

THESIS

ACCELERATING NOVA'S FELDMAN-COUSINS PROCEDURE USING HIGH
PERFORMANCE COMPUTING PLATFORMS

Submitted by

Derek Doyle

Department of Physics

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Spring 2019

Master's Committee:

Advisor: Norm Buchanan

John Harton

Louis-Noël Pouchet

Copyright by Derek Doyle 2019

All Rights Reserved

ABSTRACT

ACCELERATING NOVA'S FELDMAN-COUSINS PROCEDURE USING HIGH PERFORMANCE COMPUTING PLATFORMS

In order to assess the compatibility between models containing physically constrained parameters and small-signal data, uncertainties often must be calculated by Monte Carlo simulation to account for non-normally distributed errors. This is the case for neutrino oscillation experiments, where neutrino-matter weak interactions are rare and beam intensity at the far site is low. The NuMI Off-axis ν_e Appearance (NOvA) collaboration attempts to measure the parameters governing neutrino oscillations within the PMNS oscillation model by comparing model predictions to a small data set of neutrino interactions. To account for non-normality, NOvA uses the computationally intensive Feldman-Cousins (FC) procedure, which involves fitting thousands of independent pseudoexperiments to generate empirical distribution functions that are used to calculate the significance of observations. I, along with collaborators on NOvA and Scientific Discovery through Advanced Computing: High Energy Physics Data Analytics (SciDAC-4) collaborations, have implemented the FC procedure utilizing the High Performance Computing (HPC) facilities at the National Energy Research Scientific Computing Center (NERSC). With this implementation, we have successfully processed NOvA's complete FC corrections for our recent neutrino + antineutrino appearance analysis in 36 hours: a speedup factor of 50 as compared to the methods used in previous analyses.

TABLE OF CONTENTS

ABSTRACT	ii
LIST OF TABLES	v
LIST OF FIGURES	vi
LIST OF ACRONYMS	viii
Chapter 1. Introduction	1
Chapter 2. Neutrino Physics	3
2.1. The Standard Model of Particle Physics	3
2.2. Neutrino Oscillations	4
Chapter 3. The NOvA Experiment	8
3.1. Long-Baseline Neutrino Experiments	8
3.2. The Neutrino Beam	9
3.3. The NOvA Detectors	10
Chapter 4. The NOvA Neutrino Oscillation Analysis	14
4.1. Event Identification and Reconstruction	15
4.2. Event Selection and Extrapolation	17
4.3. Systematics	18
4.4. Oscillation Fits	18
Chapter 5. The Feldman-Cousins Procedure	20
5.1. Confidence Interval Construction	20

5.2.	The Unified Ordering Principle.....	24
5.3.	Significance Profiles and Confidence Regions.....	26
5.4.	Calculating Acceptance Intervals of Neutrino Oscillation Parameters.....	27
5.5.	Generating Pseudoexperiments.....	28
5.6.	Producing Corrected Intervals.....	31
5.7.	Inverted Mass Hierarchy Rejection.....	31
Chapter 6.	Results and the Role of High Performance Computing.....	35
6.1.	The Scale of NOvA's FC Corrections.....	35
6.2.	Motivation for High Performance Computing.....	36
6.3.	Utilizing NERSC Facilities.....	37
6.4.	Implementation.....	38
6.5.	Job Submission and Load Balancing.....	39
6.6.	Performance Results.....	40
6.7.	Results of the Feldman-Cousins (FC) Corrections.....	41
6.8.	Domain Decomposition with DIY.....	42
Chapter 7.	Conclusion.....	45
Bibliography	46

LIST OF TABLES

4.1	NOvA's best-fit parameters	19
6.1	Summary of FC pseudoexperiments	36
6.2	NERSC machine configurations	40
6.3	Mean fit times on NERSC systems	40
6.4	NOvA's best-fit parameters	42
6.5	Mass hierarchy and octant rejection significances	42

LIST OF FIGURES

2.1	The Standard Model of Particle Physics.....	4
2.2	Rotation in \mathbb{R}^2	5
2.3	Neutrino mass hierarchy	7
3.1	Sketch of the NOvA experiment	9
3.2	Neutrino production from NuMI	10
3.3	Off-axis NuMI beam distributions.....	11
3.4	Appearance probability	12
3.5	NOvA detectors	13
4.1	Muon PID Score.....	16
4.2	Extrapolation procedure	17
5.1	Confidence interval construction for a Gaussian distributed parameter.....	22
5.2	Flip-flopping and Feldman-Cousins confidence band.....	24
5.3	Significance profile for a Gaussian distributed parameter	27
5.4	Gaussian profiles.....	30
5.5	Graph of best fit nuisance parameter values.....	30
5.6	Feldman-Cousins distribution of $\Delta\chi^2$	32
5.7	Critical value difference surface	32
5.8	Inverted Hierarchy rejection distribution	34

6.1	Feldman-Cousins corrected contours	43
-----	--	----

LIST OF ACRONYMS

APD Avalanche Photo Diode

CVN Convolutional Vision Network

FC Feldman-Cousins

FD Far Detector

Fermilab Fermi National Accelerator Laboratory

HDF5 Hierarchical Data Format

HPC High Performance Computing

IH Inverted Hierarchy

IHLO Inverted Hierarchy Lower Octant

IHUO Inverted Hierarchy Upper Octant

LO Lower Octant

LRT Likelihood Ratio Test

MLE Maximum Likelihood Estimator

MPI Message Passing Interface

ND Near Detector

NERSC National Energy Research Scientific Computing Center

NH Normal Hierarchy

NHLO Normal Hierarchy Lower Octant

NOvA NuMI Off-axis ν_e Appearance

NuMI Neutrinos at the Main Injector

PDG Particle Data Group

PID Particle Identity

SciDAC-4 Scientific Discovery through Advanced Computing: High Energy Physics Data
Analytics

VM Virtual Machine

CHAPTER 1

INTRODUCTION

"In the beginning there was nothing, which exploded."

Terry Pratchett

The NOvA (NuMI Off-axis ν_e Appearance) experiment is a long-baseline neutrino oscillation experiment designed to measure the oscillation parameters Δm_{32}^2 , θ_{23} , θ_{13} , and δ_{CP} of the PMNS neutrino oscillation matrix, as well as the neutrino mass hierarchy, through observation of ν_e appearance ($\nu_\mu \rightarrow \nu_e$). Employing a two-detector design, the neutrino content of the NuMI (Neutrinos at the Main Injector) beam observed in the Near Detector (ND) is used to predict what would be observed in the Far Detector (FD) under oscillation hypotheses. The best-fit PMNS parameters are determined by comparing FD prediction to observation by likelihood maximization. Statistical power of the parameter estimation is quantified using the computationally intensive unified ordering principle from Feldman and Cousins [1], which corrects for highly non-Gaussian (non-normal) uncertainties.

The unified ordering principle, or Feldman-Cousins (FC) procedure, involves generating millions of statistically-fluctuated FD prediction spectra that are used to populate distributions of the $-2 \log$ -likelihood test-statistic resulting from NOvA's fitting procedure. The distributions that are produced are used to calculate the statistical power of NOvA's parameter estimation based on the observed data. Recent inclusion of antineutrino data increases the computational demand of the fitting procedure, motivating NOvA to expand their pool of computational resources to include the High Performance Computing facilities (HPC) at

NERSC. This thesis presents an overview of the NOvA experiment, the FC procedure, and the modifications to NOvA's existing FC infrastructure required for HPC environments like NERSC. Also included are the performance results achieved during two dedicated computing resource reservations at NERSC during which NOvA's full FC corrections were processed, implications for future analyses, and future and on-going work to improve and expand the newly developed FC infrastructure.

CHAPTER 2

NEUTRINO PHYSICS

*"We cannot solve our problems with the same thinking
we used when we created them."*

Albert Einstein

2.1. THE STANDARD MODEL OF PARTICLE PHYSICS

The Standard Model (Fig. 2.1) of particle physics summarizes our current understanding of the fundamental particles in the universe and how they interact. It is made up of two categories of particles: matter and force mediators. Matter is further divided into quarks and leptons. Quarks are charged particles that make up larger particles called hadrons, like protons, neutrons, kaons, and pions. They interact via the strong, weak, and electromagnetic force mediated by the gluon (g), the W and Z bosons, and the photon, respectively. Leptons include the electron (e), muon (μ), and tau (τ) particles, and their much lighter partners, the electron (ν_e), muon (ν_μ), and tau (ν_τ) neutrinos. Neutrinos are light, neutral particles that are named after their more massive partners with which they exclusively interact via charged current (CC) interactions mediated by the W^\pm bosons. Neutrinos can also interact with any flavor of lepton or quark through a neutral current (NC) process mediated by the Z^0 . Although first thought to be massless, in 1957, Bruno Pontecorvo [2] hypothesized that if neutrinos did have mass, they could change type, or oscillate, during propagation over long distances, eg. $\nu_\mu \rightarrow \nu_e$. Finally, in 1998, the Super-Kamiokande collaboration confirmed

observation of oscillation due to non-zero masses while studying the time-of-day-varying flux of neutrinos produced in the atmosphere from cosmic ray interactions [3].

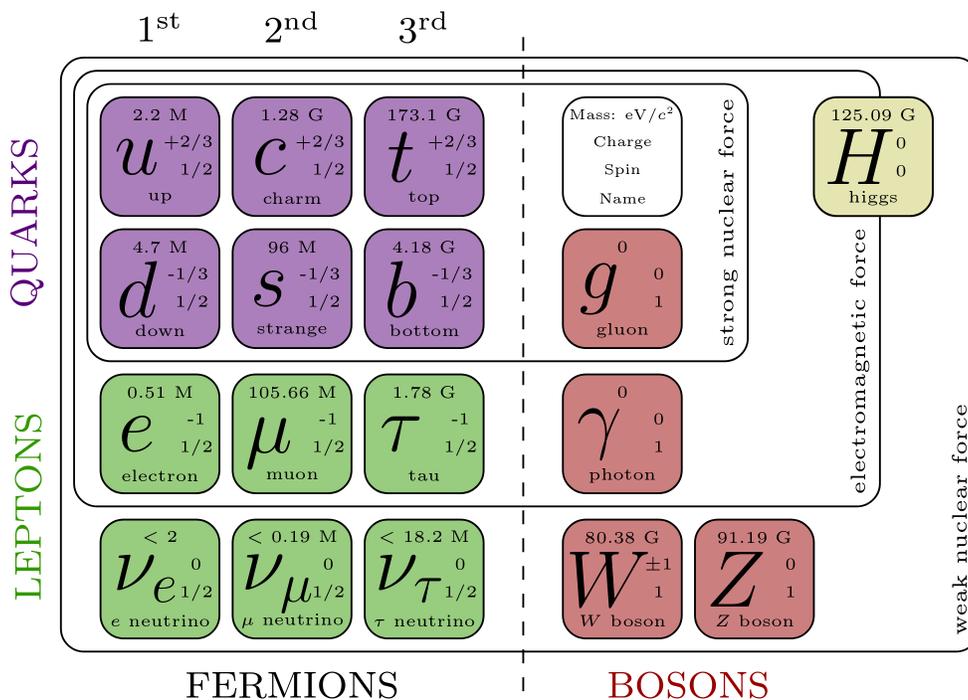


FIGURE 2.1. The Standard Model of Particle Physics

2.2. NEUTRINO OSCILLATIONS

Neutrinos oscillate because they travel through space in different quantum mechanical states than the ones in which they interact with matter. These states are related by a unitary transformation that transforms an energy representation into an interaction representation, and vice versa. Neutrino mass is associated with energy and type, or flavor, is associated with interaction.

For the case of two dimensional spaces, or two-flavor oscillations, the unitary transformation is a simple coordinate rotation (Fig. 2.2). The flavor states are labeled ν_e and ν_μ because the neutrinos interact with the muon and electron via the weak force. The mass

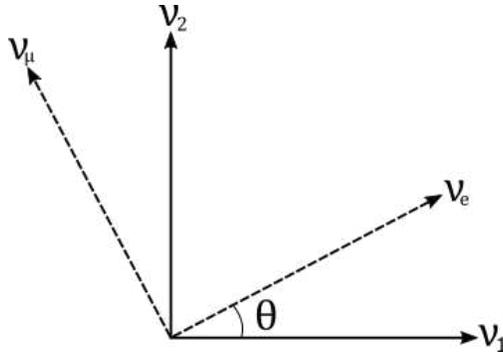


FIGURE 2.2. Transformation from energy to weak representations as a rotation in \mathbb{R}^2 .

states are labeled ν_1 and ν_2 . These energy states time-evolve quantum mechanically according to Schrödinger's equation. In natural units[†], this leads to,

$$(2.1) \quad |\nu_i(t)\rangle = e^{-iE_i t} |\nu_i(0)\rangle$$

Then the flavor states, ν_e and ν_μ , can be expressed as time dependent superpositions of the mass states.

$$(2.2) \quad \begin{aligned} |\nu_e(t)\rangle &= \cos \theta |\nu_1(t)\rangle - \sin \theta |\nu_2(t)\rangle \\ |\nu_\mu(t)\rangle &= \sin \theta |\nu_1(t)\rangle + \cos \theta |\nu_2(t)\rangle \end{aligned}$$

Oscillations occur as this superposition evolves through space and time. The quantum mechanical probability of oscillating from flavor α to flavor β is the inner-product $|\langle \nu_\alpha(t) | \nu_\beta(0) \rangle|^2$.

This formalism generalizes to three dimensions by including one additional flavor and mass state, ν_τ and ν_3 . The analogous transformation is the 3×3 matrix, U_{PMNS} , which is

[†] $\hbar = c = 1$

parameterized by three mixing angles, θ_{13} , θ_{23} , θ_{12} , and a complex phase δ_{CP} (Eq. 2.2),

$$\begin{aligned}
U_{PMNS} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{CP}} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta_{CP}} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{CP}} \\ -s_{12}c_{23} - c_{12}s_{13}s_{23}e^{i\delta_{CP}} & c_{12}c_{23} - s_{12}s_{13}s_{23}e^{i\delta_{CP}} & c_{13}s_{23} \\ s_{12}s_{23} - c_{12}s_{13}c_{23}e^{i\delta_{CP}} & -c_{12}s_{23} - s_{12}s_{13}c_{23}e^{i\delta_{CP}} & c_{13}c_{23} \end{pmatrix}
\end{aligned}$$

where $s_{ij} = \sin \theta_{ij}$ and $c_{ij} = \cos \theta_{ij}$. Now, the probability of oscillating from type α to β is,

$$\begin{aligned}
(2.3) \quad P(\nu_\alpha \rightarrow \nu_\beta) &= |\langle \nu_\alpha | \nu_\beta \rangle|^2 = \delta_{\alpha\beta} - 4 \sum_{i < j} \text{Re}[U_{\alpha i} U_{\beta i}^* U_{\alpha j}^* U_{\beta j}] \sin^2 \left(\frac{\Delta m_{ij}^2 L}{4E} \right) \\
&\quad + 2 \sum_{i < j} \text{Im}[U_{\alpha i} U_{\beta i}^* U_{\alpha j}^* U_{\beta j}] \sin^2 \left(\frac{\Delta m_{ij}^2 L}{4E} \right)
\end{aligned}$$

where the time dependence has been replaced by length L assuming the neutrinos are traveling close to the speed of light. From (2.3) it can be seen that the oscillation probabilities are in general dependent on the mass-squared differences ($\Delta m_{ij}^2 = m_i^2 - m_j^2$), and that oscillations will not occur if the neutrinos are massless or have the same mass.

Experiments studying neutrino oscillations aim to measure the values of these parameters; those of particular interest being θ_{13} , θ_{23} , and δ_{CP} . A non-zero measurement of δ_{CP} would support theories concerning the matter/antimatter symmetry in the universe; and a measurement of $\theta_{23} = 45^\circ$ could reveal a symmetry between ν_μ and ν_τ and how they contribute to the mass states.

Though Equation 2.3 was derived for neutrinos traveling in vacuum, an analogous expression exists for neutrinos traveling through matter. The addition of the matter potential

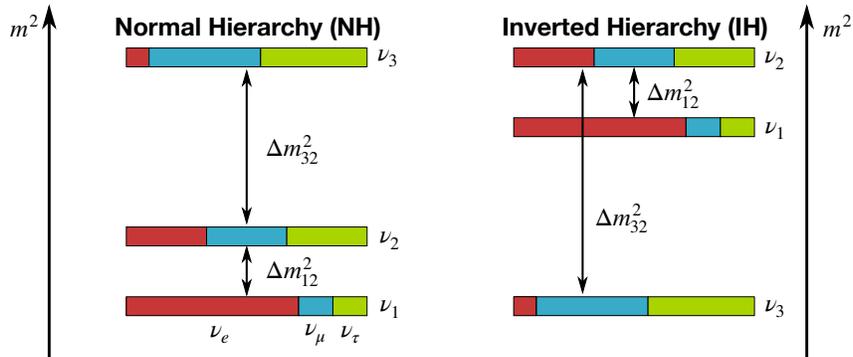


FIGURE 2.3. Normal (left) and inverted (right) neutrino mass hierarchies. Colors indicate neutrino interaction states ν_e (red), ν_μ (blue), and ν_τ (green).

perturbative Hamiltonian results from the forward scattering of neutrinos off of matter, disproportionately effecting electron neutrinos due to an added charged-current contribution from electrons within the matter. It is because of this perturbation that the sign of Δm_{32}^2 can be measured by terrestrial long-baseline experiments, whereas the vacuum probability (2.3) is only dependent on $|\Delta m_{32}^2|$.

Whether $\Delta m_{32}^2 > 0$ or $\Delta m_{32}^2 < 0$ is known as the mass ordering, problem, and making this determination is one of the main physics goals of long-baseline oscillation experiments. The former would imply that the neutrino mass states are ordered similarly to the charged-lepton cousins (Fig. 2.3), and therefore is referred to as a Normal Hierarchy (NH); the latter, Inverted Hierarchy (IH). The mass hierarchy has important implications relevant to experiments like EXO [4] that are designed to put constraints on the absolute neutrino masses and determine whether neutrinos are their own antiparticles (Majorana particles) through the observation of neutrino-less double beta decay ($0\nu\beta\beta$). If the hierarchy is found to be normal, the half-life of a $0\nu\beta\beta$ decay may be well outside the sensitivity of current experiments [5].

CHAPTER 3

THE NOvA EXPERIMENT

*"Science is the organized skepticism in the reliability
of expert opinion."*

Richard Feynman

3.1. LONG-BASELINE NEUTRINO EXPERIMENTS

Long-baseline neutrino experiments are designed for maximum sensitivity to ν_μ disappearance ($\nu_\mu \rightarrow \nu_\mu$) and ν_e appearance ($\nu_\mu \rightarrow \nu_e$) oscillations with the primary goals of measuring the oscillation parameters including the mass hierarchy. The ν_μ disappearance is dominated by the parameters $\sin^2 \theta_{23}$ and $|\Delta m_{32}^2|$, while the ν_e appearance provides constraints on δ_{CP} and the sign of Δm_{32}^2 . Long-baseline neutrino experiments, like the NuMI Off-axis ν_e Appearance (NOvA) experiment, measure these parameters with a joint fit to the appearance and disappearance probabilities, while constraining the remaining oscillation parameters, θ_{13} , θ_{12} , Δm_{12}^2 , and Δm_{13}^2 , to their values measured by other experiments [6] as compiled by the Particle Data Group (PDG).

The measurement is made possible by a formulaic design common to all long-baseline neutrino oscillation experiments (Fig. 3.1). A beam of neutrinos is created by the decay of kaons and pions, which for example can be produced by proton-graphite collisions. The beam is then sampled by a Near Detector (ND) in order to characterize its composition. Continuing through the Earth's crust, the neutrinos oscillate over a large distance, or baseline, as described in Section 2.2. Finally, some of the neutrinos interact with the matter in

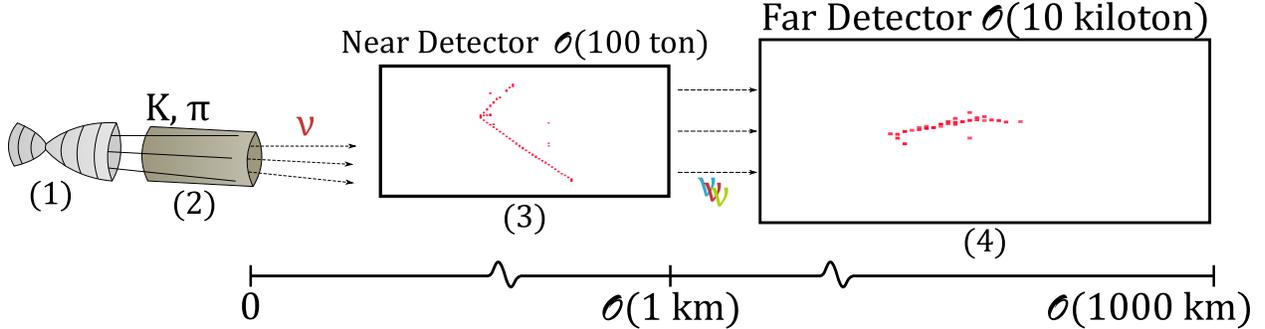


FIGURE 3.1. Cartoon of a long-baseline neutrino oscillation experiment. Mesons, mostly kaons and pions, are focused by focusing horns (1) and left to decay into neutrinos while traveling through a decay pipe (2). The neutrinos travel to the near detector (3) used to characterize the beam. They then oscillate while traveling large distances through Earth’s crust and some are measured in the Far Detector (4).

the Far Detector (FD) and are identified as either ν_e , ν_μ , or ν_τ based on the detection of the interactions caused by their heavier partners. The resulting number of identified neutrinos (ν_μ and ν_e separately) are compared to what would be expected under different parameters of the PMNS matrix (2.2), and the best-fit parameters are extracted.

3.2. THE NEUTRINO BEAM

The neutrino beam utilized by NOvA is provided by the Neutrinos at the Main Injector (NuMI) beam facility [7] at Fermi National Accelerator Laboratory (Fermilab). The NuMI facility accelerates protons from the Fermilab Main Injector to 120 GeV. These protons are extracted and directed toward a graphite target. The collisions create mostly charged kaon and pions, which decay via $K^{+(-)} \rightarrow \mu^{+(-)} + \bar{\nu}_\mu$ and $\pi^{+(-)} \rightarrow \mu^{+(-)} + \bar{\nu}_\mu$. The focusing horns (Fig. 3.2) select the sign of the charged particles depending on the direction of the magnetic field, which is tunable by an applied current. The horn system acts as a lens with focal length equal to the momentum of the hadron, resulting in a mostly collimated beam of hadrons that eventually decay into neutrinos and antineutrinos in the 675 m long decay pipe.

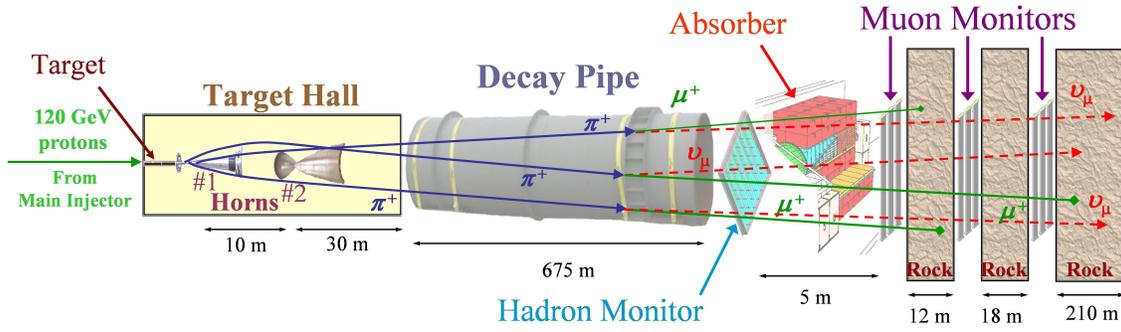


FIGURE 3.2. Neutrino production from NuMI. The magnetic horns (#1, #2) enable the selection of either positively or negatively charged hadrons, which are then allowed to decay in the Decay Pipe.

Hadron monitors are positioned downstream of the decay pipe to measure the intensity and integrity of the beam through detection of the predominately proton flux. Similarly, muon monitors measure the intensity of muon flux. Since the muons and neutrinos originate from the same decay process, muon flux can be used to characterize the neutrino flux by kinematic considerations. Hadron and muon absorbers stop the leftover beam and decay products to prevent these particles from interacting in the NOvA and MINOS [8] near detectors.

3.3. THE NOvA DETECTORS

The NOvA experiment employs a two detector design to sample the neutrinos before and after propagation. A 220 ton ND is situated 100 m underground 1 km away from the beam target and is used to characterize the neutrino beam from NuMI. The FD is located on the surface at Ash River, MN, 810 km from the beam source. Both detectors are placed 14 mrad off of the beam axis, which narrows the energy spectrum of incoming neutrinos and reduces background in the signal region from neutral current events, while also ensuring a near maximal ν_e appearance probability as shown in figure 3.4.

Muons originating from cosmic rays and muon neutrinos interacting with the rock that surrounds the detector make up a significant amount of background that must be subtracted

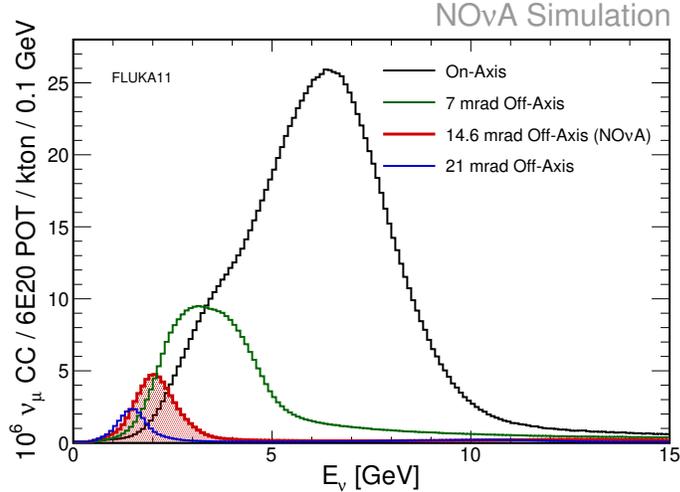


FIGURE 3.3. Off-axis neutrino energy distributions from NuMI. The NOvA near and far detectors are placed at 14 mrad off-axis so that the energy peak is at the appearance maximum and to reduce backgrounds in the signal region from neutral current events.

in order to identify the neutrino events used in the analysis. The overburden (rock and dirt) above the ND provides shielding from cosmic rays, reducing this type of background to a negligible level (30 Hz). The FD, however, sits on the surface and is exposed to cosmic rays at a rate of 130 kHz. This background is reduced by strict selection criteria, which discard events that occur outside the $10 \mu\text{s}$ beam spill, events that are not contained within the fiducial volume of the detector, and vertical tracks. An additional reduction step is performed by a boosted decision tree that is trained on cosmic ray events that occur outside the beam spill using reconstructed kinematic and detector position variables. Similar selection criteria reduce the background due to neutrinos interacting with the rock surrounding both detectors to a negligible amount.

3.3.1. DETECTOR COMPOSITION. The NOvA detectors are composed of rectangular extruded PVC tubes, called cells, that are filled with liquid scintillator and wavelength-shifting fibers as shown on the right side of Figure 3.5. Two sixteen-cell extrusions form a module, and the modules are arranged alternating between vertical and horizontal orientations (Fig.

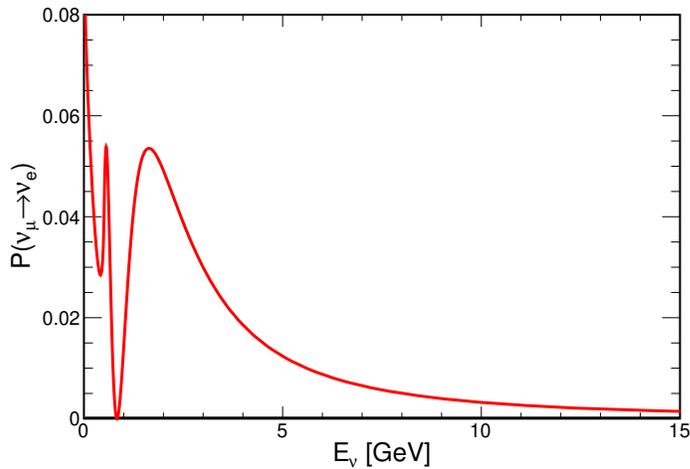


FIGURE 3.4. Oscillation probability for ν_e appearance as a function of energy at the NOvA baseline length of $L = 810$ km and $\delta_{CP} = 0$.

3.5). Charged particles resulting from neutrino-matter interactions within these cells excite the liquid scintillator resulting in emitted light that is collected in the wavelength-shifting fibers. The fibers shift the blue light (400 – 450 nm) emitted from the scintillators to green light (490 – 550 nm), some of which is directed into Avalanche Photo Diodes (APD), which have an 80% quantum efficiency in the range of the collected light. The signal from an APD is amplified, digitized, and fed into the data acquisition system, where the signals from cells containing intensity and timing information can be collected into two and three dimensional images by virtue of the alternating planes. These images facilitate particle and event identification used for the reconstruction of the neutrino properties important for the oscillation analysis.

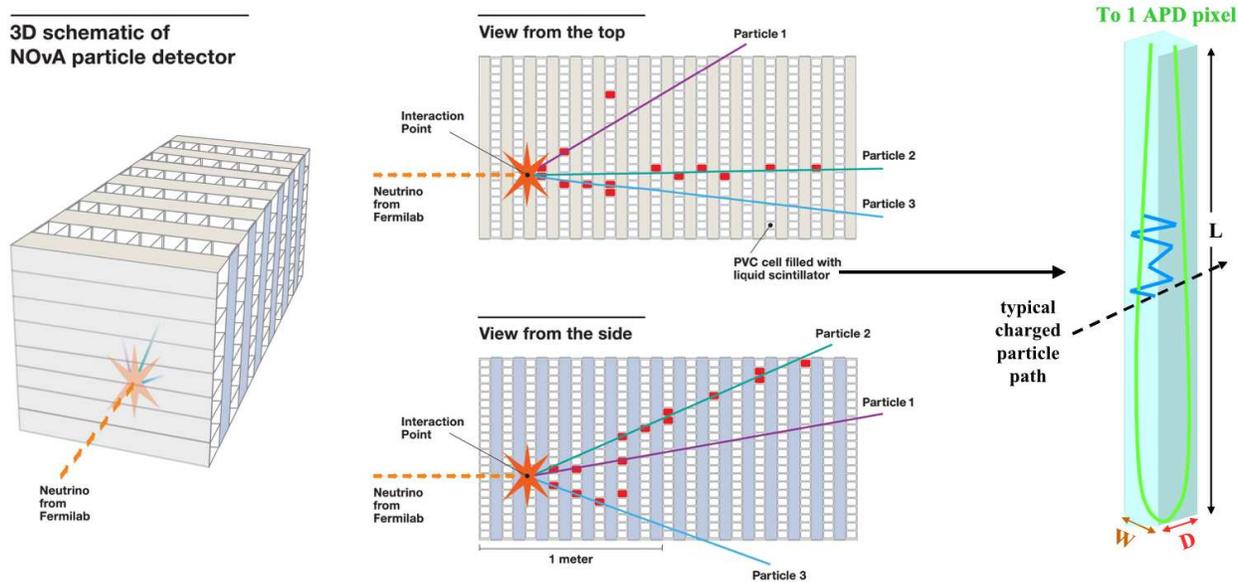


FIGURE 3.5. The NOvA Detectors. Modules of extruded PVC filled with liquid scintillator are arranged in alternating horizontal and vertical orientations (left). Neutrinos interact with the scintillator (middle), which subsequently emits light that is picked up by wavelength-shifting fiber in each cell (right). The signal from the light is digitized and sent to the Data Acquisition system to be later processed by the analysis.

CHAPTER 4

THE NOvA NEUTRINO OSCILLATION ANALYSIS

*"I was taught that the way of progress was neither
swift nor easy."*

Marie Curie

The NOvA collaboration's neutrino oscillation analysis is performed by using ND data to create predictions of the FD neutrino energy distributions. These predictions are compared to the FD data in order to extract the best-fit oscillation parameters. The main components that make the analysis possible are the following:

Identify and reconstruct candidate ND neutrino events: Particle interactions within the ND are isolated in space and time to form what are called events. Raw signals that make up the events are used to estimate the energy of the neutrinos from which the particles originate.

Extrapolate to FD: Reconstructed ND data are used to predict the neutrino energy spectrum observed in the FD.

Measure FD neutrino energy spectrum: Identical event identification and reconstruction techniques used in the ND are used to determine neutrino energy spectrum observed in the FD.

Determine oscillation parameters: The FD prediction that most closely matches the data is determined and oscillation parameters are extracted.

4.1. EVENT IDENTIFICATION AND RECONSTRUCTION

4.1.1. **PARTICLE RECONSTRUCTION.** The first step in the event identification component is to cluster cell hits in the detector in space and time into a single object called an event. The event is then processed by an algorithm based on Hough Transforms called Elastic Arms [9] to identify prominent particle trajectories and the interaction vertex. Cell hits within the event are then associated with prominent Hough lines using a FuzzyK algorithm [9] in order to cluster together hits that originated from a single particle; the resulting clusters are called prongs. This is done separately in both vertical and horizontal planes. Three dimensional clusters are reconstructed by associating 2D prongs in time, space, and energy deposition per unit length [9].

These data products can then be classified at the event level and individual particle level by a series of Convolutional Vision Networks (CVN), which are image recognition Convolutional Neural Networks developed by NOvA [10]. The "images" NOvA's CVNs process are pixel mappings of the cells within the detectors onto matrix elements containing associated deposited energy. Images are classified at the event level (EventCVN) with only these pixel mappings as the input; particles are classified with these pixel mappings as well as the isolated prongs (ProngCVN) [11]. The output from the two CVNs are vectors of "scores" corresponding to the likelihood that a given event is ν_μ -CC, ν_e -CC, or NC or a given particle is a proton, muon, electron, pion, or photon. Figure 4.1 shows an example of the muon Particle Identity (PID) score output from ProngCVN.

4.1.2. **ENERGY ESTIMATION.** The neutrinos used in NOvA's oscillation analysis are characterized by neutrino flavor and reconstructed energy. In order to accurately reconstruct neutrino energy, the detector must be calibrated against the observation of well-understood

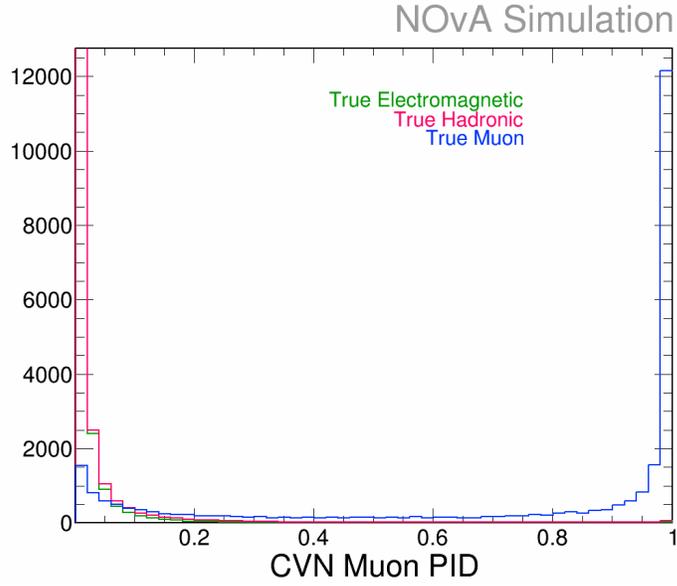


FIGURE 4.1. Sample muon PID score from simulated true muon events as classified by ProngCVN.

phenomenon, in this case cosmic muons. An absolute scaling is done by examining the energy deposited per unit length (dE/dx) at the end of a cosmic muon track and comparing with known dE/dx curves (Bethe-Bloch curves). A cell-by-cell calibration using through-going cosmic ray muons must also be performed due to the attenuation from the wavelength-shifting fibers [11]. Together, these two calibration corrections allow for the correspondence between detector signal and the energy seen in the detector, or calorimetric energy.

Due to differences in the way the detectors respond to different types of particles, the calorimetric energy must be converted into an event energy based on the type of particles interacting within the detector. For electron neutrino events, calorimetric energy is divided into electromagnetic or hadronic contributions based on the characteristics of the prongs from which the energy deposition originated. The corresponding energies are then fed into a conversion function to calculate the neutrino's "reconstructed" energy. This function is found by fitting simulated reconstructed energies of events to their true energies. Muon

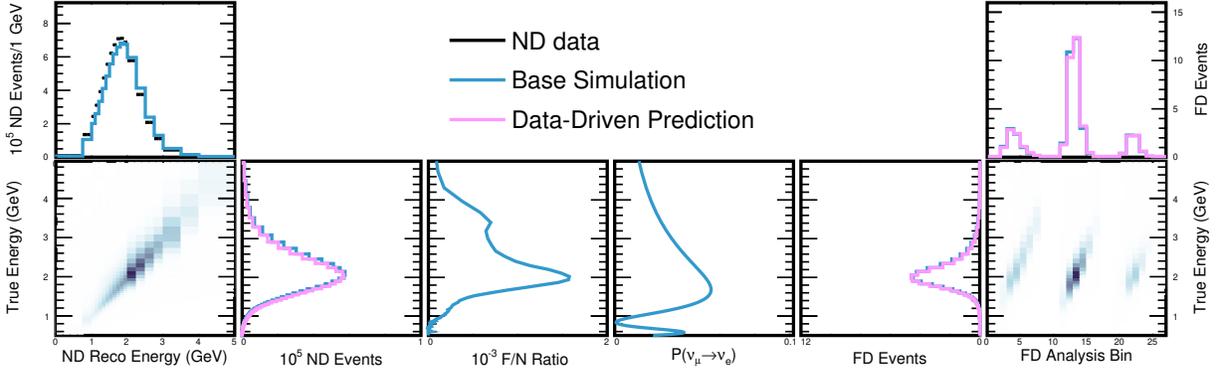


FIGURE 4.2. Extrapolation of ν_μ events in the near detector to predicted far detector spectrum of signal ν_e 's.

neutrino energy is found by identifying the daughter muon within an identified ν_μ -CC event with a Kalman filter algorithm [12] and extracting its true energy from the length of the muon track, again using the Bethe-Bloch formula. The hadronic energy is found by similarly fitting a linear function of measured hadronic energy to the hadronic portion of the true energy.

4.2. EVENT SELECTION AND EXTRAPOLATION

The ND is used to characterize the neutrino beam to produce data-driven predictions of oscillated neutrino energy distributions observed in the FD. First, ν_μ events are detected in the ND and those that pass PID and data quality cuts are divided into histogram bins of reconstructed energy. Monte Carlo simulation of the beam, neutrino interactions in the detector, and detector response are tuned according to these data. Individual components of the beam simulation (ν_e , ν_μ , ν_τ) are extrapolated to the FD according to oscillation hypotheses in true neutrino energy, detector acceptance, and efficiency. By inverting the reconstructed-to-true energy transformation, a FD prediction of signal ν_e 's (ν_e 's originating from $\nu_\mu \rightarrow \nu_e$) is produced (Fig. 4.2).

4.3. SYSTEMATICS

Systematic uncertainties are quantified using Monte Carlo and data-driven simulation. Independently-generated Monte Carlo data sets are created for each systematic uncertainty by shifting only the relevant quantity up or down by their 1σ values. These include uncertainties due to the extrapolation process, detector calibration, and interaction cross sections. The shifted data sets, which are assumed to be orthogonal to one another, are extrapolated to the FD identically to the nominal data set to quantify the effect that the individual systematic uncertainties have on the expected levels of signal and backgrounds. By taking the ratio of the systematically shifted FD predictions to the nominal predictions, the systematics' 1σ shifts are obtained and used in the fitting procedure described in the next section.

4.4. OSCILLATION FITS

The extrapolated predictions are used to predict what is expected from the data under the assumption of chosen values of oscillation parameters, $\vec{\theta}$, as well as shifted systematic uncertainties, \vec{s} . This allows for the comparison between data and prediction under different values of oscillation parameters and systematic variations, or pulls, and a determination of the best-fit parameters values.

The goodness-of-fit metric used for this comparison is the $-2 \log$ -likelihood recommended by the PDG [6] and loosely referred to here as χ^2 . It is given by

$$(4.1) \quad \chi^2 \equiv -2 \log \mathcal{L}(\vec{\theta}, \vec{s}) = 2 \sum_i^{bins} \left[e_i(\vec{\theta}, \vec{s}) - o_i + o_i \log \frac{o_i}{e_i(\vec{\theta}, \vec{s})} \right] + \sum_j^{systs} \frac{s_j}{\sigma_j}$$

where $\mathcal{L}(\vec{\theta})$ is the Poisson likelihood function, $e_i(\vec{\theta})$ and o_i are the expected and observed events in the i^{th} bin, and s_j and σ_j are the j^{th} systematic pull and its 1σ value. The last

term in (4.1) is the sum of Gaussian penalty terms for each systematic uncertainty. The penalty terms increase the value of χ^2 to prevent the systematic pulls from being large compared to their 1σ values. Minimizing this function with respect to $\vec{\theta}$ and \vec{s} yields the set of parameters that best estimate the true parameters. This minimization is done using the MINUIT2 minimization package [13] within the ROOT data analysis framework [14].

With our latest analysis on neutrino and antineutrino data [15], NOvA determines the best fit set of parameters to be the following: In addition, NOvA measures a 4.2σ observation of $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$ oscillation – the first strong evidence of $\bar{\nu}_e$ appearance measured by a long-baseline experiment.

TABLE 4.1. Best-fit oscillation parameters and associated 1σ confidence intervals with and without Feldman-Cousins (FC) corrections determined for NOvA’s neutrino + antineutrino data set [15].

Parameter	Central Value	No FC Interval	FC Interval
$\Delta m_{32}^2/10^{-3} \text{ eV}^2/c^4$	+2.51	+0.10 −0.08	+0.12 −0.08
$\sin^2 \theta_{23}$	0.58	+0.02 −0.03	± 0.03
δ_{CP}/π	0.2	+0.1 −0.5	+1.2 −0.6

CHAPTER 5

THE FELDMAN-COUSINS PROCEDURE

”There are two possible outcomes: if the result confirms the hypothesis, then you’ve made a measurement.

If the result is contrary to the hypothesis, then you’ve made a discovery.”

Enrico Fermi

When performing parameter estimation as described in Chapter 4 it is necessary to report a metric upon which to base the validity of the measurement. A popular approach, NOvA’s approach, is the Frequentist confidence interval, which describes how compatible the data are with parameter values other than those that were measured.

5.1. CONFIDENCE INTERVAL CONSTRUCTION

In Frequentist statistics, confidence intervals predict that, given an ensemble of experiments measuring the parameter μ of some model, a fraction α of those experiments will measure an interval $[\mu_1, \mu_2]$ that contains the true value μ_t . Here, the true value is fixed and the interval is a random variable based on data. In contrast, Bayesian statistics state that the probability of the true parameter lying within some interval $[\mu_1, \mu_2]$, where the true parameter is considered the random variable and the data are fixed. Where Bayesian intervals require the experimenter’s subjective prior knowledge of the possible values that μ can hold, namely $P(\mu)$, Frequentist intervals require no such assumptions. The objectivity

of a Frequentist interval is a desired trait among physicists, and is one of the reasons they are employed by NOvA.

More precisely, Frequentist confidence intervals belong to the set of all intervals that contain a fixed μ with some probability α .

$$(5.1) \quad P(\mu \in [\mu_1, \mu_2]) = \alpha$$

This is to hold true for all values of μ , therefore such a set exists for μ_t . The construction of confidence intervals was first put forward by Jerzy Neyman in 1937 [16], and is referred to as Neyman construction.

In order to determine a confidence interval, first consider the allowed values of a measurable random variable x where x is a function of μ . For a fixed μ in an ensemble of experiments, x will be measured within a so-called an acceptance region, $[x_1, x_2]$, in a fraction α of the total number of experiments. The collection of all acceptance regions is called the confidence band (Fig. 5.1). There is a freedom in the choice of x_1 and x_2 , the only restriction being,

$$(5.2) \quad \int_{x_1}^{x_2} P(x|\mu)dx = \alpha$$

Three common choices of $[x_1, x_2]$, referred to as ordering principles, are the central limit, lower limit and upper limit. Central limits satisfy $P(x < x_1) = P(x > x_2) = (1 - \alpha)/2$; upper limits satisfy $P(x < x_1) = 1 - \alpha$; and lower limits satisfy $P(x > x_2) = 1 - \alpha$. Each choice is valid as long as the choice is not made based on the results of a measurement of x .

The confidence interval associated with a measurement of x is determined by considering the values of μ that contain the measured x within their acceptance region. Since a fraction

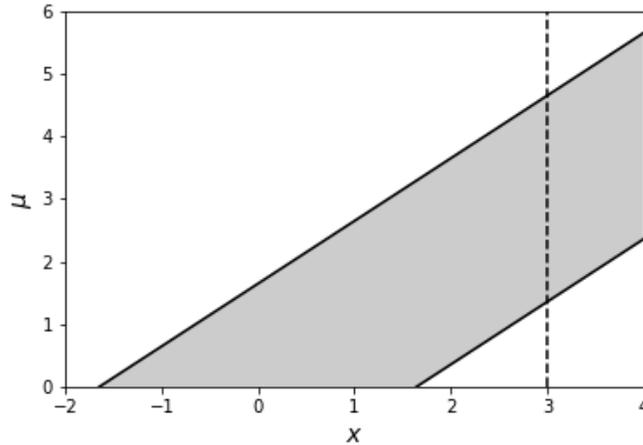


FIGURE 5.1. Confidence interval construction for the Gaussian distributed parameter μ . A vertical line is drawn at the measured $x = 3$. The intersection of the vertical line with the confidence band is the confidence interval, $[\mu_1 = 1.36, \mu_2 = 4.65]$, for the observed x .

α experiments in an ensemble will measure x within the acceptance region of the true parameter, μ_t , the intervals $[\mu_1, \mu_2]$ drawn by the experiments will contain μ_t in α of the total number of experiments, thereby satisfying Equation 5.1.

The above definitions lend themselves to a simple recipe for Neyman construction:

- (1) Find the intervals $[x_1, x_2]$, such that $P(x \in [x_1, x_2]) = \alpha$ for each fixed value of μ with an ordering principle that uniquely determines the acceptance region. This creates the confidence band.
- (2) Upon a measurement of x yielding x_o , draw a vertical line at x_o as shown in Figure 5.1 by the dotted line.
- (3) The intersection of the vertical line with the confidence band is the resulting confidence interval.

It is possible that the central, upper, and lower limit ordering principles produce intervals that do not contain certain values of μ with the reported probability. If for an ensemble of experiments measuring a parameter whose true value is μ_t , a proportion greater than α of

the measured intervals $[\mu_1, \mu_2]$ contain μ_t , μ_t is said to be over covered; if that proportion is less than α , μ_t is said to be under covered. Intervals that over cover are too wide, artificially reducing the statistical power of the experiment; intervals that under cover leave the possibility for the true value to be rejected by more than $1 - \alpha$ experiments in the ensemble. Since the parameter's true value is never known, intervals must correctly cover all values of μ .

One case that leads to improper coverage is when the choice of ordering principle is based on data – this is called flip-flopping. For example, if the experimenter decides that they will use an upper limit when a measurement of x yields a value less than some value and a central limit otherwise, the resulting intervals may under cover. The effective confidence band for a flip-flopping experiment measuring a Gaussian-distributed parameter is shown in Figure 5.2. Here, $\mu = 1.86$ is under covered by 5% when a central limit is chosen for measurements of $x > 3$ and an upper limit otherwise. Since confidence intervals depend on the acceptance region for each μ , these regions must be consistently constructed to avoid excluding values of μ that are α -consistent with the measurement of x .

Problems also arise if μ is physically bounded, eg. $\mu > 0$, even though a measurement may result in a value of μ less than zero. When measurement yields $x < 0$, the experimenter can report those values to be zero, resulting in an interval that significantly over covers (Fig. 5.2), or report that their measurement fell outside the bounds of all physical acceptance regions (eg, a vertical line at $x = -1.9$ in Figure 5.1). In the latter case, the experimenter can only conclude that their measurement fell outside the α acceptance region of the true parameter.

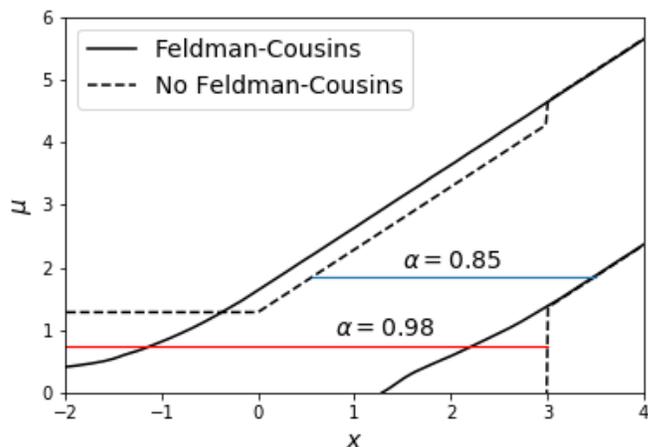


FIGURE 5.2. Dashed lines show the effective confidence band due to flip-flopping for the Gaussian distributed parameter μ bounded by $\mu > 0$. The red and blue lines show representative values of μ that are over and under covered, respectively. The solid line shows the confidence band found using the Feldman-Cousins unified ordering principle.

5.2. THE UNIFIED ORDERING PRINCIPLE

The shortcomings of the central, upper, and lower limit ordering principles mentioned in the previous section stem from choices made by the experimenter that artificially modify the acceptance bands, leaving parameter values that are not correctly covered by the resulting confidence intervals. The ordering principle put forth by Feldman and Cousins [1] alleviates the experimenter from the burden of such choices, thereby constructing confidence intervals with correct coverage by design. Instead of ordering the values of x to include in an interval $[x_1, x_2]$ in an increasing sense between the bounds that satisfy Equation 5.2, the FC procedure orders the values of x in increasing order of the ratio,

$$(5.3) \quad R = \frac{\mathcal{L}(\mu|x)}{\mathcal{L}(\mu_{\text{best}}|x)}$$

where μ_{best} , called the Maximum Likelihood Estimator (MLE), is the value of μ that maximizes the likelihood function, $\mathcal{L}(\mu|x)$, for a given value of x . Values of x are included in the interval until the sum of $P(x|\mu) \geq \alpha$.

Often it is more useful to use the log of this ratio because of the logarithm's arithmetic properties and instead order in an increasing sense. Although recast into a slightly different form, the ordering principle is unchanged since the logarithm is a monotonic function. The ratio then becomes:

$$\begin{aligned}
 R' &= -2 \ln \mathcal{L}(\mu|x) + 2 \ln \mathcal{L}(\mu_{best}|x) \\
 (5.4) \quad &= \chi_{\text{fixed}}^2 - \chi_{\text{best}}^2 \\
 &= \Delta\chi^2
 \end{aligned}$$

where χ^2 is Equation 4.1 for the NOvA analysis. The use of $\Delta\chi^2$ is justified based on the Neyman-Pearson lemma [17], which states that the most powerful test for hypothesis rejection is constructed using this test statistic; this is known as the Likelihood Ratio Test (LRT). The $\Delta\chi^2$ test statistic is zero at the best fit point since $\mu = \mu_{best}$. Large values of $\Delta\chi^2$ correspond to locations of parameter space that are far from the best fit point.

Ordering based on Equation 5.4 assures that the resulting intervals never undercover by construction, and automatically determines whether the intervals are upper, lower, or central limits. Figure 5.2 shows the confidence band drawn by the unified ordering principle in the case of a Gaussian distributed parameter μ bounded by $\mu > 0$. Coverage is maintained for values of μ close to the physical boundary where the standard ordering principles over cover. While over coverage is not a desirable characteristic since it degrades the statistical power

of the experiment, it may be unavoidable due to discrete measurements or the numerical methods used within the analysis.

5.3. SIGNIFICANCE PROFILES AND CONFIDENCE REGIONS

Significance profiles show how a confidence interval changes for different values of α , and offer a more intuitive way to visualize probability statements regarding the value of some parameter given measured data. Significance profiles require the data to be fit at every fixed value of μ to place the data somewhere in the ordering. Figure 5.3 shows a significance profile for a Gaussian-distributed parameter, μ , based on a measurement $x_o = 3$. The plot on the left shows the likelihood ratio, R' , as a function of μ given x_o . A p -value, or $1 - \alpha$, is calculated as the proportion of $\Delta\chi^2$ values beyond the value resulting from data. Using a standard normal distribution, this p -value is canonically transformed into σ , the number of standard deviations from zero, that correspond to the p -value. The right hand plot in Figure 5.3 shows the significance, σ , derived from R' . Confidence intervals corresponding to different values of α are obtained by finding the intersection points of a horizontal line and the significance profile. For example, a dotted line is drawn at 1σ ($\alpha = 0.68$), from which the interval $[2, 4]$ is obtained. Moreover, it can be seen that $\mu > 8$ is rejected at the 5σ -level.

Another useful tool for visualizing confidence intervals in more detail is the confidence region. A confidence region is the extension of a confidence interval to higher dimension, which yields rejected *regions* in parameter space. Additionally, confidence regions facilitate the visualization of an interval's dependence on two or more parameters, display properties of the parameter space, and make statements about the combined values of those parameters.

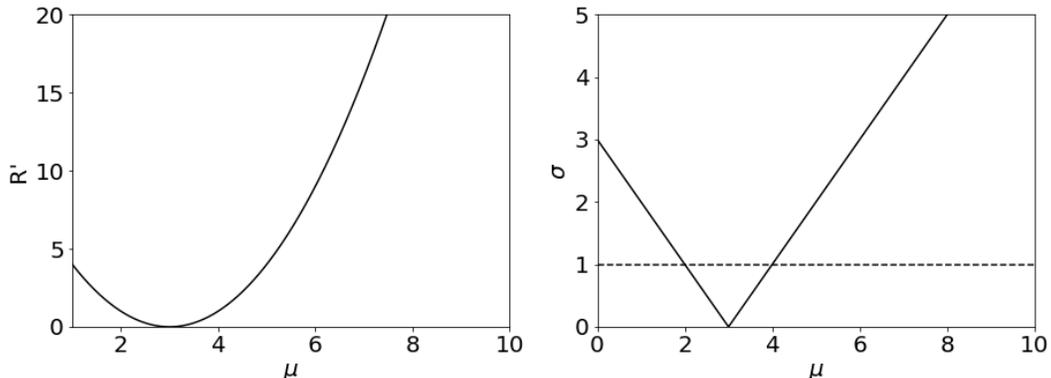


FIGURE 5.3. Significance profile for a Gaussian distributed parameter. On the left, the ratio R' is plotted for $x_o = 3$ as a function of μ . The right plot shows the significance calculated from R' . The intersection points of the dotted and solid lines correspond to the bounds of the 1σ confidence interval.

5.4. CALCULATING ACCEPTANCE INTERVALS OF NEUTRINO OSCILLATION PARAMETERS

Figure 5.2 shows the unified ordering principle applied to a toy example of a bounded Gaussian-distributed parameter, μ . The likelihood function is the well-known Gaussian distribution function with a simple solution for the MLE, which allows for the acceptance regions ordered by the likelihood ratio to be calculated analytically. The same cannot be said for neutrino oscillation physics, requiring Monte Carlo simulation to be employed for the task of calculating these acceptance regions.

A result from Wilks published in 1938 [18] shows that the distribution of the likelihood ratio (5.4) generally approaches a χ^2 -distribution when the MLEs are normally distributed. For this to be true in the context of long-baseline experiments like NOvA, event counts per bin of neutrino energy must be large ($\mathcal{O}(30)$), and parameters must be far from physical boundaries. Due to the neutrino's weak coupling to matter and low neutrino intensity at the FD, event counts are typically $\mathcal{O}(10)$ per bin of neutrino energy. Furthermore, many parameters in the PMNS oscillation model (2.2) are physically bounded between zero and

one. Therefore, the $\Delta\chi^2$ test-statistic is not guaranteed to be χ^2 -distributed, and Monte Carlo simulation is required to empirically generate the "true" distributions.

Filling acceptance regions of $\Delta\chi^2$ (5.4) requires knowledge of all possible measurements given the values of oscillation parameters. It is known that the detection of a neutrino event within a certain energy range is Poisson-distributed, but due to the large dimensionality and trigonometric dependence of the oscillation probability, analytically solving Equation 5.4 for the set oscillation parameters that simultaneously maximize the likelihood is not possible. Monte Carlo simulation offers an efficient way to approximate the acceptance regions by simulating statistically independent data, or pseudoexperiments, and performing the multidimensional likelihood maximization with common numerical fitting techniques.

5.5. GENERATING PSEUDOEXPERIMENTS

The NOvA experiment employs Monte Carlo simulation of data to fill acceptance regions of oscillation parameters according to the unified ordering principle. These simulations, referred to as pseudoexperiments, are used to calculate a large number of values of $\Delta\chi^2$ to approximate the statistical distributions to be ordered. Once a distribution is generated, significance profiles can be calculated for the observed data. Although a computationally intensive process, calculating values of $\Delta\chi^2$ to fill these empirical distributions can be done in parallel since each pseudoexperiment, and hence value of $\Delta\chi^2$, is statistically independent from all others.

The acceptance intervals of parameters, $\vec{\mu}$, are filled by generating pseudoexperiments that are consistent with the parameters $\vec{\mu}$. Pseudoexperiments are created by first generating FD predictions at oscillation parameters, $\vec{\mu}$, then applying Poisson fluctuations to the number

of events detected in each neutrino energy bin to simulate the Poisson-distributed detection rate.

The pseudoexperiments are subjected to NOvA's fitting procedure to calculate values of $\Delta\chi^2$. For the multidimensional case, Equation 5.4 becomes

$$(5.5) \quad \Delta\chi^2 = -2 \log \frac{\mathcal{L}(\vec{\mu}, \vec{\lambda}_{\vec{\mu}}|x)}{\mathcal{L}(\vec{\mu}_{\text{best}}, \vec{\lambda}_{\text{best}}|x)}$$

where $\vec{\lambda}_{\theta}$ is a set of nuisance parameters that maximize the likelihood function at parameters θ . Nuisance parameters are parameters that enter the fit but whose values are not of particular interest; these can be systematic pulls or other oscillation parameters. For example, δ_{CP} is a nuisance parameter when constructing intervals of $\sin^2 \theta_{23}$. This treatment of nuisance parameters is called likelihood profiling and is meant to remove the dependence of nuisance parameter values on the confidence intervals of the interesting parameters. Likelihood profiling ensures that the values of the nuisance parameters are at least covered [19].

Likelihood profiling requires the pseudoexperiments are generated as being consistent with the nuisance parameters as well as the interesting oscillation parameters. To extract appropriate values for these nuisance parameters, fits are performed to the data in the parameter spaces of interest before the process of generating pseudoexperiments. This step is also necessary for calculating the measured values of $\Delta\chi^2$ that will be placed in the acceptance interval ordering for calculating significance profiles. These fits to data result in collections of best-fit nuisance parameter and $\Delta\chi^2$ values as a function of the interesting parameter(s) (Fig. 5.4 and 5.5). The collections of $\Delta\chi^2$ values are referred to as Gaussian profiles, for in the asymptotic limit of large numbers of expected events, acceptance intervals could be calculated analytically from a χ^2 -distribution [18].

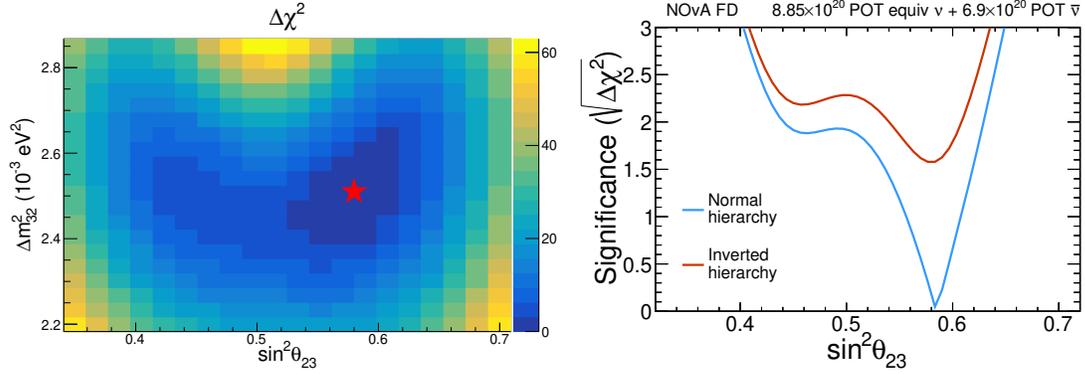


FIGURE 5.4. Gaussian profiles for the $\sin^2 \theta_{23}$ - Δm_{32}^2 plane in the Normal Hierarchy showing the best fit point to data as a star (left) and $\sin^2 \theta_{23}$ axis (right).

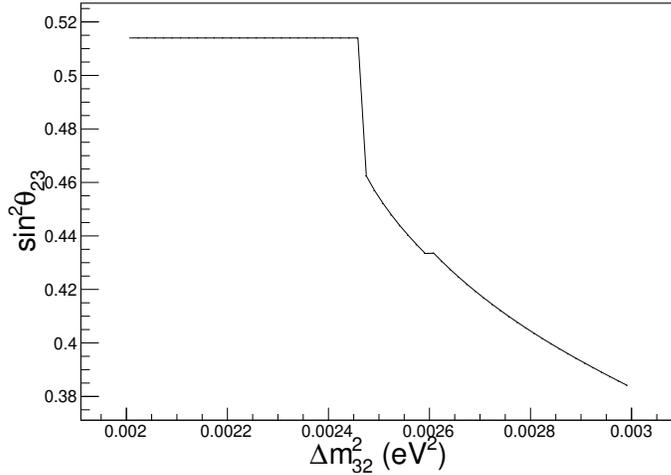


FIGURE 5.5. During the calculation of the Gaussian profiles, the best fit values of all nuisance parameters, here $\sin^2 \theta_{23}$, are saved to be used later during the generation of pseudoexperiments. Here, $\sin^2 \theta_{23}$ was constrained to lower octant.

The FC procedure begins once the nuisance parameter and Gaussian profiles are calculated. For each 1D and 2D significance profile, the parameter space is stepped through bin-by-bin, wherein each bin, a $\Delta\chi^2$ distribution is generated by fitting ~ 2500 pseudoexperiments. Details of the computational challenges and load balancing will be explained in Chapter 6. Once the distributions are filled, Gaussian profiles are used to determine confidence intervals and regions.

5.6. PRODUCING CORRECTED INTERVALS

Distributions of $\Delta\chi^2$ resulting from fitting ensembles of pseudoexperiments are used to calculate significance profiles in parameter spaces of interest. As previously discussed, p -values are found as the proportion of $\Delta\chi^2$ beyond the value resulting from the observed data, $\Delta\chi_o^2$. That is,

$$(5.6) \quad p - \text{value} = \frac{\# \text{ of values } > \Delta\chi_o^2}{\# \text{ of experiments in the ensemble}}$$

These p -values are transformed into values of σ , which are reported as significance profiles.

In the case of a 2-dimensional surface, confidence regions are drawn by first calculating a surface of critical values, $\Delta\chi_c^2$, within the empirical distributions that correspond to $\sigma = 1, 2,$ and 3 . These critical values mark the point at which the data rejects the corresponding parameter values at a $1, 2,$ or 3σ significance. The intersection of these surfaces with the Gaussian profiles are the $1, 2,$ and 3σ contours, respectively.

Figure 5.6 shows how the 90% critical values can differ between the FC and asymptotic χ^2 -distributions. Here, since $\Delta\chi_c^2 > \chi_c^2$, power of rejection against these parameters is decreased with respect to the Gaussian profiles. Similarly, Figure 5.7 shows the difference between $\Delta\chi_c^2$ and χ_c^2 and how the 90% confidence region is affected in the $\sin^2\theta_{23}$ - Δm_{32}^2 plane.

5.7. INVERTED MASS HIERARCHY REJECTION

As mentioned in the first chapter, the NOvA collaboration is interested in determining the neutrino mass hierarchy: the sign of Δm_{32}^2 . Since the best fit to NOvA's data falls in

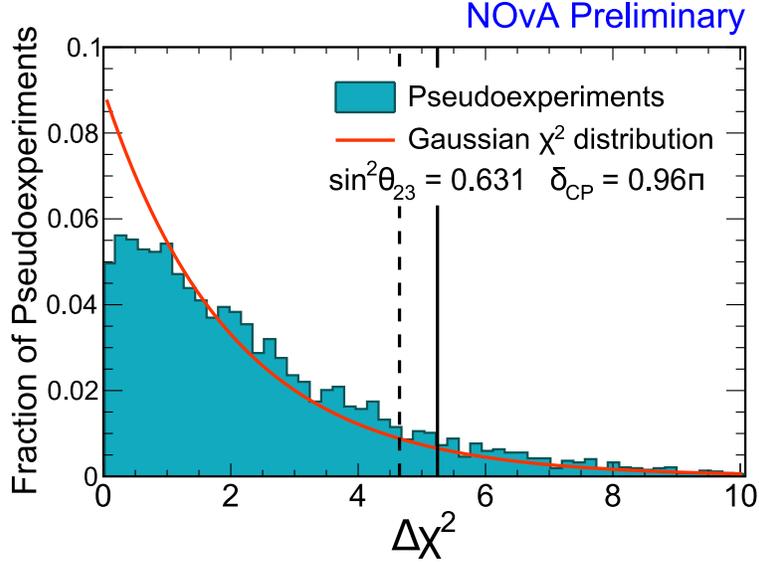


FIGURE 5.6. FC distribution of $\Delta\chi^2$ at $\sin^2\theta_{23} = 0.631$ and $\delta_{CP} = 0.96\pi$ compared to a χ^2 of two degrees of freedom. In this case, the 90% critical values from FC and χ^2 distributions are $\Delta\chi_c^2 = 5.25$ (solid line) and $\chi_c^2 = 4.60$ (dashed line), respectively.

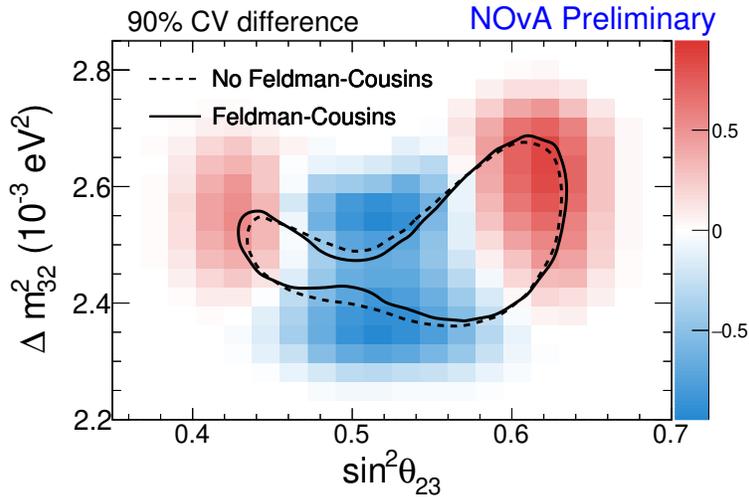


FIGURE 5.7. This surface shows the difference between the FC-calculated 90% critical value, $\Delta\chi_c^2$, subtracted by χ_c^2 . The solid and dashed contours are the FC corrected and uncorrected 90% CL regions in the $\sin^2\theta_{23}$ - Δm_{32}^2 Normal Hierarchy plane. Areas in blue and red show where NOvA's power of rejection is increased and decreased by the FC procedure, respectively.

the Normal Hierarchy (NH), the significance with which the Inverted Hierarchy (IH) can be ruled out is calculated.

To generate the acceptance interval associated with the $\Delta m_{32}^2 < 0$, pseudoexperiments are generated as being consistent with the best fit point in the IH ($\Delta m_{32}^2 = -2.55 \times 10^{-3} \text{ eV}^2$, $\sin^2 \theta_{23} = 0.58$, and $\delta_{CP} = 1.5\pi$) since in this case, all oscillation parameters are nuisance parameters. The test statistic $\Delta\chi^2$ is calculated as

$$(5.7) \quad \Delta\chi^2 = \chi_{\text{IH}}^2 - \chi_{\text{best}}^2$$

where in Equation 5.4, χ_{fixed}^2 becomes χ_{IH}^2 . Here, oscillation parameters are allowed to settle to their minimum values with the constraint that Δm_{32}^2 be less than zero. Best fit points that fall in the inverted hierarchy result in $\Delta\chi^2 = 0$, causing a pile up at zero in the FC distribution. It is evident from Figure 2.3 that the resulting FC distribution is far from a χ^2 -distribution caused by the highly non-Gaussian distribution of the binary mass hierarchy estimator.

A similar analysis can be performed for combinations of mass hierarchy and octant of θ_{23} constraints. For example, the rejection significance of the union of the IH and Upper Octant (UO; $\theta_{23} > \pi/4$) parameter spaces can be calculated from the $\Delta\chi^2$ distribution where χ_{IH}^2 is replaced by $\chi_{\text{IH,UO}}^2$.

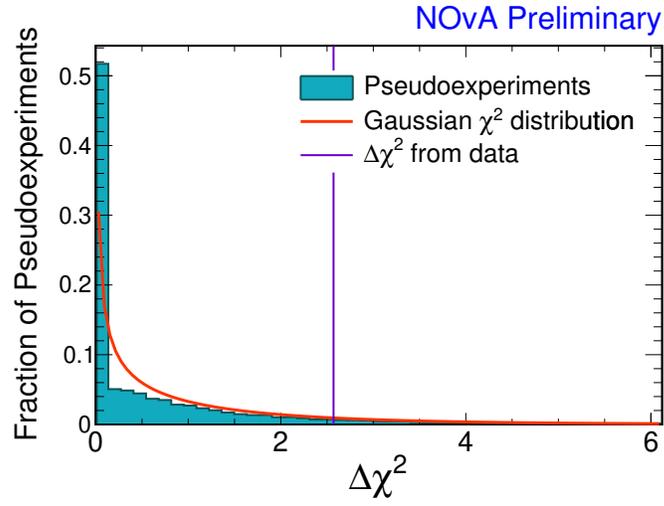


FIGURE 5.8. Feldman-Cousins distribution of the inverted neutrino mass hierarchy hypothesis. The measured $\Delta\chi^2$ is 2.47 corresponding to a IH rejection significance of 1.8σ .

CHAPTER 6

RESULTS AND THE ROLE OF HIGH PERFORMANCE

COMPUTING

*"The science of operations, as derived from
mathematics more especially, is a science of itself, and has
its own abstract truth and value."*

Ada Lovelace

A measurement of the neutrino oscillation parameters θ_{23} , $|\Delta m_{32}^2|$, δ_{CP} , and the mass hierarchy requires the ordering principle of Feldman and Cousins to correct for low statistics and physical boundaries, ensuring that confidence intervals possess the correct coverage. Producing these intervals is a computationally demanding task that requires the use of High Performance Computing (HPC) facilities to obtain results on a reasonable time-scale.

6.1. THE SCALE OF NOVA'S FC CORRECTIONS

To thoroughly analyze properties of the oscillation parameter space, NOvA produces significance profiles and confidence regions, which are constrained to a mass hierarchy, an octant of θ_{23} , or both. Additionally, rejection significances are calculated for the IH, Lower Octant (LO), Inverted Hierarchy Lower Octant (IHLO), Inverted Hierarchy Upper Octant (IHUO), and Normal Hierarchy Lower Octant (NHLO). Table 6.1 shows a summary of the number of $\Delta\chi^2$ values calculated broken down by individual significance profile, confidence region, and rejection significance for NOvA's neutrino + antineutrino analysis. In total, over

9 million values of $\Delta\chi^2$ were calculated utilizing the National Energy Research Scientific Computing Center (NERSC)'s HPC facilities and the methods described in this chapter.

TABLE 6.1. Summary of the psuedoexperiments generated for NOvA's 2018 FC corrections with the help from NERSC's HPC facilities. *NBins* and *NPts* refer to the number FC distributions generated for a given plot and the number of $\Delta\chi^2$ values calculated for a given plot, respectively. The first section contains a summary of the 2D confidence regions. The second section contains 1D significance profiles. The third section contains rejection significances.

Plot		NBins	NPts	Max NPts/Bin	Min NPts/Bin
δ_{CP} vs $\sin^2 \theta_{23}$	NH	598	3258041	20219	3226
δ_{CP} vs $\sin^2 \theta_{23}$	IH	357	2213382	18498	4994
$\sin^2 \theta_{23}$ vs Δm_{32}^2	NH	186	901250	8185	2724
$\sin^2 \theta_{23}$ vs Δm_{32}^2	IH	203	1010956	7738	2946
δ_{CP}	LOIH	60	257712	5691	2698
δ_{CP}	UOIH	60	162000	2700	2698
δ_{CP}	LONH	60	162000	2700	2697
δ_{CP}	UONH	60	162000	2700	2697
$\sin^2 \theta_{23}$	IH	60	322380	5373	5363
$\sin^2 \theta_{23}$	NH	60	323820	5397	5382
Δm_{32}^2	LOIH	60	162000	2700	2693
Δm_{32}^2	UOIH	60	162000	2700	2689
Δm_{32}^2	LONH	60	162000	2700	2687
Δm_{32}^2	UONH	60	162000	2700	2692
IH Rejection			45000		
LO			37705		
IHLO			40000		
IHUO			15000		
NHLO			47500		
Total			9,606,746		

6.2. MOTIVATION FOR HIGH PERFORMANCE COMPUTING

As mentioned in Chapter 5, each value of $\Delta\chi^2$ is calculated by the fitting procedure performed on pseudoexperiments that are completely independent from one another. This

type of problem lends itself to the technical term "Embarrassingly Parallel", which means that the level of concurrency – the number of simultaneous $\Delta\chi^2$ calculations – is limited by the number of available machines. This enables the computation time of FC calculations to scale perfectly with the number of available machines.

In the past, NOvA was using the limited resources of the Fermigrid. The Fermigrid, which is a shared computing cluster at Fermilab, allocates a limited number of concurrent processes for each experiment at Fermilab. NOvA is limited to 2500 concurrent processes, which caused the FC corrections for the 2017 analysis on neutrino-only data [20] to run over a four week period that was largely spent waiting in queue. Based on performance benchmarks of the neutrino + antineutrino FC code, it was estimated that it would take ~ 4 -5 months to complete NOvA's complete set of neutrino + antineutrino FC corrections. Given the need for production of results on a reasonable time scale, NOvA began working with the Scientific Discovery through Advanced Computing: High Energy Physics Data Analytics (SciDAC-4) collaboration to develop an infrastructure to run their FC on the large HPC machines at NERSC. The rest of this chapter is devoted to summarizing this work [21] and highlighting the results achieved because of it.

6.3. UTILIZING NERSC FACILITIES

The Fermigrid is a Unix cluster operating with the Scientific Linux Fermilab (SLF7) operating system and network-mounted repositories containing the code base used by each experiment. This is by design the same environment in which most Fermilab-based experiments develop their code on interactive Virtual Machines (VM). Since both platforms, development machines and the Fermigrid, are functionally identical at the operating system

and filesystem level, developers do not need to worry about the underlying architecture on which they are running code.

The machines at NERSC are designed to accommodate a wide range of applications, and so are not restricted to running the SLF7 operating system. Instead, NERSC users are encouraged to package their individual software into Docker images [22]. Docker images provide a way for developers to containerize their software applications and dependencies for portability and scalability. Similar to a VM, Docker containers can emulate operating systems and run programs loaded into the image on the virtual operating system. The advantage of using a Docker image over a VM is a Docker image instance can communicate directly with the host machine rather than having to virtualize hardware as VMs do. This results in less overhead and better overall performance. Docker containers also support mounted volumes that act as local filesystems within the container. Therefore, to run NOvA software at NERSC, the code, all dependencies, and the SLF7 operating system must be built into a Docker image.

Using the Slurm workload manager [23], jobs (units of execution) are submitted to NERSC machines by specifying a container, the command to be executed within the container, and resources requested for the job. Slurm distributes the job among the number of requested nodes, a homogeneous collection of computer processors running the number of processes on each node specified in the submission command. Results are returned to the user's or group's filesystem.

6.4. IMPLEMENTATION

To run NOvA's FC code at NERSC, the FC software and its dependencies, as well as the SLF7 operating system, are installed in a Docker image that is hosted on a private repository.

This image can be downloaded to NERSC where it is converted into a Shifter image format. Shifter was developed by NERSC to convert Docker images, VMs, and the like to a common container format that can be scaled on their machines. This conversion is transparent to the user and Shifter images can be executed in the same way as Docker images.

For the neutrino + antineutrino results, the generation of and fitting to an individual FC pseudoexperiment are implemented in a compiled C++ ROOT macro. The macro runs over a specified bin range in the selected parameter space, wherein each bin, a specified number of pseudoexperiments are generated and a fit is performed. A data object containing the pseudoexperiment's $\Delta\chi^2$ and parameter values are added to a collection that holds all such data objects created from an individual instance of the macro. The results from each instance are aggregated separately from the macro resulting in file sizes between 200 MB and 1 GB.

In addition to the FC software, a Python driver script is installed in the Docker image that assigns bin ranges to each processor running the FC job. The assignment is done in a round-robin fashion to ensure an even distribution of work among the job's resources.

This container implementation on NERSC was validated against NOvA's neutrino-only results [20], where the one and two dimensional profiles were reproduced in less than 12 hours during a reservation of dedicated resources.

6.5. JOB SUBMISSION AND LOAD BALANCING

The NERSC facilities are composed of three different architectures, each of which were utilized for the NOvA corrections. Table 6.2 shows a break down of the number of nodes, cores per node, and memory per node available per architecture. The available memory per core limits the concurrency of FC jobs on each architecture. Considering that each fit requires ~ 2 GB of memory, concurrency was limited on the Ivy Bridge and KNL architectures to 45

TABLE 6.2. NERSC machine configurations.

Machine	Processor Architecture	No. Nodes	No. Cores/Node	Memory/Node
Cori	Intel Haswell	2388	32	128 GB
Cori	Intel Knights Landing (KNL)	9688	64	96 GB
Edison	Ivy Bridge	5586	48	64 GB

and 32, respectively. Table 6.3 shows how the mean fit times vary across the architecture flavors. The Haswell nodes achieved the best performance, while the KNL nodes were the most abundant, and so processed a large proportion of the FC jobs.

TABLE 6.3. Mean fit times (minutes) per parameter and architecture achieved during the neutrino + antineutrino NOvA analysis.

Fit	δ_{cp}	Δm_{32}^2	$\sin^2 \theta_{23}$	Δm_{32}^2 vs. $\sin^2 \theta_{23}$	$\sin^2 \theta_{23}$ vs. δ_{CP}
Cori - KNL	-	25.41	60.99	59.17	55.18
Cori - Haswell	-	5.10	8.36	13.04	-
Edison - Ivy Bridge	50.79	17.02	14.20	-	-

Jobs were submitted to the NERSC machines based on architecture average run time, concurrency level, and price per CPU hour for efficient distribution and use of time during which the machines were reserved for our use [21]. HEPCloud [24], developed by Fermilab scientists, was used for the submission process. HEPCloud weights the requested resources to what are available across many different platforms and intelligently submits jobs wherever resources can be used most efficiently.

6.6. PERFORMANCE RESULTS

To avoid time in queue, NOvA and SciDAC-4 were able to reserve NERSC resources for allotted amounts of time; one reservation lasted 16 hours and another 36. During the first reservation, 17 million CPU-hours were consumed for a maximum concurrency of 1.1 million.

This concurrency level set a record as the largest ever pool of resources running simultaneously. The results from that reservation showed that some of the fits were unexpectedly minimizing to local minima in unphysical regions of phase space. A second reservation was made after the code was modified to rectify the error. During this reservation, 20 million CPU hours were consumed over 36 hours, with concurrency peaking at 0.7 million. These results were validated and compiled in time to present at the Neutrino 2018 Conference [15].

During the validation process of the neutrino-mode results, it was found that fits to those data averaged two to three minutes of computation time per fit on the Haswell nodes. Using this as a baseline, it is estimated that NERSC enabled the NOvA and SciDAC-4 team to process Felman-Cousins corrections 50 times faster than before. This dramatic improvement alone opens the door for more precise and complicated analyses to be done in the future.

6.7. RESULTS OF THE FC CORRECTIONS

Figure 6.1 shows a collection of the FC corrected results from the second NERSC reservation. The solid lined-contours correspond to the FC corrected acceptance regions and the dashed lined-contours are drawn from the Gaussian profiles without FC corrections applied. Generally, it can be seen that the FC procedure slightly reduced NOvA’s statistical power of a $\sin^2 \theta_{23}$ measurement while increasing the power of a Δm_{32}^2 measurement. Table 6.4 shows the best fit values for Δm_{32}^2 , $\sin^2 \theta_{23}$, and δ_{CP} with associated 1σ confidence intervals before and after FC corrections. Table 6.5 shows rejection significances for the various mass hierarchy and θ_{23} octant combination hypotheses.

TABLE 6.4. Best-fit oscillation parameters and associated 1σ confidence intervals with and without FC corrections determined for NOvA’s neutrino + antineutrino data set [15].

Parameter	Central Value	No FC Interval	FC Interval
$\Delta m_{32}^2/10^{-3} \text{ eV}^2/c^4$	+2.51	+0.10 -0.08	+0.12 -0.08
$\sin^2 \theta_{23}$	0.58	+0.02 -0.03	± 0.03
δ_{CP}/π	0.2	+0.1 -0.5	+1.2 -0.6

TABLE 6.5. Rejection significances for combined mass hierarchy and θ_{23} octant hypotheses.

	IH	IHUO	IHLO	NHLO	NHUO
Significance (σ)	1.77	1.72	2.18	2.29	-

6.8. DOMAIN DECOMPOSITION WITH DIY

The results achieved for NOvA’s first neutrino + antineutrino analysis using the HPC facilities at NERSC were a dramatic improvement over the Fermigrid, but there is room for further improvement by continuing to optimize the FC infrastructure for HPC machines. Work is ongoing to disentangle NOvA’s FC procedure from the ROOT analysis framework and move toward an implementation that can take full advantage of HPC machines with multithreading and vectorization support, as well as modularity to support custom analyses.

The first step in this direction was to implement domain decomposition in a Message Passing Interface (MPI) based framework using DIY [25], rather than the Python driver script used during the May reservations. DIY is a collection of libraries developed to add a layer of abstraction onto MPI. With it, arbitrary problem spaces, or domains, can be decomposed into units called "blocks", where MPI communication between neighboring blocks is mediated behind the scenes by the DIY libraries. This provides a natural mapping of the

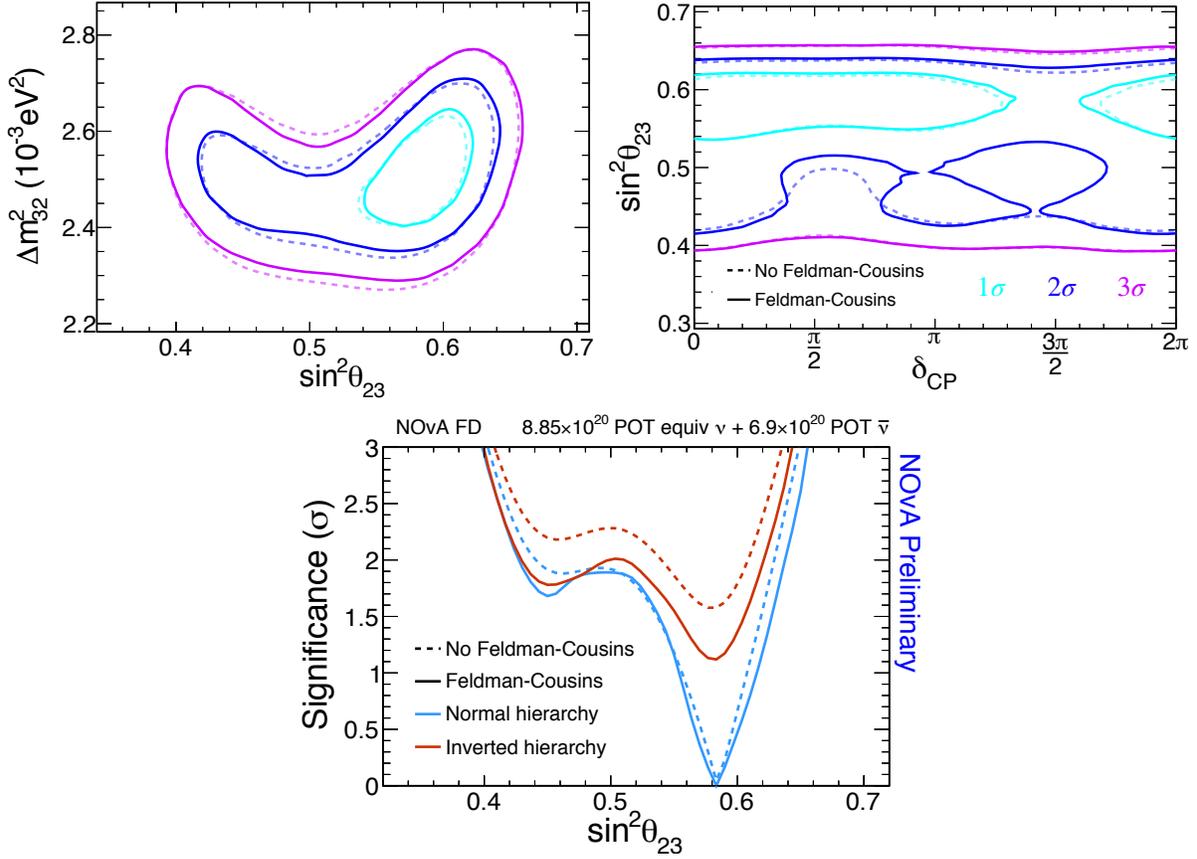


FIGURE 6.1. Fully corrected profiles of $\sin^2 \theta_{23}$, Δm_{32}^2 , and δ_{CP} . The dashed and solid lines show significances derived assuming Gaussian errors and from FC distributions, respectively.

one and two dimensional parameter spaces onto DIY domains and facilitates communication between points in parameter space.

Within this DIY framework, the user specifies a total number of blocks and a desired topology (number of blocks spanning each dimension), as well as oscillation parameter specifications and the required number of pseudoexperiments per bin. Each block is assigned an equal fraction of the parameter space according to DIY block variables such as x , y , and z positions in the block-space geometry, and converting that to a list of bins for which that

block should process. The job can be easily scaled up by requesting multiple copies of the block-space geometry.

Future work by myself and the SciDAC-4 collaboration will expand on this framework with the end-goal of making it possible for NOvA to optimize our event selection criteria based on FC-corrected sensitivities. This will require the development of NOvA's event selection within the DIY framework; a transition to an HPC-optimized data format based on the read-parallel Hierarchical Data Format (HDF5) file format [26]; and a removal of the dependence on ROOT entirely by implementing a multi-threaded fitting procedure for a dramatic increase in efficiency. These improvements have the potential to greatly decrease the time to final results and improve statements made about the physics of neutrino oscillations.

CHAPTER 7

CONCLUSION

High performance computing facilities have been used to greatly improve the final step of the NOvA oscillation analysis, the Feldman-Cousins (FC) procedure, which is used to calculate statistical power by large-scale Monte Carlo simulation. Increased complexity of the fitting procedure introduced by the combined fitting of neutrino and antineutrino data prompted a greater demand for computing resources than those previously sufficient for NOvA's neutrino-only analyses. Justification for NOvA's employment of the FC procedure was demonstrated and work done to execute FC implementation on HPC machines at the NERSC facilities was discussed.

As a result of this work, NOvA's FC corrections were processed in 36 hours, consuming 20 million CPU-hours, for an estimated speedup factor of 50. On-going work is should improve this speedup further by optimizing the procedure for High Performance Computing platforms, as well as improve the infrastructure to enable other scientists to use these tools for their own custom analyses.

BIBLIOGRAPHY

- [1] G. J. Feldman and R. D. Cousins, “A Unified approach to the classical statistical analysis of small signals,” *Phys. Rev.*, vol. D57, pp. 3873–3889, 1998.
- [2] V. Gribov and B. Pontecorvo, “Neutrino astronomy and lepton charge,” *Physics Letters B*, vol. 28, no. 7, pp. 493 – 496, 1969.
- [3] Y. Fukuda *et al.*, “Evidence for oscillation of atmospheric neutrinos,” *Phys. Rev. Lett.*, vol. 81, pp. 1562–1567, Aug 1998.
- [4] J. B. Albert *et al.*, “Search for Majorana neutrinos with the first two years of EXO-200 data,” *Nature*, vol. 510, pp. 229–234, 2014.
- [5] L. Cardani, “Neutrinoless double beta decay overview,” *SciPost Physics Proceedings*, 02 2019.
- [6] M. Tanabashi *et al.*, “Review of particle physics,” *Phys. Rev. D*, vol. 98, p. 030001, Aug 2018.
- [7] P. Adamson *et al.*, “The NuMI Neutrino Beam,” *Nucl. Instrum. Meth.*, vol. A806, pp. 279–306, 2016.
- [8] S. De Rijck, “Latest results from minos and minos+,” *Journal of Physics: Conference Series*, vol. 873, p. 012032, 07 2017.
- [9] E. Niner, *Observation of Electron Neutrino Appearance in the NuMI Beam with the NOvA Experiment*. PhD thesis, Indiana University, 2015.
- [10] A. Aurisano *et al.*, “A Convolutional Neural Network Neutrino Event Classifier,” *JINST*, vol. 11, no. 09, p. P09001, 2016.
- [11] F. Psihas, *Measurement of Long Baseline Neutrino Oscillations and Improvements From Deep Learning*. PhD thesis, Indiana University, 2018.

- [12] R. E. Kalman, “A new approach to linear filtering and prediction problems,” *Transactions of the ASME—Journal of Basic Engineering*, vol. 82, no. Series D, pp. 35–45, 1960.
- [13] F. James and M. Winkler, *Minuit2 Manual*. <https://root.cern.ch/guides/minuit2-manual>.
- [14] R. Brun and F. Rademakers, “ROOT: An object oriented data analysis framework,” *Nucl. Instrum. Meth.*, vol. A389, pp. 81–86, 1997.
- [15] M. Sanchez, “NOvA results and prospects.” Neutrino18 Conference, 2018.
- [16] J. Neyman, “Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability,” *Phil. Trans. Roy. Soc. Lond.*, vol. A236, no. 767, pp. 333–380, 1937.
- [17] N. Jerzy *et al.*, “IX. on the problem of the most efficient tests of statistical hypotheses,” *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, vol. 231, pp. 289–337, 2019/01/16 1933.
- [18] S. S. Wilks, “The large-sample distribution of the likelihood ratio for testing composite hypotheses,” *Ann. Math. Statist.*, vol. 9, pp. 60–62, 03 1938.
- [19] G. Feldman, “Notes on the inclusion of nuisance parameters in the unified approach.” NOvA Internal Document: **DocDB: 15884**.
- [20] NOvA Collaboration, “New constraints on oscillation parameters from ν_e appearance and ν_μ disappearance in the NOvA experiment,” *ArXiv e-prints*, p. arXiv:1806.00096, May 2018.
- [21] A. Sousa *et al.*, “Implementation of feldman-cousins corrections and oscillation calculations in the hpc for the nova experiment,” *to appear in Proc. 23rd International Conference on Computing in High Energy and Nuclear Physics (CHEP 2018)*.

- [22] Docker, Inc., “Docker: Enterprise container platform.” <https://www.docker.com>, 2010.
- [23] SchedMD LLC, “Slurm workload manager.” <https://slurm.schedmd.com>.
- [24] M. Altunay *et al.*, “Intelligently-automated facilities expansion with the HEPCloud Decision Engine,” *arXiv e-prints*, p. arXiv:1806.03224, June 2018.
- [25] D. Morozov and T. Peterka, “Block-parallel data analysis with diy2,” in *Proceedings of the 2016 IEEE Large Data Analysis and Visualization Symposium LDAV’16*, (Baltimore, MD), 2016.
- [26] The HDF Group, “Hierarchical Data Format, version 5,” 1997-NNNN. <http://www.hdfgroup.org/HDF5/>.