-hit, SLE 31-hit,

and WIT datasets respectively. With the motivation of having the BDT perform the bulk of the event selection, these cuts are not tightened any further in order to include as much data as possible. Also note that all of these preselection cuts are applied to reconstructed quantities assuming a conservatively low 90 m water transparency following the procedure used for actual data. Events are subsequently refit with the correct water transparency measured later, and reconstructed event energies are ~ 0.2 MeV lower after refitting for a typical water transparency of 140 m.

For the background sample, due to the abundance of radioactive background data and the extremely low signal-to-background ratio ($\sim 10^{-4}$) following only the loose preselection cuts, real data was used as a stand-in for proper background simulation. We assume that the BDT will not be largely influenced by the small fraction of real signal events that are labeled as background for training. This approach eliminates the need to consider systematic errors in background simulation, limiting such errors to signal simulation only. 1% of the available data was randomly selected to use as the BDT training dataset and another 1% for the testing dataset.

11.1.3 Final Selection

Following the implementation of the respective BDTs on the SLE and WIT datasets, selection cuts from Section 10.2 are applied that are intended to remove sources of background other than the dominating ^{214}Bi and ^{214}Pb decay background events. These include the flashing PMT cut, the calibration source cut, the bad run and convection period cuts, and, most notably, the cosmic ray muon spallation cut. The cut on the BDT output score is applied last, and the value of this cut is tuned to optimize the $\frac{S}{\sqrt{S+B}}$ figure of merit for signal MC and data.

As WIT was implemented after the switch to the 31-hit trigger threshold, the WIT dataset does not overlap with the 34-hit SLE dataset. For the 31-hit period, only runs in which

WIT data was not available were included. The livetimes of these three SK-IV datasets are summarized in Table 11.2 and total 2970 days.

Dataset	Start	End	Livetime (days)
34-hit SLE	6 October 2008	1 May 2015	2047
31-hit SLE	1 May 2015	30 May 2018	305 ⁶
WIT	10 July 2015	30 May 2018	618

Table 11.2: Livetimes of the datasets used in this selection.

11.2 Signal Extraction

Following the event selection, the approach for the previous solar analysis described in Section 10.3 is largely followed to measure the number of solar neutrino events. However, while the standard method for generating the solar angle background shape is typically used as the primary method, we choose to do the opposite here due to the very poor signal-to-background ratio in this very low energy range, designating the scramble method as the default and the standard as the alternative for systematic error determination.

The binning of [27] is also followed and is simply extended to this energy region of interest, maintaining the 0.5 MeV bins and bins of 0-0.35, 0.35-0.45, and 0.45-1 in MSG. The fit is applied simultaneously over each MSG bin, and the solar angle distributions with the signal and background shapes in these bins are shown in Figures 11.9 and 11.10.

⁶Only includes livetime when WIT was not available.

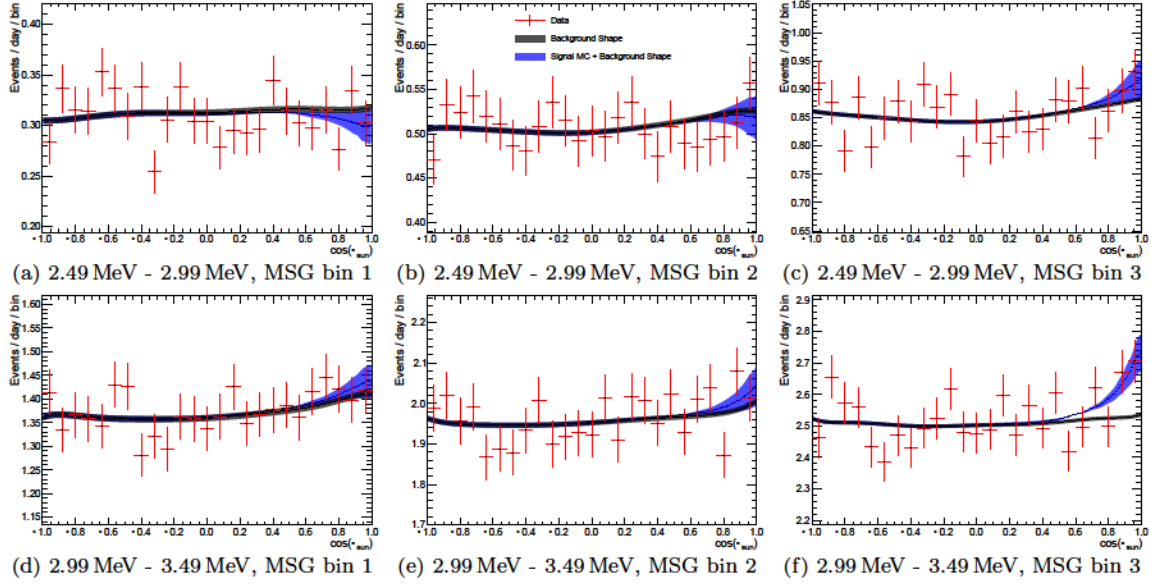


Figure 11.9: Solar angle distributions for the WIT dataset in all MSG bins for data selection (red), calculated background shape (black), and calculated signal shape added to background shape (blue) all with 1σ statistical error bands.

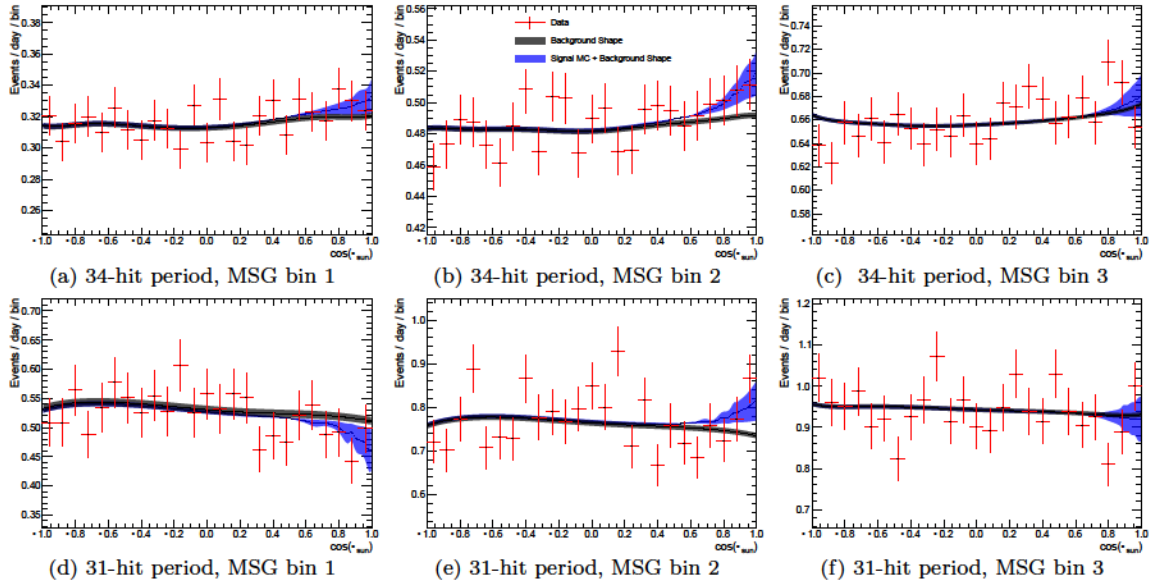


Figure 11.10: Solar angle distributions for the SLE dataset in all MSG bins separated by trigger threshold period. The definition of colors is the same as in Figure 11.9.

In order to demonstrate the presence of solar neutrinos, the signal and background shapes are aggregated over the three MSG bins by weighting each bin according to its associated statistical error of a solar signal observation. Figures 11.11 and 11.12 show these aggregated solar angle distributions.

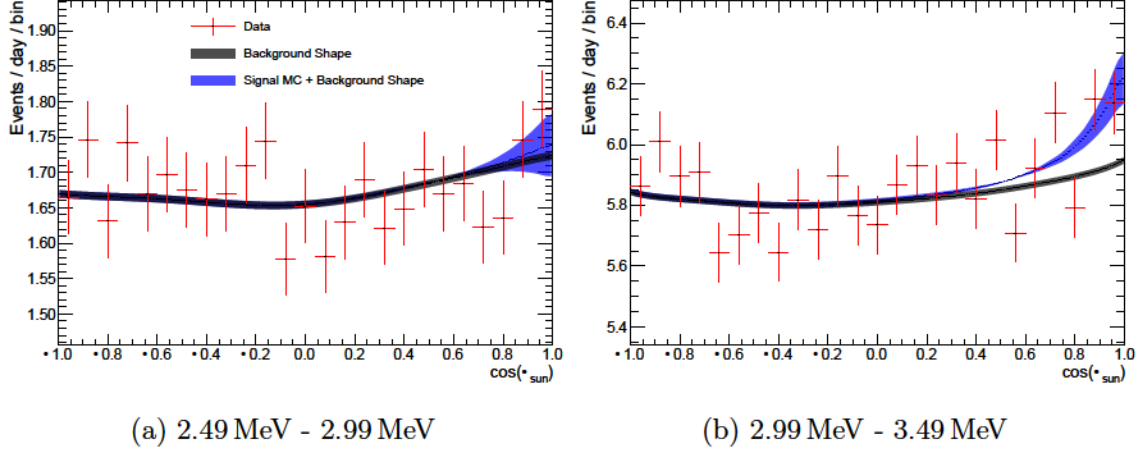


Figure 11.11: Solar angle distributions aggregated over MSG bins for the WIT dataset. The definition of colors is the same as in Figure 11.9.

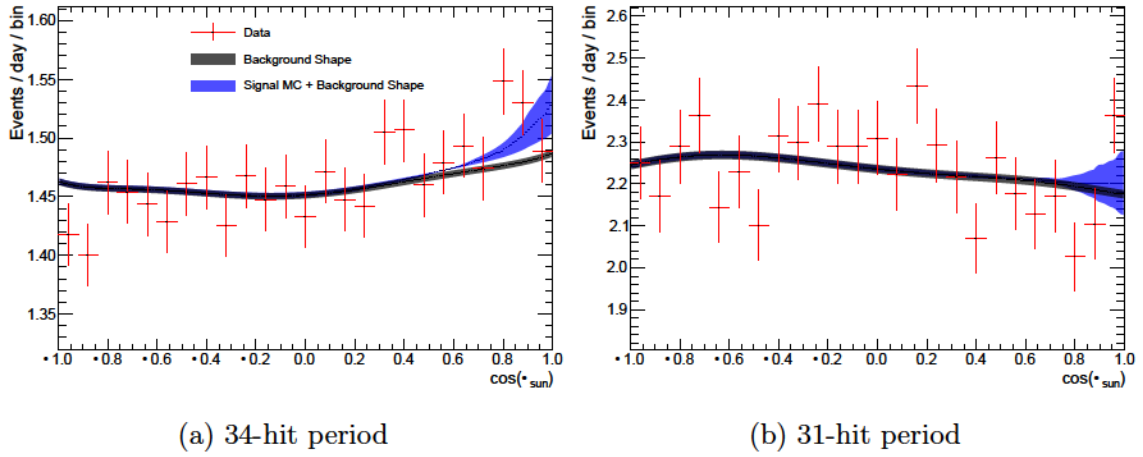


Figure 11.12: Solar angle distributions aggregated over MSG bins for the SLE datasets in 2.99 MeV - 3.49 MeV separated by trigger threshold period. The definition of colors is the same as in Figure 11.9.

11.3 Systematic Uncertainties

Many of the systematic uncertainty calculation methods for this analysis are unchanged compared to previous Super-K solar analyses, especially those relating to aspects of the analysis that are also unchanged. Changes that were necessary for the BDT-based selection or WIT dataset are highlighted in this section.

11.3.1 Background Shape

The uncertainty in the background shape calculation is the largest source of systematic uncertainty. While this error is only 2.7% in the lowest 3.49 MeV - 3.99 MeV energy bin of [27], the signal-to-background ratio is considerably worse in the next lowest energy bin here. Similarly small fluctuations in the background shape in the region of the solar peak can result in large differences in the fit number of signal events when the solar peak is also small.

We compare the resulting number of signal events when using the two background shape methods, `scramble` and `standard`, and calculate the error as

$$\frac{\text{scramble} - \text{standard}}{\text{scramble}}$$

taking the `scramble` method as nominal (Section 10.3.1). Figure 11.13 shows the comparison of the `scramble` and `standard` shapes for all bins. The background shape uncertainties in 2.99 MeV - 3.49 MeV for the full MSG 0-1 range were 11.2% for WIT and 41.8% for 34-hit SLE. We do not calculate an error for the 31-hit period (no WIT) or the WIT 2.49 MeV - 2.99 MeV bin as no significant signal was observed Section 12.1.

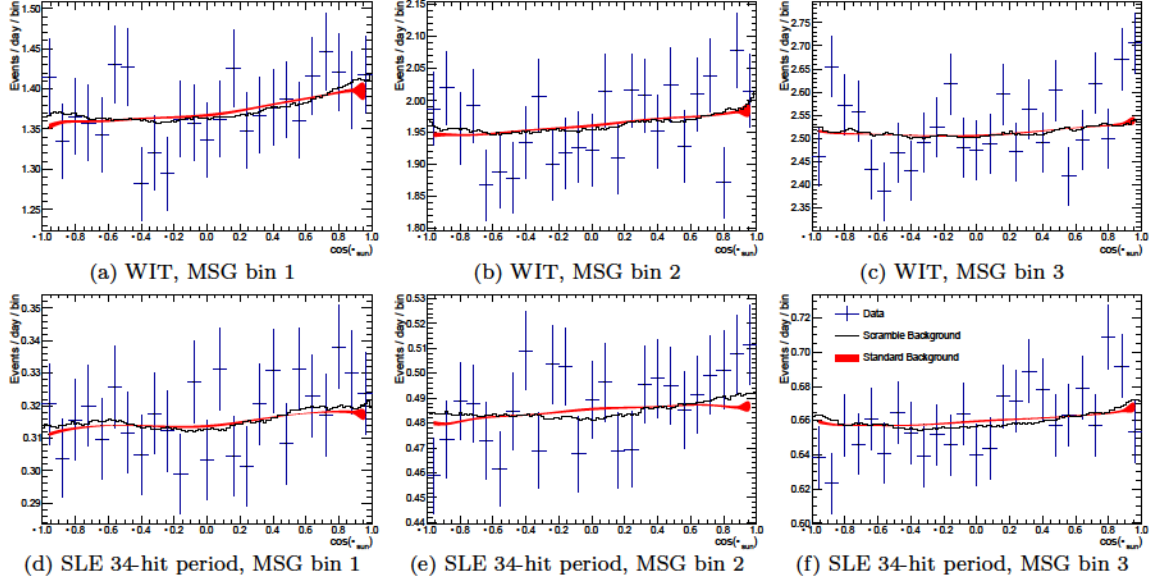


Figure 11.13: Comparison of background shape calculation methods for systematic error calculation in each MSG bin for WIT and SLE 34-hit period in 2.99 MeV - 3.49 MeV. The blue points show the data, the black line shows the scramble method, and the red band shows the standard method. The width of the of the red line shows the standard method's 1σ uncertainty (see Section 10.3.1).

11.3.2 Signal Extraction

The signal extraction uncertainty is calculated by generating fake datasets. After running the nominal signal extraction and measuring the number of signal and background events present in real data, fake datasets are generated by filling $\cos\theta_{\text{sun}}$ distributions with this number of signal and background events, randomly drawing from the signal and background shapes. 500 fake datasets are generated in this way, and signal extraction is then run over all 500 of these datasets. We then compare the average number of fit signal events to the number measured in real data. For the WIT dataset, we observed 457.3 signal events in data and an average of 458.2 in the fake datasets, giving a systematic uncertainty of 0.19%. We therefore dropped this systematic due to the small value.

11.3.3 Trigger Efficiency

The calculation of the trigger efficiency uncertainty in SK-IV is based on extensive calibration data taken in 2015 with the Ni gamma ray source. The trigger efficiency of Ni events at various points is compared between data and MC, and the trigger efficiency uncertainty at each point is $\pm 1\%$. In the published solar analysis, which uses the TFV cut, the overall trigger efficiency uncertainty is calculated as $\frac{\Delta \epsilon}{\epsilon} = \frac{\Delta V}{V}$ where V is the volume of the corresponding region within the TFV.

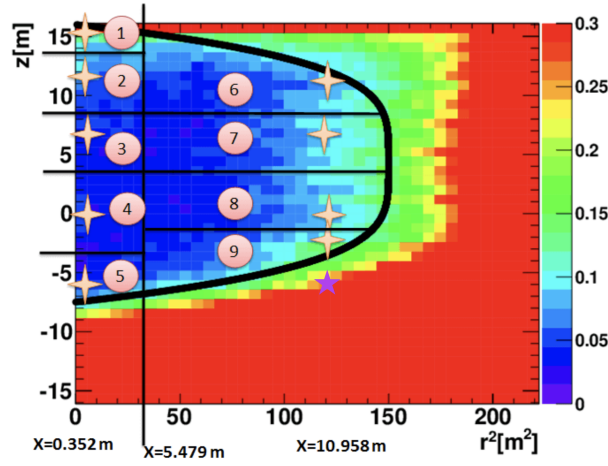


Figure 11.14: Ni calibration positions and their corresponding regions in the TFV. Positions marked with beige 4-pointed stars indicate a position inside of the TFV and are assigned a subregion of the TFV indicated by thin black lines. The purple 5-pointed star indicates another position outside of the TFV. There are four more positions outside of the range of this plot at $r^2 = 231 \text{ m}^2$.

The BDT selection approach, however, does not use an explicit TFV or even 2 m fiducial volume cut. We therefore needed to modify this approach to be compatible with a BDT selection. The BDTs used for the 34-hit and 31-hit SLE datasets were applied to all Ni data and MC from the 2015 calibration runs, including those outside of the TFV region. The same BDT cut used during solar event selection was then applied to data, giving the number of surviving events N_{surv} , adjusting for the livetime of each Ni run. This replaces the geometric volume weights V_i with the BDT acceptance. We determine the overall trigger efficiency systematic uncertainty as $\frac{\Delta \epsilon}{\epsilon} = \frac{\Delta N_{\text{surv}}}{N_{\text{surv}}}$. This uncertainty is 16.2% for the 34-hit threshold

and 7.7% for the 31-hit threshold.

For the WIT dataset, trigger efficiency uncertainty should be small because the WIT pre-trigger efficiency is near 100%. The efficiency of the online WIT cuts from Figure 8.8 in the relevant bins is given in Table 11.3.

Range (MeV)	3.49–3.99	2.99–3.49	2.49–2.99
Efficiency	94.3%	91.4%	83.0%

Table 11.3: Trigger efficiency of the online WIT cuts for solar MC.

While the systematic uncertainty of cuts on online reconstructed quantities is expected to be lower than that of a simple hit coincidence trigger, and such systematics should be covered by shifts to the relevant variables in Section 11.3.4, there is no extensive set of calibration data available during the time WIT was online during SK-IV, so the method applied for SLE data cannot be used for WIT to verify these expectations. A very conservative trigger efficiency error is therefore determined by assigning the actual trigger inefficiency as the systematic error. This inefficiency is 8.6% in 2.99 MeV - 3.49 MeV (Table 11.3).

11.3.4 BDT Input Variables

The previous solar analysis assigns systematic errors to variables used in the selection cuts or to bin the selection that may arise from discrepancies between solar signal MC and data [27]. This is done by comparing solar flux results or simply the efficiencies of these cuts after shifting the relevant variable(s) by an amount motivated by calibration data. The discrepancy in a particular variable between LINAC or Ni-Cf source data and MC determines the amount of the shift, which is energy-dependent. Both positive and negative shifts in these variables are tested, and the systematic error is chosen as the one with the larger effect.

A similar approach for the BDT selection would be to shift the input variables by the same amount and then observe the change in efficiency of the BDT output cut. However, the BDT considers all input variables simultaneously, and many of the BDT inputs are highly

correlated. Therefore, the BDTs were reevaluated after shifting variables one at a time for those variables that are not correlated with any others to first order. However, the shifts were propagated for highly-related variables; particularly the effective wall, vertex and direction variables were treated as a single systematic.

The fit goodness (Index 17) is one of the most important variables (Figure 11.4). We apply the same LINAC-motivated positive and negative shifts used to calculate the systematic of the “event quality cut” in the previous analysis. At the lowest energy LINAC setting 3.8 MeV, this shift is . After applying the BDT output cut on solar MC, the systematic error on the signal efficiency is 1.2%, 2.1%, and 2.1% for SLE 34-hit, SLE 31-hit, and WIT respectively.

Special care is needed for the MSG (Index 18) and energy (Index 14) inputs because they are also used to define the bins for the final fits. (See Section 11.3.5 for the energy-related systematics.) We therefore calculate the uncertainty using two different approaches corresponding to these two uses of these variables and assign the greater error. The first method is the same approach of shifting the input variable for the BDT and observing the change in signal efficiency. The second method is to apply these shifts to the final selection before repeating the signal extraction and comparing the fit number of signal events to the result of the unshifted sample. The latter is the method used in the previous analysis.

Shift method	WIT
Shift BDT input up	0.19%
Shift BDT input down	-0.19%
Shift final evt. MSG up	-0.19%
Shift final evt. MSG down	0.26%

Table 11.4: MSG systematic error in signal MC efficiency for the WIT dataset using plus and minus shifts under various methods. The method with the highest error is in bold.

The error for MSG is therefore 0.26%. Due to the small value, this error was not calculated for the other datasets and was dropped for this analysis.

The previous solar analysis assigns one error for the fiducial volume cut using vertex shifts from Ni data, one error for the direction reconstruction using LINAC data, and one error for the external event cut (effective wall) by simultaneously applying the shifts to the vertices and directions and recalculating effective wall. One characteristic of this approach when dealing with multi-dimensional variables is that even for the traditional cut-based analysis, shifting the variables in one direction does not strictly lead to an increase or decrease in the selected number of events, e.g., shifting all the vertices by a constant vector will cause some events to enter the fiducial volume and some to leave. This may lead to an underestimation of error compared to the above single-dimensional variables.

For the BDT, we similarly apply the simultaneous vector shifts to the vertex (Index 2-3) and direction (Index 4-6) and propagate them to the other related variables (Index 7-13). We then run the BDT again after shifting these variables. The resulting shifts to the BDT score are much more symmetric, leading to a negligible efficiency error of 0.024%, 0.25%, and 0.054% for SLE 34-hit, SLE 31-hit, and WIT respectively.

11.3.5 Energy-Correlated Uncertainties

Unlike the above energy-uncorrelated uncertainties, the energy-correlated uncertainties cannot be added in quadrature when performing a fit to the energy spectrum. These errors include the electron energy scale, electron energy resolution, and B solar neutrino energy spectrum. In [27], these were determined with a theoretical calculation considering the detector resolutions, and we simply extend the calculations down to 2.49 MeV. Table 11.5 with these errors is used for the fits to the energy spectrum in Section 12.2.

Range	B spectrum	Energy scale	Energy resolution
2.99–3.49 MeV	0.065%	0.27%	0.018%
2.49–2.99 MeV	0.084%	0.21%	0.016%

Table 11.5: Correlated energy error table extended to 2.49 MeV.

For the measurement of the solar interaction rate in Section 12.1, we follow the same methods for the γ B spectrum and energy resolution systematic errors in which signal MC events are reweighted according to the relevant error in Table 11.5 before repeating signal extraction. The effect of this reweighting in the fit number of signal events was nonzero but insignificant, for both errors in all bins and datasets.

Since the energy scale error is the largest of the energy-correlated errors, we give this error special treatment for the interaction rate measurements. The energy scale error is treated in a similar fashion as the systematic errors on the cut variables in the previous analysis, applying a shift to the energy values. Since the electron energy in this analysis is used as an input to the BDT as well as to define the energy boundaries of the fit bins, we again evaluate these effects separately. The error in the absolute energy scale is estimated to be as determined from LINAC and DT data [27]. We apply this shift in several different ways and assign that with the largest as the systematic uncertainty. The first method is again to shift the energy BDT input variable (Index 14) and observe the change in BDT cut signal efficiency. The second is to shift the energy values of the events in the final sample and count the number in 2.99 MeV - 3.49 MeV after this shift. We also consider a third approach that uses the shifted energy for both the BDT input and energy binning boundary.

Shift method	WIT	34-hit	31-hit
Shift BDT input up	3.7%	4.0%	3.9%
Shift BDT input down	-3.6%	-3.9%	-4.0%
Shift final evt. up	-4.4%	-9.4%	-6.2%
Shift final evt. down	4.5%	10.4%	6.1%
Shift both up	-0.7%	-6.3%	-2.1%
Shift both down	0.9%	5.8%	2.5%

Table 11.6: Energy scale error in signal MC efficiency for all datasets using plus and minus shifts under various methods. The method with the highest error is in bold.

The two shift methods have opposite effects on the signal efficiency, so they cancel out when both the BDT inputs and final event energies are shifted in the third method. Although we

believe that the third method is the most appropriate since the shifts are self-consistent, we conservatively assign the largest error in each dataset as the energy scale systematic error as indicated in bold in Table 11.6.

11.3.6 Error Summary

Table 11.7 shows a summary of the systematic uncertainties for the 2.99 MeV - 3.49 MeV WIT and SLE 34-hit datasets in which a solar signal was observed in Chapter 12. Sources of error that were less significant including signal extraction, γ B spectrum, energy resolution, and the rest of the BDT input variables are not listed. The “Total” systematic error is quoted for the interaction rate measurements (Section 12.1). The “Total uncorrelated” errors along with the energy-correlated errors as calculated in [27] extended down to 2.99 MeV are used for the energy spectrum fit (Section 12.2).

Error Source	WIT	34-hit SLE
Background shape	11%	42%
Trigger efficiency	8.6%	16%
fit goodness	2.1%	2.1%
Energy scale	4.5%	10.4%
Total uncorrelated	14%	45%
Total	15%	46%

Table 11.7: Systematic error summary with significant error sources. “Total” includes all errors added in quadrature, and “Total uncorrelated” excludes the energy scale error.

Chapter 12

Solar Energy Spectrum Results

12.1 Interaction Rate Measurement in 3.49 MeV

Table 12.1 shows the results for each energy and MSG bin for all datasets as well as the results aggregated over MSG bins and combined over multiple datasets. The number of observed signal events is compared to the expected number of unoscillated signal events calculated using the same approach as [27] based on the neutral current results from SNO [125].

No statistically significant indication of a signal is found in the 2.49 MeV - 2.99 MeV bin due to the large amount of radioactive background and in the 31-hit period 2.99 MeV - 3.49 MeV bin since there is little livetime left after removing runs that are available in WIT. In the 2.99 MeV - 3.49 MeV bin of the other datasets, there is a significance for a solar signal in WIT and in the 34-hit period after accounting for systematic errors. Together, these results give a indication for a solar signal in this energy bin for SK-IV. The event rate measured in this electron kinetic energy bin corresponds to a solar neutrino flux of for the entire B neutrino energy range.

Dataset	E_{kin} Bin	MSG Bin	Exp. (Unosc.)	Obs.	Obs./Exp.
WIT	2.49–2.99 MeV	0–0.35	41		
	2.49–2.99 MeV	0.35–0.45	71		
	2.49–2.99 MeV	0.45–1	115		
	2.99–3.49 MeV	0–0.35	247		
	2.99–3.49 MeV	0.35–0.45	362		
	2.99–3.49 MeV	0.45–1	467		
	2.49–2.99 MeV	0–1	227		
	2.99–3.49 MeV	0–1	1077		
SLE 34-hit	2.99–3.49 MeV	0–0.35	230		
	2.99–3.49 MeV	0.35–0.45	356		
	2.99–3.49 MeV	0.45–1	506		
	2.99–3.49 MeV	0–1	1092		
SLE 31-hit	2.99–3.49 MeV	0–0.35	70		
	2.99–3.49 MeV	0.35–0.45	99		
	2.99–3.49 MeV	0.45–1	121		
	2.99–3.49 MeV	0–1	291		
WIT+34-hit	2.99–3.49 MeV	0–1	2169		

Table 12.1: Expected number of unoscillated MC signal events, fit number of signal events, and their ratio in each energy and MSG bin. Measurements are given as measurement statistical $\pm$ systematic error. Systematic errors are only calculated for the 0-1 MSG bins, and only if a signal is observed. When aggregating across all MSG bins, bins are weighted by their statistical error. When aggregating across the datasets, datasets are weighted by their total error.

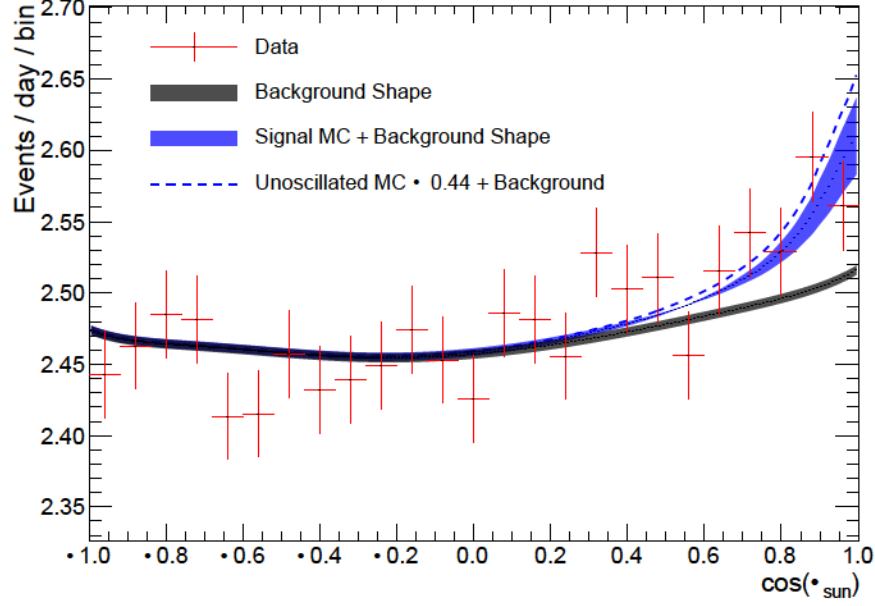


Figure 12.1: Solar angle distribution in 2.99 MeV - 3.49 MeV aggregated over the WIT and SLE 34-hit datasets. The definition of colors is the same as in Figure 11.9. The additional blue dashed line indicates the signal shape if the average interaction rate in 3.49 MeV - 19.49 MeV measured in [27] is assumed.

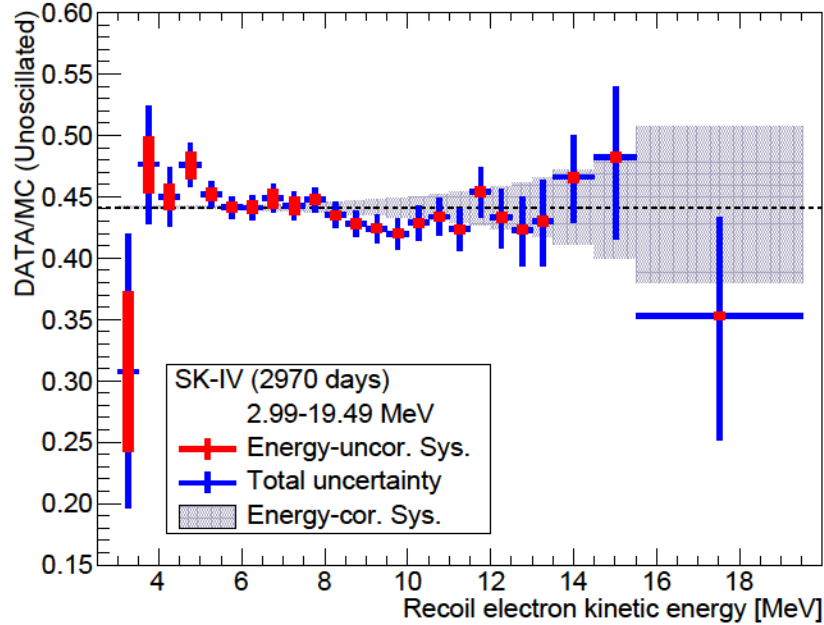


Figure 12.2: The measured electron kinetic energy spectrum in SK-IV with the addition of the lowest energy bin. The blue points and bars show the observed rate divided by the expected event rate assuming no neutrino oscillation, with statistical and energy-uncorrelated uncertainties. The red bars and gray bands show the energy-uncorrelated and energy-correlated systematic uncertainties. The dashed black line shows the average ratio.

In comparing the observed number of signal events to the expected unoscillated number, Figure 12.2 shows the measured ratio compared to those of higher energy bins. The combined ratio in this energy bin has a slight deficit compared to both the 0.44 ratio expected for an energy-independent survival probability () and with the previous 0.47 best-fit prediction for the case of an upturn () [27]. Figure 12.1 shows the combined solar angle distribution along with the distribution implied in the 0.44 ratio case. Individual datasets as well as their individual MSG bins generally have similar or better agreement with the expected ratio. However, there is a notable exception in the highest MSG bin of the 34-hit dataset (Table 12.1 and Figure 11.10c). The lower two MSG bins are consistent with the 0.44 ratio, but the highest bin has a deficit considering statistical error only.

12.2 Survival Probability Fit

Following the fits to the solar angle distributions, we repeat the global solar oscillation analysis described in [27] after adding the combined 2.99 MeV - 3.49 MeV measurement in Table 12.1 to the existing observed to expected ratio spectrum of SK-IV. Fits to the survival probability as a function of neutrino energy are then conducted using the quadratic and exponential parameterizations. The resulting coefficients are shown in Tables 12.2 and 12.3.

quad		—		—	
exp		—		—	
SK I-IV [27]			—		
Add	3.49 MeV				

Table 12.2: Resulting coefficients comparisons for the quadratic parameterization fit.

We then compare these fits along with their intervals to the contours of [27] in Figure

SK I-IV [27]					
Add	3.49 MeV		0.04		

Table 12.3: Resulting coefficients and χ^2 comparisons for the exponential parameterization fit.

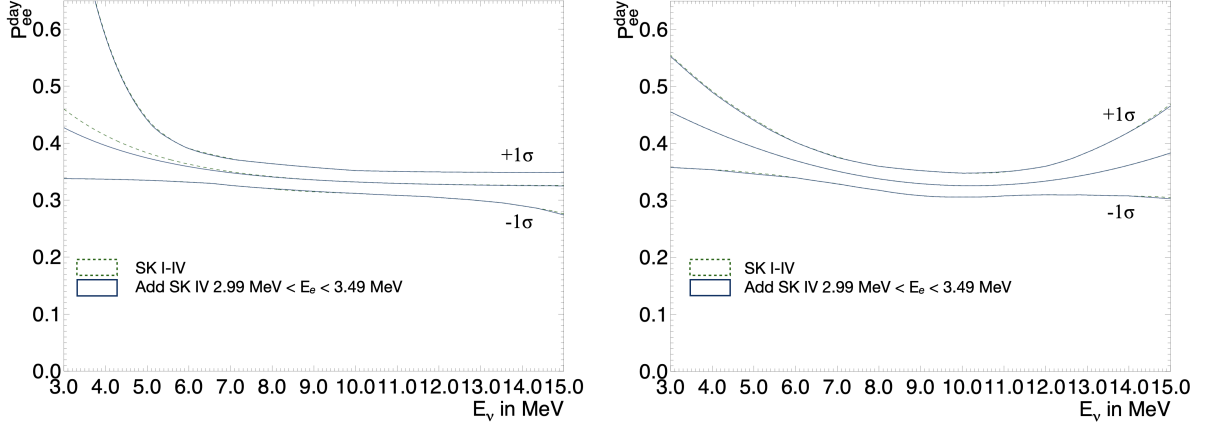


Figure 12.3: ν_e survival probability spectra as a function of neutrino energy obtained from the exponential (left) and quadratic (right) fits to SK recoil electron data. The green dashed lines use all SK-IV and earlier data from [27], and the blue solid lines add an additional 2.99 MeV - 3.49 MeV bin for SK-IV. The best fits and 1σ confidence intervals are shown.

12.3. The additional bin has a low impact on the fits due to the better signal-to-background ratio in the higher energy bins. The χ^2 of the fit in [27] is 65.56 (64.97) for the exponential (quadratic) fit, and these values change to 65.60 (65.03) with the additional bin. There is a slight decrease in the upper confidence interval at lower energies, and there is a noticeable shift in the exponential best fit due to the lower observed to expected measurement compared to higher energy bins.

Chapter 13

Summary and Future Sensitivity

The addition of boosted decision trees for event selection in this analysis of the solar neutrino events during SK-IV has improved background rejection in the radioactive background dominated energy region below the previous energy threshold. This analysis utilizes the WIT system implemented during SK-IV, which features improved trigger efficiency at these lower energies and whose data has not yet been used as a source of solar signal events. The results of fits to the solar angle gives an indication for a solar neutrino signal in 2.99 MeV - 3.49 MeV with a ratio of observed interaction rate to expected without oscillations of . Due to large uncertainties, however, the result is consistent with both the presence and absence of a solar upturn, and the next higher energy bins still provide the bulk of the sensitivity for fits to the survival probability spectrum.

This new energy bin as well as the next higher energy bins are still statistically limited (Figure 12.2). Due to the long livetime and low background levels of SK-IV compared to other phases, the SK-IV dataset ending in 2018 has the greatest weight in the study of the upturn. The lowest level of background was achieved near the end of SK-V, but the usable livetime of this low background period is about 100 days. While the addition of gadolinium in

SK-VI and later phases greatly benefits $\bar{\nu}_e$ analyses, the improved neutron tagging does not benefit the solar ν_e analysis. On the contrary, the dissolved gadolinium sulfate introduces additional radioactive background. Assuming the same SK-IV detector conditions, these measurements will approach the systematics limit with a few more years: the significance of an upturn compared with an energy-independent θ_{12} would increase from 2.1σ at the end of SK-IV to 3.1σ in 2029⁷. This prediction does not include the additional low energy bin considered in this dissertation, but a solar neutrino signal observation will be difficult at 3.99 MeV and is likely not possible at 3.49 MeV in these later phases, which reduces sensitivity to the upturn. As selection methods continue to evolve, however, the significance and impact of a solar neutrino signal at these energies may improve.

As for other aspects of the solar analysis, Figure 13.1 shows the expected improvement in the measurements of the θ_{12} and δ oscillation parameters under similar assumptions. These measurements are largely systematically limited at the end of SK-IV, but the precision for δ can improve with a more precise measurement of the day/night asymmetry. The day/night asymmetry significance is expected to improve from 1.5σ to 2.1σ in 2029.

In addition to the expected research into the fundamental oscillation parameters, Hyper-K expects to contribute to all of these questions on solar neutrinos. While low energy analyses will greatly benefit from a fiducial volume that is eight times bigger than Super-K's and PMTs with double the photon detection efficiency, Hyper-K has a lower photocoverage (20% vs 40%) and rock overburden to protect it from spallation background (650 m vs 1000 m). The leading radioactive background level from radon ^{222}Rn Bi decay is expected to be similar, but the PMTs will feature approximately half the ^{232}Th Tl contamination. Altogether, Hyper-K expects to have a 4.5 MeV selection threshold in neutrino energy, and the expected sensitivity to the upturn under this assumption is 3.1σ after 10 years (2038) [93].

⁷All future sensitivities for the Super-K solar analysis in this chapter are presented as individual student work and are not considered official by the collaboration.

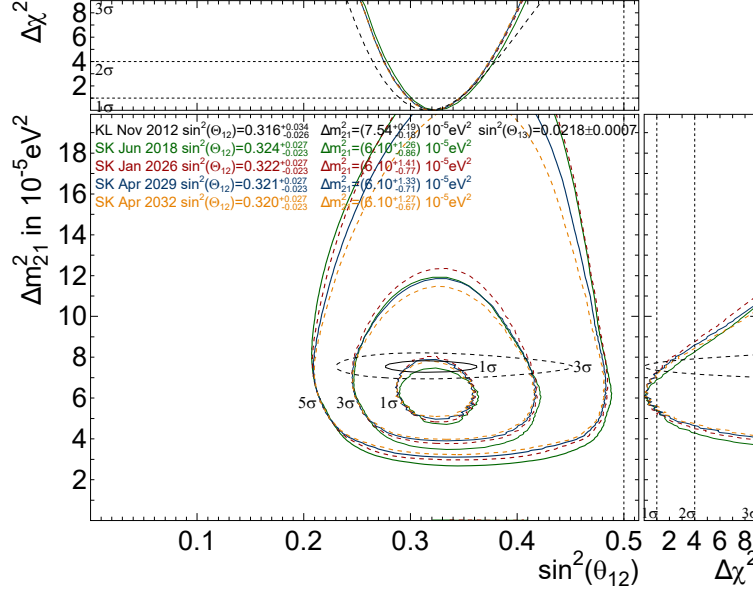


Figure 13.1: Predicted confidence level contours for Δm_{21}^2 and $\sin^2(\theta_{12})$ assuming SK-IV conditions combined with previous phases (green). Other colors indicate various future dates assuming SK-IV central values. KamLAND contours are shown in black.

JUNO recently provided world-leading measurements for $\sin^2(\theta_{12})$ and Δm_{21}^2 at $\sim 10^{-5} \text{ eV}^2$ [126]. Due to the high standards of maintaining ultralow levels of radioactive impurities, high light yield due to the scintillator detector technology, and 78% photocoverage, JUNO expects to have a 2 MeV energy threshold for neutrinos to explore this solar upturn transition region. At this $\sim 2 \text{ MeV}$ value, JUNO expects to provide a $\sim 3\sigma$ upturn significance in 10 years (2035) [127].

Bibliography

- [1] W. Pauli. Dear radioactive ladies and gentlemen. *Phys. Today*, 31(9):27, 1978.
- [2] F. Reines and C. L. Cowan. The neutrino. *Nature*, 178:446–449, 1956.
- [3] G. Danby et al. Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos. *Phys. Rev. Lett.*, 9:36–44, Jul 1962.
- [4] Relativistic Beaming. Wikimedia Commons, 2023.
- [5] Relativistic Aberration of Light. University of Texas at Austin.
- [6] Proton proton cycle. Wikimedia Commons, 2012.
- [7] F. Villante and A. Serenelli. The Relevance of Nuclear Reactions for Standard Solar Models Construction. *Frontiers in Astronomy and Space Sciences*, 7, Mar 2021.
- [8] K. Hirata et al. Observation of a neutrino burst from the supernova SN1987A. *Phys. Rev. Lett.*, 58(14):1490–1493, Apr 1987.
- [9] R. M. Bionta et al. Observation of a neutrino burst in coincidence with supernova 1987A in the Large Magellanic Cloud. *Phys. Rev. Lett.*, 58:1494–1496, Apr 1987.
- [10] E. N. Alekseev et al. Detection of the Neutrino Signal from Supernova 1987A Using the INR Baksan Underground Scintillation Telescope. In *European Southern Observatory Conference and Workshop Proceedings*, 26:237, Jan 1987.
- [11] Y. Suwa et al. Observing Supernova Neutrino Light Curves with Super-Kamiokande: Expected Event Number over 10 s. *The Astrophysical Journal*, 881:139, Aug 2019.
- [12] K. Abe et al. Search for Diffuse Supernova Neutrino Background with 956.2 days of Super-Kamiokande Gadolinium Dataset, 2025. arXiv:2511.02222.
- [13] P. Mészáros. Astrophysical Sources of High-Energy Neutrinos in the IceCube Era. *Annual Review of Nuclear and Particle Science*, 67:45–67, 2017.
- [14] S. Aiello et al. Observation of an ultra-high-energy cosmic neutrino with KM3NeT. *Nature*, 638:376–382, Feb 2025.
- [15] R. Davis, D. S. Harmer, and K. C. Hoffman. Search for Neutrinos from the Sun. *Phys. Rev. Lett.*, 20:1205–1209, May 1968.

- [16] Y. Fukuda et al. Evidence for Oscillation of Atmospheric Neutrinos. *Phys. Rev. Lett.*, 81:1562–1567, Aug 1998.
- [17] Q. R. Ahmad et al. Measurement of the Rate of $d + p \rightarrow p + e$ Interactions Produced by Solar Neutrinos at the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 87:071301, Jul 2001.
- [18] S. Navas et al. Review of Particle Physics. *Phys. Rev. D*, 110:030001, 2024. and May 2025 update.
- [19] M. Fukugita and T. Yanagida. Baryogenesis Without Grand Unification. *Phys. Lett. B*, 174:45–47, 1986.
- [20] W. Buchmuller and M. Plumacher. Baryon asymmetry and neutrino mixing. *Phys. Lett. B*, 389:73–77, 1996.
- [21] S. P. Mikheev and A. Y. Smirnov. Resonance amplification of oscillations in matter and spectroscopy of solar neutrinos. *Sov. J. Nucl. Phys.*, 42:913–, Jan 1985.
- [22] L. Wolfenstein. Neutrino oscillations in matter. *Phys. Rev. D*, 17:2369–2374, May 1978.
- [23] K. Zuber. *Neutrino Physics*. CRC Press, 2020.
- [24] M. Agostini et al. Simultaneous precision spectroscopy of ^7Be , ^8B , and ^9Be solar neutrinos with Borexino Phase-II. *Phys. Rev. D*, 100:082004, Oct 2019.
- [25] A. Gando et al. ^7Be solar neutrino measurement with KamLAND. *Phys. Rev. C*, 92:055808, Nov 2015.
- [26] G. L. Fogli, E. Lisi, D. Montanino, and A. Palazzo. Quasivacuum solar neutrino oscillations. *Phys. Rev. D*, 62:113004, Oct 2000.
- [27] K. Abe et al. Solar neutrino measurements using the full data period of Super-Kamiokande-IV. *Phys. Rev. D*, 109:092001, May 2024.
- [28] D. S. Ayres et al. The NOvA Technical Design Report. *OSTI.GOV*, Oct 2007.
- [29] K. Yonehara. Megawatt upgrade of NuMI target system. In *NuFact 2021*, 402:107, 2022.
- [30] M. Convery, M. Lindgren, S. Nagaitsev, and V. Shiltsev. Fermilab Accelerator Complex: Status and Improvement Plans. *OSTI.GOV*, Dec 2018.
- [31] P. Adamson et al. The NuMI neutrino beam. *Nucl. Instrum. Methods Phys. Res. A*, 806:279–306, 2016.
- [32] K.S. Hirata et al. Observation of a small atmospheric $\bar{\nu}_\mu/\nu_\mu$ ratio in Kamiokande. *Physics Letters B*, 280(1):146–152, 1992.

- [33] B. Bogart. Exploring nucleon-nucleon forces in-medium with the GiBUU model and MicroBooNE neutrino-nuclei data. University of Washington Rising Researchers Seminar, May 2025.
- [34] J. Wolcott. Cross section models & uncertainties in NOvA. NuSTEC cross-target generator working group seminar, Dec 2022.
- [35] M. Baird et al. Event Reconstruction Techniques in NOvA. *J. Phys.: Conf. Ser.*, 664(7):072035, Dec 2015.
- [36] R. Krishnapuram and J. M. Keller. A possibilistic approach to clustering. *IEEE Transactions on Fuzzy Systems*, 1(2):98–110, 1993.
- [37] R. E. Kalman. A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, 82(1):35–45, Mar 1960.
- [38] L. Aliaga et al. Neutrino Flux Predictions for the NuMI Beam. *Phys. Rev. D*, 94(9):092005, 2016.
- [39] C. Andreopoulos et al. The GENIE neutrino Monte Carlo generator. *Nucl. Instrum. Methods Phys. Res. A*, 614(1):87–104, 2010.
- [40] S. Agostinelli et al. Geant4—a simulation toolkit. *Nucl. Instrum. Methods Phys. Res. A*, 506(3):250–303, 2003.
- [41] A. Aurisano et al. A Convolutional Neural Network Neutrino Event Classifier. *JINST*, 11(09):P09001, 2016.
- [42] P. Adamson et al. Constraints on Oscillation Parameters from Appearance and Disappearance in NOvA. *Phys. Rev. Lett.*, 118:231801, Jun 2017.
- [43] A. G. Howard et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *CoRR*, abs/1704.04861, 2017.
- [44] C. Szegedy et al. Going deeper with convolutions. In *CVPR*, pages 1–9, 2015.
- [45] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger. Densely Connected Convolutional Networks. In *CVPR*, pages 2261–2269, Jul 2017.
- [46] A. Vaswani et al. Attention is All you Need. In *Advances in Neural Information Processing Systems*, 30, 2017.
- [47] T. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar. Focal Loss for Dense Object Detection. In *ICCV*, pages 2999–3007, Los Alamitos, CA, USA, Oct 2017. IEEE Computer Society.
- [48] C. Sweeney. *Measurement of the muon neutrino charged-current single charged pion production cross-section in the NOvA Near Detector*. PhD thesis, University Coll. London, 2023.

- [49] A. Chatla. *Classification of particles with a convolutional neural network for neutrino and antineutrino events in the NOVA experiment*. PhD thesis, University of Hyderabad, 2024.
- [50] K. Simonyan, A. Vedaldi, and A. Zisserman. Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps. In *Workshop at ICLR*, 2014.
- [51] J. Wolcott. New neutrino oscillation results from nova with 10 years of data, Jun 2024.
- [52] S. Abubakar et al. Precision Measurement of Neutrino Oscillation Parameters with 10 Years of Data from the NOvA Experiment. *Phys. Rev. Lett.*, 136:011802, Jan 2026.
- [53] M. A. Acero et al. Improved measurement of neutrino oscillation parameters by the NOvA experiment. *Phys. Rev. D*, 106:032004, Aug 2022.
- [54] S. Abubakar et al. Joint neutrino oscillation analysis from the T2K and NOvA experiments. *Nature*, 646(8086):818–824, 2025.
- [55] F. P. An et al. Precision Measurement of Reactor Antineutrino Oscillation at Kilometer-Scale Baselines by Daya Bay. *Phys. Rev. Lett.*, 130(16):161802, 2023.
- [56] A. I. Kalitkina. Estimating Neutrino Energy for the NOvA 3-Flavor Analysis. In *AYSS 2024*, 22(5):1119–1121, Oct 2025.
- [57] A. C. Booth. *Electron neutrino appearance at the NOvA experiment*. PhD thesis, Sussex U., Oct 2021.
- [58] M. A. Acero et al. New constraints on oscillation parameters from $\bar{\nu}_e$ appearance and disappearance in the NOvA experiment. *Phys. Rev. D*, 98:032012, Aug 2018.
- [59] P. Singh. *Extraction of neutrino oscillation parameters using a simultaneous fit of disappearance and appearance data with the NOvA experiment*. PhD thesis, Delhi University, Sep 2019.
- [60] M. A. Acero et al. Measurement of the double-differential muon-neutrino charged-current inclusive cross section in the NOvA near detector. *Phys. Rev. D*, 107:052011, Mar 2023.
- [61] M. A. Acero et al. Measurement of the $\bar{\nu}_e$ -Nucleus Charged-Current Double-Differential Cross Section at $\sqrt{s} = 1.96$ TeV Using NOvA. *Phys. Rev. Lett.*, 130:051802, Feb 2023.
- [62] M. A. Acero et al. Measurement of $\bar{\nu}_e$ charged-current inclusive μ^+ production in the NOvA near detector. *Phys. Rev. D*, 107:112008, Jun 2023.
- [63] Z. Kohley et al. Modeling interactions of intermediate-energy neutrons in a plastic scintillator array with Geant4. *Nucl. Instrum. Methods Phys. Res. A*, 682:59–65, 2012.

- [64] M. A. Acero et al. Expanding neutrino oscillation parameter measurements in NOvA using a Bayesian approach. *Phys. Rev. D*, 110(1):012005, 2024.
- [65] M. A. Acero et al. Monte Carlo method for constructing confidence intervals with unconstrained and constrained nuisance parameters in the NOvA experiment. *JINST*, 20(02):T02001, 2025.
- [66] G. J. Feldman and R. D. Cousins. A Unified approach to the classical statistical analysis of small signals. *Phys. Rev. D*, 57:3873–3889, 1998.
- [67] J. S. Speagle. A Conceptual Introduction to Markov Chain Monte Carlo Methods, 2020. arXiv:1909.12313.
- [68] S. Brooks, A. Gelman, G.L. Jones, and X.-L Meng. *Handbook of Markov Chain Monte Carlo*. May 2011.
- [69] J. Lee, W. Sung, and J.-H. Choi. Metamodel for Efficient Estimation of Capacity-Fade Uncertainty in Li-Ion Batteries for Electric Vehicles. *Energies*, 8(6):5538–5554, 2015.
- [70] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21:1087–1092, 1953.
- [71] W. K. Hastings. Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometrika*, 57:97–109, 1970.
- [72] M. Betancourt. A Conceptual Introduction to Hamiltonian Monte Carlo. 2017. arXiv:1701.02434.
- [73] J. M. Dickey. The Weighted Likelihood Ratio, Linear Hypotheses on Normal Location Parameters. *The Annals of Mathematical Statistics*, 42(1):204–223, 1971.
- [74] J. Mulder, E. J. Wagenmakers, and M. Marsman. A Generalization of the Savage-Dickey Density Ratio for Testing Equality and Order Constrained Hypotheses. Apr 2020.
- [75] R. L. Workman et al. Review of Particle Physics. *PTEP*, 2022:083C01, 2022.
- [76] K. Abe et al. First Joint Oscillation Analysis of Super-Kamiokande Atmospheric and T2K Accelerator Neutrino Data. *Phys. Rev. Lett.*, 134(1):011801, 2025.
- [77] T. Wester et al. Atmospheric neutrino oscillation analysis with neutron tagging and an expanded fiducial volume in Super-Kamiokande I–V. *Phys. Rev. D*, 109(7):072014, 2024.
- [78] P. Dunne et al. *Oscillation results from Neutrino 2022*, Jul 2022. Zenodo.6908532.
- [79] R. Abbasi et al. Measurement of Atmospheric Neutrino Oscillation Parameters Using Convolutional Neural Networks with 9.3 Years of Data in IceCube DeepCore. *Phys. Rev. Lett.*, 134(9):091801, 2025.

- [80] P. Adamson et al. Precision Constraints for Three-Flavor Neutrino Oscillations from the Full MINOS+ and MINOS Dataset. *Phys. Rev. Lett.*, 125(13):131802, 2020.
- [81] D. Carabadjac. T2K Neutrino Oscillation Analysis. In *ICHEP*, Jul 2024.
- [82] S. Jeon et al. Measurement of reactor antineutrino oscillation parameters using the full 3800-day dataset of the reno experiment. *Phys. Rev. D*, 111:112006, Jun 2025.
- [83] F. P. An et al. Measurement of Electron Antineutrino Oscillation Amplitude and Frequency via Neutron Capture on Hydrogen at Daya Bay. *Phys. Rev. Lett.*, 133(15):151801, 2024.
- [84] I. Yu. New Results from RENO. In *Neutrino*, Jun 2018.
- [85] M. Gonchar, L. Kolupaeva, and Y. Malyshkin. NuOsc. <https://git.jinr.ru/nu/osc>, May 2024.
- [86] P. Baldi, J. Bian, L. Hertel, and L. Li. Improved energy reconstruction in NOvA with regression convolutional neural networks. *Phys. Rev. D*, 99:012011, Jan 2019.
- [87] B. J. Jargowsky. *A Measurement of Appearance and Disappearance Using 10 Years of Data from the NOvA Experiment*. PhD thesis, UC Irvine, 2024.
- [88] S. Wu. *Machine Learning-Based Energy Estimation for Sterile Neutrino Searches in the NOvA Experiment*. PhD thesis, Minnesota U., Aug 2025.
- [89] D. Torbunov. *Improving Energy Estimation at NOvA with Recurrent Neural Networks*. PhD thesis, Minnesota U., May 2021.
- [90] M. Wallbank. The NOvA Test Beam Program. In *ICHEP*, page 188, 2021.
- [91] A. Abusleme et al. Sub-percent precision measurement of neutrino oscillation parameters with JUNO. *Chin. Phys. C*, 46(12):123001, Dec 2022.
- [92] B. Abi et al. Deep Underground Neutrino Experiment (DUNE), Far Detector Technical Design Report, Volume II: DUNE Physics, 2020. arXiv:2002.03005.
- [93] K. Abe et al. Hyper-Kamiokande Design Report, 2018.
- [94] S. Fukuda et al. The Super-Kamiokande detector. *Nucl. Instrum. Methods Phys. Res. A*, 501(2):418–462, 2003.
- [95] Y. Nakano et al. Measurement of the radon concentration in purified water in the Super-Kamiokande IV detector. *Nucl. Instrum. Methods Phys. Res. A*, 977:164297, 2020.
- [96] H. Nishino et al. High-speed charge-to-time converter ASIC for the Super-Kamiokande detector. *Nucl. Instrum. Methods Phys. Res. A*, 610(3):710–717, 2009.
- [97] K. Abe et al. First gadolinium loading to Super-Kamiokande. *Nucl. Instrum. Methods Phys. Res. A*, 1027:166248, 2022.

- [98] G. Carminati. The new Wide-band Solar Neutrino Trigger for Super-Kamiokande. In *TAUP*, 61:666–672, 2015.
- [99] M. Elnimr. Low Energy $\bar{\nu}_\mu$ Solar Neutrinos with the Wideband Intelligent Trigger at Super-Kamiokande. *J. Phys. Conf. Ser.*, 888(1):012189, Sep 2017.
- [100] S. Locke et al. New methods and simulations for cosmogenic induced spallation removal in Super-Kamiokande-IV. *Phys. Rev. D*, 110:032003, Aug 2024.
- [101] S. Abe et al. Combined Pre-supernova Alert System with KamLAND and Super-Kamiokande. *ApJ*, 973:140, 2024.
- [102] M. Nakahata et al. Calibration of Super-Kamiokande using an electron LINAC: The Super-Kamiokande Collaboration. *Nucl. Instrum. Methods Phys. Res. A*, 421(1):113–129, 1999.
- [103] E. Blaufuss et al. ^{252}Cf as a calibration source for Super-Kamiokande. *Nucl. Instrum. Methods Phys. Res. A*, 458(3):638–649, 2001.
- [104] J. P. Cravens et al. Solar neutrino measurements in Super-Kamiokande-II. *Phys. Rev. D*, 78:032002, Aug 2008.
- [105] M. Smy. Low Energy Event Reconstruction and Selection in Super-Kamiokande-III. In *ICRC-30*, 5:1279–1282, Jul 2007.
- [106] J. Hosaka et al. Solar neutrino measurements in Super-Kamiokande-I. *Phys. Rev. D*, 73:112001, Jun 2006.
- [107] K. Abe et al. Solar neutrino measurements in Super-Kamiokande-IV. *Phys. Rev. D*, 94:052010, Sep 2016.
- [108] Y. Nakano et al. Measurement of radon concentration in super-Kamiokande’s buffer gas. *Nucl. Instrum. Methods Phys. Res. A*, 867:108–114, 2017.
- [109] K. Abe et al. Calibration of the Super-Kamiokande detector. *Nucl. Instrum. Methods Phys. Res. A*, 737:253–272, 2014.
- [110] Y. Zhang et al. First measurement of radioactive isotope production through cosmic-ray muon spallation in Super-Kamiokande IV. *Phys. Rev. D*, 93:012004, Jan 2016.
- [111] Y. Maekawa et al. Measurement of the branching ratio of ^{10}B , ^{10}C , ^{10}N , and ^{10}O isotopes through the nuclear muon capture reaction in the Super-Kamiokande detector, 2025.
- [112] N. Prouse, P. de Perio, and W. Fedorko. Machine Learning Techniques to Enhance Event Reconstruction in Water Cherenkov Detectors. In *NuFact*, 8(1), 2023.
- [113] K. He et al. Deep Residual Learning for Image Recognition. In *CVPR*, pages 770–778, 2016.

- [114] WatChMaL. <https://github.com/WatChMaL/WatChMaL>.
- [115] L. Berns. Machine Learning based ring reconstruction for water Cherenkov detectors. Neutrinos to Open Elementary Particles and Space Meeting, Dec 2020.
- [116] D. Maksimović, M. Nieslony, and M. Wurm. CNNs for enhanced background discrimination in DSNB searches in large-scale water-Gd detectors. *JCAP*, (11):051, 2021.
- [117] T. N. Kipf and M. Welling. Semi-Supervised Classification with Graph Convolutional Networks. In *ICLR*, 2017.
- [118] K. Xu, W. Hu, J. Leskovec, and S. Jegelka. How Powerful are Graph Neural Networks? In *ICLR*, 2019.
- [119] Y. Wang et al. Dynamic Graph CNN for Learning on Point Clouds. *ACM Trans. Graph.*, 38(5), Oct 2019.
- [120] C. R. Qi et al. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. In *CVPR*, Dec 2016.
- [121] T. Chen and C. Guestrin. XGBoost: A Scalable Tree Boosting System. In *KDD*, pages 785–794, New York, NY, USA, 2016. ACM.
- [122] C. Wang, C. Deng, and S. Wang. Imbalance-XGBoost: leveraging weighted and focal losses for binary label-imbalanced classification with XGBoost. *Pattern Recognit. Lett.*, 136:190–197, 2020.
- [123] J. Tian et al. Recall Loss for Imbalanced Image Classification and Semantic Segmentation. In *ICLR*, 2021.
- [124] R. Brun. *GEANT: Detector Description and Simulation Tool*. CERN Program Library. CERN, Geneva, 1993. Long Writeup W5013.
- [125] Q. R. Ahmad et al. Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 89:011301, Jun 2002.
- [126] A. Abusleme et al. First measurement of reactor neutrino oscillations at JUNO, 2025. arXiv:2511.14593.
- [127] A. Abusleme et al. Feasibility and physics potential of detecting $\bar{\nu}_\beta$ solar neutrinos at JUNO. *Chin. Phys. C*, 45(2):023004, 2021.