

Electron Neutrino Appearance in the MINOS Experiment

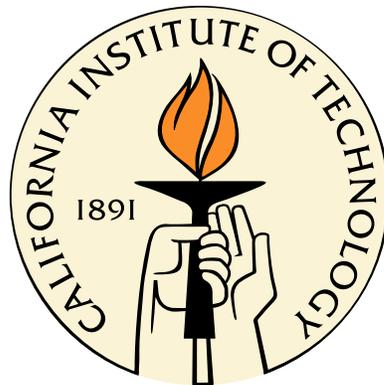
Thesis by

Mhair-Armen Hagop Orchanian

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



California Institute of Technology

Pasadena, California

2012

(Defended May 9, 2012)

©2012

Mhair-Armen Hagop Orchanian

All Rights Reserved

To my parents, Anna and Hagop

Acknowledgements

With the successful completion of this thesis, I would like to express my deep gratitude to the people who helped make it a reality.

Despite his many commitments, my advisor, Harvey Newman, skillfully guided me through the graduate program and made sure I stayed on track. His advice along the way – particularly his careful reading of and comments on drafts of this thesis – was absolutely invaluable. I am deeply indebted to Ryan Patterson; his patience, incredible generosity with his time, and great skill in explaining difficult concepts made my education as a physicist not only possible but also quite enjoyable. My fellow graduate students on MINOS, Alex Himmel and Pedro Ochoa, were always happy to answer my questions about programming and physics and to share with me their experiences in the graduate program. I am also very grateful to Leon Mualem for helping me take over the Caltech Monte Carlo generation responsibilities and to Jason Trevor for training me in the $\text{NO}\nu\text{A}$ R&D lab in the year before I started the graduate program.

Working with my colleagues on MINOS has been a very rewarding experience. With her indefatigable patience and guidance, Lisa Whitehead was directly responsible for the timely completion of the work I did for my candidacy; her continued assistance since then has been greatly appreciated. My fellow graduate students on the ν_e appearance analysis – João Coelho, Ruth Toner, and Adam Schreckenberger – have been wonderful collaborators; their camaraderie has made the joy of scientific discovery even sweeter. I am also thankful to all of the ν_e group conveners, past and present – including Tricia Vahle, Mayly Sanchez, Jeff Nelson, and Greg Pawloski – for their able leadership through our challenging undertakings.

My family has been a constant source of inspiration throughout my life and has served as a solid support system for as long as I can remember. My elder cousin, Karoun Bagamian, blazed the trail for graduate education in our family, setting an example that I was honored to follow. My paternal grandfather, Sarkis, was an ardent proponent of education, despite never making it past middle school in the tumult following the Armenian Genocide. He missed this milestone in my life by just four years, but I just know that he would have been so proud of me. My parents, Hagop and Anna, instilled in me the importance of education and character. They have never thought twice about sacrificing for my ultimate benefit, and their unwavering support and encouragement in all of my endeavors – particularly throughout my writing of this thesis – has been utterly indispensable. I consider myself incredibly fortunate to be their son.

Abstract

This thesis describes a search for ν_e appearance in the two-detector long-baseline MINOS neutrino experiment at Fermilab, based on a data set representing an exposure of 8.2×10^{20} protons on the NuMI target. The analysis detailed herein represents an increase in sensitivity to the θ_{13} mixing angle of approximately 25% over previous analyses, due to improvements in the event discriminant and fitting technique. Based on our observation, we constrain the value of θ_{13} further, finding $2 \sin^2 \theta_{23} \sin^2 2\theta_{13} < 0.12(0.20)$ at the 90% confidence level for $\delta_{\text{CP}} = 0$ and the normal (inverted) neutrino mass hierarchy. The best-fit value is $2 \sin^2 \theta_{23} \sin^2 2\theta_{13} = 0.041_{-0.031}^{+0.047}(0.079_{-0.053}^{+0.071})$ under the same assumptions. We exclude the $\theta_{13} = 0$ hypothesis at the 89% confidence level.

Contents

Acknowledgements	v
Abstract	vii
List of Figures	xiii
List of Tables	xvii
1 Neutrinos: The Standard Model and Beyond	1
1.1 An Overview of the Standard Model	2
1.1.1 Theoretical Foundation	3
1.1.2 Generating Neutrino Masses	6
1.2 Getting to Massive Neutrinos	8
1.2.1 Solar Neutrinos	9
1.2.2 Atmospheric Neutrinos	12
1.2.3 Accelerator Neutrinos	19
1.2.4 Reactor Neutrinos	20
1.3 Describing Massive Neutrinos	23
1.3.1 Formalism of Neutrino Mixing	23
1.3.1.1 General Probability of Flavor Transition	25
1.3.1.2 The Mixing Matrix	28
1.3.2 Two-Neutrino Oscillations	30
1.3.2.1 The Two-Neutrino Framework	30
1.3.2.2 Oscillation Probabilities and Parameter Measurements	31

1.3.2.3	Matter Effects	38
1.3.3	Three-Neutrino Oscillations	45
1.3.3.1	Understanding the Two-Neutrino Approximation	45
1.3.3.2	$\nu_\mu \rightarrow \nu_e$ Oscillation in the Three-Neutrino Framework	48
2	Experimental Setup	51
2.1	Neutrinos at the Main Injector (NuMI)	51
2.2	The MINOS Detectors	59
2.2.1	The Magnetic Field	61
2.2.2	The Scintillator Strips	62
2.2.2.1	The Light Injection System	65
2.2.3	The Photodetectors	67
2.2.4	The Electronics	69
2.2.5	The Data Acquisition System (DAQ)	71
2.3	Calibration	72
2.4	Event Types of Interest	78
2.5	Reconstruction	81
2.6	Simulation	83
2.6.1	Simulating the Beam	83
2.6.2	Simulating Neutrino Interactions	85
2.6.3	Simulating the Detectors	85
3	Overview of the Analysis	89
4	Event Selection	95
4.1	Fiducial Volume	95
4.2	ν_e -like Preselection	96
4.3	Library Event Matching	99
4.3.1	The Library	101
4.3.2	Preparing Events for Matching	104
4.3.2.1	Pulse Height Measurement	104

4.3.2.2	Attenuation and Interdetector Corrections	105
4.3.2.3	Compacting an Event	107
4.3.3	The Likelihood Formula and the Comparison	110
4.3.4	Gauging the Signal-Like Nature of the Best-Matched Events	114
4.3.5	The LEM PID	118
4.3.6	LEM in the Near Detector	121
5	The Near Detector Background Decomposition	127
5.1	The PORP Method	128
5.2	The HOOHE Method	130
6	The Far Detector Prediction	139
6.1	Predicting the ND-Type Background Components	139
6.2	Predicting the Signal ν_e CC and Background ν_τ CC Components	140
6.2.1	Selection Efficiency Considerations	143
6.3	Far Detector Predictions	145
7	Systematic Errors	151
7.1	Effects Common to Both Detectors	153
7.1.1	Neutrino Flux Uncertainties	153
7.1.2	Neutrino Interaction Modeling Uncertainties	154
7.1.3	Energy Scale Uncertainties	155
7.2	Effects Differing Between the Detectors	156
7.2.1	Normalization Uncertainties	156
7.2.2	Calibration Uncertainties	156
7.2.3	PMT Crosstalk Modeling Uncertainties	158
7.3	Effects Unique to ν_τ CC and Signal ν_e CC Events	159
7.4	Systematics and the HOOHE Decomposition	159
7.5	Summary	160
8	Sensitivity, Sidebands, and Results	165
8.1	Sensitivity to θ_{13}	165

8.2	Validating the Analysis with Sidebands	168
8.3	The Far Detector Data	171
8.4	Fitting the Data	178
9	Conclusions and a Look at the Future of Neutrino Physics	187
A	Reducing the Sensitivity of LEM to Calibration Systematics	195
A.1	Options for Pulse Height Measurement in LEM	196
A.2	Poisson Distribution Substitute for Reconstructed Energy Option . .	198
A.3	Comparison of the Two Options	201
B	PID Cut Optimization and Analysis Binning Studies with PORP	203
C	Mathematical Details of the Systematic Error Evaluation	207
D	Oscillation Parameter Uncertainties and Pseudoexperiment Generation	211
	Bibliography	213

List of Figures

1.1	Solar neutrino energy spectra	10
1.2	Homestake chlorine detector results	11
1.3	SSM predictions vs. experimental results	13
1.4	SNO results	14
1.5	Super-K atmospheric neutrino zenith angle distributions	17
1.6	Super-K data and MC L/E curves	18
1.7	K2K visible energy distribution	20
1.8	MINOS reconstructed neutrino energy distribution	21
1.9	Summary of reactor experiment observations	24
1.10	Atmospheric oscillation parameter contours	33
1.11	Global solar fit	35
1.12	KamLAND oscillation curve and results	36
1.13	KamLAND plus global solar fit	37
1.14	CHOOZ exclusion contour	39
1.15	MINOS ν_e appearance results	50
2.1	Overhead and cross-sectional views of the MINOS experiment	52
2.2	Plan and elevation views of the NuMI and MINOS facilities at Fermilab	52
2.3	Protons-on-target vs. date and run periods	53
2.4	Illustration of meson dispersal out of target; technical drawing of the NuMI target	54
2.5	Technical drawing of Horn 2	56
2.6	Cartoon illustration of focusing horns	56
2.7	Neutrino spectra in “true” and “pseudo” target/horn configurations . .	57

2.8	Main components of NuMI beamline	59
2.9	Photographs of the Near and Far Detectors	60
2.10	Arrangement of MINOS planes in the detectors	60
2.11	Cross-sectional illustrations of the magnet coils	63
2.12	Magnetic field maps in the ND and FD	63
2.13	Cutaway drawing of scintillator strip	64
2.14	Typical MINOS module manifold assembly	65
2.15	Layout of modules on Far Detector planes	66
2.16	Layout of modules on Near Detector planes	66
2.17	Cutaway view of a light injection module	67
2.18	Schematic drawing of the scintillator system readout for a module . . .	68
2.19	PMT mounting assembly and fiber “cookie” layouts	69
2.20	Calibration chains	77
2.21	Event types in the MINOS detectors	80
2.22	Multiple events in the ND from one beam spill	82
2.23	ν_μ CC data, raw MC, and tuned MC in the Near Detector	84
2.24	Decalibration scheme for simulating data	87
4.1	Preselection cuts in the FD	98
4.2	Sample input event and matches	100
4.3	Overview of the LEM procedure	102
4.4	<i>FOM</i> vs. library ν_e fraction and library size	103
4.5	Detector light levels before and after corrections	108
4.6	Strip grouping pattern in LEM	110
4.7	Compacting of sample event	111
4.8	$-\Delta\mathcal{L}$ distributions for Far Detector Monte Carlo	114
4.9	f_{50} for signal and background	115
4.10	y_{50} for signal and background	116
4.11	q_{50} for signal and background	117
4.12	Correlations among f_{50} , y_{50} , and q_{50}	119

4.13	LEM PID distributions in Far Detector Monte Carlo	122
4.14	LEM distribution for ND data and MC	123
4.15	Correlation between y_{50} and true hadronic character	124
4.16	LEM PID in Near Detector: before and after “Baryon x_f Selection” reweighting	125
4.17	y_{50} in Near Detector: before and after “Baryon x_f Selection” reweighting	126
5.1	True neutrino energy spectra in different beam configurations	131
5.2	Near Detector event type fractions in different beam configurations . .	132
5.3	Near Detector data components vs. reconstructed energy	136
5.4	Near Detector data components vs. LEM PID	137
6.1	MRE selection efficiencies and correction factor	146
6.2	Predicted LEM PID distribution	147
6.3	Predicted reconstructed energy distribution for LEM PID > 0.7	149
6.4	Predicted reconstructed energy distributions in analysis region	150
8.1	90% C.L. sensitivity to $\theta_{13} = 0$	167
8.2	Sensitivity improvements from analysis changes	168
8.3	LEM PID distribution for anti-LEM sideband	170
8.4	LEM PID distribution for MRCC sideband	172
8.5	Far-to-Near ratio in data vs. time	173
8.6	Timing distribution of Far Detector events	174
8.7	Far Detector event vertex coordinate distributions	175
8.8	LEM PID distribution for Far Detector data	176
8.9	Vertex distributions in x - y plane for all Far Detector events	177
8.10	Reconstructed energy distribution for Far Detector data	178
8.11	Extended reconstructed energy distribution for Far Detector data . . .	179
8.12	Vertex distributions in x - y plane for 5–6 GeV Far Detector events . . .	179
8.13	Likelihood contour for normal mass hierarchy at $\delta_{\text{CP}} = 0$	182
8.14	Feldman–Cousins confidence intervals	184

8.15	Best-fit signal spectra in LEM PID analysis bins	185
9.1	MINOS and T2K results overlaid	189
9.2	Summary and combination of results for θ_{13}	191
9.3	Combined-result signal spectra in LEM PID analysis bins	192
A.1	Reconstructed energy vs. true deposited energy in Far Detector Monte Carlo	199
A.2	Reconstructed energy distributions for fixed true deposited energy . . .	200

List of Tables

1.1	Standard Model fermions	3
1.2	Standard Model gauge bosons	3
2.1	Decay processes contributing to neutrino flux	58
2.2	Proton exposures and hardware notes for each run period	59
6.1	Far Detector predictions	148
7.1	Systematic errors on ND-type backgrounds for LEM PID > 0.7	161
7.2	Systematic errors on ν_τ CC background and signal ν_e CC for LEM PID > 0.7	162
D.1	Oscillation parameter uncertainties	211

Chapter 1

Neutrinos: The Standard Model and Beyond

Despite the abundance of neutrinos in the universe, these elusive particles have challenged physicists to study their behavior experimentally and to fit their properties within a theoretical framework. In 1930, Wolfgang Pauli postulated the existence of the neutrino [1] as a “desperate remedy” to save the laws of conservation of energy and angular momentum, following the discovery of the continuous nature of the energy distribution of electrons emitted in nuclear beta decay [2] and the subsequent discovery that the apparent energy deficit could not be accounted for as energy lost by the electrons to the surrounding material between emission and detection [3]. Enrico Fermi proposed the first theory of beta decay in 1934 [4], and soon thereafter, Hans Bethe and Rudolf Peierls estimated the cross section for neutrino–nucleus interactions to be on the order of 10^{-44} cm², concluding that “there is no practically possible way of observing the neutrino” [5]. Nevertheless, Clyde Cowan and Frederick Reines set out to attempt it anyway, designing an experiment to detect the antineutrinos produced by the reactions taking place in a nuclear reactor. Following an inconclusive attempt in 1953 [6] at the Hanford reactor in Washington state, they totally redesigned the detector and tried again at the Savannah River Plant in South Carolina, and in 1956, they announced the unambiguous detection of antineutrinos [7].¹

We have since continued to expand our knowledge of neutrino properties at a brisk pace. In Section 1.1, we provide an overview of the Standard Model of Parti-

¹A personal account by Reines of this endeavor is given in [8].

cle Physics – our current theory of the constituents of matter (including neutrinos) and the interactions (not including gravity) among these constituents. The Standard Model assumes neutrinos to be massless; as discussed in Section 1.2, a long series of experiments have consistently disagreed with predictions based on this assumption, inexorably leading us to the conclusion that neutrinos are in fact massive, thereby providing the first (and so far, only) clear evidence for physics beyond the Standard Model. Section 1.3 details the theoretical framework describing the behavior of massive neutrinos and the experimental results providing measurements of the parameters involved in this framework. We conclude that section with a discussion of the basic goal of this thesis, including a description of the aforementioned framework’s application to this goal.

1.1 An Overview of the Standard Model

According to the Standard Model [9], matter consists of twelve spin-1/2 fermions (six quarks² and six leptons) – listed in Table 1.1 along with some key properties – and their antiparticles. Interactions among these particles are mediated by the exchange of spin-1 bosons, listed in Table 1.2 along with some of their key properties.³ All quarks and leptons interact via the weak interaction, mediated by the W^\pm and Z^0 bosons. One could say that the weak vector bosons couple to fermion flavor (u, d, c, s, t , and b for quarks; e, μ , and τ for leptons). All quarks and all charged leptons interact via the electromagnetic interaction, mediated by the photon – that is, the photon couples to any particle with an electric charge, which is why neutrinos are the only fundamental particles that do not participate in the electromagnetic interaction. Finally, only particles with color participate in the strong interaction, mediated by the gluon,⁴ and

²Technically, there are 18 quarks – three colors (r, b, g) for each flavor (u, d, c, s, t, b).

³Gravity, not included in the Standard Model, is expected to be mediated by the exchange of the graviton, a spin-2 boson.

⁴Gluons are also colored particles. In particular, every gluon carries one unit of color and one unit of anticolor. One may expect $3 \times 3 = 9$ different gluons, then, but there are only eight, an octet. The ninth constitutes a singlet and is colorless; the principle of confinement requires that all naturally occurring particles be color singlets. Thus, we would expect to find the ninth gluon in great abundance, but we do not. Hence, we conclude that it must not exist [9].

the only fermions with color are the quarks.

Generation	Quarks			Leptons		
	Name	Charge	Mass (MeV)	Name	Charge	Mass (MeV)
First	u	$+2/3$	$2.49^{+0.81}_{-0.79}$	e	-1	$0.511 \pm 1.3 \times 10^{-8}$
	d	$-1/3$	$5.05^{+0.75}_{-0.95}$	ν_e	0	$< 2 \times 10^{-6}$
Second	c	$+2/3$	$(1.27^{+0.07}_{-0.09}) \times 10^3$	μ	-1	$105.658 \pm 3.8 \times 10^{-6}$
	s	$-1/3$	101^{+29}_{-21}	ν_μ	0	< 0.19
Third	t	$+2/3$	$(172.0 \pm 1.6) \times 10^3$	τ	-1	1776.82 ± 0.16
	b	$-1/3$	$(4.19^{+0.18}_{-0.06}) \times 10^3$	ν_τ	0	< 18.2

Table 1.1: Fermions of the Standard Model. All are spin-1/2, and charges are in units of the proton charge. Masses are as reported in [10]. In particular, masses for the u, d , and s quarks are “current-quark masses,” and masses for the c and b quarks are “running masses,” all in the $\overline{\text{MS}}$ mass-independent subtraction scheme. The neutrinos are massless in the Standard Model by assumption; the given masses are the best available experimental limits on their effective masses. Cosmological constraints are much stricter: data on the Cosmic Microwave Background (CMB) from the WMAP experiment, combined with supernovae data and data on galaxy clustering, provide an upper limit of approximately 0.68 eV (1.7 eV if a more conservative estimate of the uncertainties on the astrophysical data is used) at 95% C.L. on the *sum* of the neutrino masses [10].

Name	Charge	Mass (GeV)	Interaction
γ	0	0	electromagnetic
g	0	0	strong
W^\pm	± 1	80.399 ± 0.023	weak
Z^0	0	91.1876 ± 0.0023	weak

Table 1.2: Gauge bosons of the Standard Model, responsible for mediating the strong, weak, and electromagnetic interactions. All are spin-1, and charges are in units of the proton charge. Masses are as reported in [10].

1.1.1 Theoretical Foundation

In the language of group theory, the Standard Model is based on the symmetry group $SU(3)_{\text{color}} \otimes SU(2)_L \otimes U(1)$ [11]. $SU(3)_{\text{color}}$ is the symmetry group of quantum chromodynamics (QCD), the physics of particles with color (i.e., quarks). $SU(2)_L \otimes U(1)$ is the symmetry group of electroweak theory, the unified theory of the electromagnetic

and weak interactions formulated by Glashow [12], Weinberg [13], and Salam [14] during the 1960s.⁵ In particular, $SU(2)_L$ refers to weak isospin, where the subscript indicates that only left-handed chiral fields are involved; $U(1)$ refers to weak hypercharge, which involves fields of both chiralities (i.e., both left-handed and right-handed fields).

Interactions in the Standard Model arise from requiring local gauge invariance of the free Lagrangian of the fermions among which the interaction is to occur [9]. Theories that result from this requirement are known as gauge theories, which have been proven to be renormalizable [15, 16]. In the case of QCD, we begin with the free Lagrangian for quarks and require that the Lagrangian be invariant under local (i.e., space-time-dependent) $SU(3)_{\text{color}}$ transformations on the quark fields. This necessitates the introduction of eight massless gauge fields, which we identify with the eight gluons that mediate the strong interaction. The situation is slightly different for local $SU(2)_L \otimes U(1)$ transformations on the fermion fields in electroweak theory, as the requirement of local gauge invariance alone would require that all four associated gauge fields be massless. The Standard Model postulates the existence of an additional field – the scalar Higgs field – whose potential attains a minimum at a nonzero value (the vacuum expectation value, or “vev”) of the field; spontaneous symmetry breaking in this field imparts mass to the W^\pm and Z^0 bosons but leaves the photon massless. This combination of local gauge invariance and spontaneous symmetry breaking [17, 18] is referred to as the Higgs mechanism, a crucial ingredient of the electroweak theory. As this thesis focuses on the physics of neutrinos, we will largely set QCD aside and focus on electroweak theory.

As suggested earlier, electroweak theory is typically formulated in terms of chiral (i.e., left-handed and right-handed) fermion fields. In particular, the left-handed fields are arranged into doublets:

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L, \begin{pmatrix} u \\ d' \end{pmatrix}_L, \begin{pmatrix} c \\ s' \end{pmatrix}_L, \begin{pmatrix} t \\ b' \end{pmatrix}_L. \quad (1.1)$$

⁵This formulation is often referred to as the “GWS model,” for its authors.

Note that the quark doublet fields d', s', b' are of definite flavor and thus correspond to the weak eigenstates $|d'\rangle, |s'\rangle, |b'\rangle$. They are not identical to the physical fields d, s, b – i.e., those of definite mass – which correspond to the mass eigenstates $|d\rangle, |s\rangle, |b\rangle$. Rather, they are related via the unitary Cabibbo–Kobayashi–Maskawa (CKM) mixing matrix V [19, 20]:⁶

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \quad (1.2)$$

Thus, the fields of definite flavor are linear combinations of the fields of definite mass. The right-handed fermion fields, on the other hand, form singlets:

$$e_R, \mu_R, \tau_R, u_R, d_R, c_R, s_R, t_R, b_R. \quad (1.3)$$

Local gauge invariance under $SU(2)_L \otimes U(1)$ requires the introduction of four vector fields; W_μ^1, W_μ^2 , and W_μ^3 correspond to the three generators of $SU(2)_L$, and B_μ corresponds to the single generator of $U(1)$. The W^\pm bosons are linear combinations of W_μ^1 and W_μ^2 :

$$W_\mu^\pm = \frac{1}{\sqrt{2}} (W_\mu^1 \mp iW_\mu^2). \quad (1.4)$$

In Glashow’s theory, the neutral fields W_μ^3 and B_μ mix to produce A_μ , the photon field, and Z_μ , the Z^0 field:

$$A_\mu = B_\mu \cos \theta_w + W_\mu^3 \sin \theta_w, \quad (1.5)$$

$$Z_\mu = -B_\mu \sin \theta_w + W_\mu^3 \cos \theta_w, \quad (1.6)$$

where θ_w is the “weak mixing angle” or “Weinberg angle,” a fundamental parameter of the Standard Model. The interaction Lagrangian – originally written in terms of the W_μ^i and B_μ and associated coupling constants g_w and g' , respectively (using the

⁶As determined by experiment, the CKM matrix is nearly diagonal, but not quite [10]; that is, “cross-generational” weak interactions (e.g., $s \rightarrow u + W^-$) do occur in the quark sector, albeit much less often than intragenerational ones (e.g., $d \rightarrow u + W^-$).

notation of [9]) – can now be rewritten in terms of W_μ^\pm , Z_μ , and A_μ and coupling constants g_w , g_z , and g_e , where

$$g_e = g_w \sin \theta_w = g' \cos \theta_w, \quad (1.7)$$

$$g_z = \frac{g_w}{\cos \theta_w} = \frac{g'}{\sin \theta_w}. \quad (1.8)$$

Thus, the weak and electromagnetic coupling constants are not independent, and electroweak unification is made manifest.

1.1.2 Generating Neutrino Masses

We have already pointed out that local gauge invariance and spontaneous symmetry breaking in the Higgs field impart mass to the W^\pm and Z^0 bosons⁷ while leaving the photon massless. The Higgs field also imparts mass to the fermions, via additional terms in the Lagrangian that couple the Higgs field to the various fermion fields. To conserve isospin invariance in these couplings, however, the Higgs field must be combined with a fermion doublet and the corresponding singlet in each such term in the Lagrangian, thereby coupling a left-handed field to a right-handed field in a so-called Yukawa coupling. Once the Higgs field acquires its vacuum expectation value, this term splits into two terms: one is a Yukawa coupling of the fermion field to the physical Higgs field, and the other is a pure fermion mass term.

Note the absence of right-handed neutrinos in Equation 1.3, as the particle content of the Standard Model includes only left-handed neutrinos. Indeed, it has been experimentally verified [21, 22, 23] that only left-handed neutrinos (and right-handed antineutrinos) participate in the weak interaction. Without these neutrino singlets, one cannot form Yukawa couplings of the neutrino to the Higgs field, so neutrinos remain massless in the Standard Model. The experiments discussed in Section 1.2,

⁷It is worth mentioning that Fermi's original theory of beta decay [4] – essentially the first formulation of the charged-current weak interaction – treated the process as a contact interaction, i.e., one that did not require a mediating particle. Indeed, Fermi's theory is the low-energy limit of the GWS model, yielding excellent approximations when the momentum transfer is small compared to the W^\pm mass.

however, leave us no choice but to conclude that neutrinos are in fact massive, and we are thus faced with the question of how to add neutrino masses to the Standard Model.

The simplest approach would be to introduce right-handed neutrinos, adding

$$\nu_{eR}, \nu_{\mu R}, \nu_{\tau R} \tag{1.9}$$

to Equation 1.3. We would then add the appropriate Higgs-to-neutrino Yukawa couplings to the Standard Model Lagrangian, and the neutrinos would acquire mass following spontaneous symmetry breaking in the Higgs field. The most general way to do this would be in a manner similar to that in the quark sector: assume that the fields of definite flavor are not identical to the fields of definite mass and write the mass term as a matrix product. Thus, if we define

$$\nu_{\ell L} = \begin{pmatrix} \nu_{eL} \\ \nu_{\mu L} \\ \nu_{\tau L} \end{pmatrix}, \nu_{\ell R} = \begin{pmatrix} \nu_{eR} \\ \nu_{\mu R} \\ \nu_{\tau R} \end{pmatrix}, \tag{1.10}$$

and let M be a 3×3 complex matrix, then the mass term in the Lagrangian will be

$$- \bar{\nu}_{\ell L} M \nu_{\ell R} + h.c. \tag{1.11}$$

We can diagonalize M using two 3×3 unitary matrices U and V :

$$M = U m V^\dagger, \tag{1.12}$$

where m is a 3×3 diagonal matrix whose nonzero elements are positive. We recognize these diagonal elements as the masses of the physical neutrino fields; denote these

fields as ν_1, ν_2, ν_3 . If we define

$$\nu_{iL} = \begin{pmatrix} \nu_{1L} \\ \nu_{2L} \\ \nu_{3L} \end{pmatrix}, \nu_{iR} = \begin{pmatrix} \nu_{1R} \\ \nu_{2R} \\ \nu_{3R} \end{pmatrix}, \quad (1.13)$$

substitute Equation 1.12 into Equation 1.11, and rearrange terms, we find that

$$\nu_{iL} = U^\dagger \nu_{\ell L}, \nu_{iR} = V^\dagger \nu_{\ell R}. \quad (1.14)$$

Inverting gives

$$\nu_{\ell L} = U \nu_{iL}, \nu_{\ell R} = V \nu_{iR}. \quad (1.15)$$

The unitary matrix U is known as the Pontecorvo–Maki–Nakagawa–Sakata (PMNS) mixing matrix and is the leptonic analogue of the CKM matrix. We will explore the PMNS matrix in much greater detail in Section 1.3.

The preceding derivation implicitly assumed that neutrinos are Dirac particles and are thus distinct from antineutrinos. This is obviously the case for the quarks and charged leptons, but as neutral particles, neutrinos could be Majorana particles – i.e., their own antiparticles. This possibility would introduce Majorana mass terms – distinct from the Dirac mass terms in Equation 1.11 – involving the charge conjugates of the left- and right-handed neutrino fields. The details of the derivation performed above would change [24], but the final result – that the fields of definite flavor are linear combinations of the physical fields – would remain. In Section 1.3, we will point out the minor change that must be made to the mixing matrix if neutrinos happen to be Majorana particles.

1.2 Getting to Massive Neutrinos

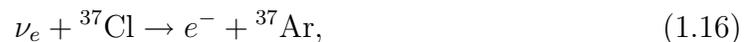
The discovery of mixing in the lepton sector due to nonzero neutrino masses and the study of the resulting phenomenon of neutrino oscillations have been the subjects of a

wide variety of experiments since the late 1960s. In this section, we provide a summary of many of these experiments and their key findings, organized into subsections by the sources of the neutrinos studied. We postpone discussion of these experiments' measurements of the oscillation model's parameters to Section 1.3.

1.2.1 Solar Neutrinos

The Standard Solar Model (SSM) describing the Sun's means of energy generation took form in the 1960s. In particular, the reactions hypothesized by the SSM to occur in the Sun's core produced, among other products, electron neutrinos (ν_e); Figure 1.1 shows the various neutrino spectra, as computed in a recent refinement of the SSM, with each spectrum labeled to identify the reaction that produces it. Since the neutrino interacts with matter via the weak interaction only – i.e., very rarely – neutrinos produced in the Sun's core would stream rapidly out of the Sun's interior, interacting rarely with the solar matter and arriving at Earth approximately eight minutes later,⁸ thereby providing a current picture of the reactions occurring in the Sun.

In 1964, John Bahcall and Raymond Davis proposed an experiment [27] to detect some of these neutrinos (primarily those from the ^8B spectrum) by radiochemical means, whereby a ν_e from the Sun would react with a chlorine-37 nucleus:



and the resulting radioactive argon atoms would be collected and counted. The experiment started running in 1968 in the Homestake mine in South Dakota, using 615 tons of perchloroethylene (a chlorine-rich dry-cleaning fluid) as the target medium. Initial results implied approximately one-third the expected number of ν_e s, and the “solar neutrino problem” was born. Continued data taking at Homestake through 1995 [28] showed an average neutrino capture rate consistent with the initial result, as shown

⁸Photons generated in the Sun's core, on the other hand, are constantly absorbed and reemitted, taking at least tens of thousands of years to exit the Sun.

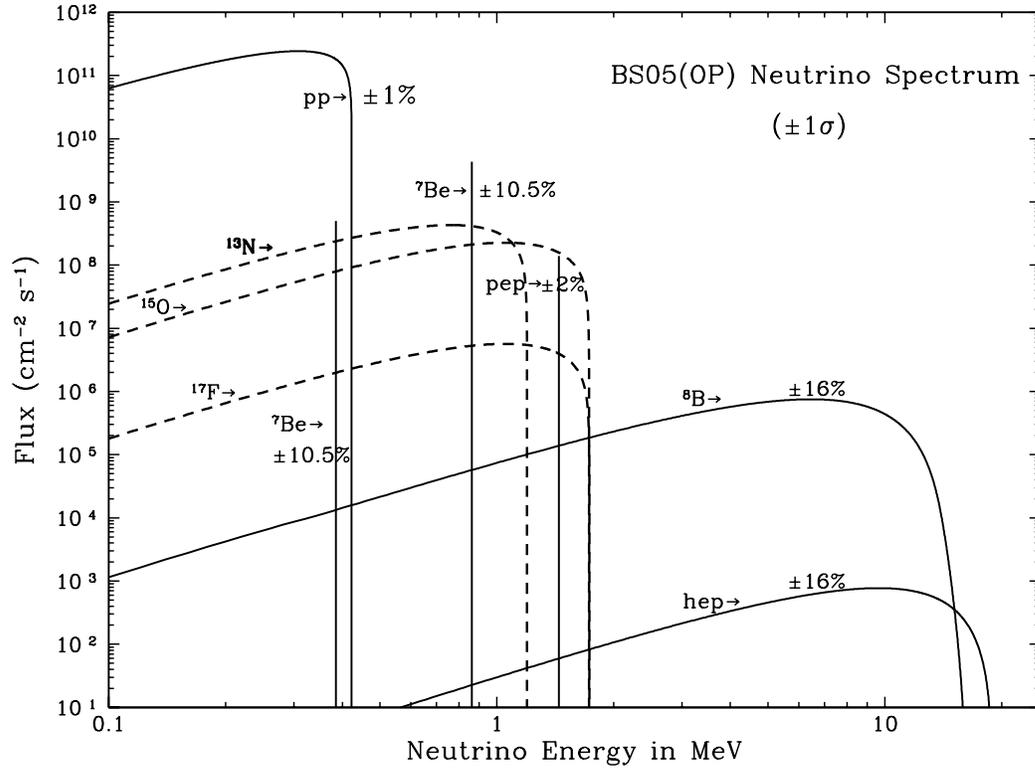


Figure 1.1: Solar neutrino energy spectrum as computed by the BS05(OP) solar model [25]. The reactions of interest – mostly because the neutrino fluxes they produce can be reliably measured by experiments – are pp ($p + p \rightarrow d + e^+ + \nu_e$), pep ($p + e^- + p \rightarrow d + \nu_e$), ${}^7\text{Be}$ ($e^- + {}^7\text{Be} \rightarrow {}^7\text{Li} + \nu_e$), ${}^8\text{B}$ (${}^8\text{B} \rightarrow {}^8\text{Be}^* + e^+ + \nu_e$), and parts of the CNO cycle (${}^{13}\text{N} \rightarrow {}^{13}\text{C} + e^+ + \nu_e$, ${}^{15}\text{O} \rightarrow {}^{15}\text{N} + e^+ + \nu_e$, and ${}^{17}\text{F} \rightarrow {}^{17}\text{O} + e^+ + \nu_e$). The units on the vertical axis are correct for line sources; for continuum sources, the units are understood to be $\text{cm}^{-2}\text{s}^{-1}\text{MeV}^{-1}$. The uncertainties on the spectra are from Table 8 of [26]. Figure taken from [25].

in Figure 1.2.

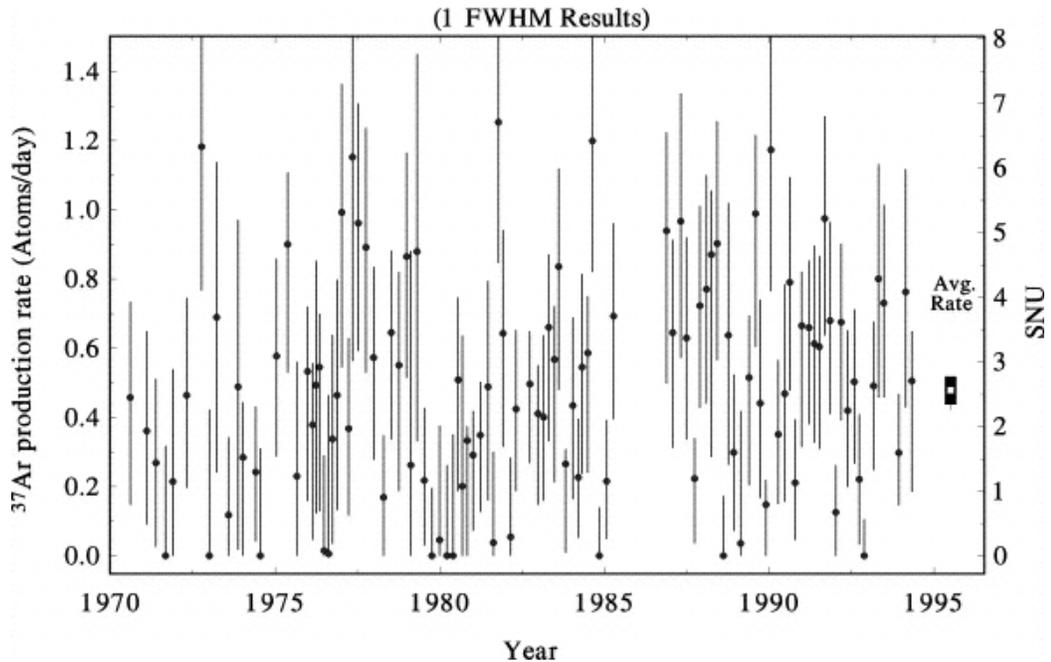


Figure 1.2: Observed daily rates of ^{37}Ar production for 108 individual solar neutrino observations made with the Homestake chlorine detector. The vertical axis on the right associates the ^{37}Ar production rate with the implied solar neutrino flux in Solar Neutrino Units (SNU), where 1 SNU is equivalent to 10^{-36} neutrino interactions per target atom per second. The production rates of ^{37}Ar shown have all backgrounds subtracted, and the errors shown for individual measurements are statistical only. The combined result is a solar neutrino-induced ^{37}Ar production rate of $2.56 \pm 0.16(\text{stat.}) \pm 0.16(\text{syst.})$ SNU, with a SSM-based expectation of 8.1 ± 1.2 SNU; the error shown on the combined result in the plot is statistical and systematic summed in quadrature. Taken from [29].

Other experiments – Kamiokande [30, 31], SAGE [32], GALLEX/GNO [33], and Super-Kamiokande [34] – utilized different detection methods with sensitivity to different parts of the solar neutrino spectrum and found varying deficits relative to SSM predictions (but deficits nonetheless). By the late 1990s, the solar neutrino problem was well established, and the most popular explanation – neutrino oscillations resulting from mixing of massive neutrinos – had been around even before the Homestake experiment [35, 36]. Specifically, it was thought that the SSM predictions were accurate, that neutrino mixing was responsible for transitions from ν_e to ν_μ and ν_τ , and that the deficit observed in every experiment thus far was a consequence of each experiment’s inability to detect the full flux of neutrinos from the Sun. The Sud-

bury Neutrino Observatory (SNO) experiment was designed to solve the problem in a model-independent way. In particular, the target medium in SNO was heavy water (D_2O), which enabled three modes of detection:

$$\nu_e + d \rightarrow e^- + p + p, \quad (1.17)$$

$$\nu + e^- \rightarrow \nu + e^-, \quad (1.18)$$

$$\nu + d \rightarrow \nu + p + n. \quad (1.19)$$

Equation 1.17 is a charged-current interaction, sensitive only to ν_e . Elastic neutrino–electron scattering (Equation 1.18) can occur via Z^0 exchange for all neutrino flavors as well as by W exchange for ν_e s only; as a result, this detection channel is sensitive to all neutrino flavors, but the cross section for ν_μ and ν_τ is approximately 16% that for ν_e [24]. Finally, the neutral current interaction in Equation 1.19 is equally sensitive to all neutrino flavors. SNO was thus able to measure the solar neutrino flux in three separate ways: ν_e only, ν_e plus 16% of ν_μ and ν_τ , and all three flavors equally. Their results [37], published in 2002, showed that while the solar ν_e flux is indeed less than expected, the *total* solar neutrino flux is well within the SSM prediction. These results are summarized in Figure 1.3 alongside the results of the other solar neutrino experiments discussed thus far.

An updated (2005) version of the SNO results is shown in Figure 1.4. The solar neutrino problem was thus conclusively solved. Furthermore, these results fit the neutrino oscillation model well, as we will see in Section 1.3.

1.2.2 Atmospheric Neutrinos

The Earth’s upper atmosphere is constantly bombarded by cosmic rays – mostly high-energy protons from outer space. These collisions produce primarily pions, which decay to produce what we refer to as “atmospheric” neutrinos (in a muon-to-electron

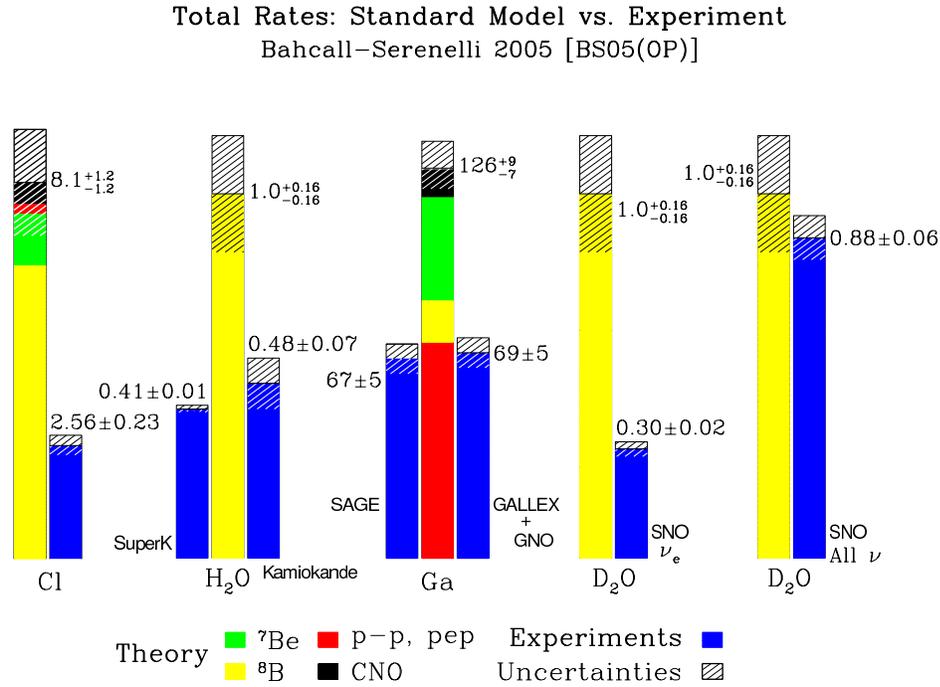


Figure 1.3: Comparison of SSM-predicted solar neutrino fluxes (yellow, green, red, and black bars) for different target media (left to right: chlorine, water, gallium, and heavy water) to the results of experiments using those target media (blue bars). All fluxes are given in SNU. Note the significant deficit relative to the theoretical prediction for the chlorine, water, and gallium experiments as well as for the SNO ν_e measurement (first four sets of bars, starting from the left), as contrasted with the excellent agreement for the SNO measurement of the total neutrino flux (rightmost set of bars). Taken from [38], with theoretical predictions updated to [25] and with experimental data including results reported up to April 20, 2005.

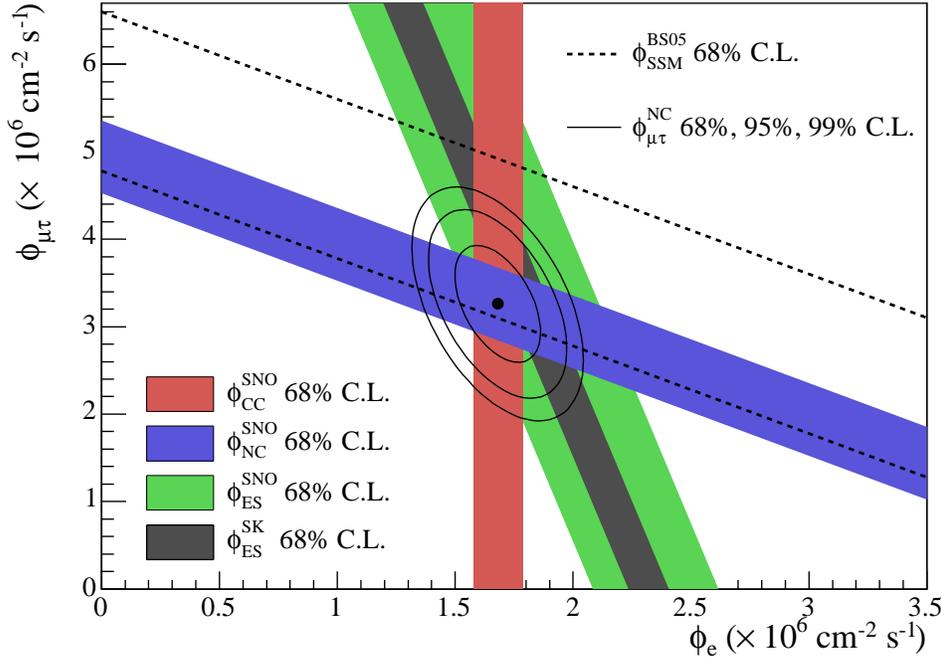


Figure 1.4: Results from the SNO experiment. Total ν_μ and ν_τ flux is plotted on the vertical axis, and ν_e flux is plotted on the horizontal axis. The red band represents the ν_e flux measurement made via the reaction in Equation 1.17; note its independence of the total ν_μ and ν_τ flux. The blue band represents the total neutrino flux measurement made by observing the neutral-current reaction (Equation 1.19) rate; note its slope, reflecting the equal sensitivity of this detection mode to all neutrino flavors. The total neutrino flux measurement made by observing the elastic neutrino–electron scattering (Equation 1.18) rate is shown in green; the result from the Super-Kamiokande experiment for the same detection mode is shown in black. Note the sharper slope of these bands compared to that of the blue band, reflecting the smaller cross section for elastic scattering of ν_μ and ν_τ relative to that of ν_e . The dashed lines indicate the 1σ SSM prediction of the total solar neutrino flux – in good agreement with the SNO neutral-current measurement. Finally, the solid lines are the 1σ , 2σ , and 3σ confidence level contours for the common intersection of all three bands. In particular, the fact that this intersection occurs at a value of total ν_μ and ν_τ flux that is significantly above zero can be viewed as the very solution to the solar neutrino problem. Taken from [39].

flavor ratio of 2-to-1):

$$\pi^+ \rightarrow \mu^+ + \nu_\mu, \quad \mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu, \quad (1.20)$$

$$\pi^- \rightarrow \mu^- + \bar{\nu}_\mu, \quad \mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu. \quad (1.21)$$

This ratio varies with neutrino energy, as higher-energy cosmic ray collisions also produce kaons, for which the relevant decays are

$$K^+ \rightarrow \mu^+ + \nu_\mu, \quad (1.22)$$

$$K^- \rightarrow \mu^- + \bar{\nu}_\mu, \quad (1.23)$$

$$K_L^0 \rightarrow \pi^+ + e^- + \bar{\nu}_e, \quad (1.24)$$

$$K_L^0 \rightarrow \pi^- + e^+ + \nu_e. \quad (1.25)$$

A variety of Monte Carlo simulations predict the flux of muon-flavor ($\nu_\mu/\bar{\nu}_\mu$) and electron-flavor ($\nu_e/\bar{\nu}_e$) atmospheric neutrinos, but these absolute fluxes have uncertainties on the order of 20%–30% [11]. The *ratio* of these fluxes, on the other hand, can be predicted with an accuracy of $\pm 5\%$ as a result of cancellations between uncertainties in the absolute fluxes. In an appropriately designed experiment, one could count the number $N(\mu)$ of $\nu_\mu/\bar{\nu}_\mu$ -induced charged-current events⁹ and the number $N(e)$ of $\nu_e/\bar{\nu}_e$ -induced charged current events:

$$\nu_\mu/\bar{\nu}_\mu + N \rightarrow \mu^-/\mu^+ + X, \quad (1.26)$$

$$\nu_e/\bar{\nu}_e + N \rightarrow e^-/e^+ + X, \quad (1.27)$$

where N represents a nucleus in the target medium and X is the hadronic recoil system, and compute their ratio. One could then use the aforementioned fluxes to accurately simulate the value of this ratio for the conditions of this experiment. A

⁹This is possible because atmospheric ν_μ s, unlike solar neutrinos, have sufficiently high energies to produce muons in collisions.

significant deviation of R , the ratio of these ratios:

$$R = \frac{[N(\mu)/N(e)]_{\text{observed}}}{[N(\mu)/N(e)]_{\text{expected}}} \quad (1.28)$$

from unity would imply new physics [40]. This is precisely what was observed in the Kamiokande [41] experiment using a “water Cherenkov” detector – a tank containing 2,140 tons of ultrapure water viewed by photomultiplier tubes that collect the Cherenkov radiation produced by muons and electrons traveling through the water. In particular, they found $R = 0.57^{+0.08}_{-0.07}(\text{syst.}) \pm 0.07(\text{stat.})$. Kamiokande ran from 1987 until 1995, at which time it was upgraded to Super-Kamiokande (Super-K), a 50-kiloton water Cherenkov detector. Super-K expanded on the work performed in Kamiokande by collecting data on the zenith angle distribution of ν_μ and ν_e events – in effect, looking for variation in R with zenith angle θ . That is, Super-K was able to discriminate among events based on the direction in which the neutrino producing the event was traveling: $\cos \theta = 1$ for neutrinos incident directly from above, $\cos \theta = -1$ for those incident directly from below (i.e., produced on the opposite side of Earth). In 1998, they reported [42] that the ν_μ flux appeared to decline with decreasing $\cos \theta$, i.e., with increasing distance traveled from production to detection. This observed “atmospheric neutrino anomaly” provided model-independent evidence for neutrino (specifically, ν_μ) oscillations [24] – neutrino survival probability dependent on distance traveled. These results are shown in Figure 1.5. Note that a similar effect is not seen for ν_e s, as ν_e oscillations are governed by the oscillation model parameters in such a way that they do not occur appreciably at these length scales and neutrino energies.

In Section 1.3, we will see that the neutrino survival probability oscillates as a function of L/E – distance traveled by the neutrino divided by its energy. In an analysis reported in 2004 [43], Super-K found that their data clearly displayed this behavior, as shown in Figure 1.6.

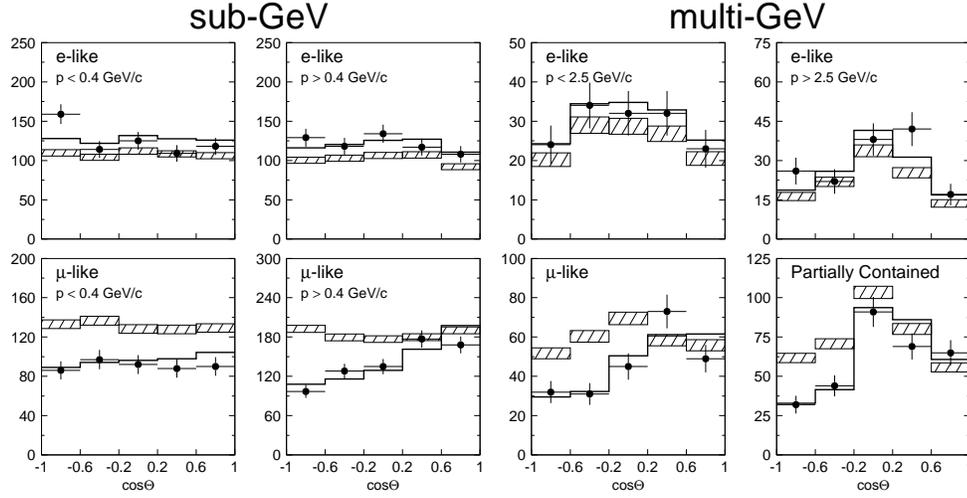


Figure 1.5: Distributions of zenith angle for μ -like (ν_μ CC) and e -like (ν_e CC) events in Super-K, divided into sub-GeV (events with visible energy less than 1.3 GeV) and multi-GeV (events with visible energy greater than 1.3 GeV) data sets. Upward-going particles have $\cos\theta < 0$, and downward-going particles have $\cos\theta > 0$. Sub-GeV data are shown separately for $p < 0.4$ GeV/c and $p > 0.4$ GeV/c. Multi-GeV e -like distributions are shown for $p < 2.5$ GeV/c and $p > 2.5$ GeV/c, and the multi-GeV μ -like are shown separately for FC and PC events (events fully contained and partially contained, respectively, within the fiducial volume of the detector). The hatched regions show the Monte Carlo expectations for no oscillations, normalized to the data live-time with statistical errors. The bold lines are the best-fit expectations under the oscillation model (in particular, for ν_μ to ν_τ oscillations) with the overall normalization fitted as a free parameter. Taken from [42].

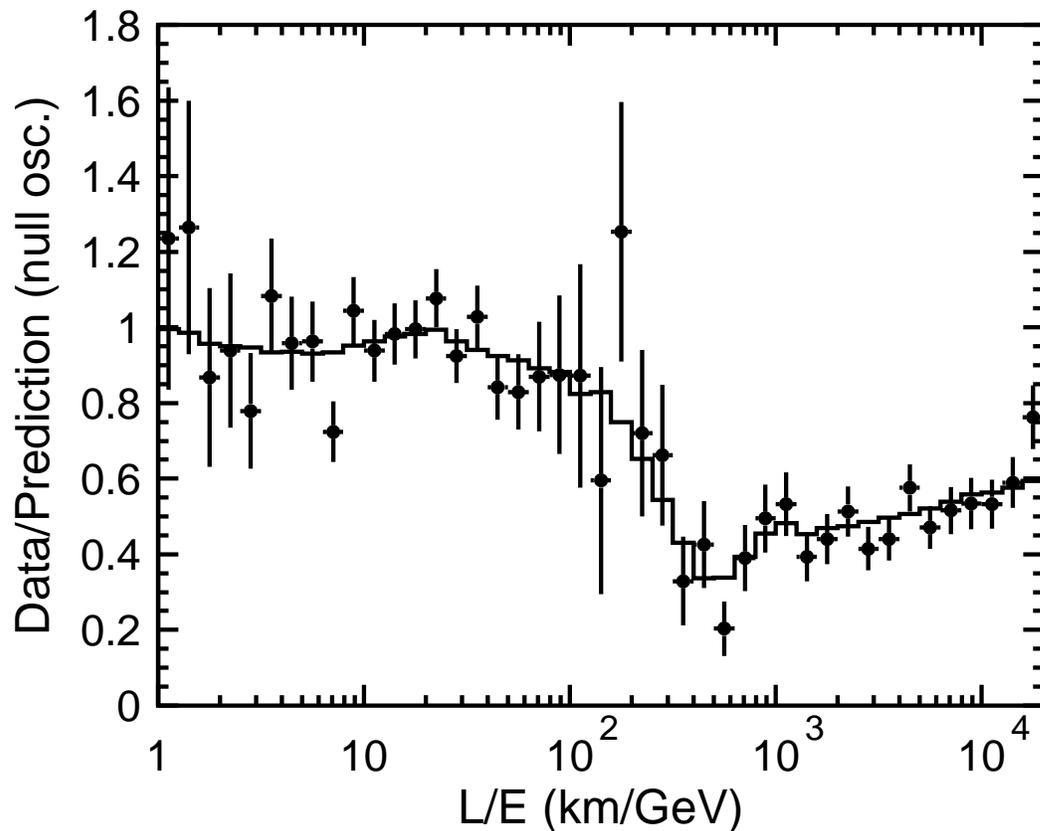


Figure 1.6: Ratio of the number of μ -like (ν_μ CC) events observed in Super-K to the number of events (as determined by the Monte Carlo) expected without neutrino oscillations as a function of L/E (points). Error bars are statistical only. The solid curve is the best-fit expectation under the oscillation model (in particular, for ν_μ to ν_τ oscillations). Adapted from [43].

1.2.3 Accelerator Neutrinos

Colliding a beam of high-energy protons from a particle accelerator with a suitable target material produces a large flux of pions as well as some kaons. By focusing the positively charged secondary particles, allowing them to decay in flight, and absorbing any remaining charged particles, one obtains a ν_μ beam with a small $\bar{\nu}_\mu$ contamination and an even smaller $\nu_e/\bar{\nu}_e$ contamination. These are referred to as “accelerator neutrinos” and are akin to atmospheric neutrinos due to the similarities in the underlying production mechanism and the resulting neutrino flavor compositions.

Two key accelerator neutrino experiments designed to study the atmospheric neutrino anomaly are K2K and MINOS. In K2K, the neutrino beam is generated at KEK in Japan by colliding 12-GeV protons with an aluminum target. The beam passes through a detector located 300 m downstream of the target and then continues on to the Super-Kamiokande detector 250 km away. MINOS is similarly designed: the neutrino beam is generated at the Fermi National Accelerator Laboratory (Fermilab) in Batavia, Illinois, by colliding 120-GeV protons with a graphite target. The beam passes through a detector located 1 km downstream of the target and then continues on to a second detector located 735 km downstream in the Soudan mine in northern Minnesota.¹⁰ Thus, neutrinos produced in the same location travel the same distance between two detectors; in atmospheric neutrino experiments, by contrast, neutrinos produced in different locations travel different distances to a single detector. Nonetheless, both types of experiments involve the same sort of observation – change in ν_μ flux with distance traveled. In Section 1.3, we will see that accelerator neutrino and atmospheric neutrino experiments complement each other in the precision they provide in measurements of the relevant underlying oscillation parameters.

In its final result [44], K2K expected to observe $158.1_{-8.6}^{+9.2}$ ν_μ events at Super-K based on the ν_μ flux measured at the near detector but observed only 112 events. The deficit, visible in Figure 1.7, was found to be energy dependent in a manner consistent with neutrino oscillations. As shown in Figure 1.8, the most recent results

¹⁰We will, of course, discuss the MINOS experiment in much greater detail shortly.

from MINOS [45] indicate a deficit of a similar nature. T2K in Japan and NO ν A based at Fermilab are the next-generation accelerator experiments sensitive to this phenomenon and are discussed in Chapter 9.

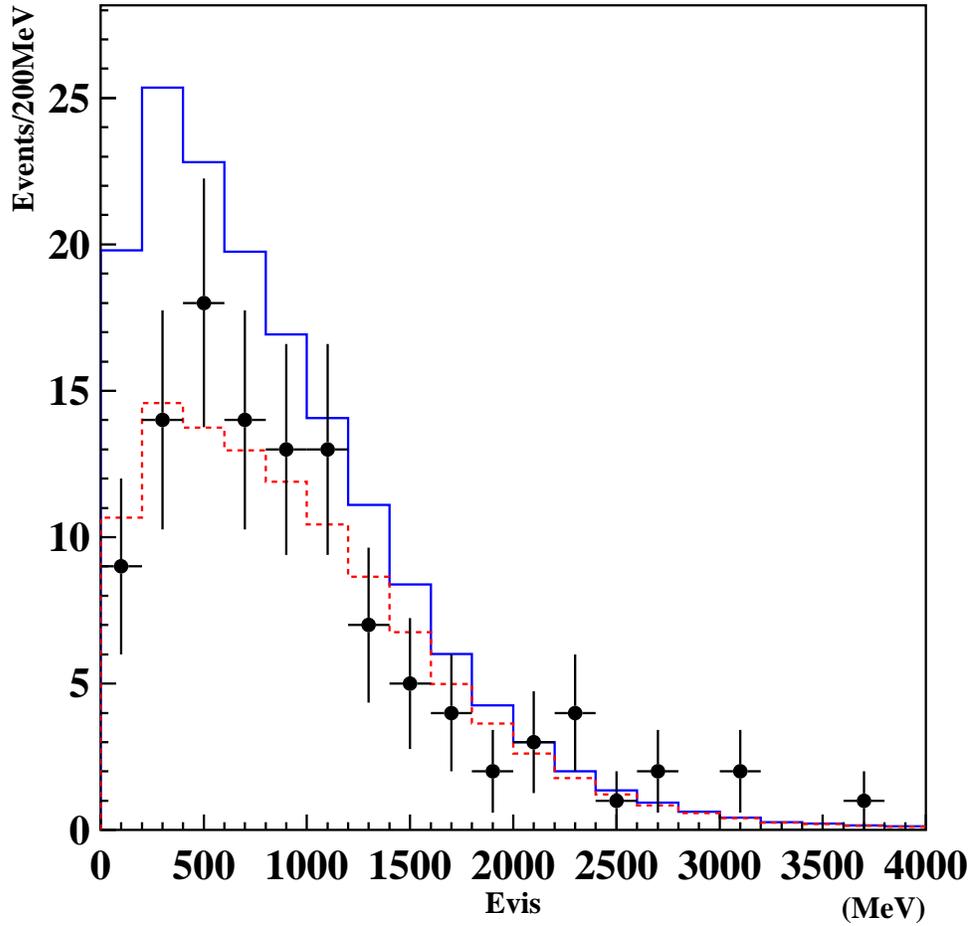


Figure 1.7: K2K visible energy distribution for fully contained μ -like fiducial volume events in Super-K, which serves as K2K’s far detector. The points are the observed data. The solid blue histogram is the Monte Carlo expectation based on the near detector measurement assuming no neutrino oscillation; the dashed red histogram is the best-fit expectation with neutrino oscillations. Taken from [44].

1.2.4 Reactor Neutrinos

As we discussed in the introduction to this chapter, nuclear reactors are a source of antineutrinos – specifically, $\bar{\nu}_e$ s from the beta decays of unstable neutron-rich fission

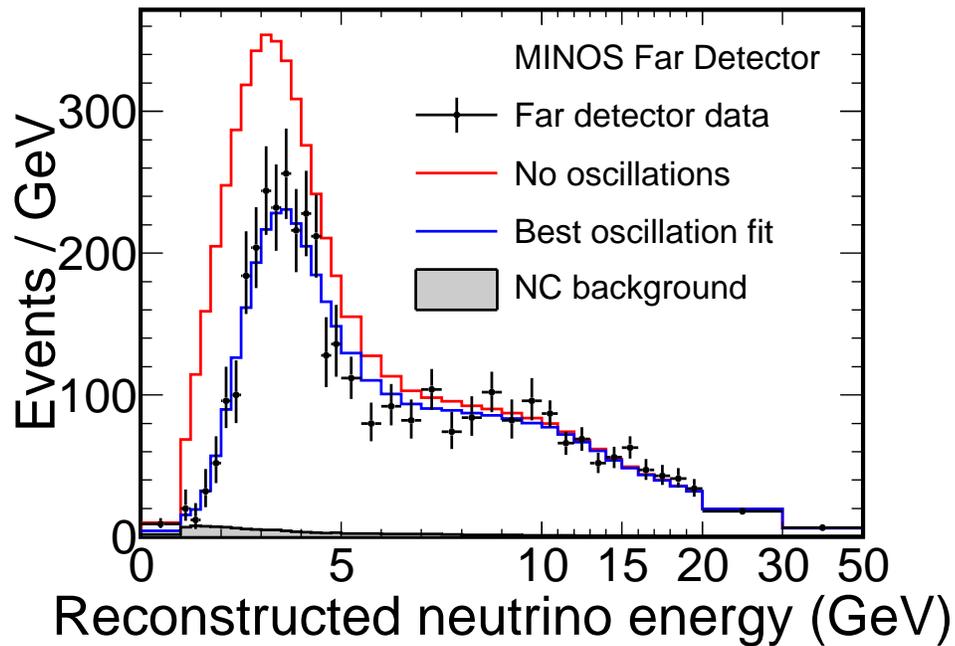


Figure 1.8: Energy spectra of fully reconstructed events in the MINOS Far Detector classified as ν_μ charged-current interactions. The red histogram represents the spectrum predicted from measurements in the Near Detector assuming no oscillations; the blue histogram represents the best fit under the oscillation model. The shaded area shows the predicted background from neutral-current events. Taken from [45].

fragments. The energy spectrum for these $\bar{\nu}_e$ s peaks at 2–3 MeV and extends up to about 8 MeV. Thus, if a $\bar{\nu}_e$ happens to oscillate to a different flavor, it would be impossible to detect in a charged-current interaction, as its energy is far below the threshold for μ or τ production. Thus, so-called reactor experiments always look for $\bar{\nu}_e$ disappearance; that is, the expected $\bar{\nu}_e$ flux from the reactor at the detector’s location is computed and compared to the measured flux. The mode of detection used most often is inverse beta decay:

$$\bar{\nu}_e + p \rightarrow e^+ + n. \quad (1.29)$$

In particular, the positron promptly annihilates with a nearby electron, producing “prompt” gamma radiation. The neutron diffuses and thermalizes within 10–100 μ s; it is then captured by a nucleus, which emits “delayed” gamma radiation. The coincidence between these two pulses is used to identify a $\bar{\nu}_e$ interaction. Furthermore, if the small neutron recoil energy is neglected, the $\bar{\nu}_e$ ’s energy is readily obtained from the positron’s measured energy.

A variety of small (less than 0.5-ton fiducial mass) reactor experiments located no farther than 250 m from any individual reactor have found no evidence of $\bar{\nu}_e$ disappearance; these are ILL through Krasnoyarsk in Figure 1.9. More recently (late 1990s), motivated by the possible involvement of ν_e or $\bar{\nu}_e$ in the atmospheric neutrino anomaly, the CHOOZ and Palo Verde experiments were constructed in France and Arizona, respectively, about 1 km away from the reactors being observed. The reason for the increased distance was to have a value of L/E comparable to that of the atmospheric experiments; as we have previously noted, the neutrino survival probability oscillates as a function of this quantity. Neither of these experiments found any evidence for oscillation [46, 47]; indeed, both found the energy-averaged ratio of observed to expected numbers of events to be unity within errors, suggesting that any involvement of ν_e or $\bar{\nu}_e$ in the atmospheric neutrino anomaly must be very small. The next generation of reactor neutrino experiments investigating that involvement include Daya Bay in China, RENO in Korea, and Double Chooz in France; these

experiments and their observations will be discussed in Chapter 9. In Section 1.3, we will see that the observations of such experiments are intimately associated with the observations analyzed in this thesis.

In Section 1.3 we will see that the oscillations experienced by solar neutrinos are associated with a larger value of L/E . The KamLAND reactor experiment, located in the Kamioka mine in Japan (also home to Super-Kamiokande), was built to have such a value. A total of 55 commercial nuclear power plants in Japan and South Korea deliver 155 GW of power, corresponding to a $\bar{\nu}_e$ flux of $4 \times 10^6 \text{ cm}^{-2}\text{s}^{-1}$. Slightly less than half (70 GW) of that power is produced by reactors located $175 \pm 35 \text{ km}$ from KamLAND, a distance two orders of magnitude greater than the distance between CHOOZ or Palo Verde and their associated reactors. As indicated in Figure 1.9, KamLAND obtained an observed-to-expected ratio of $0.593 \pm 0.020(\text{stat.}) \pm 0.026(\text{syst.})$ – a significant deficit consistent with oscillation model expectations.

1.3 Describing Massive Neutrinos

In Section 1.1.2, we showed that the neutrino fields of definite flavor will be linear combinations of the physical neutrino fields (Equation 1.15). Furthermore, the transformation from the latter to the former is given by U , the unitary PMNS mixing matrix. We now proceed to develop the formalism of neutrino mixing and of the phenomenon of neutrino oscillations that arises from it. We will then explore oscillations in greater depth, including our current state of knowledge regarding the parameters that govern them.

1.3.1 Formalism of Neutrino Mixing

Let us begin by assuming that there are n neutrino fields, which can be written either as n physical neutrino fields ν_i or as n neutrino fields ν_α of definite flavor. These sets

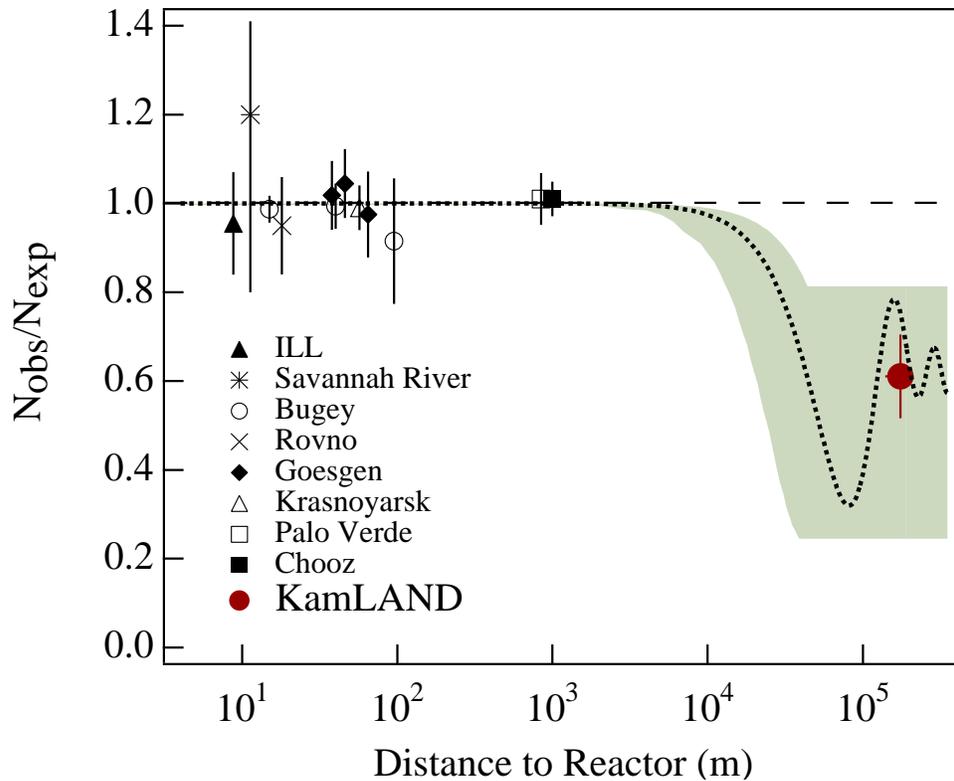


Figure 1.9: The ratio of measured to expected $\bar{\nu}_e$ flux versus distance from reactor for various reactor experiments. The solid red dot is the KamLAND result, plotted at a flux-weighted average distance. The dotted curve is the best fit to global solar neutrino data under the oscillation model; the dashed curve is the no-oscillation case (observation equal to expectation). Note that CHOOZ and Palo Verde were too close to their respective reactors to observe the “solar oscillation” seen in KamLAND; we will see in Section 1.3, however, that they were in the right range to set limits on $\bar{\nu}_e$ disappearance due to “atmospheric oscillation.” Taken from [48].

of fields are related via a unitary mixing matrix U :

$$\nu_\alpha = \sum_i U_{\alpha i} \nu_i. \quad (1.30)$$

The weak eigenstate $|\nu_\alpha\rangle$, then, is given in terms of the mass eigenstates $|\nu_i\rangle$ by

$$|\nu_\alpha\rangle = \sum_i U_{\alpha i}^* |\nu_i\rangle, \quad (1.31)$$

where the complex conjugation of $U_{\alpha i}$ is due to the creation operator contained in ν_α^\dagger that creates $|\nu_\alpha\rangle$ from the vacuum state $|0\rangle$. One can show that the each set of eigenstates forms an orthonormal eigenbasis, so $\langle\nu_\beta|\nu_\alpha\rangle = \delta_{\alpha\beta}$ and $\langle\nu_j|\nu_i\rangle = \delta_{ij}$.

1.3.1.1 General Probability of Flavor Transition

Note that the mass eigenstates are, by definition, stationary states, so they have the usual time dependence given by the Schrödinger equation. Furthermore, we will assume that the physical neutrino can be approximated as a plane wave.¹¹ Thus, we have

$$\begin{aligned} |\nu_i(x, t)\rangle &= e^{-iE_i t} |\nu_i(x, 0)\rangle \\ &= e^{-iE_i t} e^{ip_i x} |\nu_i(0, 0)\rangle \\ &= e^{i(p_i x - E_i t)} |\nu_i\rangle, \end{aligned} \quad (1.32)$$

where we have taken the plane wave to be traveling in the x -direction without loss of generality and have written $|\nu_i\rangle$ instead of $|\nu_i(0, 0)\rangle$. We now combine Equations 1.31 and 1.32 to obtain the space and time dependence of the flavor eigenstate:

$$\begin{aligned} |\nu_\alpha(x, t)\rangle &= \sum_i U_{\alpha i}^* |\nu_i(x, t)\rangle \\ &= \sum_i U_{\alpha i}^* e^{i(p_i x - E_i t)} |\nu_i\rangle. \end{aligned} \quad (1.33)$$

¹¹A proper treatment would use a wave packet description of the neutrino; see the classic paper by Boris Kayser [49]. The simplified plane wave treatment presented here, however, provides the same results.

Note that the process of interest to us is the transition $\nu_\alpha \rightarrow \nu_\beta$ – i.e., a neutrino of flavor α is produced and after traveling a distance x in time t is detected as a neutrino of flavor β . The amplitude for this transition is

$$\begin{aligned}
A(\nu_\alpha \rightarrow \nu_\beta) &= \langle \nu_\beta(x, t) | \nu_\alpha(0, 0) \rangle \\
&= \left(\sum_j \langle \nu_j | U_{\beta j} e^{i(p_j x - E_j t)} \right) \left(\sum_i U_{\alpha i}^* | \nu_i \rangle \right) \\
&= \sum_{i,j} U_{\beta j} U_{\alpha i}^* e^{i(p_j x - E_j t)} \langle \nu_j | \nu_i \rangle \\
&= \sum_i U_{\beta i} U_{\alpha i}^* e^{i(p_i x - E_i t)},
\end{aligned} \tag{1.34}$$

where we have used $\langle \nu_j | \nu_i \rangle = \delta_{ij}$ to perform the sum over j . The probability of a transition is given by the squared magnitude of its amplitude, so:

$$\begin{aligned}
P(\nu_\alpha \rightarrow \nu_\beta) &= |A(\nu_\alpha \rightarrow \nu_\beta)|^2 \\
&= \left(\sum_j U_{\beta j}^* U_{\alpha j} e^{-i(p_j x - E_j t)} \right) \left(\sum_i U_{\beta i} U_{\alpha i}^* e^{i(p_i x - E_i t)} \right) \\
&= \sum_{i,j} U_{\beta j}^* U_{\alpha j} U_{\beta i} U_{\alpha i}^* e^{i[(p_i - p_j)x - (E_i - E_j)t]}.
\end{aligned} \tag{1.35}$$

At this point, we note that neutrinos are highly relativistic ($p_i \gg m_i$) and make the following approximations:

$$E_i = \sqrt{p_i^2 + m_i^2} \approx p_i + \frac{m_i^2}{2p_i}, \tag{1.36}$$

$$x \approx t, \tag{1.37}$$

$$p_i \approx p_j = p \approx E. \tag{1.38}$$

Substituting these approximations into Equation 1.35 and taking $x = L$ to be the distance traveled by the neutrino between production and detection, we obtain

$$P(\nu_\alpha \rightarrow \nu_\beta) = \sum_{i,j} U_{\beta j}^* U_{\alpha j} U_{\beta i} U_{\alpha i}^* e^{i \frac{\Delta m_{ji}^2 L}{2E}}, \tag{1.39}$$

where $\Delta m_{ji}^2 \equiv m_j^2 - m_i^2$. We now perform a series of algebraic manipulations on Equation 1.39, taking advantage of complex number identities, the unitarity of U , and trigonometric identities:

$$\begin{aligned}
P(\nu_\alpha \rightarrow \nu_\beta) &= \sum_i |U_{\alpha i}|^2 |U_{\beta i}|^2 + 2\Re \sum_{j>i} U_{\beta j}^* U_{\alpha j} U_{\beta i} U_{\alpha i}^* e^{i \frac{\Delta m_{ji}^2 L}{2E}} \\
&= \delta_{\alpha\beta} - 2\Re \sum_{j>i} U_{\beta j}^* U_{\alpha j} U_{\beta i} U_{\alpha i}^* + 2\Re \sum_{j>i} U_{\beta j}^* U_{\alpha j} U_{\beta i} U_{\alpha i}^* e^{i \frac{\Delta m_{ji}^2 L}{2E}} \\
&= \delta_{\alpha\beta} - 2\Re \sum_{j>i} U_{\beta j}^* U_{\alpha j} U_{\beta i} U_{\alpha i}^* \left(1 - e^{i \frac{\Delta m_{ji}^2 L}{2E}}\right) \\
&= \delta_{\alpha\beta} - 2 \sum_{j>i} \Re (U_{\beta j}^* U_{\alpha j} U_{\beta i} U_{\alpha i}^*) \left(1 - \cos \frac{\Delta m_{ji}^2 L}{2E}\right) \\
&\quad - 2 \sum_{j>i} \Im (U_{\beta j}^* U_{\alpha j} U_{\beta i} U_{\alpha i}^*) \sin \frac{\Delta m_{ji}^2 L}{2E} \\
&= \delta_{\alpha\beta} - 4 \sum_{j>i} \Re (U_{\beta j}^* U_{\alpha j} U_{\beta i} U_{\alpha i}^*) \sin^2 \frac{\Delta m_{ji}^2 L}{4E} \\
&\quad - 2 \sum_{j>i} \Im (U_{\beta j}^* U_{\alpha j} U_{\beta i} U_{\alpha i}^*) \sin \frac{\Delta m_{ji}^2 L}{2E}.
\end{aligned} \tag{1.40}$$

We see that in order to have transitions between neutrino flavors, neutrinos must not only be massive but must also have *different* masses. Otherwise, the sine terms in Equation 1.40 (for which these transitions are dubbed ‘‘oscillations’’) vanish, and no flavor change is observed.

The transition probability $P(\bar{\nu}_\alpha \rightarrow \bar{\nu}_\beta)$ for antineutrinos can be obtained by exploiting the CP-conjugate of Equation 1.31:

$$|\bar{\nu}_\alpha\rangle = \sum_i U_{\alpha i} |\bar{\nu}_i\rangle, \tag{1.41}$$

and noting that since CPT invariance requires $|\bar{\nu}_i\rangle$ to have the same mass as $|\nu_i\rangle$, $\Delta \bar{m}_{ji}^2 = \Delta m_{ji}^2$. Thus, all that is necessary is to replace the U matrix elements in Equation 1.40 with their complex conjugates. This would leave the first two terms

unchanged and would change the sign on the last term, so we can write

$$\begin{aligned}
 P(\bar{\nu}_\alpha \rightarrow \bar{\nu}_\beta) &= \delta_{\alpha\beta} - 4 \sum_{j>i} \Re(U_{\beta j}^* U_{\alpha j} U_{\beta i} U_{\alpha i}^*) \sin^2 \frac{\Delta m_{ji}^2 L}{4E} \\
 &\quad + 2 \sum_{j>i} \Im(U_{\beta j}^* U_{\alpha j} U_{\beta i} U_{\alpha i}^*) \sin \frac{\Delta m_{ji}^2 L}{2E}.
 \end{aligned} \tag{1.42}$$

Note that, in general, $P(\nu_\alpha \rightarrow \nu_\beta) \neq P(\bar{\nu}_\alpha \rightarrow \bar{\nu}_\beta)$, so neutrino oscillations are, in general, CP-violating. However, if the mixing matrix is real ($U = U^*$), the last term in Equations 1.40 and 1.42 vanishes, $P(\nu_\alpha \rightarrow \nu_\beta) = P(\bar{\nu}_\alpha \rightarrow \bar{\nu}_\beta)$, and CP is conserved. We will address leptonic CP violation in Section 1.3.3.2.

1.3.1.2 The Mixing Matrix

One can show that an $n \times n$ unitary matrix can be parametrized by $\frac{n(n-1)}{2}$ angles and $\frac{n(n+1)}{2}$ complex phases [24]. Not all of these phases are physical, however, as the mixing matrix enters into the charged current Lagrangian together with fields of charged leptons and neutrinos. In the case of Dirac neutrinos, there are a total of $2n - 1$ relative phases among the n neutrino fields and n charged lepton fields. These phases can be chosen in such a way as to make $2n - 1$ of the $\frac{n(n+1)}{2}$ phases in U vanish. Of course, the total Lagrangian (free Lagrangian of leptons plus the Lagrangians of the electromagnetic and neutral current interactions) is invariant under such phase transformations. Thus, the number of physical phases in the case of Dirac neutrinos is given by $\frac{n(n+1)}{2} - (2n - 1) = \frac{(n-1)(n-2)}{2}$. For Majorana neutrinos, however, the phases of the neutrino fields are fixed by the so-called Majorana condition, and only the n phases associated with the charged lepton fields can be chosen in such a way as to make n phases in U vanish. This leaves $\frac{n(n+1)}{2} - n = \frac{n(n-1)}{2}$ physical phases in the case of Majorana neutrinos – $n - 1$ more than in the Dirac case. The presence of these complex phases guarantees that U is not real and, thus, that CP is violated. Accordingly, they are referred to as CP-violating phases; in particular, the $\frac{(n-1)(n-2)}{2}$ phases from the case of Dirac neutrinos are referred to as the Dirac CP-violating phases, while the additional $n - 1$ phases from the case of Majorana neutrinos are

referred to as the Majorana CP-violating phases.

We have already noted that there are neutrinos of three different flavors in the Standard Model, so there must be three physical neutrinos. In the 1990s, the ALEPH, DELPHI, L3, and OPAL experiments at the Large Electron–Positron (LEP) collider at CERN confirmed this after observing approximately 17 million decays of the Z^0 boson and finding the number of light neutrinos (specifically, the number of neutrinos that couple to the Z^0 boson) to be 2.9840 ± 0.0082 [50]. Thus, we will need to write U for the case $n = 3$, but we first consider the instructive and relatively simpler case of two neutrinos.

For $n = 2$, U is parametrized by one mixing angle (θ) and one Majorana CP-violating phase (α); there are no Dirac CP-violating phases in the $n = 2$ case. We write U as the product of a rotation matrix and a diagonal matrix containing the Majorana phase:

$$U = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & 1 \end{pmatrix}. \quad (1.43)$$

In the $n = 3$ case, we have three mixing angles (θ_{12} , θ_{23} , θ_{13}), one Dirac CP-violating phase (δ_{CP}), and two Majorana CP-violating phases (α_1 , α_2). We write U as the product of three Euler rotation matrices and a diagonal matrix containing the Majorana phases. In particular, the three rotations are as follows: about $|\nu_3\rangle$ by θ_{12} , about the rotated $|\nu_2\rangle$ by θ_{13} , and about the doubly rotated $|\nu_1\rangle$ by θ_{23} . We include δ_{CP} in the rotation by θ_{13} . Defining $c_{ij} \equiv \cos \theta_{ij}$ and $s_{ij} \equiv \sin \theta_{ij}$, we have

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{\text{CP}}} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta_{\text{CP}}} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} e^{i\alpha_1} & 0 & 0 \\ 0 & e^{i\alpha_2} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (1.44)$$

Before we proceed, we note that if one multiplies out the matrices in Equation 1.43 or 1.44, each entry in a given column will be multiplied by the same Majorana phase factor. Thus, in every term of the sum in Equation 1.34, a given Majorana phase factor

and its complex conjugate will cancel, and the total amplitude will be independent of the Majorana phases. Of course, the transition probability computed from this amplitude will also be independent of the Majorana phases, and we conclude that these Majorana phases – and thus, the Dirac or Majorana nature of the massive neutrinos – cannot be determined by an oscillation experiment. Consequently, we omit the matrices containing the Majorana phases in Equations 1.43 and 1.44 from this point forward.

1.3.2 Two-Neutrino Oscillations

We now explore two-neutrino oscillations in some detail and proceed to discuss the observations made by various experiments as well as the measurements of oscillation parameters made based on these observations. We will also consider the effect of the presence of matter on oscillation probabilities, as the currently accepted interpretation of solar neutrino oscillations relies on it.

1.3.2.1 The Two-Neutrino Framework

Substituting Equation 1.43 into Equations 1.40 and 1.42, we find

$$P(\nu_\alpha \rightarrow \nu_\alpha) = P(\nu_\beta \rightarrow \nu_\beta) = P(\bar{\nu}_\alpha \rightarrow \bar{\nu}_\alpha) = P(\bar{\nu}_\beta \rightarrow \bar{\nu}_\beta) = 1 - \sin^2 2\theta \sin^2 \frac{\Delta m^2 L}{4E},$$
(1.45)

$$P(\nu_\alpha \rightarrow \nu_\beta) = P(\nu_\beta \rightarrow \nu_\alpha) = P(\bar{\nu}_\alpha \rightarrow \bar{\nu}_\beta) = P(\bar{\nu}_\beta \rightarrow \bar{\nu}_\alpha) = \sin^2 2\theta \sin^2 \frac{\Delta m^2 L}{4E}.$$
(1.46)

Note that the probabilities of “survival” (Equation 1.45) and of “appearance” (Equation 1.46) sum to unity, as expected. Furthermore, since there is no Dirac CP-violating phase in the two-neutrino case, the antineutrino transition probabilities are identical to those for neutrinos.¹² Under CPT invariance, CP invariance implies T invariance,

¹²This implies that any CP violation in the lepton sector cannot be explained in a two-neutrino framework. Rather, a full three-neutrino analysis is warranted.

which is evident in the flavor symmetry of Equation 1.46.

Since there are only two mass eigenstates, there is only one mass-squared difference Δm^2 . The frequency of the oscillation is determined by Δm^2 , while the amplitude is determined by the mixing angle θ . In particular, it is helpful to see how the “baseline” L and typical neutrino energy E of an experiment and the mass-squared difference Δm^2 governing the particular oscillation being investigated come together. The oscillatory behavior of the survival and appearance probabilities becomes prominent when

$$\frac{\Delta m^2 L}{4E} = 1.267 \frac{(\Delta m^2/\text{eV}^2)(L/\text{km})}{E/\text{GeV}} \sim 1. \quad (1.47)$$

When this quantity is much larger than unity, the oscillatory term can effectively be replaced by $1/2$, and the oscillatory behavior is lost. By rearranging this approximation, we can obtain insight into the design of neutrino oscillation experiments. For example, an experiment with a given value of L/E will be sensitive to oscillations driven by

$$\Delta m^2 \sim \frac{1}{1.267 \frac{L/\text{km}}{E/\text{GeV}}} \text{eV}^2. \quad (1.48)$$

Conversely, if an experiment is to probe oscillations driven by a certain Δm^2 , its baseline and average neutrino energy should satisfy

$$\frac{L}{E} \sim \frac{1}{1.267 \Delta m^2/\text{eV}^2} \frac{\text{km}}{\text{GeV}}. \quad (1.49)$$

The two-neutrino framework thus provides a simple yet insightful introduction to understanding neutrino oscillations. In fact, as we will now see, many neutrino experiments have historically used a two-neutrino approximation to analyze their data and measure the values of various oscillation parameters.

1.3.2.2 Oscillation Probabilities and Parameter Measurements

Atmospheric experiments observe neutrinos with energies around 1 to 10 GeV and have a baseline on the order of Earth’s diameter (~ 1000 km). According to Equation 1.48, such experiments are sensitive to oscillations driven by Δm^2 around 10^{-4} –

10^{-3} eV^2 . Accelerator experiments involve neutrinos with energies on the order of a few gigaelectronvolts and have baselines of several hundred kilometers, which implies sensitivity to $\Delta m^2 \sim 10^{-3} \text{ eV}^2$. Thus, in addition to the fact that both types of experiments observe the behavior of ν_μ s, they are both sensitive to oscillations driven by similar values of Δm^2 . Consequently, a hypothesized oscillation mode can be investigated by two rather different approaches.

Recall from Figure 1.5 that Super-K saw a deficit of μ -like events but not of e -like events. In the two-neutrino framework, this situation cannot be explained by oscillations between ν_μ and ν_e , since if such oscillations did occur, $P(\nu_\mu \rightarrow \nu_e) = P(\nu_e \rightarrow \nu_\mu)$ and the atmospheric ν_μ/ν_e flux ratio together would imply an observed deficit of ν_μ and an observed excess of ν_e . Thus, we expect that ν_μ oscillates to ν_τ , but ν_e does not appear to undergo oscillations driven by the same value of Δm^2 . By fitting the μ -like data shown in Figure 1.6 to $P(\nu_\mu \rightarrow \nu_\mu)$, Super-K measured the “atmospheric” oscillation parameters Δm_{atm}^2 and θ_{atm} ; the corresponding contour in the $\Delta m_{\text{atm}}^2 - \sin^2 2\theta_{\text{atm}}$ plane is shown in Figure 1.10.

The accelerator experiments observe a similar disappearance of ν_μ but under different conditions, as described in Section 1.2.3. Measurements of the oscillation parameters are obtained by fitting the energy spectrum to $P(\nu_\mu \rightarrow \nu_\mu)$ (with the baseline L fixed) as shown in Figures 1.7 and 1.8. Oscillation parameter contours from K2K and MINOS are shown in Figure 1.10. Note how the Super-K and MINOS results complement each other. As an atmospheric experiment, Super-K collects much greater statistics from the pervasive atmospheric neutrino flux and can thus measure the oscillation amplitude ($\sin^2 2\theta_{\text{atm}}$) with smaller uncertainty, while MINOS has better energy resolution and a neutrino source with a much better understood energy spectrum and can thus measure the oscillation “frequency” (Δm_{atm}^2) with smaller uncertainty. These measurements are

$$\sin^2 2\theta_{\text{atm}} > 0.92 \text{ (90\% C.L.) [10]} \quad \left| \Delta m_{\text{atm}}^2 \right| = 2.32_{-0.08}^{+0.12} \times 10^{-3} \text{ eV}^2 \text{ [45]}. \quad (1.50)$$

We must now face the question of ν_e disappearance in the various solar experi-

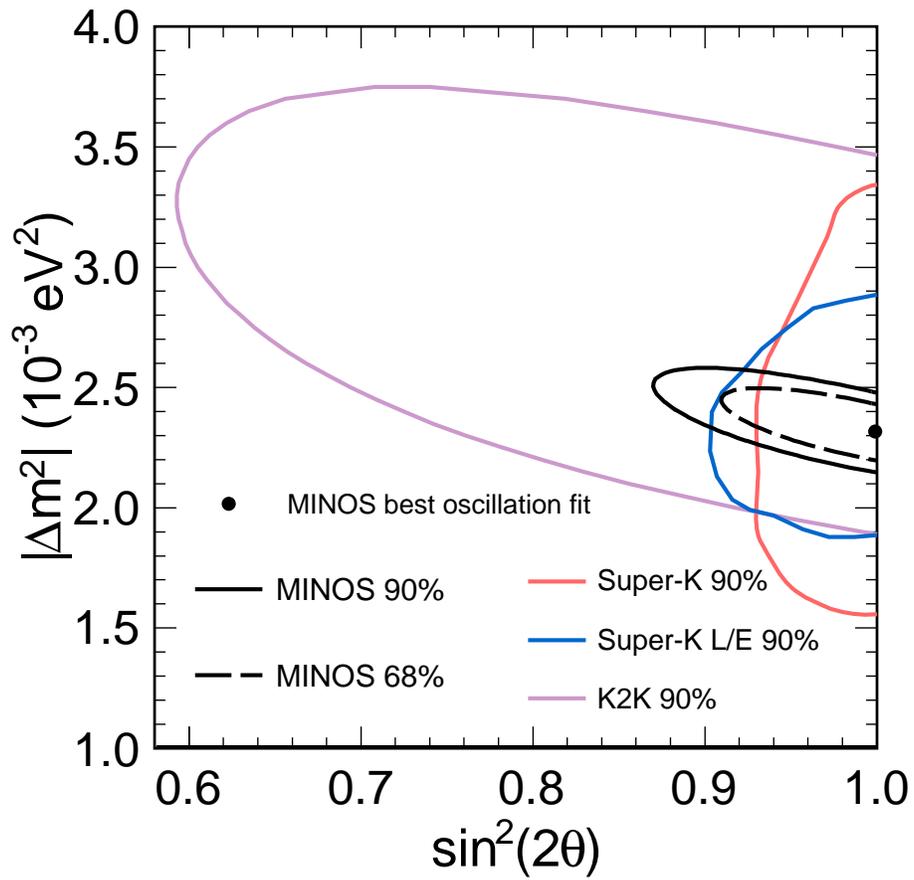


Figure 1.10: Contours for the atmospheric oscillation parameters from Super-K and the accelerator experiments K2K and MINOS. Taken from [45].

ments and $\bar{\nu}_e$ disappearance in the KamLAND experiment. As discussed in previous sections, oscillations driven by Δm_{atm}^2 do not appear to underlie these phenomena. One might expect that they are the result of the same underlying oscillation, driven by a new set of oscillation parameters and related via $P(\nu_e \rightarrow \nu_e) = P(\bar{\nu}_e \rightarrow \bar{\nu}_e)$.¹³ A two-neutrino oscillation analysis [51] of data from multiple solar neutrino experiments (SNO, Super-K, the chlorine and gallium experiments, and Borexino¹⁴) reveals that these “solar” oscillations are driven by $\Delta m_{\text{sol}}^2 \sim 4.90 \times 10^{-5} \text{ eV}^2$, with an associated mixing angle θ_{sol} satisfying $\tan^2 \theta_{\text{sol}} = 0.437$.¹⁵ This analysis, whose results are shown in Figure 1.11, and the interpretation of solar neutrino oscillations implied by it involve so-called matter effects, which we will discuss in Section 1.3.2.3.

Recall that KamLAND was located 180 km from the reactors producing the $\bar{\nu}_e$ s it observed and that those neutrinos had energies on the order of 1 MeV. Thus, by Equation 1.48, KamLAND was sensitive to $\bar{\nu}_e$ oscillations driven by $\Delta m^2 \sim 10^{-6} - 10^{-5} \text{ eV}^2$ – within reach of Δm_{sol}^2 as determined by the solar experiments. Their best-fit oscillation curve and the oscillation parameter contours implied by their analysis are shown in Figure 1.12; note the overlap with the solar experiments’ contours, confirming our expectation that the solar experiment and KamLAND observations are related.

Note the complementarity between the solar experiments (greater statistics) and KamLAND (better understood neutrino source and experimental energy resolution) in their measurements of the oscillation parameters, akin to the complementarity between Super-K and MINOS. A combined fit of the solar and KamLAND data [51], shown in Figure 1.13, finds the following values for the solar oscillation parameters:

$$\sin^2 2\theta_{\text{sol}} = 0.87 \pm 0.03 \quad \Delta m_{\text{sol}}^2 = 7.59_{-0.21}^{+0.19} \times 10^{-5} \text{ eV}^2. \quad (1.51)$$

Thus far, we have seen that ν_{μ} s disappear via Δm_{atm}^2 -driven oscillations, presu-

¹³The comparison is actually more subtle, as we will soon see.

¹⁴Borexino is a relatively new solar neutrino experiment located in the Gran Sasso Laboratory in Italy.

¹⁵These are best-fit values, quoted without errors.

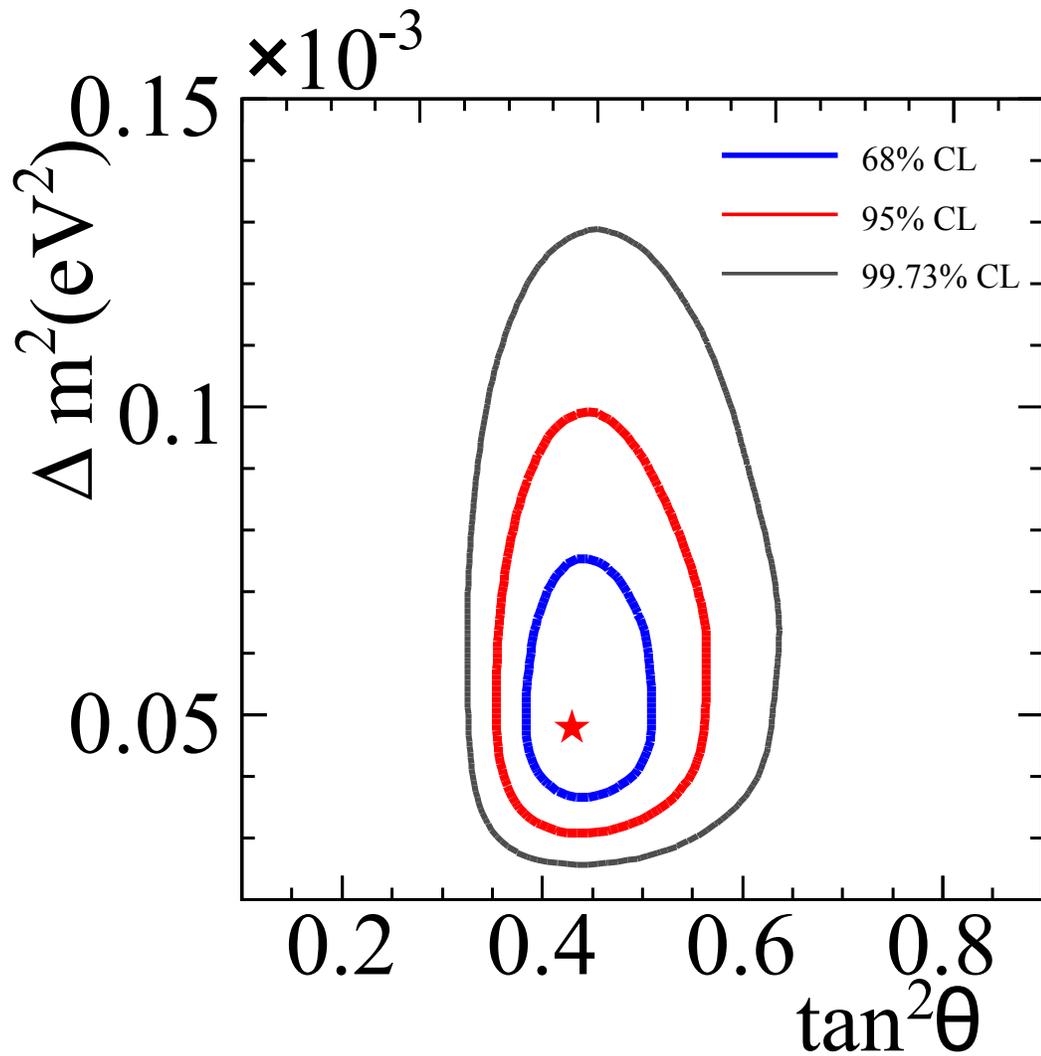


Figure 1.11: Solar oscillation parameter contours resulting from a fit to data from multiple solar neutrino experiments. Taken from [51].

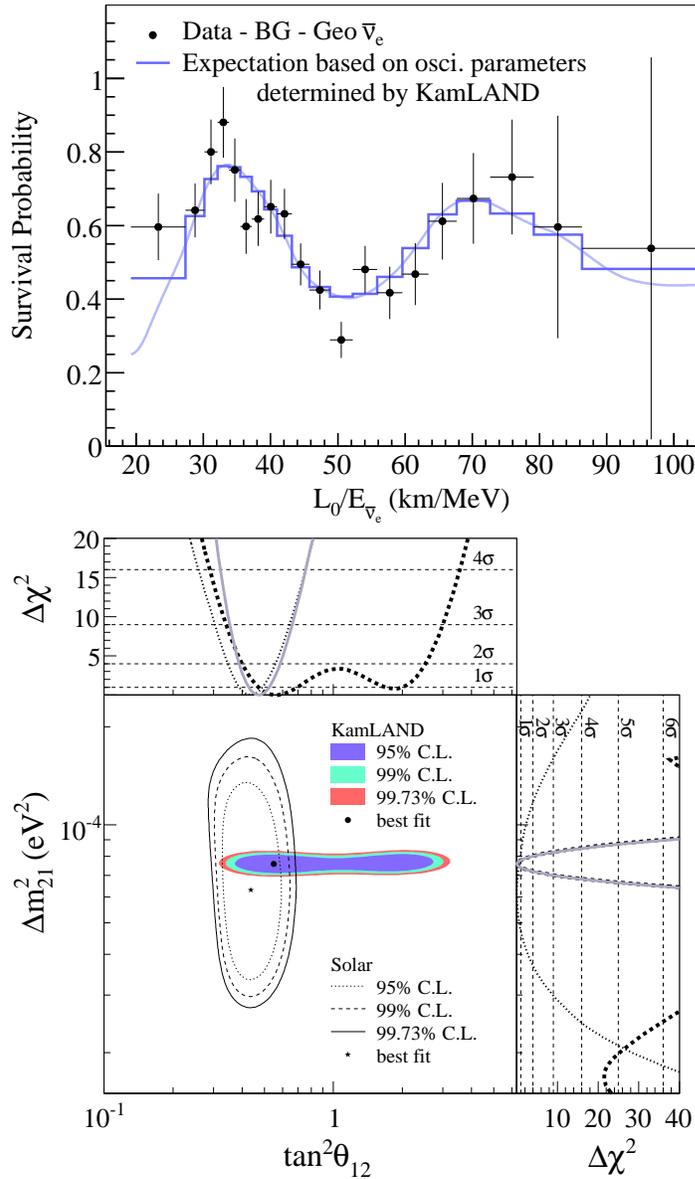


Figure 1.12: (Top) KamLAND data, displayed as the ratio of the background- and geo-neutrino-subtracted $\bar{\nu}_e$ spectrum to the expectation for no oscillation as a function of L_0/E . L_0 is the effective baseline taken as a flux-weighted average ($L_0 = 180$ km). The histogram and curve represent the best fit to the data and account for the distances to the individual reactors, time-dependent flux variations, and efficiencies. Error bars are statistical only. Note the clearly visible oscillatory pattern in the data. (Bottom) Allowed regions for solar oscillation parameters from KamLAND and solar neutrino experiments. The side panels show the $\Delta\chi^2$ profiles for KamLAND (dashed) and solar experiments (dotted) individually, as well as the combination of the two (solid). Note that the set of solar neutrino data used in producing these contours differs slightly from the one used in producing the contours shown in Figure 1.11. Taken from [52].

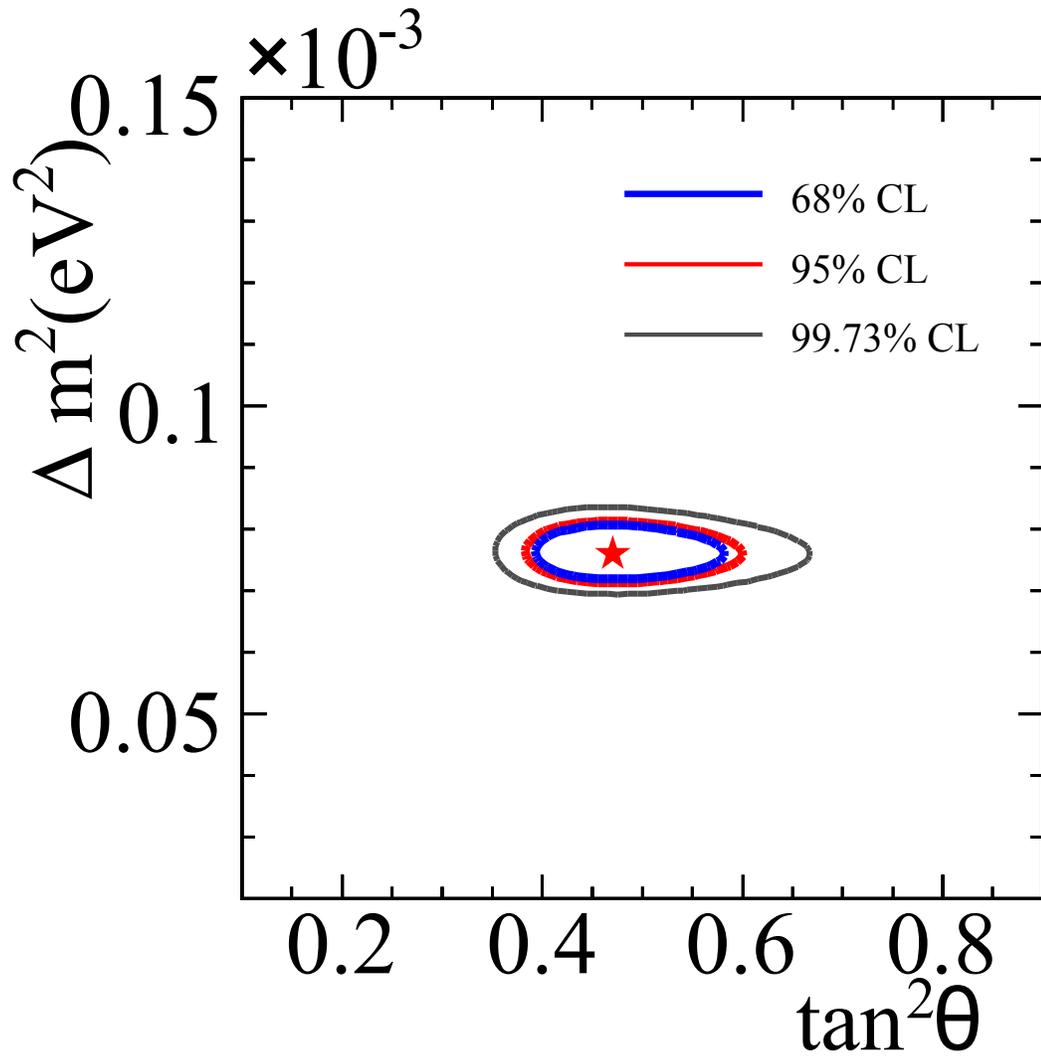


Figure 1.13: Solar oscillation parameter contours resulting from a fit to data from multiple solar neutrino experiments (the same as those used for the contours in Figure 1.11) and the KamLAND results reported in [52]. Taken from [51].

ably oscillating to ν_τ , and that ν_e s disappear via Δm_{sol}^2 -driven oscillations, oscillating to ν_μ and ν_τ . The remaining question, then, is whether ν_μ s *ever* oscillate to ν_e s via Δm_{atm}^2 -driven oscillations – i.e., is there a subdominant $\nu_\mu \rightarrow \nu_e$ oscillation mode in the atmospheric sector? The CHOOZ and Palo Verde experiments had L/E values that made them sensitive to Δm_{atm}^2 -driven oscillations, and neither saw any $\bar{\nu}_e$ disappearance. Under the assumption that oscillations were in fact occurring, they used their null results to set constraints on the oscillation parameters governing those oscillations. CHOOZ obtained stricter limits; its results are shown in Figure 1.14. By Equation 1.46, $P(\bar{\nu}_e \rightarrow \bar{\nu}_\mu) = P(\nu_\mu \rightarrow \nu_e)$, so ν_e appearance in a ν_μ beam is closely related to $\bar{\nu}_e$ disappearance. In Figure 1.14, $\sin^2 2\theta < 0.16$ at 90% C.L. when $\Delta m^2 = \Delta m_{\text{atm}}^2 = 2.32 \times 10^{-3} \text{ eV}^2$. This is referred to as the “CHOOZ limit” and sets an upper limit on any possible atmospheric $\nu_\mu \rightarrow \nu_e$ oscillation. The MINOS experiment has thus far performed two searches for this particular oscillation by looking for an excess of ν_e in its Far Detector; the analysis presented in this thesis constitutes the third such search. These analyses are performed in the full three-neutrino framework, however, so we will postpone discussion to Section 1.3.3.

1.3.2.3 Matter Effects

All of the experiments discussed thus far, with the exception of the solar experiments, made the reasonable approximation that the oscillations they observed could be analyzed as if they were occurring in vacuum. In environments such as the solar interior, however, the matter density is significant, and we should consider the possible effects on oscillations. In particular, we are interested in the effect of coherent scattering of the neutrino by components of the medium (neutrons, protons, electrons), for only in this case can interference occur between scattered and unscattered neutrino wave functions.

In 1978, Lincoln Wolfenstein showed that coherent neutrino scattering gives rise to an additional potential energy in the Hamiltonian of the neutrino [53]. In particular, it is the sum of potential energies due to charged-current and neutral-current interactions. The latter occurs equally for all neutrino flavors, via Z^0 exchange, and

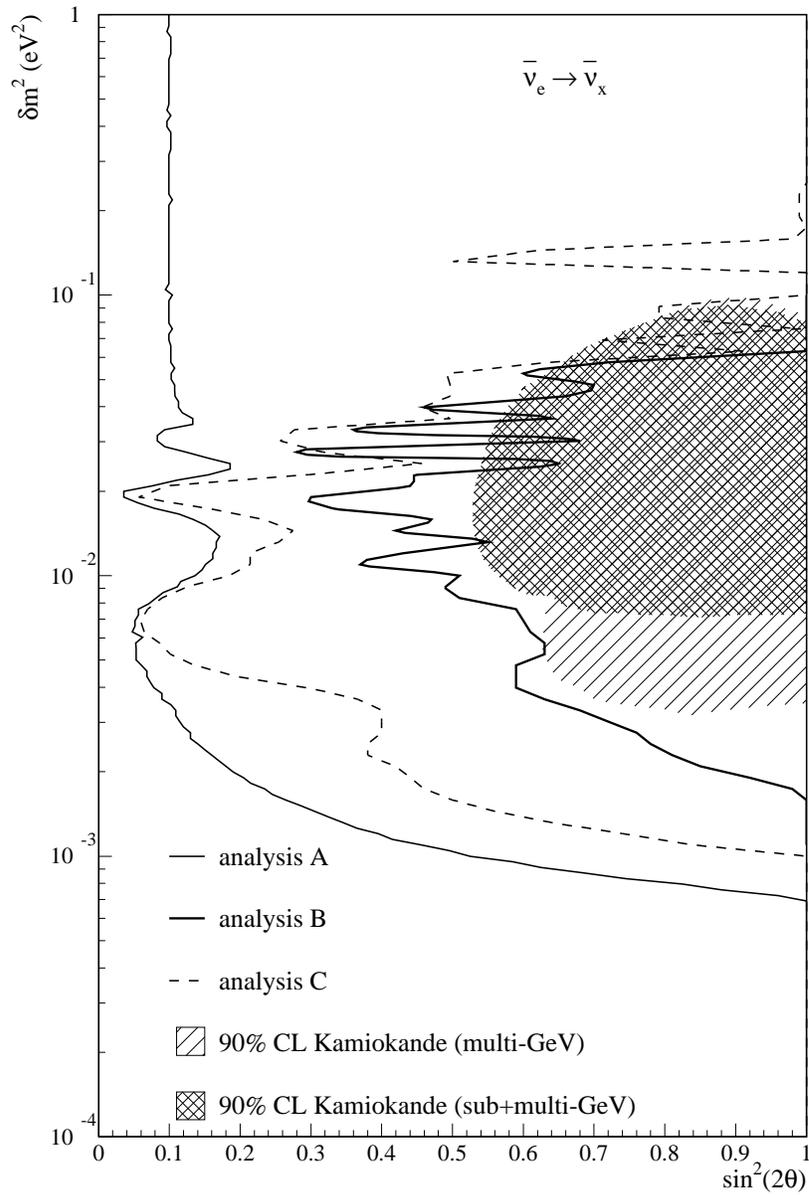


Figure 1.14: 90% C.L. constraints on oscillation parameters governing $\bar{\nu}_e$ disappearance in the CHOOZ experiment. The constraints were obtained via three different analyses (A, B, and C) described in [46], all based on a two-neutrino oscillation model. Parameter values above and to the right of the contours are excluded at 90% C.L. It is interesting to note that the solar oscillation parameters are well within the allowed region – off the plot, below the $\sin^2 2\theta$ axis. The Kamiokande allowed region for atmospheric $\nu_\mu \rightarrow \nu_e$ oscillations is shown for comparison. Of course, Super-K and MINOS have since shrunk and shifted that region. Taken from [46].

contributes the same phase shift to each flavor component of the neutrino state, resulting in an overall phase shift. In the context of neutrino oscillations, only *relative* phases between components matter, so we discard the neutral-current contribution for our purposes. The charged-current interaction, on the other hand, occurs only for ν_e , via W^\pm exchange, and produces a phase shift on the ν_e component only. This contribution to the potential energy is

$$V = \pm\sqrt{2}G_F N_e, \quad (1.52)$$

where G_F is the Fermi constant, N_e is the density of electrons in the medium, and the upper (lower) sign is for ν_e ($\bar{\nu}_e$).

Let us consider $\nu_e \rightarrow \nu_x$ ($x = \mu, \tau$ or some mixture thereof), the oscillation responsible for solar neutrino disappearance. We write the neutrino state as

$$|\Psi(t)\rangle = a_e(t) |\nu_e\rangle + a_x(t) |\nu_x\rangle, \quad (1.53)$$

where $a_e(t)$ and $a_x(t)$ are the time-dependent amplitudes of the ν_e and ν_x components, respectively, of $|\Psi(t)\rangle$. The time evolution of $|\Psi(t)\rangle$ in vacuum is given by the Schrödinger equation:

$$i\frac{d}{dt} |\Psi(t)\rangle = H_0^{\text{weak}} |\Psi(t)\rangle, \quad (1.54)$$

where H_0^{weak} is the vacuum Hamiltonian in the weak basis,¹⁶ related to the vacuum Hamiltonian in the mass basis H_0^{mass} by a unitary transformation:

$$H_0^{\text{weak}} = U H_0^{\text{mass}} U^\dagger, \quad (1.55)$$

where U is the 2×2 mixing matrix given in Equation 1.43. Numbering the mass

¹⁶While we are used to working with the vacuum Hamiltonian in the *mass* eigenbasis, it is easier to work in the *flavor* eigenbasis for the purposes of matter effects, for the interactions producing these effects are flavor dependent.

eigenstates 1 and 2 with masses m_1 and m_2 , we have

$$\begin{aligned}
H_0^{\text{weak}} &= \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \frac{m_1^2}{2E} & 0 \\ 0 & \frac{m_2^2}{2E} \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \\
&= \frac{1}{2E} \begin{pmatrix} m_1^2 \cos^2 \theta + m_2^2 \sin^2 \theta & (m_2^2 - m_1^2) \sin \theta \cos \theta \\ (m_2^2 - m_1^2) \sin \theta \cos \theta & m_2^2 \cos^2 \theta + m_1^2 \sin^2 \theta \end{pmatrix} \\
&= \frac{m_1^2 + m_2^2}{4E} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{m_2^2 - m_1^2}{4E} \begin{pmatrix} -\cos 2\theta & \sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix}.
\end{aligned} \tag{1.56}$$

We drop the first matrix on the right of Equation 1.56, as it is a multiple of the identity and thus represents a constant overall shift in the Hamiltonian, and define $\Delta m^2 \equiv m_2^2 - m_1^2$ to obtain

$$H_0^{\text{weak}} = \frac{\Delta m^2}{4E} \begin{pmatrix} -\cos 2\theta & \sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix}. \tag{1.57}$$

Equation 1.54 can now be written as

$$i \frac{d}{dt} \begin{pmatrix} a_e(t) \\ a_x(t) \end{pmatrix} = \frac{\Delta m^2}{4E} \begin{pmatrix} -\cos 2\theta & \sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix} \begin{pmatrix} a_e(t) \\ a_x(t) \end{pmatrix}. \tag{1.58}$$

The contribution of the effective matter potential V can be written in matrix form as

$$\begin{pmatrix} V & 0 \\ 0 & 0 \end{pmatrix}, \tag{1.59}$$

and the evolution equation in matter is

$$\begin{aligned} i \frac{d}{dt} \begin{pmatrix} a_e(t) \\ a_x(t) \end{pmatrix} &= \frac{\Delta m^2}{4E} \begin{pmatrix} -\cos 2\theta & \sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix} \begin{pmatrix} a_e(t) \\ a_x(t) \end{pmatrix} + \begin{pmatrix} V & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a_e(t) \\ a_x(t) \end{pmatrix} \\ &= \frac{\Delta m^2}{4E} \begin{pmatrix} \frac{4EV}{\Delta m^2} - \cos 2\theta & \sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix} \begin{pmatrix} a_e(t) \\ a_x(t) \end{pmatrix}. \end{aligned} \quad (1.60)$$

Note that the potential term was added to the ee term of the Hamiltonian. The matter Hamiltonian H_m^{weak} is thus

$$H^{\text{weak}} = \frac{\Delta m^2}{4E} \begin{pmatrix} \frac{4EV}{\Delta m^2} - \cos 2\theta & \sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix}. \quad (1.61)$$

We find that the eigenvalues of the Hamiltonian in matter differ from those in vacuum:

$$E_{1,2}^m = \frac{V}{2} \pm \frac{\Delta m^2}{4E} \sqrt{\sin^2 2\theta + \left(\frac{2EV}{\Delta m^2} - \cos 2\theta \right)^2}. \quad (1.62)$$

Also, H^{weak} can be diagonalized by a new unitary matrix U^m :

$$U^m = \begin{pmatrix} \cos \theta^m & \sin \theta^m \\ -\sin \theta^m & \cos \theta^m \end{pmatrix}, \quad (1.63)$$

where the new mixing angle θ^m satisfies

$$\tan 2\theta^m = \frac{\sin 2\theta}{\cos 2\theta - \frac{2EV}{\Delta m^2}}. \quad (1.64)$$

We can now compute the $\nu_e \rightarrow \nu_x$ oscillation probability in matter, which we denote as $P_m(\nu_e \rightarrow \nu_x)$. This requires a simple modification of Equation 1.46, where $\frac{\Delta m^2}{2E}$ is replaced by the difference of the eigenvalues given in Equation 1.62 and θ is

replaced by θ^m . We find

$$\begin{aligned}
 P_m(\nu_e \rightarrow \nu_x) &= \sin^2 2\theta^m \sin^2 \left(\frac{\Delta m^2 L}{4E} \sqrt{\sin^2 2\theta + \left(\frac{2EV}{\Delta m^2} - \cos 2\theta \right)^2} \right) \\
 &= \frac{\sin^2 2\theta}{\sin^2 2\theta + \left(\frac{2EV}{\Delta m^2} - \cos 2\theta \right)^2} \sin^2 \left(\frac{\Delta m^2 L}{4E} \sqrt{\sin^2 2\theta + \left(\frac{2EV}{\Delta m^2} - \cos 2\theta \right)^2} \right).
 \end{aligned}
 \tag{1.65}$$

The presence of matter has several interesting effects. First, the term $\frac{2EV}{\Delta m^2} - \cos 2\theta$ allows one to determine the sign of Δm^2 , thereby establishing which of the two mass eigenstates is of greater mass. In addition, the difference in the sign of V for ν_e and $\bar{\nu}_e$ implies that the effect of matter on the oscillation probability is different for neutrinos and antineutrinos. Next, if $\frac{2EV}{\Delta m^2} - \cos 2\theta = 0$, the mixing becomes maximal – i.e., $\sin^2 2\theta^m = 1$, and the oscillation amplitude is unity.¹⁷ This condition is equivalent to

$$N_e = \frac{\Delta m^2 \cos 2\theta}{2G_F E \sqrt{2}}.
 \tag{1.66}$$

Using the KamLAND best-fit values $\Delta m^2 = 7.59 \times 10^{-5} \text{ eV}^2$ and $\theta = 0.28$ [10] and $E \sim 1 \text{ MeV}$, we find $N_e \sim 8 \times 10^{25} \text{ cm}^{-3}$. Assuming about 1 nucleon per electron in the Sun, this electron density corresponds to a mass density of 140 g cm^{-3} , which is in fact achieved in the solar interior.

This resonant mixing due to the presence of matter is not the full explanation for solar neutrino disappearance. The mass – and thus, electron – density in the Sun slowly decreases from the core to the surface. As a result, the neutrino mass eigenstates evolve as neutrinos propagate from the core to the surface – that is, the mass eigenstates are local. We know that neutrinos produced in the Sun’s core are ν_e s. Taking into account that the density there is very high ($N_e \gg \frac{\Delta m^2 \cos 2\theta}{2G_F E \sqrt{2}}$), we can show that the higher local mass eigenstate $|\nu_2\rangle$ is approximately equal to $|\nu_e\rangle$ and that

¹⁷Note that this phenomenon can occur even if the vacuum mixing angle θ is very small. Also, because of the difference in sign of V for ν_e and $\bar{\nu}_e$, it can occur either for neutrinos (if $\cos 2\theta > 0$) or for antineutrinos (if $\cos 2\theta < 0$) but not for both.

the lower local mass eigenstate $|\nu_1\rangle$ is approximately equal to $|\nu_x\rangle$. Thus, neutrinos produced in the Sun’s core are effectively born in the local $|\nu_2\rangle$ state. In fact, the solar density changes slowly enough for them to propagate adiabatically through the Sun, remaining in the local $|\nu_2\rangle$ state at every point throughout their journey to the surface. Thus, solar neutrinos exit the Sun in the vacuum $|\nu_2\rangle$ state (i.e., the local $|\nu_2\rangle$ state in vacuum). Since this is a stationary state, the solar neutrinos remain in that state throughout their journey to Earth. The probability of detecting solar neutrinos as ν_e s on Earth, then, is

$$P(\nu_e \rightarrow \nu_e) = |\langle \nu_e | \nu_2 \rangle|^2 = \sin^2 \theta. \quad (1.67)$$

This effect was first described by Stanislav Mikheyev and Alexei Smirnov in 1985 [54] and is known as the MSW effect, where the “W” is for Wolfenstein. The solar neutrino data is interpreted and analyzed in this framework. Our arrangement of the mass eigenstates in the above explanation implies that Δm_{sol}^2 is positive. In Section 1.3.3, we will find that $\Delta m_{\text{sol}}^2 = \Delta m_{21}^2$, the mass-squared difference between the second and first of the three neutrino mass eigenstates. This tells us that the mass eigenstate we call “second,” a specified superposition of weak eigenstates, is higher in mass than the “first,” a different superposition of weak eigenstates. In addition, the general observation by solar neutrino experiments that $P(\nu_e \rightarrow \nu_e) \lesssim \frac{1}{2}$, together with Equation 1.67, implies that $0 < \theta_{\text{sol}} < \frac{\pi}{4}$.¹⁸ This is why the global solar contours in Figure 1.11 or 1.12 lie mostly to the left of $\tan^2 \theta = 0.5$. Meanwhile, the KamLAND experiment called for a simple two-neutrino vacuum oscillation analysis, which cannot distinguish between $0 < \theta < \frac{\pi}{4}$ and $\frac{\pi}{4} < \theta < \frac{\pi}{2}$; this results in contours, as shown in Figure 1.12, that are symmetric in $\log \tan^2 \theta$ about $\tan^2 \theta = 1$.

In the atmospheric sector, no similar statement can be made about the sign of

¹⁸Note that in Figure 1.3, not all experiments show the same level of solar ν_e flux suppression, while Equation 1.67 implies that the suppression factor should always be the same. The explanation presented above is somewhat simplified. The MSW effect is in fact nearly negligible for lower-energy solar neutrinos; the differences among different types of experiments (chlorine, gallium, water), then, is a consequence of their differing sensitivities to different ranges of neutrino energies. For a thorough overview, see [10].

Δm_{atm}^2 ; we only know its absolute value. Since the resonance density (Equation 1.66) is proportional to Δm^2 and $\Delta m_{\text{atm}}^2 \approx 30\Delta m_{\text{sol}}^2$, the mass density required to observe resonant atmospheric oscillation enhancement similar in scale to that in the solar sector is $\approx 4000 \text{ g cm}^{-3}$. This does not mean, however, that no matter effects whatsoever are observed in the atmospheric sector; in fact, the analysis presented in this thesis – a search for $\nu_\mu \rightarrow \nu_e$ oscillations in the atmospheric sector – includes these small matter effects in its calculations.

1.3.3 Three-Neutrino Oscillations

1.3.3.1 Understanding the Two-Neutrino Approximation

The major results obtained through two-neutrino analyses have now been laid out. The next step is to see how these results fit into the complete three-neutrino framework. We begin by noting that there are two distinct mass-squared differences, Δm_{atm}^2 and Δm_{sol}^2 , driving neutrino oscillations. With three physical neutrinos, there are three masses and thus two independent mass-squared differences, for

$$\Delta m_{21}^2 + \Delta m_{13}^2 + \Delta m_{32}^2 = 0. \quad (1.68)$$

(Note the particular ordering of the indices.) We expect that Δm_{atm}^2 and Δm_{sol}^2 are those two independent mass-squared differences. As noted in Section 1.3.2.3, we identify $\Delta m_{\text{sol}}^2 = \Delta m_{21}^2$; this is a convention and is without loss of generality, as we are simply numbering the mass eigenstates such that $|\nu_1\rangle$ and $|\nu_2\rangle$ are the more closely spaced mass eigenstates and $|\nu_2\rangle$ is the more massive of the two. By Equation 1.68, it does not matter much whether we identify $|\Delta m_{31}^2| = |\Delta m_{\text{atm}}^2|$ or $|\Delta m_{32}^2| = |\Delta m_{\text{atm}}^2|$, as $|\Delta m_{31}^2 - \Delta m_{32}^2| = \Delta m_{21}^2 \ll |\Delta m_{31}^2|, |\Delta m_{32}^2|$.¹⁹ Note the absolute value bars, indicating that the sign of Δm_{atm}^2 – known as the “neutrino mass hierarchy” – remains unknown. In particular, the case where $|\nu_3\rangle$ is the most massive physical neutrino is known as the “normal” hierarchy, while the “inverted” hierarchy has $|\nu_3\rangle$ as the least

¹⁹Indeed, all atmospheric and accelerator experiments to date measure an unresolved mixture of $|\Delta m_{31}^2|$ and $|\Delta m_{32}^2|$.

massive physical neutrino.

One can see that the primary reason for the validity of the two-neutrino approximation is that $\Delta m_{\text{atm}}^2 \approx 30\Delta m_{\text{sol}}^2$. For example, if a given experiment satisfies

$$\frac{\Delta m_{\text{atm}}^2 L}{4E} \sim 1, \quad (1.69)$$

then it also satisfies

$$\frac{\Delta m_{\text{sol}}^2 L}{4E} \ll 1. \quad (1.70)$$

Thus, in an experiment sensitive to Δm_{atm}^2 -driven oscillations, any possible Δm_{sol}^2 -driven oscillations will not have had a chance to develop. If, on the other hand, an experiment satisfies

$$\frac{\Delta m_{\text{sol}}^2 L}{4E} \sim 1, \quad (1.71)$$

then it also satisfies:

$$\frac{\Delta m_{\text{atm}}^2 L}{4E} \gg 1. \quad (1.72)$$

Thus, in an experiment sensitive to Δm_{sol}^2 -driven oscillations, any possible Δm_{atm}^2 -driven oscillations will have already occurred multiple times, and the oscillatory term can be replaced by its average value of $\frac{1}{2}$.

The atmospheric/accelerator experiments' observations of ν_μ disappearance is described by $P(\nu_\mu \rightarrow \nu_\mu)$. Substituting Equation 1.44 into Equation 1.40 with $\alpha = \beta = \mu$, neglecting any terms involving Δm_{21}^2 , and taking $\Delta m_{31}^2 \approx \Delta m_{32}^2$, we obtain

$$\begin{aligned} P(\nu_\mu \rightarrow \nu_\mu) &= 1 - 4 \cos^2 \theta_{13} \sin^2 \theta_{23} (\cos^2 \theta_{23} + \sin^2 \theta_{13} \sin^2 \theta_{23}) \sin^2 \frac{\Delta m_{32}^2 L}{4E} \\ &= 1 - (\cos^2 \theta_{13} \sin^2 2\theta_{23} + \sin^4 \theta_{23} \sin^2 2\theta_{13}) \sin^2 \frac{\Delta m_{32}^2 L}{4E}. \end{aligned} \quad (1.73)$$

Note that the second summation in Equation 1.40 vanishes, taking with it the sine term that could reveal the sign of Δm_{32}^2 . While matter effects could achieve this goal in principle, very long baselines (longer than in any existing experiments) are required

to make the effect large enough to be observable.²⁰

While all reactor experiments study $\bar{\nu}_e$ disappearance, KamLAND was sensitive to solar oscillations, and CHOOZ and Palo Verde were sensitive to atmospheric oscillations. For both cases, we substitute Equation 1.44 into Equation 1.42 with $\alpha = \beta = e$ and take $\Delta m_{31}^2 \approx \Delta m_{32}^2$ to obtain

$$P(\bar{\nu}_e \rightarrow \bar{\nu}_e) = 1 - \cos^4 \theta_{13} \sin^2 2\theta_{12} \sin^2 \frac{\Delta m_{21}^2 L}{4E} - \sin^2 2\theta_{13} \sin^2 \frac{\Delta m_{32}^2 L}{4E}. \quad (1.74)$$

For KamLAND, $\sin^2 \frac{\Delta m_{32}^2 L}{4E} \approx \frac{1}{2}$, so

$$P_{\text{KamLAND}}(\bar{\nu}_e \rightarrow \bar{\nu}_e) = 1 - \cos^4 \theta_{13} \sin^2 2\theta_{12} \sin^2 \frac{\Delta m_{21}^2 L}{4E} - \frac{1}{2} \sin^2 2\theta_{13}. \quad (1.75)$$

For CHOOZ and Palo Verde, $\sin^2 \frac{\Delta m_{21}^2 L}{4E} \approx 0$, so

$$P_{\text{CHOOZ}}(\bar{\nu}_e \rightarrow \bar{\nu}_e) = 1 - \sin^2 2\theta_{13} \sin^2 \frac{\Delta m_{32}^2 L}{4E}. \quad (1.76)$$

We computed Equation 1.76 in the three-neutrino framework, but it has the same form as the two-neutrino survival probability of Equation 1.45. The CHOOZ limit is thus a direct limit on the θ_{13} mixing angle. In particular, $\sin^2 2\theta_{13} < 0.16$ at 90% C.L. – that is, θ_{13} is expected to be quite small.²¹ If we take $\theta_{13} \approx 0$, Equation 1.73 becomes

$$P(\nu_\mu \rightarrow \nu_\mu) = 1 - \sin^2 2\theta_{23} \sin^2 \frac{\Delta m_{32}^2 L}{4E}, \quad (1.77)$$

and Equation 1.75 becomes

$$P_{\text{KamLAND}}(\bar{\nu}_e \rightarrow \bar{\nu}_e) = 1 - \sin^2 2\theta_{12} \sin^2 \frac{\Delta m_{21}^2 L}{4E}. \quad (1.78)$$

²⁰One can see the general principle behind this statement by expanding the argument of the oscillatory term in Equation 1.65 about $V = 0$. To first order in V , that argument is $\frac{\Delta m^2 L}{4E} - \frac{VL \cos 2\theta}{2}$. Thus, for densities far below the resonant density for $\Delta m^2 = \Delta m_{\text{atm}}^2$, which is certainly the case in the Earth's interior, L must be sufficiently large for matter effects to reveal the sign of Δm^2 .

²¹Recent results from Daya Bay, RENO, and Double Chooz – discussed in Chapter 9 – have significantly changed the state of our knowledge of θ_{13} , showing in particular that θ_{13} is actually not all that small.

Thus, under the assumptions made thus far, the probabilities computed in the three-neutrino framework reduce to the appropriate two-neutrino probabilities. Finally, we can identify $\theta_{\text{atm}} = \theta_{23}$ and $\theta_{\text{sol}} = \theta_{12}$.

1.3.3.2 $\nu_\mu \rightarrow \nu_e$ Oscillation in the Three-Neutrino Framework

This thesis describes a search for ν_e appearance in the MINOS experiment. Thus, we are looking for atmospheric $\nu_\mu \rightarrow \nu_e$ oscillations. As mentioned in Section 1.3.2.2, the extent of such oscillations is constrained by the CHOOZ limit, so the effect being sought is small. We compute the vacuum oscillation probability in full, without discarding Δm_{21}^2 -driven oscillations and taking $\Delta m_{31}^2 \approx \Delta m_{32}^2$ only where appropriate, and find

$$\begin{aligned}
P(\nu_\mu \rightarrow \nu_e) &= P_{\text{atm}} + P_{\text{sol}} + P_{\text{CP-even}} + P_{\text{CP-odd}} \\
P_{\text{atm}} &= \sin^2 \theta_{23} \sin^2 2\theta_{13} \sin^2 \frac{\Delta m_{32}^2 L}{4E} \\
P_{\text{sol}} &= \cos^2 \theta_{23} \sin^2 2\theta_{12} \sin^2 \frac{\Delta m_{21}^2 L}{4E} \\
P_{\text{CP-even}} &= 8J_{\text{CP}} \cot \delta_{\text{CP}} \sin \frac{\Delta m_{32}^2 L}{4E} \cos \frac{\Delta m_{32}^2 L}{4E} \sin \frac{\Delta m_{21}^2 L}{4E} \\
P_{\text{CP-odd}} &= \mp 8J_{\text{CP}} \sin^2 \frac{\Delta m_{32}^2 L}{4E} \sin \frac{\Delta m_{21}^2 L}{4E} \\
J_{\text{CP}} &= \Im (U_{e3}^* U_{\mu 3} U_{e2} U_{\mu 2}^*) = \frac{1}{8} \cos \theta_{13} \sin 2\theta_{12} \sin 2\theta_{23} \sin 2\theta_{13} \sin \delta_{\text{CP}}.
\end{aligned} \tag{1.79}$$

For experiments sensitive to Δm_{32}^2 -driven oscillations, the leading term in this expression is P_{atm} , and we often make the approximation $P(\nu_\mu \rightarrow \nu_e) \approx P_{\text{atm}}$. P_{sol} makes a small contribution for such experiments. The remaining terms, $P_{\text{CP-even}}$ and $P_{\text{CP-odd}}$, depend on the Dirac CP-violating phase δ_{CP} and are so named for their respective symmetries under $\delta_{\text{CP}} \rightarrow -\delta_{\text{CP}}$. The upper (lower) sign in $P_{\text{CP-odd}}$ is for neutrinos (antineutrinos). The expression J_{CP} is the ‘‘rephasing invariant’’ or ‘‘Jarlskog invariant’’ and controls the overall magnitude of any CP-violating effects that may be observed in neutrino oscillations. In particular, because J_{CP} is a product involving the sines of all of the mixing angles, the observation of leptonic CP violation in an oscillation experiment requires that all of the angles be nonzero. While θ_{23} and θ_{12}

are known to be large, θ_{13} is known to be smaller and must be measured with greater precision (and, of course, found to be sufficiently larger than zero) in order to make any measurement of CP-violating effects possible. Note, in addition, that the sign of $P_{\text{CP-even}}$ is equal to the sign of Δm_{32}^2 , so $P(\nu_\mu \rightarrow \nu_e)$ also depends on the mass hierarchy. The inclusion of matter effects [55] renders P_{atm} and $P_{\text{CP-odd}}$ sensitive to the mass hierarchy as well. The analysis described in this thesis takes matter effects into consideration, using the oscillation probabilities presented in [55] throughout our calculations.

Thus, our goal in this thesis is to make a measurement of θ_{13} or, at least, to improve the upper limit on its value. As neither δ_{CP} nor the mass hierarchy is known today, we present our results as a function of δ_{CP} , for each hierarchy. The analysis presented herein represents a significant improvement over previous MINOS analyses [56, 57] in sensitivity to nonzero θ_{13} . Results from the second of those two previous analyses [57] are shown in Figure 1.15.

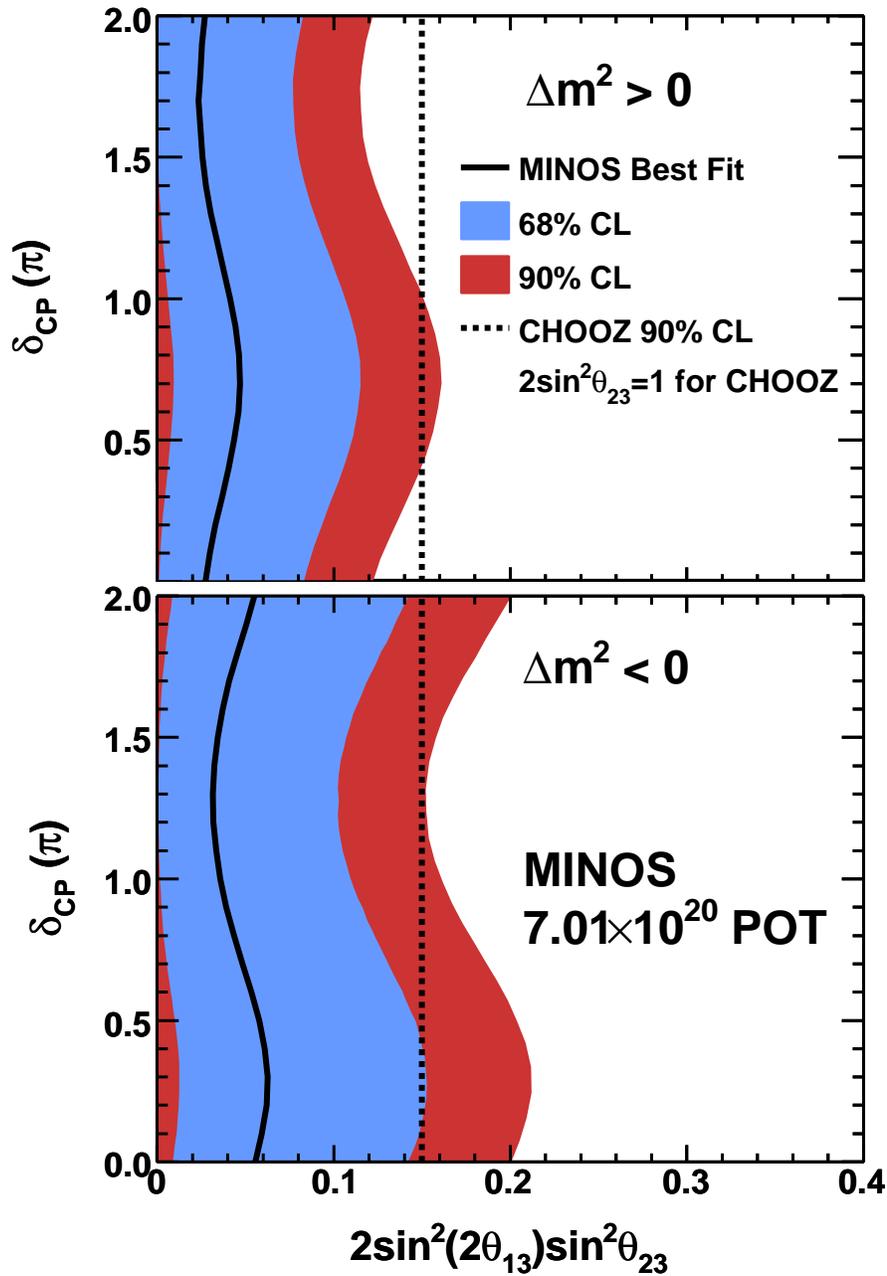


Figure 1.15: Results from the most recent MINOS search for ν_e appearance, shown as confidence intervals for $2\sin^2\theta_{23}\sin^2 2\theta_{13}$ ($=\sin^2 2\theta_{13}$ for maximal atmospheric mixing) as a function of δ_{CP} for the normal (top) and inverted (bottom) mass hierarchies. The CHOOZ limit shown ($\sin^2 2\theta_{13} < 0.15$ at 90% C.L.) assumes $\Delta m_{32}^2 = 2.43 \times 10^{-3} \text{ eV}^2$, the accepted value of Δm_{32}^2 at the time this analysis was performed. Taken from [57].

Chapter 2

Experimental Setup

The MINOS experiment is composed of three key components: the Neutrinos at the Main Injector (NuMI) beamline at the Fermi National Accelerator Laboratory (Fermilab) in Batavia, IL, the MINOS Near Detector (ND) located 1 km downstream of the NuMI beam source (also at Fermilab), and the MINOS Far Detector (FD) located 735 km downstream in the Soudan Underground Laboratory in northern Minnesota. Neutrinos from the NuMI beamline interact with the ND, which measures the beam's initial composition and energy spectrum, and then with the FD, which measures the beam's composition and energy spectrum once again. Essentially, our goal is to look for differences between these compositions and spectra that cannot be explained by beam divergence and other established effects but by the phenomenon of neutrino oscillation. We presently will describe each component of the experiment at a level of detail suitable for this work; further information can be found in the references.

2.1 Neutrinos at the Main Injector (NuMI)

To produce the NuMI beam [58], we begin by extracting 120-GeV protons from the Main Injector accelerator at Fermilab. Next, a series of magnets focus the proton beam, bend it downwards, and bring it in line with the detectors. Overall, the proton beam is bent 58 mrad downward into the Earth relative to its direction of travel immediately after being extracted from the Main Injector; this is necessary to

aim the beam towards the FD in Soudan, as depicted in Figure 2.1. A final set of quadrupole magnets controls the size and angular dispersion of the beam such that the beam spot size matches the diameter of the graphite target.

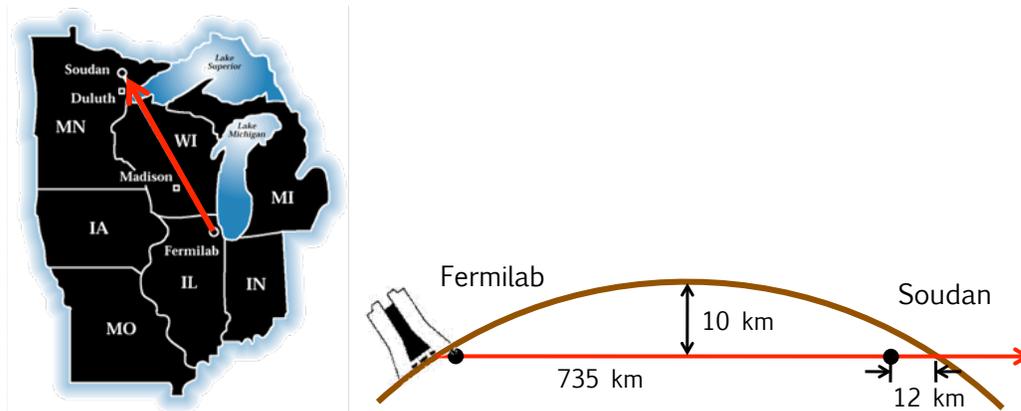


Figure 2.1: Overhead and cross-sectional views of the MINOS experiment

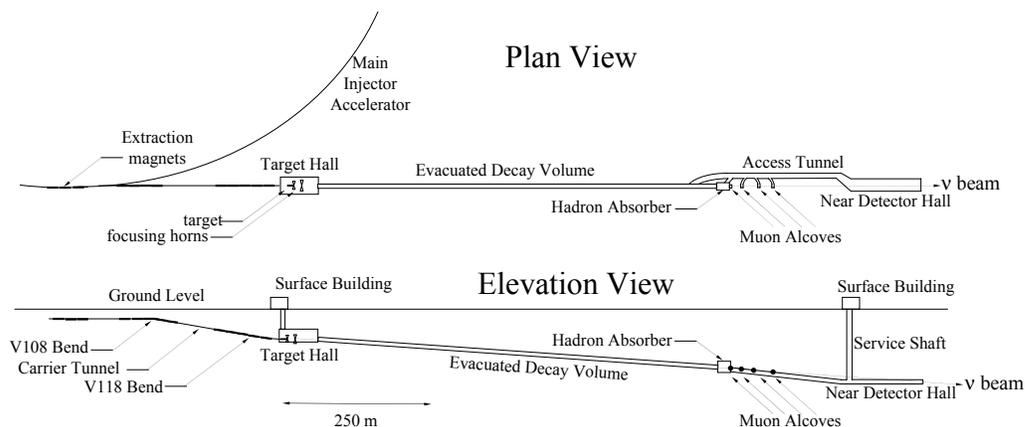


Figure 2.2: Plan and elevation views of the NuMI and MINOS facilities at Fermilab. Taken from [59].

When the Main Injector was shared between NuMI and the Tevatron [60] (“mixed mode”), the target receives 36×10^{12} protons every 2.2 seconds (the duty cycle of NuMI); in “NuMI-only mode,” all protons from the Main Injector – 40×10^{12} or more per pulse – are delivered to the target. Figure 2.3 shows the rate of proton delivery and the total protons-on-target (“PoT”) delivered as functions of time. Once in the target, most of the protons interact with the carbon nuclei present, producing pions and kaons with a wide range of momenta. While the total length of the target is approximately

95 cm (about 1.9 hadronic interaction lengths), it is actually segmented into 48 thin “fins,” each one 0.3 mm away from its neighbors. Thus, a large fraction of the incident protons get a chance to interact with the target, while the secondary pions and kaons, endowed with some transverse momentum upon production, are allowed to disperse quickly out of the target radius, thus avoiding additional interactions with the target and getting the opportunity to decay farther downstream and produce neutrinos. A short fin length has the added benefit of minimizing stresses induced by the high intensity proton beam. Figure 2.4 provides a visualization of this process and a technical drawing of the target.

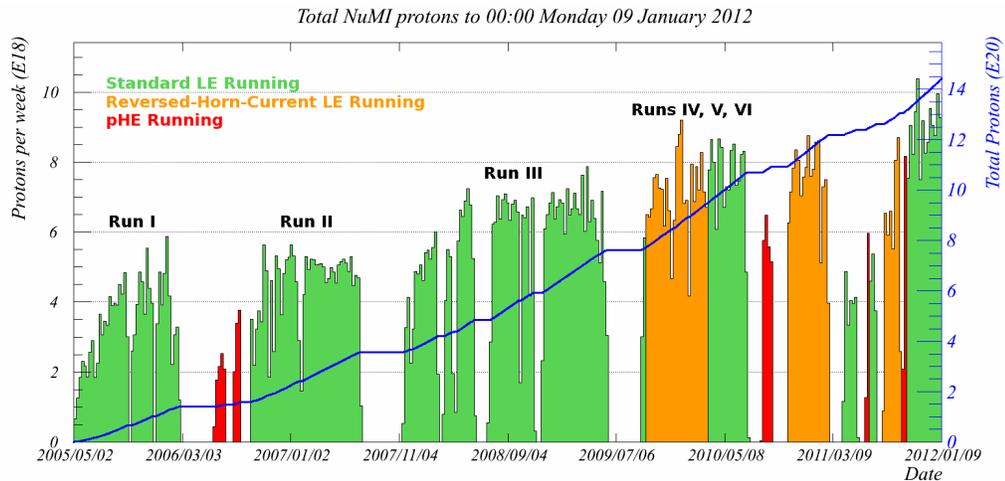


Figure 2.3: The left vertical axis provides the scale for the bar graph (protons delivered per week), and the right vertical axis provides the scale for the blue curve (total protons delivered). Extended gaps in proton delivery indicate technical shutdowns. Run numbering distinguishes periods of running with distinctly different hardware configurations, as summarized in Table 2.2 toward the end of this section. The data from the period marked “Runs IV, V, VI” that were used in this analysis excludes the data collected in reversed-horn-current LE running.

The secondary π and K are produced with a typical transverse momentum of ~ 300 MeV, which corresponds to production with an angular distribution about some mean. Without a system to gather and focus these secondaries, the flux of beam neutrinos, resulting from the decay of only the most forward-going mesons, would be quite small and would have a relatively flat energy spectrum with a broad maximum beyond our energy range of interest. To remedy this, we employ a set of two magnetic

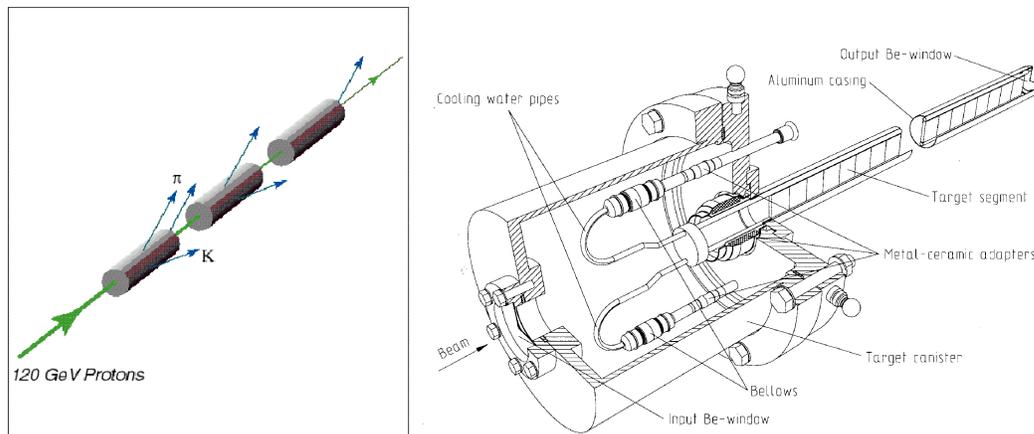


Figure 2.4: The cartoon on the left illustrates how secondary pions and kaons, due to their transverse momentum, disperse out of the target radius while primary protons continue to travel through the target. The technical drawing on the right shows the NuMI target and its enclosure, including an electrically isolated cooling water system. Both figures taken from [61].

horns, located just downstream of the target, that serve as a focusing system [62]. Each horn is 3 m long, is made of aluminum sheeting a few (2 to 4.5, depending on the horn and the constituent part of the horn) millimeters thick, and consists of a cylindrical outer conductor and an inner conductor in the shape of two paraboloids placed almost end-to-end with a small gap (“neck”) in between; see Figure 2.5. The horns are pulsed with electrical current (variable up to 200 kA) running axially down the inner conductor and back through the outer conductor. This produces a toroidal magnetic field between the conductors that sign-selects and focuses the secondary mesons streaming out of the target, as illustrated in Figure 2.6. In effect, the horns act as a point-to-parallel lens; in particular, the horns’ focal length is approximately proportional to the pion momentum, so one can ultimately select the neutrino energy spectrum by varying the distance of the target from Horn 1 and of Horn 1 from Horn 2. Varying the horn current provides another means of tuning the neutrino energy spectrum.

Note that since there is no magnetic field along the horns’ axis, particles traveling along that line avoid being focused or defocused. Such π^- and K^- ultimately decay to produce a small $\bar{\nu}_\mu$ contamination in the beam with a broad energy distribution. Also, such π^+ and K^+ of higher energies, which the horns would normally not focus,

remain in the beam and decay into higher-energy neutrinos, resulting in the observed neutrino spectrum's significant high-energy tail. In the original design of NuMI, the position of the horns relative to each other, as well as the design of the target, could be changed to produce neutrino spectra of different shapes with maxima at different energies. Ultimately, however, this approach was abandoned early in the experiment; instead, the position of the (original-design) target relative to the first horn and the current pulsing through the horns were varied to produce such spectra. Simulated spectra and further details for each technique are shown in Figure 2.7. The analysis described in this thesis uses data collected primarily¹ in the standard LE10/185kA configuration, in which the upstream end of the target is placed 10 cm upstream of the upstream end of the first horn, and the horns are pulsed with a 185-kA current oriented such that positively charged mesons are focused and negatively charged mesons are defocused. The energy spectrum of the resulting neutrino beam peaks in the range 1–6 GeV; compared to other spectra, this is the most desirable, as the probability of Δm_{32}^2 -driven oscillations away from ν_μ (i.e., $1 - P(\nu_\mu \rightarrow \nu_\mu)$) for the MINOS baseline of 735 km attains its final maximum at a neutrino energy of about 1.4 GeV (i.e., all other maxima occur at lower energies, specifically 0.5 GeV and below).

Upon exiting Horn 2, the secondary mesons drift a distance of 30 m in the Target Hall before entering the decay pipe, a 2-m-diameter, 675-m-long steel pipe surrounded by concrete shielding. Here, the mesons drift freely,² and some decay.³ Tertiary μ^\pm produced in the decay of these mesons subsequently decay into $e^+ \bar{\nu}_\mu \nu_e$ or $e^- \nu_\mu \bar{\nu}_e$; the neutrinos from these decays also contribute to the overall neutrino flux. Table 2.1 summarizes the processes that contribute to the neutrino flux and the frequency with which they contribute, as estimated by the data-tuned simulation (see Section 2.6 for

¹See Section 5.2 for important exceptions.

²Until the end of Run II, the interior of the decay pipe was kept at a pressure of 1 torr – nearly evacuated. At that time, the thin aluminum window at the upstream end of the decay pipe was found to have sustained serious radiation damage and was at risk of breaching. High radiation levels in the Target Hall ruled out repairing or replacing the window. Instead, before the start of Run III, the decay pipe was filled with helium to a pressure of 0.9 atm to relieve the stress on the window.

³675 m is approximately the mean decay length of a 12-GeV π^\pm , a 90-GeV K^\pm , or a 22-GeV K_L^0 . Thus, mesons of higher energies are less likely to decay before reaching the end of the decay pipe.

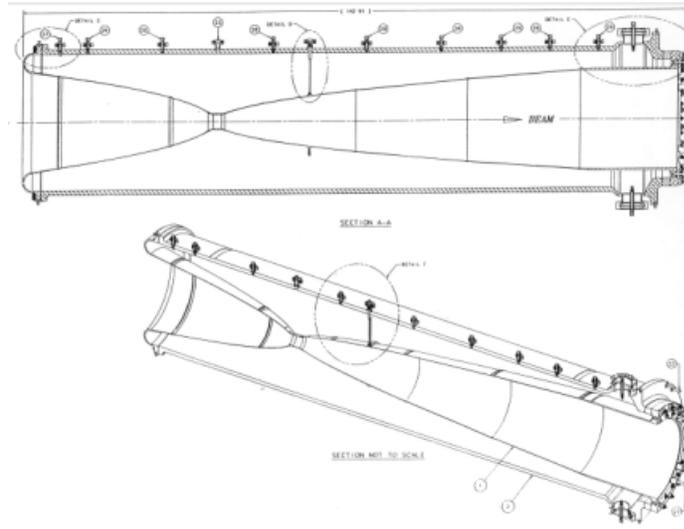


Figure 2.5: Two cross-sectional views of a technical drawing of Horn 2. Note the direction of the beam and the parabolic shape of the inner conductor. Taken from [61].

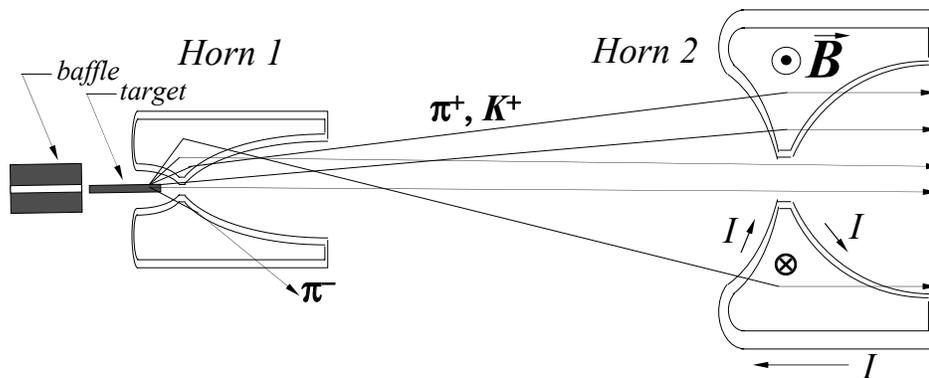


Figure 2.6: A cartoon illustrating the sign selection and focusing performed by the horns on the secondary mesons streaming out of the target. Note that hadrons traveling along the horns' axis are not affected by the magnetic field and thus remain in the beam, regardless of charge-sign or momentum. Taken from [61].

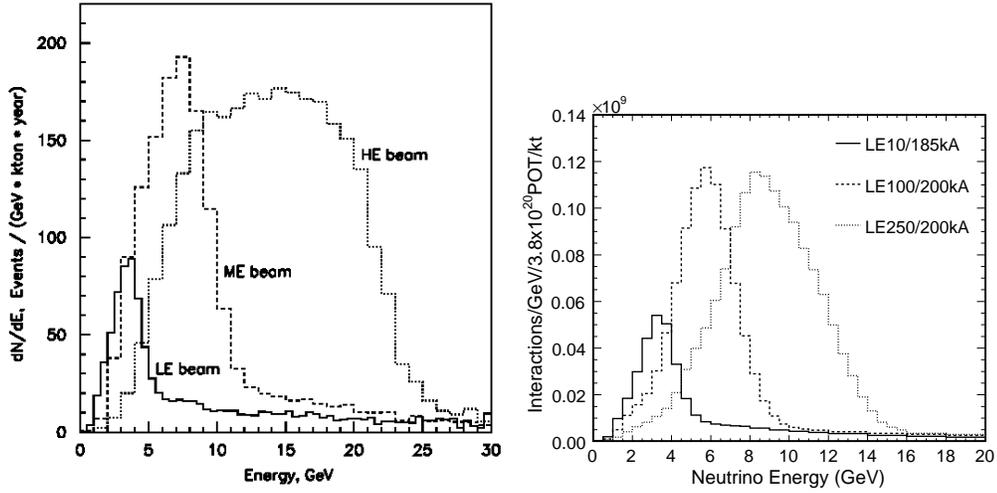


Figure 2.7: The simulated Far Detector ν_μ charged-current (CC) interaction spectra on the left – low energy (LE), medium energy (ME), and high energy (HE) – assume the use of an individually designed target; positioning of this target a distance of 0.34, 1.40, or 4.00 m, respectively, upstream of Horn 1; and positioning of Horn 2 a distance of 10, 23, or 40 m, respectively, downstream of Horn 1. Taken from [62]. The simulated Near Detector ν_μ CC interaction spectra on the right – LE10/185kA (LE), LE100/200kA (pseudo-ME, or “pME”), and LE250/200kA (pseudo-HE, or “pHE”) – assume the use of the LE target shown in Figure 2.4; positioning of this target a distance of 0.10, 1.00, or 2.50 m, respectively, upstream of Horn 1; and Horns 1 and 2 fixed in position relative to each other but pulsed with 185, 200, and 200 kA, respectively. Taken from [59].

further explanation).

Process	Frequency
$\pi^+ \rightarrow \nu_\mu + \mu^+$	78.2%
$\pi^- \rightarrow \bar{\nu}_\mu + \mu^-$	6.3%
$K^+ \rightarrow \nu_\mu + \mu^+$	12.5%
$K^+ \rightarrow \nu_e + \pi^0 + e^+$	0.6%
$K^+ \rightarrow \nu_\mu + \pi^0 + \mu^+$	0.3%
$K^- \rightarrow \bar{\nu}_\mu + \mu^-$	0.7%
$K^- \rightarrow \bar{\nu}_e + \pi^0 + e^-$	0.03%
$K^- \rightarrow \bar{\nu}_\mu + \pi^0 + \mu^-$	0.02%
$K_L^0 \rightarrow \nu_e + \pi^- + e^+$	0.3%
$K_L^0 \rightarrow \bar{\nu}_e + \pi^+ + e^-$	0.1%
$K_L^0 \rightarrow \nu_\mu + \pi^- + \mu^+$	0.2%
$K_L^0 \rightarrow \bar{\nu}_\mu + \pi^+ + \mu^-$	0.1%
$\mu^+ \rightarrow \bar{\nu}_\mu + \nu_e + e^+$	1.0%
$\mu^- \rightarrow \nu_\mu + \bar{\nu}_e + e^-$	0.1%

Table 2.1: Decay processes contributing to the neutrino flux seen at the Near Detector, as estimated by the data-tuned simulation. Note that pion decays account for more than 80% of the neutrino flux. Kaon and muon decays are responsible for the small $\nu_e/\bar{\nu}_e$ component that constitutes an irreducible background to the analysis presented in this thesis.

Before the beam is allowed to proceed to the ND, we must eliminate as many remaining hadrons and muons as possible; our goal is to observe interactions of neutrinos with the detectors, and these charged particles would interact extensively with the ND and overwhelm the data acquisition system. All hadrons remaining in the beam (primary protons that did not interact in the target, secondary mesons that did not decay) are stopped by a hadron absorber, consisting of a water-cooled aluminum central core and a surrounding steel jacket, located just downstream of the decay pipe. Tertiary muons remaining in the beam are negligibly impeded by the hadron absorber. As the NuMI beamline is located in dolomite, a dense rock, the beam is simply allowed to travel through 240 m of unexcavated dolomite downstream of the hadron absorber, in which the muons dissipate their energy primarily via ionization. A series of ionization chambers placed along the beamline monitor the beam throughout this process. Figure 2.8 provides an illustrated summary of the main components

of the NuMI beamline. Ultimately, the composition of the NuMI beam is 92.9% ν_μ , 5.8% $\bar{\nu}_\mu$, and 1.3% ν_e and $\bar{\nu}_e$.

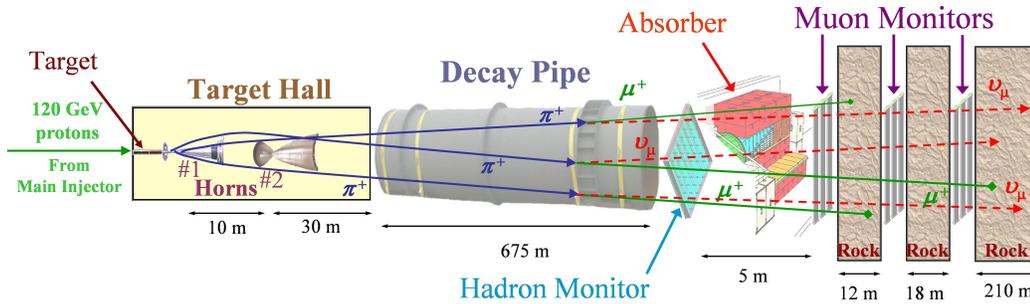


Figure 2.8: Illustration of the main components of the NuMI beamline. Note the placement and number of the ionization chambers (Hadron Monitor, Muon Monitors) monitoring the beam beyond the decay pipe.

Run	Exposure (10^{20} PoT)	Comments
I	1.21	NuMI Target #1 (NT-01), evacuated decay pipe
II	1.93	NT-02, evacuated decay pipe
III	3.87	NT-02, helium in decay pipe
IV	0.09	NT-03, helium in decay pipe, just prior to running with focusing horn current reversed
V	0.45	NT-03, helium in decay pipe, just after running with focusing horn current reversed
VI	0.61	NT-03, helium in decay pipe, target skewed

Table 2.2: This table lists, by Run Period, the proton exposure corresponding to the data on which the analysis presented in this thesis is based. Notes regarding hardware changes that distinguish the Run Periods are also given.

2.2 The MINOS Detectors

The MINOS detectors, shown in Figure 2.9, are functionally identical, steel–scintillator sampling calorimeters, capable of tracking, energy, and topology measurements [63]. Each detector, consisting of alternating planes of plastic scintillator strips and 2.54-cm-thick steel plates, is magnetized by a coil running parallel to its longitudinal axis, enabling charge-sign and momentum determination of muons passing through

the detector. The steel plates make up the bulk of the detectors' mass and provide the primary target for neutrino interactions. The extruded polystyrene scintillator strips constitute the active portion of the detectors. The strips, each 1-cm-thick by 4.1-cm-wide, are read out with wavelength-shifting (WLS) fibers and multianode photomultiplier tubes (PMTs). All strips in a given plane are oriented in the same direction; however, strips in any given plane are oriented orthogonally to strips in immediately neighboring planes, thereby allowing for three-dimensional reconstruction of particle interactions occurring in the detectors. Figure 2.10 illustrates this arrangement of “U-view” and “V-view” planes, with strips at $\pm 45^\circ$ with respect to the vertical.

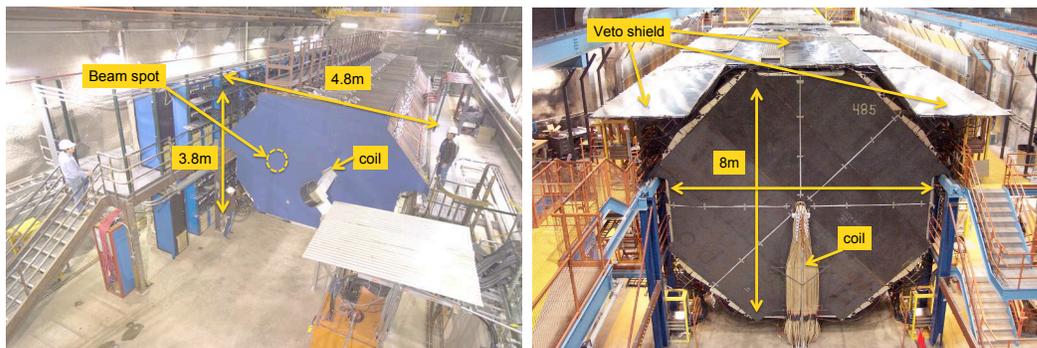


Figure 2.9: The Near (left) and Far (right) Detectors, with certain features and measurements indicated.

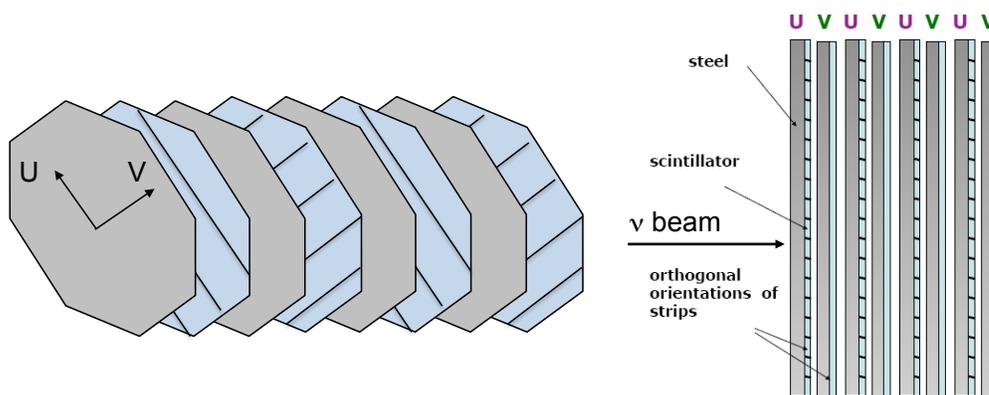


Figure 2.10: Arrangement of MINOS planes in the detectors. Note how, as described above, successive planes are arranged such that their strips are oriented orthogonally to each other.

The Far Detector is located 705 m underground (2070 meters-water-equivalent)

in a now-inactive iron mine in Soudan, MN, currently operated as a State Park. The detector consists of 486 octagonal steel planes with edge-to-edge dimension of 8 m alternating with planes of plastic scintillator strips. It has a mass of 5400 metric tons and is divided along its axis into two “supermodules” separated by a 1.15-m gap.

The Near Detector is housed in a 100-m-deep (225 meters-water-equivalent) underground cavern excavated as part of the NuMI facility. Since the Near Detector is much closer to the beam source than is the Far Detector (1 km versus 735 km), it receives a much higher neutrino flux and, despite having a mass of only 980 metric tons, has a neutrino interaction rate approximately 10^4 times that of the Far Detector. The Near Detector consists of 282 “squashed octagon”-shaped steel plates, instrumented (i.e., interleaved with planes of plastic scintillator) in a particular pattern. The first 120 upstream planes of the detector, referred to as the calorimeter, are all instrumented and contain the target fiducial volume. The downstream portion of the detector – the spectrometer – has every fifth plane instrumented and serves to measure the momenta of higher-energy muons.

A third detector, the Calibration Detector (CalDet), consisted of sixty 1-m \times 1-m planes of steel and scintillator, was unmagnetized, and had a total mass of 12 tons. It was placed in test beams of protons, pions, electrons, and muons of known momenta at the CERN PS [64] to measure the calorimetric and topological responses of the MINOS detector design. As we shall describe later, the Near and Far Detectors used different readout electronics, so the Calibration Detector recorded data with each set of electronics.

2.2.1 The Magnetic Field

The magnet system was designed to allow measurement of muon momenta with a resolution of $\sigma_p/p \sim 12\%$ for muons with energy exceeding 2 GeV. The magnetic field strength in the ND, averaged over the fiducial volume, is 1.28 T; in the FD, this quantity is 1.42 T. The similarity in the field strengths is intentional to minimize systematic uncertainties due to differences between the ND and FD. Monte Carlo

studies indicate that if the average magnetic field in each toroid is known to $\sim 3\%$, detector acceptance uncertainties will be much less than 1% in the range of muon energies of interest, and the uncertainty in the energy of muon tracks exiting the detector⁴ will be less than 2%.

The different geometries of the ND and FD impose different constraints on the designs of the magnet coils, so each coil's design was optimized independently. Each FD coil (one for each supermodule) consists of 190 turns of copper wire carrying 80 A for a total of 15.2 kA-turn. Furthermore, each coil dissipates 20 kW and is water-cooled through 15 copper tubes distributed throughout its cross section, as illustrated in Figure 2.11. As is visible in Figure 2.9, the ND coil is not centered on the plane; rather, it is offset by 55.8 cm so that the beam spot falls halfway between the coil hole and the left edge of the detector, creating an uninterrupted fiducial volume for beam neutrino events. The ND coil consists of 48 aluminum conductors arranged as shown in Figure 2.11, with 5 kA running in parallel through each row of six conductors for a total of 40 kA-turn. This coil dissipates 47 kW and is cooled by a closed-loop low-conductivity water (LCW) system that transfers the heat out of the underground cavern. The magnetic fields generated in each detector were computed by finite element analysis (FEA) using the ANSYS [65] program,⁵ the field maps are shown in Figure 2.12.

2.2.2 The Scintillator Strips

The scintillator strips – 1 cm thick, 4.1 cm wide, and up to 8 m long – constitute the core of the MINOS detectors' particle detection system. As illustrated in Figure 2.13, the strips actually consist of a clear polystyrene core and a 0.25-mm-thick TiO₂-loaded

⁴Calculations by the Bethe–Bloch equation could provide more reliable values of muon energies, but this is only possible if the muon track is fully contained within the detector. For tracks exiting the detector, we must use the track curvature, induced by the detector magnetic field, to calculate the muon momentum and, in turn, the muon energy.

⁵Test toruses were fabricated from each batch of steel used to make the detector steel planes, and the B – H curve of each such torus was measured by magnetic induction. Variations in the curves were found to be small, so a single representative B versus H relationship was provided as input to the FEA computation. The computed field is normalized to coil currents measured using power-supply current shunts.

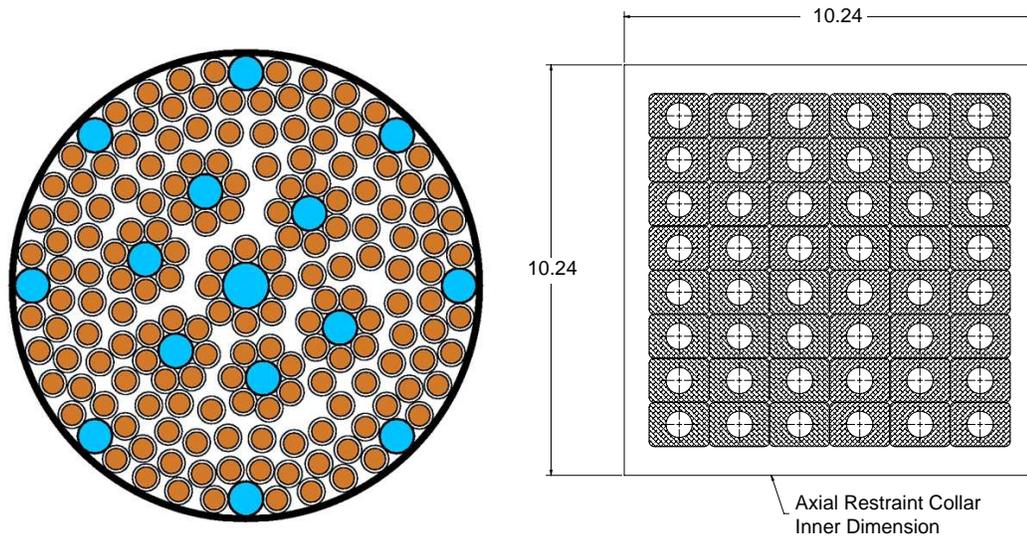


Figure 2.11: Cross-sectional illustrations of the Far Detector (left) and Near Detector (right) magnet coils. Dimensions for the ND coils are given in inches. Taken from [63].

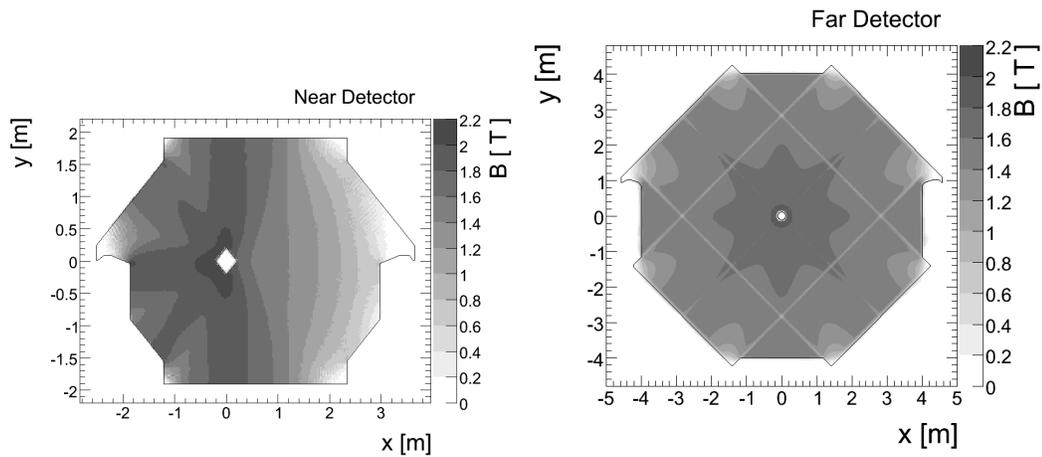


Figure 2.12: Magnetic field maps for a typical Near Detector plane (left) and Far Detector plane (right). Note that the Near Detector plane is shown looking upstream, towards the Target Hall, while in Figure 2.9, it is shown looking downstream, toward Soudan. Taken from [63].

polystyrene cap. The clear polystyrene contains 1.03% scintillating fluors by weight – 1.0% PPO (2,5-diphenyloxazole) and 0.03% POPOP (1,4-bis(5-phenyloxazol-2-yl) benzene). The TiO_2 -loaded polystyrene cap contains 15% TiO_2 by weight and is coextruded with the clear polystyrene core; it serves as a diffuse reflector for the scintillation light produced in the core and is sufficiently opaque so that ultraviolet light (similar to scintillation light) cannot shine through, thereby optically isolating each strip from its neighboring strips.

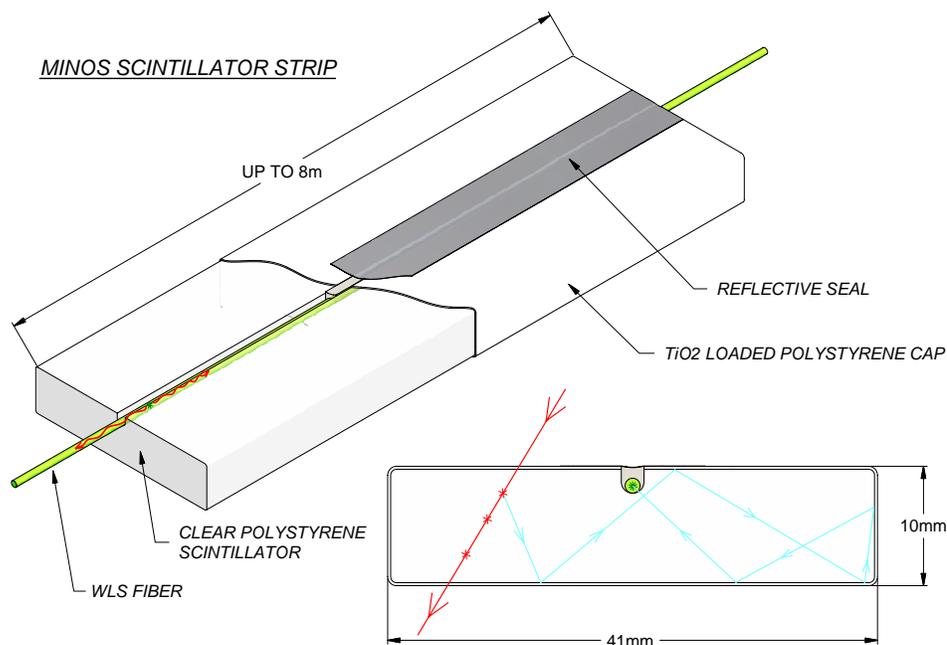


Figure 2.13: Cutaway drawing of a scintillator strip. In the cross-sectional view, an ionizing particle (red) produces light (light blue) as it crosses the strip. The light is multiply reflected by the TiO_2 -loaded cap until it is absorbed by the WLS fiber. The wavelength-shifted light is emitted isotropically, and those photons whose directions fall within the total internal reflection cones are transported along the fiber to the edges of the detector. Taken from [63].

The top of the strip has a 2.3-mm-deep by 2.0-mm-wide groove, free of the TiO_2 layer. The WLS fiber is glued in this groove and covered with a specularly reflective strip of aluminized Mylar tape. The fiber is 1.20-mm-diameter double-clad polystyrene with 175 ppm of Y11 fluor; these fibers minimize self-absorption by absorbing light with wavelength peaked at 420 nm (coincident with scintillation light wavelength) and reemitting it isotropically at 470 nm. Figure 2.13 shows the function of the WLS fiber – to collect scintillation light generated in the strip, which the fiber

then carries away to be read out. Note that while both ends of the fiber are read out in the FD, only one end is read out in the ND while the other end is cut with a hot knife and covered with reflective tape. This was possible with the shorter ND strips, as the light was not significantly attenuated in traveling to the mirrored end and returning to the readout end.

These scintillator strips, with embedded WLS fiber, are then bundled into “modules” – light-tight aluminum casings with routing and optical connectors for the fibers, as shown in Figure 2.14. This construction greatly facilitates shipping of the scintillator to the detector sites, mounting them to the steel plates, and connecting the WLS fibers to clear fibers for transmission of the light signal to the PMTs. A total of 22 different module types were constructed, as illustrated in Figures 2.15 and 2.16.

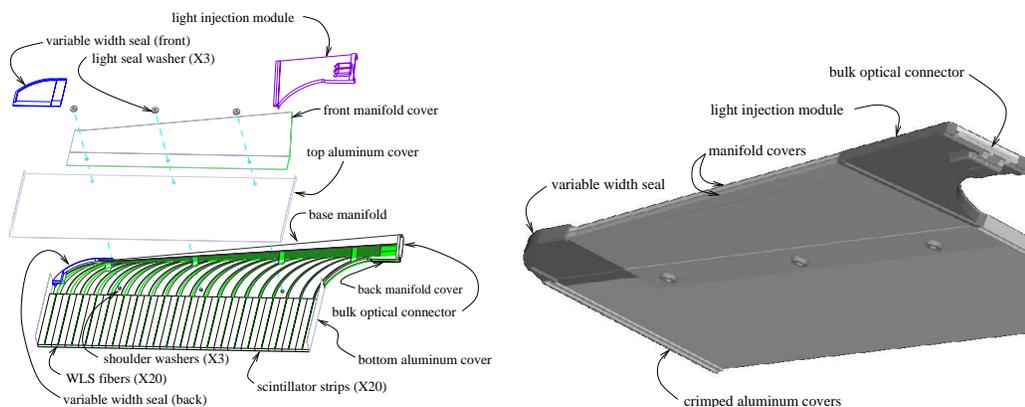


Figure 2.14: A typical MINOS module manifold assembly. WLS fibers from the scintillator strips are routed to bulk optical connectors (left) and enclosed by protective light-tight aluminum covers (right). Taken from [63].

2.2.2.1 The Light Injection System

The light injection (LI) [66, 67] system is a light-injecting optical fiber calibration system, the precise function of which is discussed in Section 2.3. Nearly identical systems are used in the ND, FD, and CalDet. Pulsed UV LEDs are housed in rack-mounted “pulser boxes,” with each box containing a set of 20 or 40 LEDs. Optical-fiber fan-outs allow each LED to illuminate multiple fibers through a set of optical connectors on the back panel of each pulser box. Optical fibers carry the light from

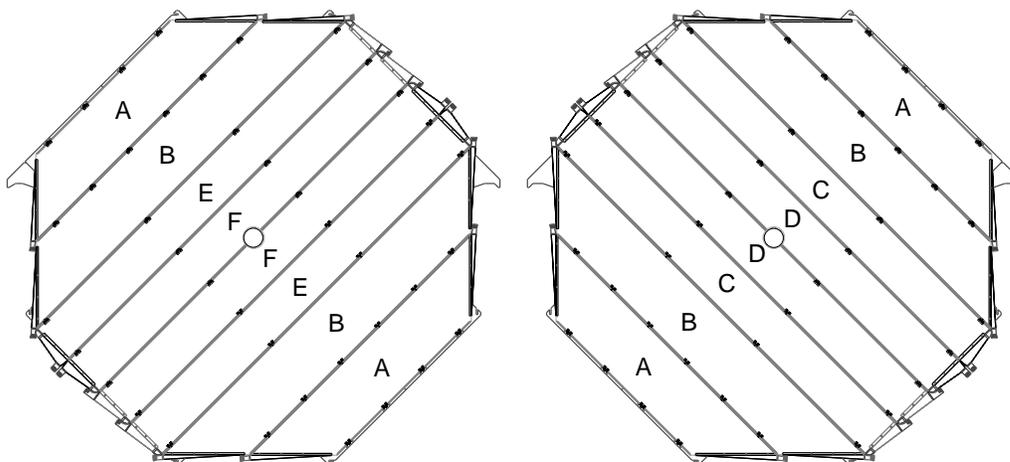


Figure 2.15: Layout of U-view (left) and V-view (right) modules on FD planes. “A” and “B” module types each contain 28 scintillator strips; all others each contain 20 strips. The first plane of each supermodule is a V-view plane. Taken from [63].

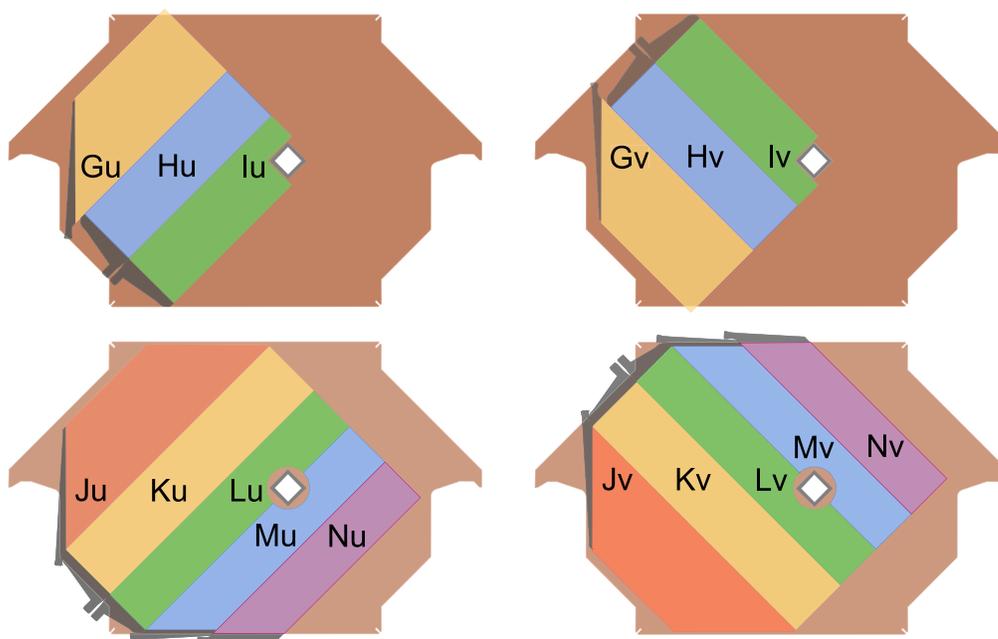


Figure 2.16: Layout of modules on ND planes: (clockwise from top left) partial U-view (PU), partial V-view (PV), full V-view (FV), full U-view (FU). Planes in the calorimeter section follow the repeating pattern FU-PV-PU-PV-PU-FV-PU-PV-PU-PV. Planes in the spectrometer section follow the same pattern but with all partial-view scintillator modules removed. That is, one out of every five planes in the spectrometer is a full-view plane, and the other four are uninstrumented. Note the location of the module ends (black). Taken from [63].

the pulser boxes to the light injection modules (LIMs), indicated in Figure 2.14. The LIMs are highly reflective cavities that allow the LED light to illuminate the scintillator modules' WLS fibers directly, mimicking scintillation light signals from scintillator strips; Figure 2.17 shows a cutaway view of a LIM. The intensity of the injected light is monitored by PIN (Positive Intrinsic Negative) photodiodes that are read out simultaneously with the PMTs.

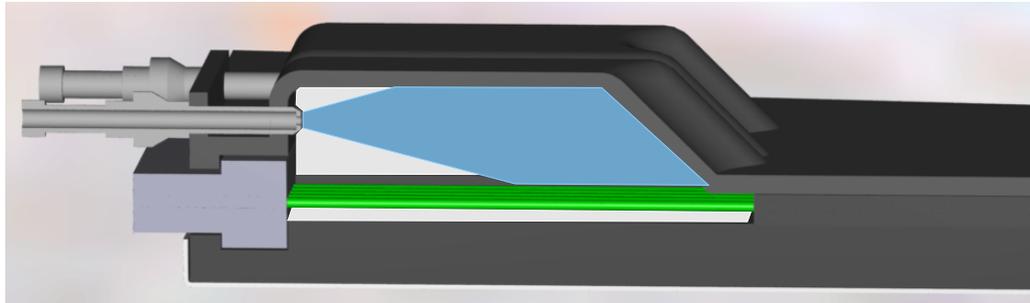


Figure 2.17: A cutaway view of a light injection module. A curtain of WLS fibers runs across the bottom of the LIM cavity, and the injected light illuminates them from above. The T-shaped component at the lower left is the bulk optical connector into which the WLS fibers are glued. In the upper left foreground is a cutaway of one of the connectors terminating the light-injection fibers; another connector is visible behind it. Taken from [63].

2.2.3 The Photodetectors

Light carried by the WLS fibers to an optical connector on a module is then carried by a cable of clear optical fibers to a PMT, as illustrated in Figure 2.18. These PMTs are housed in light-tight, steel enclosures that contain short segments of clear optical fiber leading from cable connectors on the front of the enclosure to the PMT pixels. In the case of the ND, each enclosure contains one Hamamatsu 64-anode (M64) PMT, and each clear fiber, which can be traced back to one strip in the detector, terminates on one pixel⁶ of the PMT using a machined plastic part called a “cookie,” as shown in Figure 2.19. The ND uses a total of 194 PMTs. In the FD, however, each

⁶While this is true throughout the ND, each pixel is read out individually only in the calorimeter section. In the spectrometer section of the ND – the 162 downstream planes of which every fifth plane is instrumented with scintillator modules – the signals from sets of four pixels are electrically summed to reduce the number of front-end electronics channels. In each of these sets, the pixels correspond to strips that are about 1 m apart, thereby enabling unambiguous muon tracking through this region despite the multiplexing.

enclosure contains three Hamamatsu 16-anode (M16) PMTs and is called a “MUX box,” reflecting the optical multiplexing that takes place inside: eight clear fibers, corresponding to the west ends (or east ends) of eight strips separated by about 1 m, are bundled in a close-packed arrangement and terminate on one pixel, as shown in Figure 2.19. The pixel assignment of the fibers from the opposite ends of a given set of eight comultiplexed strips is done in a way that allows for unambiguous “demultiplexing” of the signals, enabling unambiguous three-dimensional reconstruction of charged particle tracks and showers [68]. The FD uses a total of 1452 PMTs housed in 484 MUX boxes; an additional 64 PMTs serve the cosmic ray veto shield, indicated in Figure 2.9.

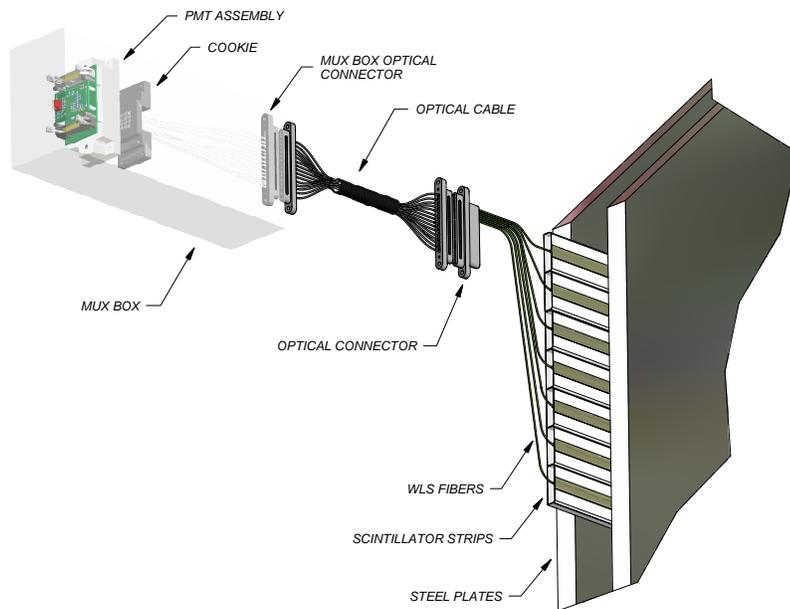


Figure 2.18: A schematic drawing showing the scintillator system readout for a module. An edge of a detector plane is on the right side of the sketch. Several strips are shown extending out of a module and beyond the edge for clarity. The light produced in a strip travels out of the module in a WLS fiber, as shown in Figure 2.13, and is then carried by a clear optical fiber (bundled into a cable) to an enclosure (a MUX box in the case of the FD) where it is routed to a pixel of the PMT assembly shown in Figure 2.19. Taken from [63].

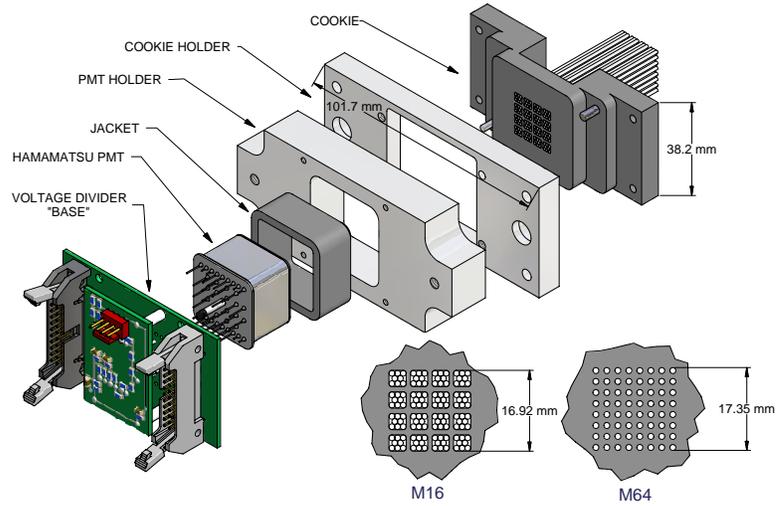


Figure 2.19: The M16 PMT mounting assembly – one of three inside each FD MUX box – that is used to bring eight fibers, as shown, to each PMT pixel. The ND M64 mounting assembly is identical except that only one fiber is placed on each pixel. The fiber cookie layouts are shown in the lower right of the figure. Taken from [63].

2.2.4 The Electronics

The environment of each detector imposes significantly different requirements on the detector’s respective front-end electronics. The FD’s deep underground location leads to a cosmic ray–induced event rate of about 0.5 Hz; its distance from the NuMI beam source and the beam’s nonnegligible divergence result in at most a few beam neutrino events per day. These rates are dominated by detector noise of 3–6 kHz per PMT. In the ND, on the other hand, there are several neutrino interactions per 8–10 μs -long spill (3.5 reconstructed events per 10^{13} PoT). Great effort was made to make the systems as similar as possible; any systematic uncertainties associated with remaining differences were minimized by measuring the Calibration Detector’s response separately with both ND and FD front-end electronics, as previously noted.

The high event rate at the ND necessitates fast front-end electronics [69] that can read out signals with little or no dead time. Each M64 pixel is connected to a MENU (MINOS Electronics for Neutrinos) circuit board consisting of a chip called

the Charge Integration Encoder (QIE), a commercial flash analog-to-digital converter (FADC), and a data buffer. The QIE [70], an ASIC (application-specific integrated circuit) developed at Fermilab, has four independent copies of the integrating circuits, allowing for continuous dead time-less operation. Sixteen MENUs reside on a printed circuit board called a MINDER (MINOS Near Detector Electronics Readout); up to eight of these, in turn, are read out by a module called a MASTER. At the end of a trigger gate, data are transferred from the MENUs/MINDERS to their respective MASTERS, where they are linearized, digitized, attached to channel-identification and time-stamp data, and read out by the data acquisition system (DAQ). Cosmic ray and PMT singles-rate data are acquired for 150 ns whenever the charge on the dynode exceeds a certain threshold (approximately one-third of a photoelectron). This internal trigger entails a ~ 600 -ns dead time that, considering the 1-kHz PMT dark current noise, yields a typical dead time of 0.5% in cosmic ray data. A beam-spill gate (typically 13 μ s, up to 18.83 μ s), an external trigger received from the Fermilab Accelerator Control Network (ACNET) signaling that protons are being delivered, supersedes dynode-triggered data and therefore experiences no dead time.

The significantly lower event rate at the FD makes commercially available components an adequate choice for the front-end electronics [71]. Each M16 is read out by a VA or Viking chip, a widely used ASIC from IDE Corp., Oslo, Norway. Each MUX box has three VA chips (one for each of the three PMTs inside) mounted on a VA front-end board (VFB) located on its exterior. The VFB provides circuitry for power distribution and biasing of the VA chips; it also houses the ASDLite ASIC, which monitors the PMT dynode signals to initiate time-stamping and readout. The analog signals from two VFBs (thus, six VA chips) are multiplexed onto an ADC, which is located on a VA Mezzanine Module (VMM). A VA readout controller (VARC) houses six VMMs and controls digitization, triggering, time-stamping, and bias of the VA chips; thus, each VARC services up to 36 PMTs. The VARC receives the discriminated dynode signal of each PMT; a dynode trigger results in a dead time of 5 μ s. To reduce the total dead time due to dark current noise in the PMTs and spontaneous emission in the WLS fibers, a “2-out-of-36” trigger is employed – the digitization se-

quence is started if a given VARC receives at least two discriminated dynode signals from different PMTs in a 400-ns window. The digitized data are written to onboard memory that is read out by the DAQ.

2.2.5 The Data Acquisition System (DAQ)

The data acquisition systems [72] at the ND and FD are functionally identical, with different front-end software to accommodate the different front-end electronics. The DAQ systems continuously read out the front-end electronics in an untriggered, dead time-free manner, but data are stored to disk only when some trigger condition is met. The system is constructed entirely from commercially available components.

The digitized data in the front-end electronics buffers are read out by a computer – the readout processor (ROP) – in each electronics crate. There are 16 ROPs serving the FD and 8 ROPs serving the ND. Data from the multiple ROPs along the detector are transferred via PCI Vertical Interconnection (PVIC) buses to an off-detector PC called the Branch Readout Processor (BRP). Another set of PVIC buses connects the BRPs to a small farm of trigger processors (TPs). One of the BRPs, acting as a master, selects one of the TPs and instructs the remaining BRPs to transfer their data to that TP. This way, data from the entire detector are available in one TP for processing. The TP applies software triggering algorithms to find events of physics interest. Multiple events gathered by the same trigger from the same input data are grouped into a “snarl.” At the ND, the beam-spill gate trigger from ACNET causes all data to be recorded starting 1.5 μs prior to the arrival of neutrinos at the detector. This trigger is not available at the FD, so a remote trigger is applied; specifically, the GPS system at the ND generates timestamps of the spill signals and transmits them to the FD over the Internet, where they are served to the TPs. The FD DAQ has sufficient buffering capability to allow the remote spill trigger enough time to arrive. Fake spill triggers are generated randomly between spills to provide random sampling of detector activity. Signal-based triggers are used to gather out-of-spill events – for example, the plane trigger requires that four out of five contiguous planes

each contain at least one hit. Finally, a variety of special triggers exist for detector and electronics calibration or debugging.

2.3 Calibration

Each energy deposition, or “hit,” recorded by the DAQ is, in essence, an ordered triple: time of the hit, identifier for the electronics channel in which the hit was recorded, and the pulse height of the hit in ADC (analog-to-digital converter) counts. A collection of these in temporal and spatial proximity constitutes an event. The calibration is a multistage procedure whose ultimate goal is to use this raw information to compute the actual energy deposited in each hit – a quantity that is independent of the hit’s time, strip, position along the strip, and detector. Symbolically, the raw pulse height $Q_{\text{raw}}(t, x, s, d)$ at time t and position x along strip s of detector d is converted to a fully corrected pulse height Q_{corr} by a series of multiplicative correction factors:

$$Q_{\text{corr}} = Q_{\text{raw}} \times D(d, t) \times L(d, s, Q_{\text{raw}}) \times S(d, s, t) \times A(d, s, x) \times M(d), \quad (2.1)$$

each of which is described below.

Drift $D(d, t)$

The drift calibration encompasses changes in the scintillator and WLS fibers caused by temperature variations (short-term) and aging (long-term) as well as time variations in the PMT and electronics responses. Each detector’s response is tracked on a daily basis using through-going cosmic ray muons. Specifically, the total pulse height per scintillator plane of these through-going muons is measured, and the median of these pulse heights is computed on a daily basis. The relative change in this quantity is used as the drift correction:

$$D(d, t) = \frac{\text{Median response}(d, t_0)}{\text{Median response}(d, t)}, \quad (2.2)$$

where t_0 is the beginning of the experiment. Although the energy deposition of

these through-going cosmic ray muons is not the same at each detector and also depends on zenith angle, the average energy deposited at each detector site is expected to be constant with time. From early 2003 to early 2008, the FD drift was $\sim 2\%$ per year. The ND, on the other hand, was stable from early 2005 until the end of Run I but began to drift $\sim 4\%$ per year thereafter. Such long-term drift can be attributed to aging of the detector components, particularly the scintillator.

The LI system is capable of measuring the drift in PMT and electronics responses. Indeed, this was the method used early in the experiment. Each FD strip end was pulsed about 300 times per hour, and every ND strip end was pulsed about 1000 times per hour. The pulse intensity was tuned such that a PMT pixel received approximately 50 photoelectrons per pulse; the PIN photodiodes monitored the pulses to correct for drift in LED intensity over time. The data collected in this manner provided corrections for transient instabilities as well as for changes in gains of PMTs and electronics. However, this approach cannot be used to monitor drift due to changes in the scintillator and WLS fiber and was ultimately superseded by the method using through-going cosmic ray muons described above.

Linearity $L(d, s, Q_{\text{raw}})$

The PMTs become $\sim 5\%$ – 10% nonlinear at light levels of ~ 100 photoelectrons; the FD electronics have a nonlinear response of a similar scale. It is convenient to linearize both components with a single correction, and this is done using the LI system. Once a month at each detector, each scintillator strip–end fiber is pulsed 1000 times at each of many different light levels – a range that ensures that the full dynamic range of interest of the strip’s response is covered. The resulting data are used to parametrize PMT response as a function of true illumination. The linearity correction is determined by extrapolating PMT response in the linear region to the nonlinear region and is applied offline. After the drift and linearity corrections have been applied to the ADC counts, the

pulse heights are said to be in units of “SigLin” (linearized signal).

Strip-to-Strip Uniformity $S(d, s, t)$

The purpose of this calibration is to correct for several detector effects that vary from channel to channel, including scintillator light yield, WLS collection efficiency, readout fiber attenuation, PMT quantum efficiency, and PMT gain. Through-going cosmic ray muon data are used to measure the strip-to-strip (channel-to-channel) variation in detector response. Once the data are linearized, known spatial and angular dependences are removed by applying WLS-fiber attenuation and scintillator path-length corrections such that $S(d, s, t)$ is calculated using the mean response of a muon normally incident to the center of the strip. The calibration relates the mean response of each strip end to the average detector response:

$$S(d, s, t) = \frac{\text{Mean response of detector}(d, t)}{\text{Mean response of strip end}(s, d, t)}. \quad (2.3)$$

Following this correction, which removes variations that can be as large as 30% from the mean detector response, pulse heights are in units of “SigCor.”⁷

Attenuation $A(d, s, x)$

Light is attenuated as it travels along the WLS fiber, so hits closer to the strip readout ends have higher pulse heights than do hits that occur closer to the center of the detector. The necessary corrections are calculated from the “module mapper” measurements conducted just after scintillator module construction. In this procedure, each strip was illuminated every 8 cm along its length with a γ beam from a 5-mCi ^{137}Cs source. These data were then fit to a double exponential:

$$A(d, s, x) = A_1(d, s) \exp\left(-\frac{x}{L_1(d, s)}\right) + A_2(d, s) \exp\left(-\frac{x}{L_2(d, s)}\right), \quad (2.4)$$

⁷Strictly speaking, calibration steps beyond this point are performed during reconstruction (Section 2.5). They are described now rather than later for the sake of coherence.

where the $L_i(d, s)$ are two attenuation lengths and the $A_i(d, s)$ are the amplitudes of the corresponding exponentials. The attenuation correction can also be performed using through-going cosmic ray muon data. A check against attenuation constants obtained this way at the ND, where the high cosmic ray statistics made the computation reliable, found that the two methods agree within about 4%. As a result, attenuation corrections at the ND are performed this way, but at the FD, where the cosmic ray rate is not high enough to obtain sufficient granularity of data along every strip, the module mapper data are used to perform the corrections. The resulting pulse heights are said to be in units of “SigMap.”

Interdetector $M(d)$

At this point, we find that while detector responses *within a detector* are temporally and spatially uniform, they still differ *among detectors*. A relative calibration using stopping muons is used to normalize the response scales at the three detectors. Specifically, we tabulate each detector’s response to muon crossings using only the portion of each muon track in which the muon momentum is between 0.5 and 1.1 GeV, deduced from the distance to the track endpoint. This so-called track window technique [73] avoids using data from the end of the track where dE/dx varies rapidly; in fact, dE/dx varies so slowly over this momentum range that a 2% uncertainty in the position of the muon track’s endpoint produces a 0.2% uncertainty in the energy deposition. For each muon, a correction is applied to account for the muon’s path length in each scintillator plane. The mean response is then calculated for each individual strip, and a single constant $M(d)$ for the entire detector is derived from the median over all strips. This method normalizes the detectors’ calorimetric responses to better than 2%. The resulting quantity Q_{corr} is in units of “MEU” (muon energy unit); based on muon stopping-power tables [74], we find that 1 MEU corresponds to 2.00 ± 0.02 MeV of muon energy loss in scintillator.

Absolute Track and Shower Energy

With all three detectors' responses on an equal footing, we begin to complete the calibration by understanding the detectors' response to muons, electrons, and hadrons below 10 GeV. As mentioned in the introduction to Section 2.2, CalDet was exposed to test beams of pions, protons, muons, and electrons at CERN. Data were collected at fixed beam momentum settings of both positive and negative polarities, covering a range of momenta from 0.2 to 10 GeV. These data were compared with events simulated using the same GEANT3-based [75] Monte Carlo used for the ND and FD. Stopping muon ranges were modeled to better than 3%, and the detector response to electrons was modeled to better than 2% [76]. Pion- and proton-induced showers were simulated with several shower codes and compared to the CalDet data. The simulation based on the GCALOR [77] shower code was found to agree best with the data and was adopted as the default shower code. The Monte Carlo reproduces the response to such showers to better than 6% at all momentum settings [78]. The simulation was able to reproduce the energy resolution, which can be parametrized as $56\%/\sqrt{E} \oplus 2\%$ for hadron showers and $21.4\%/\sqrt{E} \oplus 4\%/E$ for electrons, where the energy E is in GeV.

This completes the “energy branch” of the calibration, which is summarized in Figure 2.20. The “photoelectron branch” is another calibration chain that converts the number of ADC counts to photoelectrons, for use in Monte Carlo simulation, PMT crosstalk identification, strip counting efficiency, and PMT threshold adjustment. This conversion is performed using the LI data described in the “Drift” item above. By comparing the rms widths of many pulses to the mean, the number of ADC counts per photoelectron – i.e., the gain – is found for each channel. The number of ADC counts in a given channel is then divided by the channel's gain to obtain the number of photoelectrons.

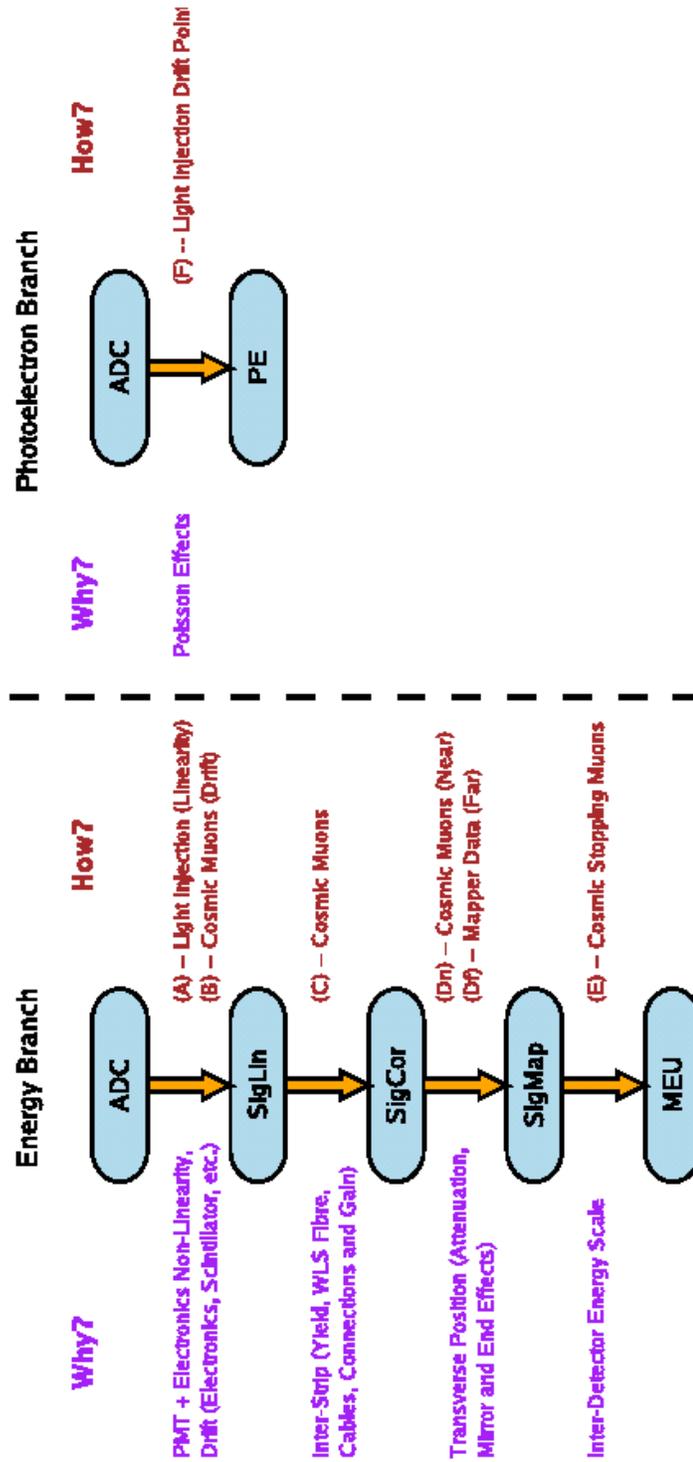


Figure 2.20: Summaries of the two calibration chains used in MINOS. Taken from [79].

2.4 Event Types of Interest

In order to perform a physics analysis with this data, we must be able to synthesize the data – the ADC counts in electronics channels, converted into energy depositions in scintillator strips – into physics objects such as tracks and particle showers. This is the goal of the reconstruction, described in Section 2.5. We shall first pause, however, to consider the variety of such tracks and showers – and combinations thereof – we expect to observe in the detectors.

Though a multipurpose experiment, MINOS was designed and optimized to study ν_μ disappearance in a ν_μ beam, via the $\nu_\mu \rightarrow \nu_\mu$ oscillation mode described in Chapter 1. The event type of interest in such an analysis is a ν_μ charged-current (CC) interaction, in which a ν_μ exchanges a W boson with a nucleon – or a quark inside a nucleon – in a target nucleus (typically iron due to its predominance in the composition of the detectors):

$$\nu_\mu + N \rightarrow \mu^- + X, \quad (2.5)$$

where N represents the nucleon or quark and X represents the resulting hadronic recoil system. The track produced by the muon typically extends far beyond the hadronic shower activity of the recoil system (see Figure 2.21), enabling straightforward identification of such an event as we will discuss shortly. Also, the magnetic field of the detectors enables us to distinguish between a ν_μ CC interaction and a $\bar{\nu}_\mu$ CC interaction:

$$\bar{\nu}_\mu + N \rightarrow \mu^+ + X, \quad (2.6)$$

as the paths of a μ^- and of a μ^+ will curve in opposite directions. Indeed, we exploit this ability to determine whether the oscillation parameters governing the $\bar{\nu}_\mu \rightarrow \bar{\nu}_\mu$ oscillation differ from those governing the $\nu_\mu \rightarrow \nu_\mu$ oscillation [80, 81, 82].

In the analysis presented in this thesis, the event type of interest is a ν_e CC

interaction, in which a ν_e exchanges a W boson with a nucleon or quark:

$$\nu_e + N \rightarrow e^- + X, \quad (2.7)$$

where N and X are as defined previously. In the MINOS detectors, electromagnetic showers have a radiation length⁸ of 4.06 cm – approximately 30% less than the distance between neighboring steel planes. The Molière radius⁹ of these showers is 3.7 cm – approximately 10% less than the width of a scintillator strip. As a result, the electron does not leave a clearly identifiable track but an electromagnetic shower that becomes fully interspersed with the hadronic shower associated with X (see Figure 2.21). The situation is essentially the same in the case of a $\bar{\nu}_e$ CC interaction, in which the resulting positron produces an electromagnetic shower. The challenge, which we will address in Chapter 4, is to distinguish this type of event from the primary background event type – the neutral-current (NC) interaction, in which a neutrino of flavor ℓ exchanges a Z boson with a nucleon or quark:

$$\nu_\ell + N \rightarrow \nu_\ell + X, \quad (2.8)$$

where N and X are as defined previously. The departing neutrino leaves no track, of course; the only activity observed is a hadronic shower. Electromagnetic showers tend to be more compact than most hadronic showers, depositing energy along a dense core. However, hadronic showers can include particles that decay electromagnetically, primarily the neutral pion ($\pi^0 \rightarrow 2\gamma$), confounding the usage of this distinction to discriminate between the two event types. Examples are shown in Figure 2.21.

As discussed in Chapter 1, the dominant oscillation mode in ν_μ disappearance is $\nu_\mu \rightarrow \nu_\tau$, thus creating the possibility of observing ν_τ CC interactions in the Far Detector. The resulting τ^- will decay in one of several ways:

- $e^- \bar{\nu}_e \nu_\tau$ (branching ratio = 17.9% [10]), which will mimic a ν_e CC event,

⁸(a) The mean distance over which a high-energy electron loses all but $1/e$ of its energy by bremsstrahlung and (b) $7/9$ of the mean free path for pair production by a high-energy photon.

⁹The radius of a cylinder centered along the shower axis that contains 90% of the shower's energy; serves as a measure of the transverse size of a shower.

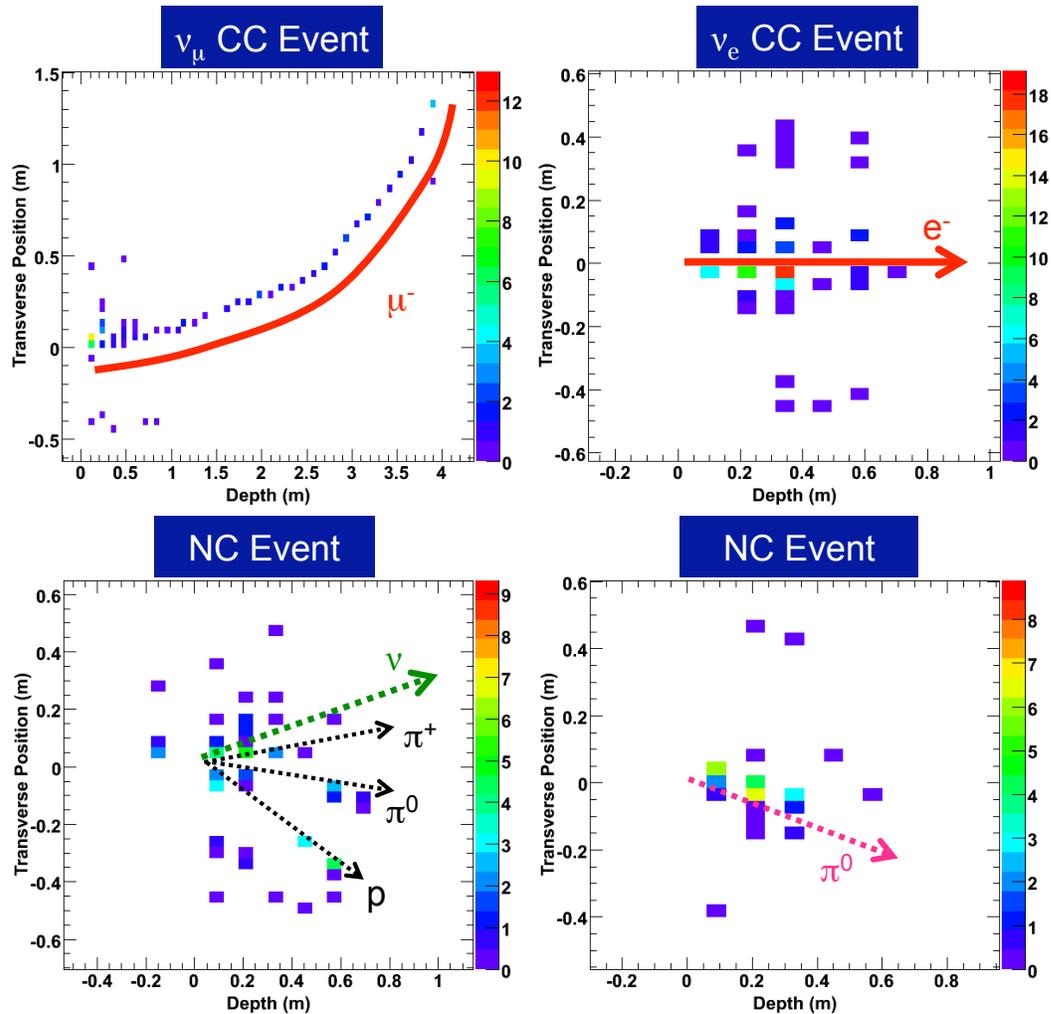


Figure 2.21: Displays of simulated events in the MINOS detectors. Pulse heights (color scale) are given in MEU. (Top left) Note the characteristic long muon track and the hadronic shower activity near the event vertex. (Top right) The electromagnetic shower tends to be compact, depositing energy in a narrow core. (Bottom left) The diffuse hadronic shower of a typical NC event. (Bottom right) When a NC event includes the production of a π^0 that carries a large fraction of the event energy, a correspondingly large fraction of the shower activity will be electromagnetic in nature.

- a low-multiplicity hadronic shower containing at least one π^0 (branching ratio = 36.5% [10]), which could mimic a ν_e CC event due to the electromagnetic shower resulting from $\pi^0 \rightarrow 2\gamma$,
- $\mu^- \bar{\nu}_\mu \nu_\tau$ (branching ratio = 17.4% [10]), which will mimic a ν_μ CC event,
- a high-multiplicity hadronic shower, which will mimic a NC event.

Thus, a ν_τ CC interaction could resemble a ν_e CC interaction approximately 54% of the time. Fortunately, the minimum ν_τ energy needed to initiate a CC interaction is about 3.4 GeV, and the interaction cross section does not become significant until around 5 GeV. At these energies, $P(\nu_\mu \rightarrow \nu_\tau) \lesssim 0.2$, and as a result, ν_τ CC events make a small (but nonnegligible) contribution to the background in this analysis.

2.5 Reconstruction

The reconstruction procedure uses the topology and timing of hits to identify neutrino interactions inside the detector as well as muon tracks from cosmic rays and neutrino interactions in the surrounding rock. The chief goal of the reconstruction is to estimate the visible energy of the interactions described in Section 2.4 as well as provide a set of quantities describing the event that can be used to discriminate among them.

As mentioned in Section 2.2.4, each beam spill can produce multiple neutrino interactions in the Near Detector; the resulting activity in the detector is referred to as a snarl. The first step in the reconstruction is to divide the snarl into one or more events, each of which contains hits that are localized in space and time. Figure 2.22 illustrates how this is possible. Although this step is, in principle, performed for Far Detector data as well, the neutrino flux at the FD is so low that a beam spill rarely produces more than one event; in fact, most spills do not produce any neutrino interactions.

Next, a track-finding algorithm is applied to each event. The algorithm uses a Hough transform [83] to find track segments, which are then chained together (taking

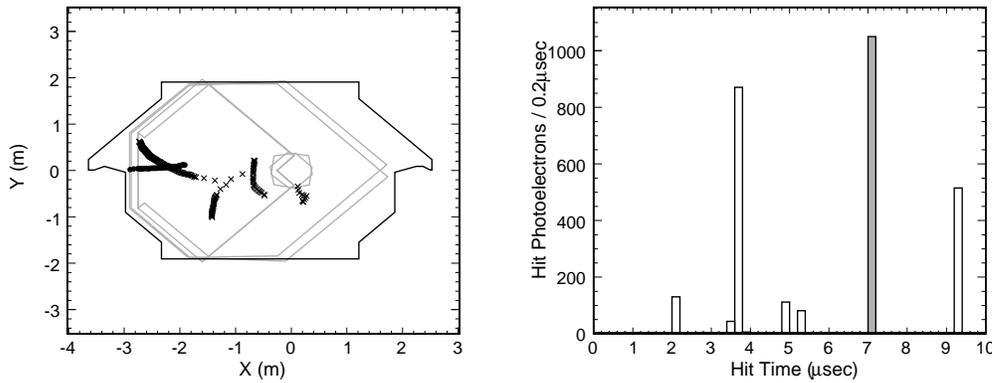


Figure 2.22: An example of a Near Detector snarl resulting from a single beam spill, with detector activity displayed in space from a beam’s-eye view (left) and in time (right). The reconstruction uses both sets of information together to break up the snarl into individual events corresponding to individual neutrino interactions. Taken from [59].

into account timing and spatial correlations) to form longer tracks. If the track stops within the detector, the track momentum is estimated from range using the range tables of [74]; otherwise, the momentum is estimated from the curvature of the track due to the detector’s magnetic field. In particular, the curvature measurement is obtained from fitting the trajectory of the track using a Kalman filter technique [84], which takes into account bending of the track from both multiple Coulomb scattering and the magnetic field. This procedure also provides an indication of the charge of the particle that left the reconstructed track. The momentum resolution for muon tracks produced in beam ν_μ CC interactions is 5% for the range-based calculation and 10% for the curvature-based calculation. It is worth noting that the track-finding algorithm also finds tracks created by other charged particles, such as π^\pm , so most reconstructed events contain at least one reconstructed track.

Showers are constructed from clusters of hits that are localized in space and time. The energy of a shower is computed from the summed pulse height of the individual hits; the pulse height contribution of any reconstructed tracks that share the same hit is subtracted. We must resort to this calorimetric method of calculating a shower’s energy – as opposed to a range- or curvature-based method, like those used for tracks – as the MINOS detectors are not sufficiently fine-grained to reliably distinguish the individual particles contributing to the shower.

2.6 Simulation

As we will see throughout this thesis, the two-detector design of MINOS reduces our dependence on a Monte Carlo (MC) simulation. A reasonably accurate simulation is still necessary, however, to perform key parts of the analysis.¹⁰ The simulation is performed in three stages, as described below.

2.6.1 Simulating the Beam

The NuMI beam is modeled in three steps:

1. a simulation of the hadrons produced by the protons striking the NuMI target,
2. the propagation of those hadrons and their progeny through the focusing horns, decay pipe, and beam absorber, allowing for the decay of unstable particles, and
3. the calculation of the probability that any produced neutrino traverses the Near and/or Far Detectors.

Step 1 is performed using the FLUKA05 [85] Monte Carlo, and particles exiting the target are recorded. In step 2, these particles are propagated in a GEANT3 [75] simulation of the NuMI beam line; this simulation models the magnetic focusing horns, surrounding shielding, decay pipe, and beam absorber. The GEANT-FLUKA code is used to describe hadronic interactions in the beam line and the associated production of secondary particles as well as the full particle decay chains. Decays in which a neutrino is produced are saved and later used as input for neutrino event simulation in the Near and Far Detectors. That simulation uses each of these decays with a probability and neutrino energy determined by the decay kinematics and the randomly chosen trajectory through the Near or Far Detector.

If we look ahead to the complete simulation, we find that the simulated Near Detector ν_μ CC spectrum differs from the observed spectrum by as much as 30%,

¹⁰The author was responsible for generating approximately half of the experiment's Monte Carlo on the computing grid at Caltech.

depending on the neutrino energy, as shown in Figure 2.23. Note, however, that the discrepancy between data and simulation moves with the spectrum’s focusing peak and is thus intimately linked to the beam configuration, suggesting that the problem lies with the beam model and not with the modeling of the detector or of neutrino interactions in it. This is not surprising, considering that secondary hadron production in the NuMI target is not strongly constrained by external data.

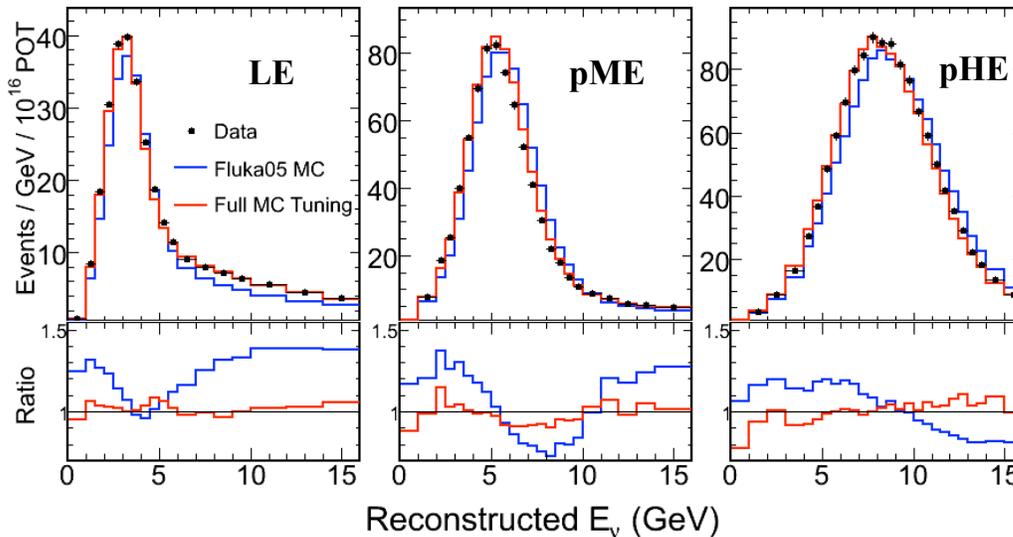


Figure 2.23: Reconstructed neutrino energy spectra in the LE, pME, and pHE beam configurations. The blue curve is the spectrum obtained by using the FLUKA05-based simulation as-is, and the red curve is the spectrum obtained after tuning the simulation.

Though not necessary for a two-detector oscillation analysis, the data are easier to analyze if the flux simulation is tuned to better describe the observed spectrum. To improve the data–MC agreement, the Near Detector data are used to constrain the beam model [86]. Specifically, we warp the flux as a function of the transverse and longitudinal momenta – p_T and p_z , respectively – of the neutrino-parent hadrons that come off the target. The warping is calculated by performing simultaneous fits to data taken in multiple beam configurations – the usual LE, pME, and pHE, as well as data taken with the focusing horns off and in LE mode but with horn current at 170 and 200 kA; doing so allows for a better tuning to the underlying hadron production, as each beam configuration samples a different region of the parent hadron p_T – p_z

space. Figure 2.23 shows that this tuning procedure, dubbed the “SKZP” technique,¹¹ results in much better agreement between data and MC, and this tuned MC is used throughout the analysis presented in this thesis.

2.6.2 Simulating Neutrino Interactions

Neutrino interactions are modeled by the NEUGEN-V3 [87] program, which simulates both quasi-elastic and inelastic neutrino scattering. NEUGEN simulates the hadronization process with the AGKY model [88], which uses PYTHIA/JETSET [89] at high hadronic invariant mass, a phenomenological model based on KNO scaling [90] at low hadronic invariant mass, and a smooth transition between the two regimes that ensures the continuity of all simulated observables as a function of the invariant mass. Hadrons produced in the neutrino scattering are allowed to interact while exiting the target nucleus; this calculation is performed using the INTRANUKE [91] intranuclear cascade model, which incorporates pion elastic and inelastic scattering, single charge exchange, and absorption.

2.6.3 Simulating the Detectors

The MINOS detector simulation, known as GMINOS, is based on GEANT3 and is used to generate raw energy depositions in the scintillator strips that serve as the input to our detector response model. The simulation randomly samples neutrinos from the flux predicted by the beam simulation (Section 2.6.1) and traces them through the Near and Far Detector halls. Events are generated inside the detectors as well as in the surrounding support structure and rock. The simulation includes a detailed geometric model of the detector that describes the material crossed by neutrinos and neutrino-induced tracks to within 1% plane-to-plane and 0.3% averaged over the detector. The position of individual scintillator strips was determined with a precision of approximately 1 mm using cosmic ray tracks. The magnetic field is modeled via finite element analysis driven by bench measurements of the steel B - H curve.

¹¹This is an acronym formed from the initials of the method’s primary creators: Sacha Kopp and Žarko Pavlović.

Figure 2.24 summarizes what happens next. First, a date from the period considered in the analysis is randomly generated, and the calibration constants associated with that date are retrieved from the database. Using these constants, the Photon-Transport software package “decalibrates” the energy deposited in each scintillator strip into MIPs (a specially defined unit of energy equal to the energy deposited by a minimum ionizing particle in a single plane) and ultimately into photons arriving at a PMT pixel. Then, using a Poisson random number generator, the number of photoelectrons produced by these photons is simulated. Note the impact of this step: without it, the calibration procedure would precisely undo the decalibration. Finally, the number of photoelectrons is passed to the DetSim software package, which simulates the PMTs and front-end electronics, producing a number of ADC counts that is recorded to disk in a manner similar to that of the DAQ. This procedure thus incorporates actual measurements of the detectors’ calorimetric response, WLS fiber attenuation, strip-to-strip variations, drift, nonlinearity, and gains, thereby producing simulated data that are faithful to the time-dependent properties of the detectors.

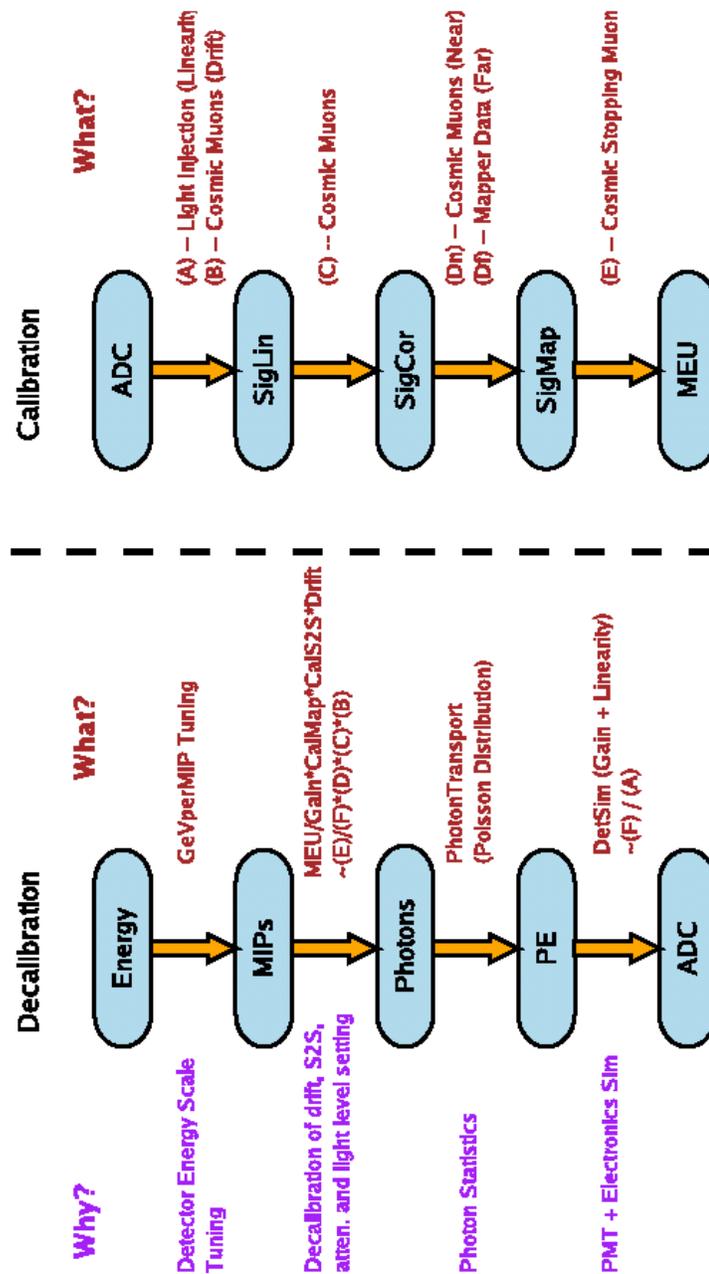


Figure 2.24: The decalibration chain (left) used in converting simulated energy depositions into ADC counts. The capital letters (A) through (F) refer to the constants associated with the calibration steps shown in Figure 2.20. Note that without the randomness involved in the “Photons \rightarrow PE” step, the calibration chain (right) would precisely undo the decalibration. Taken from [79].

Chapter 3

Overview of the Analysis

Ultimately, the goal of this analysis is to make a statement about the θ_{13} neutrino mixing angle – either to measure it or to improve the existing upper limit on its value. We recall from Section 1.3.3.2 that the $\nu_\mu \rightarrow \nu_e$ oscillation mode is governed to leading order by an oscillation probability proportional to $\sin^2 2\theta_{13}$, namely:

$$P(\nu_\mu \rightarrow \nu_e) \approx \sin^2 2\theta_{13} \sin^2 \theta_{23} \sin^2 \frac{1.27 \Delta m_{32}^2 L}{E}, \quad (3.1)$$

where Δm_{32}^2 is the atmospheric mass splitting in units of eV^2 , L is the distance in kilometers between the neutrino production and detection points, and E is the neutrino energy in gigaelectronvolts. The MINOS experiment is well positioned to probe this oscillation mode – and thus achieve the goal of this analysis – by searching for the appearance of an excess of ν_e in the primarily ν_μ NuMI beam as these neutrinos travel from the Near Detector (ND) at Fermilab to the Far Detector (FD) in the Soudan mine in northern Minnesota.

Recall from Section 2.4 the discussion of the various event types we see in the detectors. An excess in the number of ν_e CC events observed (that is, beyond the expected number) at the FD would manifest the excess of ν_e mentioned above and thus constitutes the signal in this analysis.¹ The identification of such events proves to be a challenge, as a variety of background events also occur in the detectors. A large

¹Note that the model under consideration assumes that these signal ν_e CC events arise from oscillations driven almost exclusively by Δm_{32}^2 . As such, these events are expected only in the FD, and only background events occur in the ND.

fraction of the data sample consists of events that are clearly not signal-like and can be eliminated with simple cuts on certain event properties, as discussed in Section 4.2. The remaining events are considered signal-like, as they exhibit the desired signal ν_e CC event topology. Background events found in this sample include

- NC events, particularly those containing one or more π^0 s, as this particle decays into two photons, which in turn produce electromagnetic showers;
- ν_μ CC events in which a large fraction of the neutrino’s energy has been transferred to the hadronic recoil system, resulting in a muon track buried wholly or mostly within the shower;
- ν_e CC events due to ν_e s intrinsically present in the beam (see Table 2.1);
- ν_τ CC events (in the FD only) due to ν_τ produced by the (dominant) $\nu_\mu \rightarrow \nu_\tau$ oscillation mode.

We discriminate among these and any possible remaining signal events using a sophisticated nearest-neighbors technique, developed at Caltech and refined in this thesis, called Library Event Matching (“LEM”), described in Section 4.3. As a result of this final step, a discriminant ranging from 0 to 1 (the “LEM PID”) is calculated for each remaining event to quantify how signal-like it is.

There are four event samples with which to work in this analysis: ND simulation, ND data, FD simulation, and FD data. As we have already noted, the possible signal we seek would be found in the FD data, collected over the run periods listed in Table 2.2 and representing a total exposure of 8.16×10^{20} PoT. In order to avoid any bias in designing the analysis, we follow a data blinding protocol whereby the FD data is kept off-limits until all aspects of the analysis have been finalized. We use the remaining three samples in conjunction to obtain a data-driven prediction of what we expect to observe in the FD. If we carefully consider the nature of the four types of background listed above, as well as the nature of the signal itself, we find an intricacy for which we must (and do) account in the prediction procedure: each event type is (or, in the case of NC events, is not) associated with a particular neutrino

flavor. Each neutrino flavor, in turn, is associated with different oscillation modes, each of which has a probability governed in some characteristic way by the various oscillation parameters.

Of the five event types whose rates we must predict in the FD, three are present in the ND as well: NC, ν_μ CC, and intrinsic ν_e CC. In fact, these are the *only* types of events we expect to observe in the ND. In Chapter 5, we describe two methods of decomposing the total number of events $N_{\text{total}}^{\text{data}}$ in the ND data into the respective event counts $N_{\text{NC}}^{\text{data}}$, $N_{\nu_\mu \text{ CC}}^{\text{data}}$, and $N_{\text{beam } \nu_e}^{\text{data}}$ of each of those three background components. Next, we predict the FD event count F_X^{pred} of each background component X via the following formula:

$$F_X^{\text{pred}} = N_X^{\text{data}} \times \frac{F_X^{\text{MC}}}{N_X^{\text{MC}}}, \quad (3.2)$$

where F_X^{MC} and N_X^{MC} are the number of simulated (Monte Carlo, or MC) events of type X in the FD and ND, respectively. Furthermore, the software infrastructure used to compute these predictions includes the flexibility to vary the oscillation parameters governing the transition from ND to FD, thus effecting the differing impacts of said transition on the various background components. The remaining two event types – ν_τ CC and signal ν_e CC – are the result of $\nu_\mu \rightarrow \nu_\tau$ and $\nu_\mu \rightarrow \nu_e$ oscillations, respectively. As such, they are expected in the FD only, with no analogues in the ND. Predictions for these event types are based on a sample of clearly identified ν_μ CC events in the ND data; the formula used to compute these predictions is conceptually similar to – but necessarily more sophisticated than – Equation 3.2.

The method we have outlined above, referred to as “extrapolation” and detailed in Chapter 6, is thus a versatile, data-driven technique for predicting event counts in the FD that frees us from relying solely or extensively on the Monte Carlo. In fact, a number of systematic effects (e.g., neutrino flux modeling, neutrino interaction cross section uncertainties, hadronization model uncertainties, and certain energy calibration uncertainties) are common to both detectors, and the induced systematic errors cancel to first order in the ratio $\frac{F_X^{\text{MC}}}{N_X^{\text{MC}}}$, thereby reducing the sensitivity to the details of the simulation. These and other systematic effects, as well as the uncertainties they

induce in the FD predictions, are discussed in detail in Chapter 7.

The techniques presented thus far enable us to tackle the goal of this analysis directly. We begin by asking the following question: If we happen to observe precisely the number of events predicted for $\theta_{13} = 0$, what upper limit (specifically, the 90% confidence level upper limit) can we set on the value of θ_{13} ? We refer to this upper limit as the “sensitivity” of the analysis and use it as a measure of the analysis’s strength. To compute it, we begin by generating two sets of FD predictions – one for each neutrino mass hierarchy – over a two-dimensional grid of the oscillation parameters being studied, namely θ_{13} and the CP-violating phase δ_{CP} . We then perform a statistical comparison, described in Section 8.1, between the hypothetical FD observation and the prediction at each grid point to determine the likelihood that the former is a statistical and/or systematic fluctuation of the latter.

At this point, we could proceed to unblind the FD data. However, to reassure ourselves further that everything is functioning as desired, we run the analysis in its entirety on two auxiliary data samples – one that is virtually signal-free and another that is signal-free by construction. Satisfied with the results of these “sideband” analyses, discussed in Section 8.2, we finally proceed to unblind the FD data in Section 8.3 and compute the confidence intervals implied for θ_{13} in Section 8.4.

Thus far, we have not indicated how events are classified, if at all, in this analysis. The oscillation probability responsible for the signal we seek is, like all other oscillation probabilities, energy-dependent, so it is wise to bin events by reconstructed energy. Also, performing the analysis on a high-LEM-PID subset of the data is preferable to analyzing the full data set, as signal events will tend to have values of LEM PID higher than those of background events. The optimization process performed to determine the optimal binning in reconstructed energy and LEM PID employs the full machinery of the analysis as detailed in this thesis. In the interest of presenting this analysis in a pedagogically coherent manner while keeping the objective of the analysis in clear sight, we describe the optimization process in detail in Appendix B, summarize the results here, and proceed with these results as given. We define the sensitivity at $\delta_{\text{CP}} = 0$ for the normal mass hierarchy as the figure of merit for the

optimization process. Despite the argument made above, we find that binning in reconstructed energy has a negligible effect on the figure of merit, but we choose to use bin edges of 1, 2, 3, 4, 5, and 8 GeV for reasons discussed in Section 8.3. As for LEM PID, the optimal single-bin scheme is $0.7 < \text{LEM PID} < 1.0$ (i.e., the optimal cut on LEM PID is 0.7), and the optimal multiple-bin scheme is $0.6 < \text{LEM PID} < 0.7$, $0.7 < \text{LEM PID} < 0.8$, and $0.8 < \text{LEM PID} < 1.0$.² We use the former for simple PID performance studies and the latter for computing the final confidence intervals for θ_{13} , as the multiple-bin scheme provides a considerable improvement in the figure of merit relative to the single-bin scheme. Thus, our full analysis – background decomposition, extrapolation and Far Detector prediction,³ and likelihood statistic calculation – is performed in 15 “analysis bins” (5 reconstructed energy bins by 3 LEM PID bins).

²Throughout this thesis, we refer to $\text{LEM PID} > 0.6$ as the “analysis region” of LEM PID.

³Each quantity in Equation 3.2 – N_X^{data} , F_X^{MC} , and N_X^{MC} – is binned.

Chapter 4

Event Selection

We obtain the event sample for this analysis in three stages from the data collected in the runs listed in Table 2.2. First, we reject all events that occur outside of a fiducial volume. Next, we make simple cuts on a variety of event properties to exclude events that are very likely to be background events, thus obtaining a “ ν_e -like” sample. Finally, we apply an event identification algorithm (essentially of the nearest-neighbors variety) that quantifies how signal-like each remaining event is; we can then examine the distribution of these events with respect to the algorithm’s output.

4.1 Fiducial Volume

The fiducial volume cut requires that the vertex of an event lies within a specified region of the relevant detector. The primary goal of imposing such a cut is to ensure that all activity associated with an event in a given detector is fully contained within that detector. Other, detector-specific considerations also impact the choice of fiducial volume boundaries.

The ND fiducial volume occupies a relatively small fraction of the detector’s full volume. A cylinder of radius 0.8 m, it is roughly centered horizontally between the western edge of the detector and the coil hole and vertically between the top and bottom of the detector. Longitudinally, the fiducial volume is fully contained within the calorimeter (fully instrumented) region. As discussed in Chapter 2, the ND was

designed to take advantage of the high neutrino flux at this location – specifically, to allow the definition of a small fiducial volume, uninterrupted by the magnetic coil hole, that will collect enough beam neutrino events to make feasible the extrapolation process outlined in Chapter 3.

The dramatically reduced neutrino flux at the FD motivates the choice of a fiducial volume that is as large as possible, without sacrificing event quality. Indeed, the fiducial volume is an annular cylinder of inner radius 0.50 m (thus excluding events with vertices in or near the coil hole), outer radius 3.74 m (thus excluding events with vertices less than 0.26 m from the detector edge), and length 13.80 m in the first supermodule (thus excluding events with vertices less than about 0.50 m from either end) and 11.71 m in the second supermodule (thus excluding events less than 0.34 m from the upstream end or less than 2.05 m from the downstream end).

4.2 ν_e -like Preselection

Of the events passing the fiducial volume cut, only those that are in time with a beam spill passing various quality cuts are kept. These remaining events are then subjected to a suite of “preselection” cuts, as follows:

- If there are multiple events within a beam spill window, the event with the highest total pulse height (i.e., the greatest energy) is kept and the rest are discarded. This is only done in the FD, as the event rate there is low enough that we expect no more than one beam event per spill; thus, any smaller reconstructed events within the spill window are rejected as they are likely artifacts of the reconstruction (e.g., event fragments).
- The event must contain at least one reconstructed shower. An event lacking showerlike activity is unlikely to be the result of a ν_e CC interaction, which consists of hadronic and electromagnetic showers.
- To suppress cosmogenic background, we discard any event containing

1. a track whose extent in the y (vertical) direction exceeds 2 m or whose y direction cosine is less than 0.6, or
 2. a shower that is too steep to be beam-induced.¹
- The event must contain at least five contiguous planes in each of which there is at least half as much energy as would be deposited by a minimum-ionizing particle. This requirement serves to eliminate sparse NC events – the topologies of such events differ noticeably from the expected ν_e CC event topology.
 - The reconstructed energy of the event must be between 1 and 8 GeV. While there could be a small number of signal ν_e CC events below 1 GeV, events with such energies are overwhelmingly NC and are all discarded. Based on the discussion in Section 1.3.2.2 and the oscillation probability provided in Section 1.3.3.2, we know that the $\nu_\mu \rightarrow \nu_e$ oscillation probability is negligible above 8 GeV,² so events with such energies are also discarded. See Figure 4.1.
 - To suppress obvious ν_μ CC background, we reject any event with a reconstructed track that is more than 24 planes long OR that extends more than 15 planes beyond any associated showerlike activity near the track vertex. The former are referred to as “track planes,” while the latter are referred to as “tracklike planes.” See Figure 4.1.

Clearly, these preselection cuts take advantage of various dissimilarities between signal and background events. We note in addition that while *all* events in the ND are background, these preselection cuts are applied nonetheless in the ND (unless explicitly indicated otherwise) to keep the ND and FD event samples on an equal footing.

¹More precisely, we calculate the pulse-height-weighted average transverse position of activity in each plane and perform two simple linear fits – one to average transverse position in U-view planes versus z and another to average transverse position in V-view planes versus z . If the sum of the squares of the slopes of these best-fit lines exceeds 10, the shower is considered too steep to be beam-induced.

²Such events are in fact useful for probing the large- Δm^2 region to search for nonstandard interactions, but that is beyond the scope of this thesis.

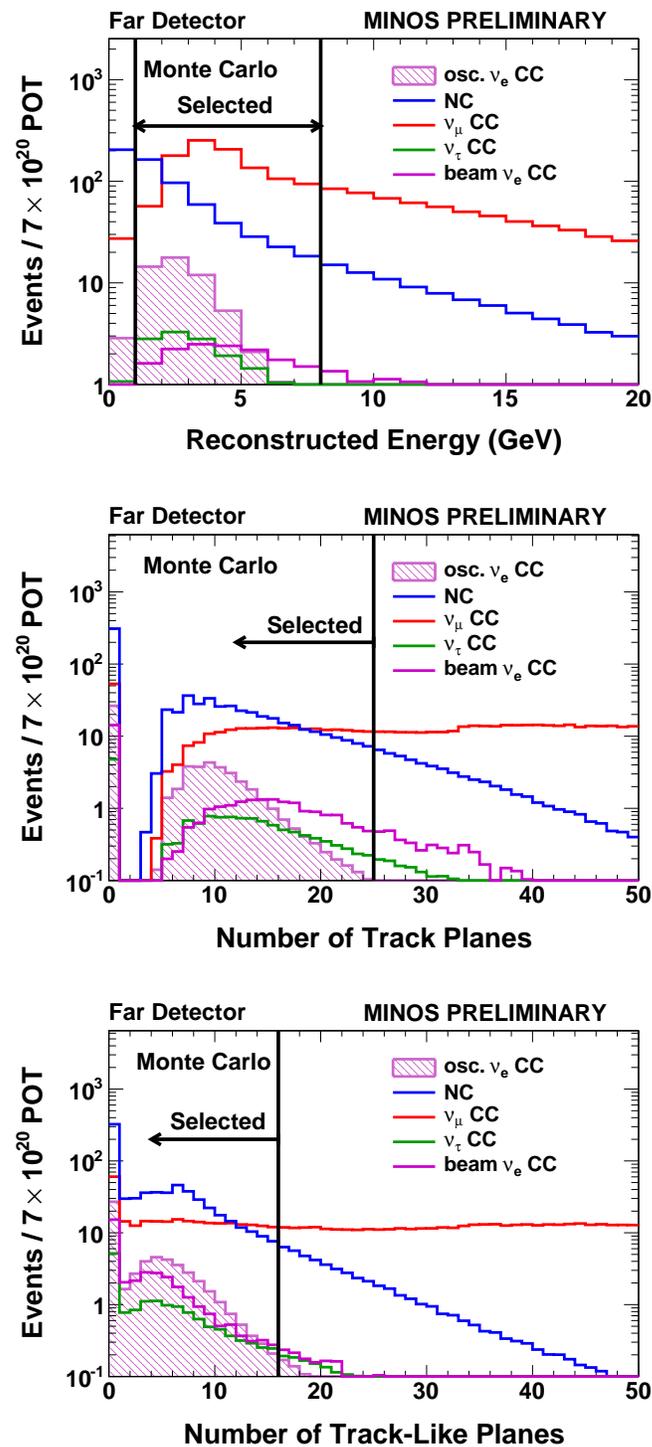


Figure 4.1: Impact of the reconstructed-energy and track-length preselection cuts on all event classes in the FD simulation. These plots were made for the analysis reported in [57], and since the preselection cuts have not changed since that analysis, the plots simply need to be scaled by an overall PoT ratio (8.2/7, the ratio between the PoT exposures in this analysis and the previous one) for this analysis.

As estimated by the FD simulation, the preselection cuts ultimately preserve 76.7% of signal ν_e CC, 38.7% of NC, 8.5% of ν_μ CC, 53.8% of ν_τ CC, and 34.6% of intrinsic ν_e CC events occurring in the fiducial volume.

4.3 Library Event Matching

As we have seen, the preselection removes a sizeable fraction of the background, but we find that the signal-to-background ratio remains low. Further improvement requires the use of a more sophisticated event identification technique that provides greater power to discriminate between signal and background. Previously [56, 57], the MINOS ν_e appearance search used a multivariate technique – the so-called Artificial Neural Network (ANN) – based on reconstructed quantities characterizing the longitudinal and transverse energy deposition profiles of an event, exploiting differences between the values of these quantities in signal and background events. This analysis uses a novel technique named Library Event Matching (LEM), first described by Ochoa in [92], that uses raw energy deposition information instead of reconstructed quantities and is based on a pattern matching algorithm. In comparison to ANN, this approach is optimal in that the event is analyzed “as-is,” discarding no information through the summarization process that is the reconstruction. We will describe a modified version of LEM that keeps the philosophy and essential architecture intact but dramatically improves its performance, sensitivity to systematic effects, and data–MC agreement.

We begin by comparing a given candidate event, using a likelihood method described in Section 4.3.3, to each of $\sim 5 \times 10^7$ simulated events – the “library” – of which $\sim 3 \times 10^7$ are background events and $\sim 2 \times 10^7$ are signal events; we discuss in Section 4.3.1 how the size of the library and the optimal mixture of signal and background events was determined. Intuitively, the comparison determines the extent to which the topologies and individual hit pulse heights of the candidate and library events are similar. Figure 4.2 provides an illustration of this idea, and in subsequent sections, we will discuss how this is actually executed.

Once this matching process is complete for a given candidate event, the library

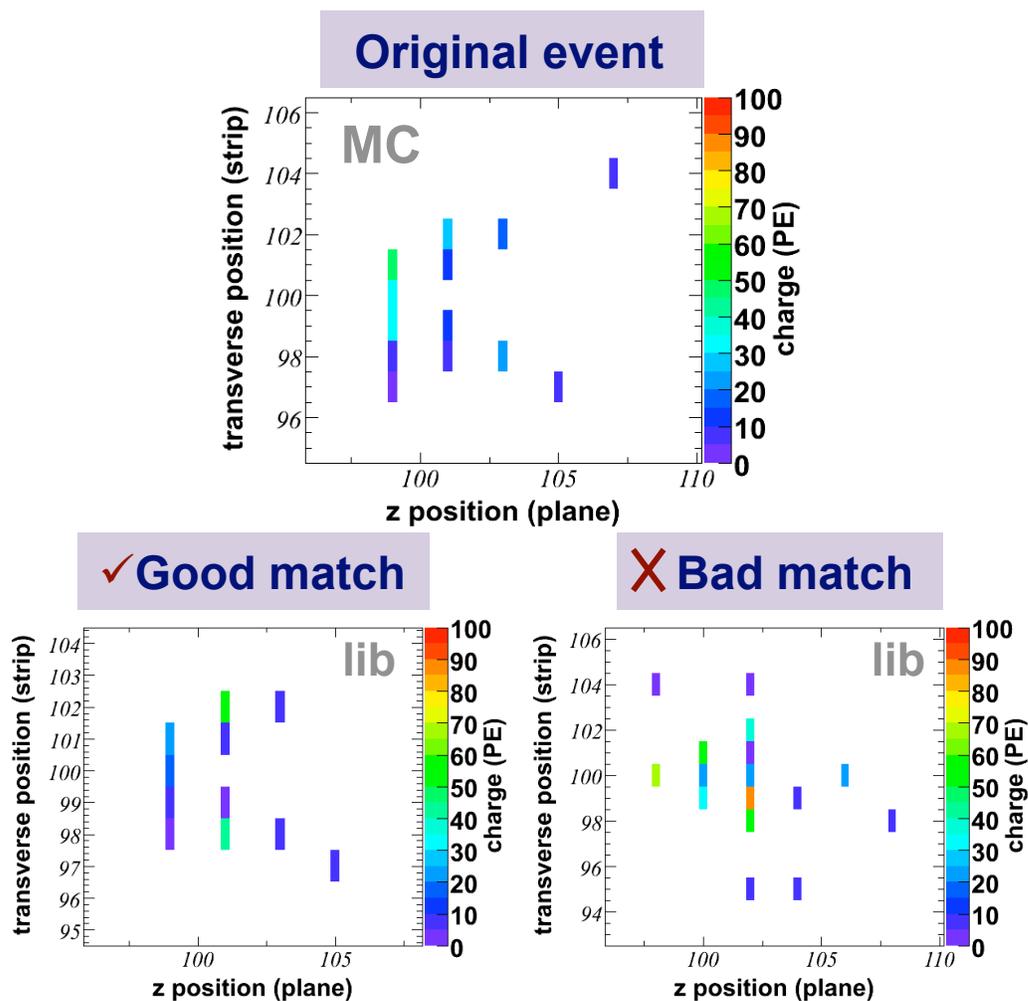


Figure 4.2: The original event is a simulated ν_e CC event. The good match is in fact the library event that best matches the original event as determined by the LEM algorithm, while the bad match is one of the many library events whose patterns do not match the pattern of the original event at all. Taken from [92].

events are ranked by how well they matched the candidate event. The top 50 of these, designated the “best matches,” are summarized by a set of three variables that, along with the reconstructed energy of the candidate event, are given as inputs to a neural network. The output of the neural network is the “LEM PID”³ (alternately referred to as the “LEM discriminant”), a number between 0 and 1 that is used to discriminate between signal and background. Figure 4.3 summarizes the process.

As we will see soon, certain free parameters in LEM were optimized by maximizing a figure of merit (*FOM*), defined as

$$FOM = \frac{N_{\text{sig}}}{\sqrt{N_{\text{bg}}}}, \quad (4.1)$$

where N_{sig} and N_{bg} are the number of signal and background events, respectively, with LEM PID above some cutoff value (so-called selected events) expected in the Far Detector assuming some nonzero value of θ_{13} . Thus, the *FOM* compares the significance of a signal relative to statistical fluctuations in the background. This simple optimization technique, as compared to performing the full analysis to optimize each free parameter, is quick and time-effective yet dependable.

4.3.1 The Library

The library introduced above is a Far Detector Monte Carlo sample generated exclusively for this purpose. In particular, the Monte Carlo version used to generate the library events and the reconstruction code version used to reconstruct them are identical to those used to generate and to reconstruct the Monte Carlo samples used in the analysis.⁴ Signal events are exclusively ν_e CC events, and background events are exclusively NC events. Though the latter may seem presumptive, it is well justified by two observations: (1) the vast majority of the background remaining upon

³Technically, PID is an acronym for “particle ID,” and “event ID” would be the more appropriate term in this case. However, “PID” is a commonly used term that conveys the desired meaning, so we will continue to use it.

⁴The entire process took about six months. Monte Carlo generation was performed at the Rutherford Appleton Laboratory on the WLCG Grid [93], and reconstruction was performed on the Fermilab Computing Grid [94].

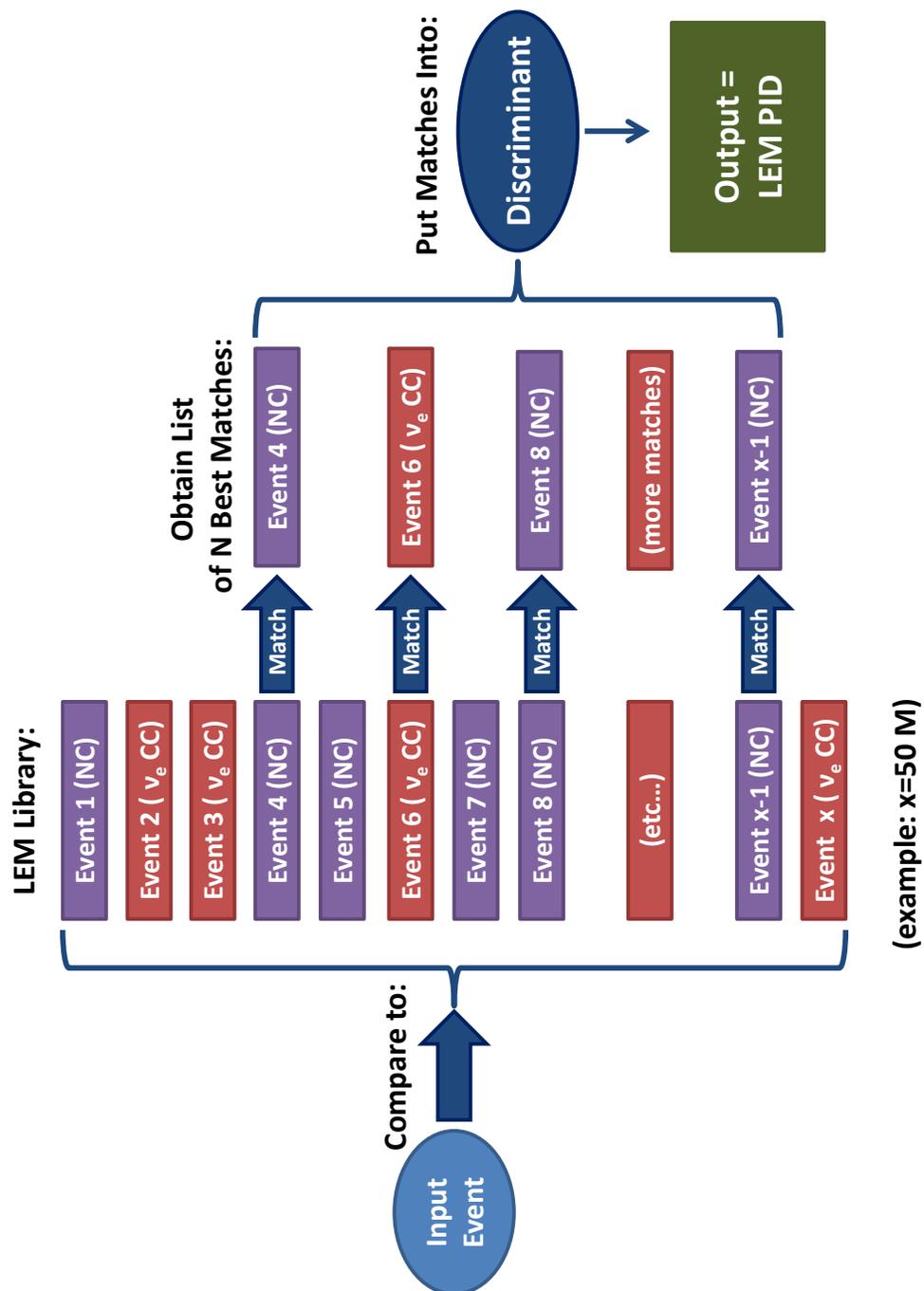


Figure 4.3: Overview of the LEM procedure. Courtesy of Ruth Toner.

application of the preselection consists of NC events and (2) the second largest source of background are ν_μ CC events with short tracks buried within the hadronic shower near the vertex and strongly resemble NC events [95]. As for the remaining ν_e CC and ν_τ CC backgrounds, the former constitutes an irreducible background that differs from the signal only in its energy spectrum, and the latter can resemble either signal or background with approximately equal probability, as discussed in Chapter 2. Fortunately, as we will see in Chapter 6, ν_e CC and ν_τ CC events contribute slightly and negligibly, respectively, to the total predicted observation in the Far Detector.

Keeping all other parameters constant, the *FOM* increases with the total number of library events and is maximized by setting the fraction of signal events in the library to ~ 0.3 – 0.4 [92], as indicated by Figure 4.4. Following these findings, we increased the size of the library by 67% relative to the original implementation of LEM while maintaining the prescribed fraction of signal events, as mentioned above.

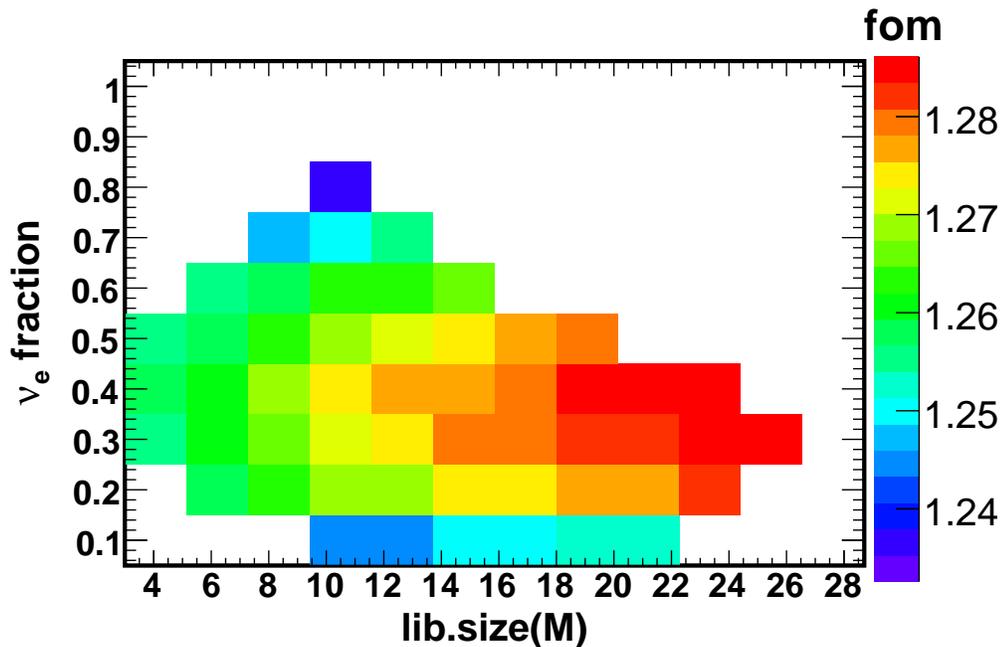


Figure 4.4: Figure of merit (*FOM*) as a function of library ν_e fraction and total library size. Taken from [92].

We note here that because the library events are generated by the standard MINOS Monte Carlo (i.e., randomly), many of them lack the properties that would make them

useful in the matching process. Thus, only those library events meeting the following requirements are actually kept in the library:

- contained within the fiducial volume,
- $5 < N_{\text{planes}} < 30$,
- $7 < N_{\text{strips}} < 80$,
- $N_{\text{strips}}^U > 2$,
- $N_{\text{strips}}^V > 2$, and
- $50 < Q_{\text{tot}} < 3000PE$,

where N_{planes} is the number of planes traversed by the event, N_{strips} is the number of strips hit in the event, N_{strips}^U and N_{strips}^V are the number of U - and V -view strips hit in the event, and Q_{tot} is the total pulse height of the event (i.e., the sum of the pulse heights of all of the event's hits) in photoelectrons. This special preselection decreases processing time and maintains the quality of the comparisons. Approximately 24 million reconstructed events remain; these constitute the LEM library.

4.3.2 Preparing Events for Matching

LEM is technically challenging and computationally intensive, as it involves a very large number of comparisons that must be performed quickly and efficiently without overwhelming the available processing and memory resources. Due to the technique's decreased reliance on the reconstruction, events must first be placed on an equal footing with respect to interdetector and intradetector variations – a task normally performed by the calibration during reconstruction. These, in addition to a few performance-improving processing steps, are described below.

4.3.2.1 Pulse Height Measurement

One of the key improvements on the original implementation of LEM, implemented in this thesis, is a change in the method of calculating a hit's pulse height as it is to

be used by the matching algorithm.⁵ Originally, LEM used pulse heights expressed in units of photoelectrons, as calculated in the photoelectron branch of the calibration described in Chapter 2, corrected for attenuation along the optical fibers and for interdetector differences in light level.⁶ This branch of the calibration is totally dependent on PMT gains calculated from light injection data; systematic uncertainties on these gains, discussed in Chapter 7, and LEM’s reliance on pulse heights calculated using them led to an unacceptably large systematic error on the predicted number of background events in the Far Detector. The version of LEM used in the present analysis addresses this problem by using pulse heights derived from the better-understood energy branch of the calibration. Specifically, the pulse height of a hit measured in SigCors (see Section 2.3) in a given detector is multiplied by the average photoelectron-to-SigCor ratio for individual hits in that detector (0.01286 for the Far Detector, 0.00802 for the Near Detector – see Appendix A). Thus, the pulse height is in photoelectron-sized units and is on an approximately equal footing with the pulse height in true photoelectrons, allowing us to use the remainder of the original LEM architecture without modification, but the dependence on the photoelectron branch of the calibration is dramatically reduced, significantly decreasing the aforementioned systematic error as we shall see in Chapter 7.

4.3.2.2 Attenuation and Interdetector Corrections

Recall that at the SigCor level, pulse heights have been corrected for drift, nonlinearity, and strip-to-strip differences. After scaling these pulse heights as described in the previous section, we perform the remaining corrections – for attenuation and for interdetector differences – by the same method used in the original implementation, which we describe here.

Part of the attenuation correction in LEM is applied using a function identical in form to Equation 2.4, with the key distinction that the attenuation lengths $L_i(d, s)$

⁵Investigations of several possible improvements, including the one ultimately adopted, were performed by the author and are discussed in Appendix A.

⁶This simply means that, even after correcting for attenuation, the same particle interactions produce, on average, different numbers of photoelectrons per channel at the Near and Far Detectors.

and the exponentials' amplitudes $A_i(d, s)$ are replaced by constants determined from a phenomenological fit to the measured attenuation [96] – that is, the same values are used for all strips in both detectors:

$$a_w(d_w) = \frac{2}{3} \exp\left(-\frac{d_w}{7.05\text{m}}\right) + \frac{1}{3} \exp\left(-\frac{d_w}{1.05\text{m}}\right), \quad (4.2)$$

where d_w is the distance traveled by light in the WLS fiber. An additional correction is applied for attenuation in the clear optical fibers leading from the scintillator modules to the PMT enclosures:

$$a_c(d_c) = \exp\left(-\frac{d_c}{10.0\text{m}}\right), \quad (4.3)$$

where d_c is the distance traveled by light in the clear fiber.

In the Far Detector, light is read out at both the east and west ends of a strip: Q_{raw}^E and Q_{raw}^W . Each of these is corrected for the attenuation experienced to obtain Q^E and Q^W ; their sum Q represents the total light actually produced at the interaction point in the strip. It is important, however, to keep the attenuation-corrected light levels on the same overall scale as the raw light levels, as the comparison metric defined in Section 4.3.3 uses Poisson distributions to model fluctuations in the number of photoelectrons generated by a PMT for a given amount of incident light. Thus, Q is scaled down by a factor $A^{FD} = 0.326$ to obtain Q_{final} :

$$Q_{\text{final}} = A^{FD} \left(\frac{Q_{\text{raw}}^E}{a_w(L - x + w^E)a_c(c^E)} + \frac{Q_{\text{raw}}^W}{a_w(x + w^W)a_c(c^W)} \right), \quad (4.4)$$

where L is the total strip length, x is the distance from the interaction point where the light is deposited to the west end of the strip, w^E (w^W) is the length of the WLS fiber connecting the east (west) end of the strip to the east (west) end of the scintillator module, and c^E (c^W) is the length of the clear fiber connecting the east (west) end of the scintillator module to the PMT enclosure. This A^{FD} factor scales light levels throughout the Far Detector, which are now flat as a function of position along a strip, to the raw light level of events occurring near the end of a Far Detector strip, as shown in Figure 4.5.

The single-ended readout in the Near Detector necessitates a different correction function. We make the approximation that light generated at the interaction point in a strip is divided into two equal portions – a west-bound portion that travels directly to the module end and an east-bound portion that is reflected at the WLS fiber’s mirrored end and travels back down the entire length of the fiber to the module end. The necessary attenuation corrections are applied to each portion, and their sum is scaled by A^{FD} . Figure 4.5 shows, however, that the resulting light level is much lower than that in the Far Detector. An interdetector correction factor $\beta = 1.59$ is computed by requiring the simulations of muons in the detectors to match on average and is applied to the corrected Near Detector light. The fully corrected pulse height is thus given by

$$Q_{\text{final}} = \beta A^{FD} \frac{2Q_{\text{raw}}^W}{[a_w(x + w^W) + a_w(2L - x + w^W)R]a_c(c^W)}, \quad (4.5)$$

where all variables are the same as defined for Equation 4.4 and $R = 0.8$ is the reflectivity of the strip’s east end. Events in both detectors are now on an equal, detector- and position-independent footing, and comparing input events from either detector to the Far Detector Monte Carlo–based library events is a reasonable thing to do.

4.3.2.3 Compacting an Event

The “compacting” procedure, whereby an event is prepared for the matching process, is motivated by the need to keep processing and memory usage within available resources while also having the algorithm achieve its purpose. We proceed as follows:

1. Any strip with a total raw charge of less than 3.0 photoelectrons is discarded, as it is almost certainly due to PMT crosstalk or poorly modeled low pulse-height hits. This cut goes slightly beyond the already existing, similarly motivated 2.0-photoelectron cut of the reconstruction, further reducing interdetector variations in event topology resulting from differences in the crosstalk effect between the detectors [97].

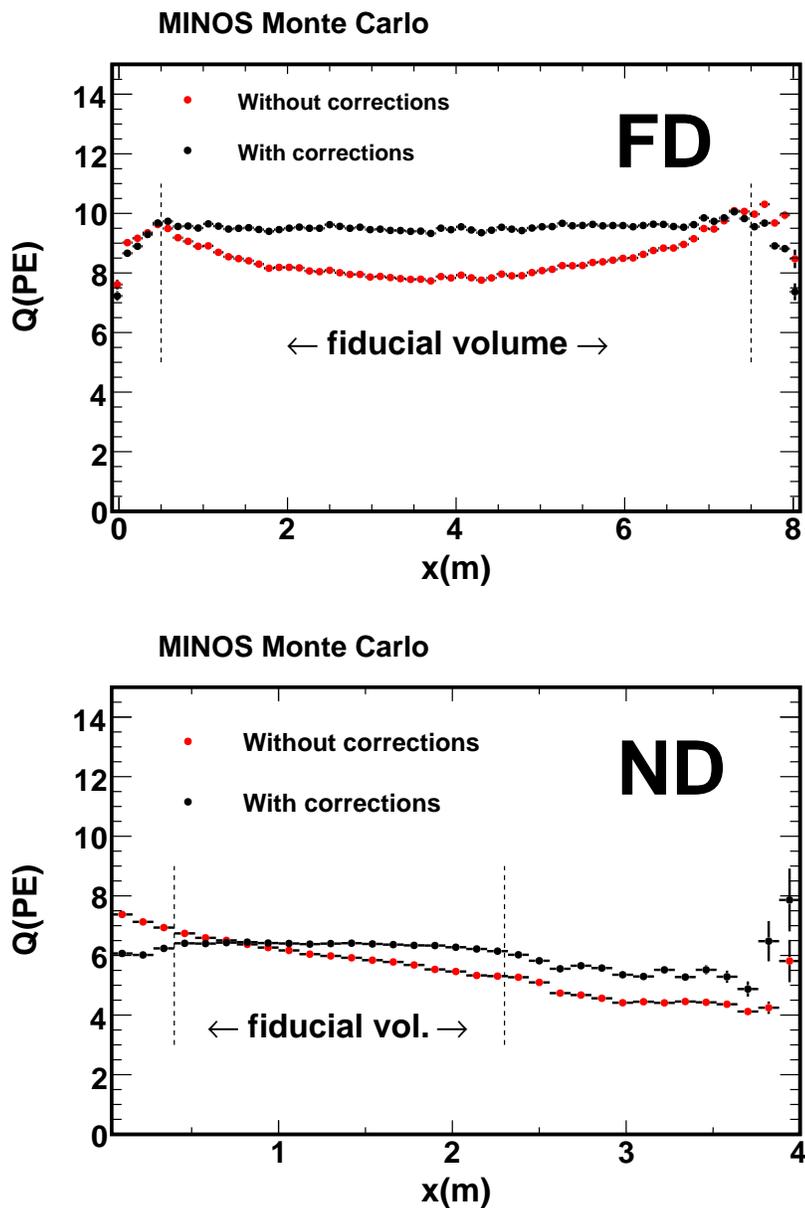


Figure 4.5: The average total pulse height per strip Q deposited by muons as a function of distance from the west end of the strip x , before (red) and after (black) being corrected for attenuation as prescribed (except for the factor of β in the case of the ND). Note that within each detector's fiducial volume (indicated by the dashed vertical lines), the corrected pulse height as a function of x is reasonably flat, implying that the attenuation correction has performed as desired. Residual differences on the order of 1% remain in the mean response along a given strip, and the attenuation systematic error discussed in Section 7.2.2 addresses the impact of these differences on the analysis. Taken from [92].

2. The pulse-height-weighted mean position (centroid) of the event is calculated in the U and V views separately, using pulse heights in SigCor. To eliminate bias resulting from high pulse-height hits at the periphery of an event that have accidentally been included by the reconstruction as part of the event, this calculation is performed using only those strips within the 8-strip window containing the highest total pulse height, ignoring any outliers.
3. The attenuation corrections described in Section 4.3.2.2 are applied to each of the strips, using the centroid calculated in the previous step as the position of the hit along the strip.
4. The event is relocated so that its centroid is in strip 100 of plane 100. As this is done to all events, comparisons can be performed in a manner that is independent of absolute position.
5. In each plane, strips are grouped about the centroid according to the pattern illustrated in Figure 4.6. Thus, hits at the core of the event – where the compact electromagnetic shower of a ν_e CC event tends to deposit much of its energy, unlike the hadronic shower of a NC or ν_μ CC event which is more diffuse – are kept as-is, but hits at the periphery of the event are combined, deemphasizing their precise topology. In fact, without this strip grouping, the performance of LEM (i.e., the *FOM*) is found to decrease by about 7% [98]. This is likely a consequence of the algorithm sacrificing match quality at the core of the event for a better match for outlier hits. Thus, strip grouping improves overall match quality by making the algorithm less sensitive to minor spatial fluctuations in energy deposition.

We should clarify that all events – library and input – undergo compacting. Compact events occupy about 60 times less hard disk space than their original reconstructed counterparts, leading to much faster I/O and thus significantly faster execution of the algorithm. Figure 4.7 illustrates the compacting procedure on a simulated Far Detector event.

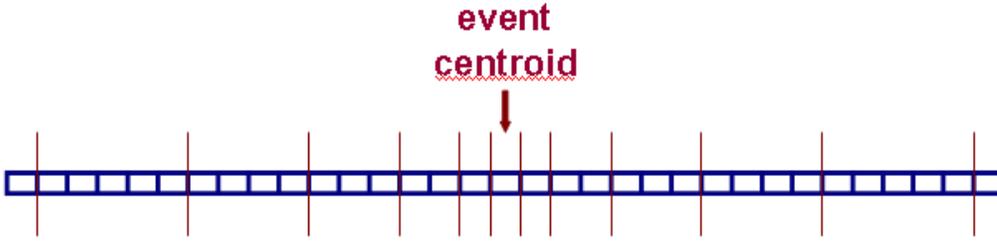


Figure 4.6: The strip grouping pattern for the strips in a single plane. Strips between thin red lines have their pulse heights summed and treated as occurring in a single strip. In this way, strips near the event core are left mostly untouched while strips farther away are grouped together, thus reducing the impact of the precise topology of outlying hits on the matching.

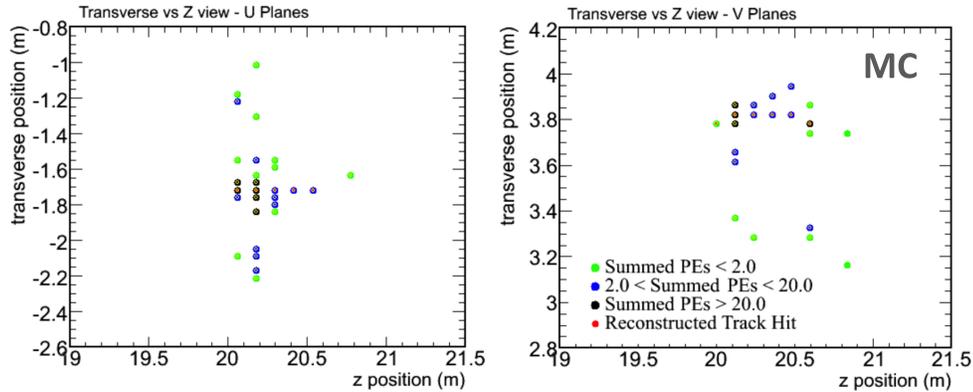
4.3.3 The Likelihood Formula and the Comparison

To quantify the extent to which an input event and a library event “match,” we ask the following question: what is the likelihood that both events, despite any apparent differences, are actually the result of identical true energy depositions in the scintillator strips? We must first be able to answer this question for the case of a single strip. Suppose a certain energy deposition in strip j of plane k in event A generates a hit with a pulse height of n_A photoelectrons, and the same energy deposition in the same strip and plane in event B has a pulse height of n_B photoelectrons. Each of the observed pulse heights is the end result of photons produced from energy deposited by a particle ultimately reaching and striking a PMT cathode to produce photoelectrons – a random process described by Poisson statistics. The given energy deposition will produce a certain mean number of photons, which in turn will produce a certain mean number of photoelectrons λ . Thus, n_A and n_B are Poisson-distributed with mean λ , and we can define the desired single-strip likelihood as

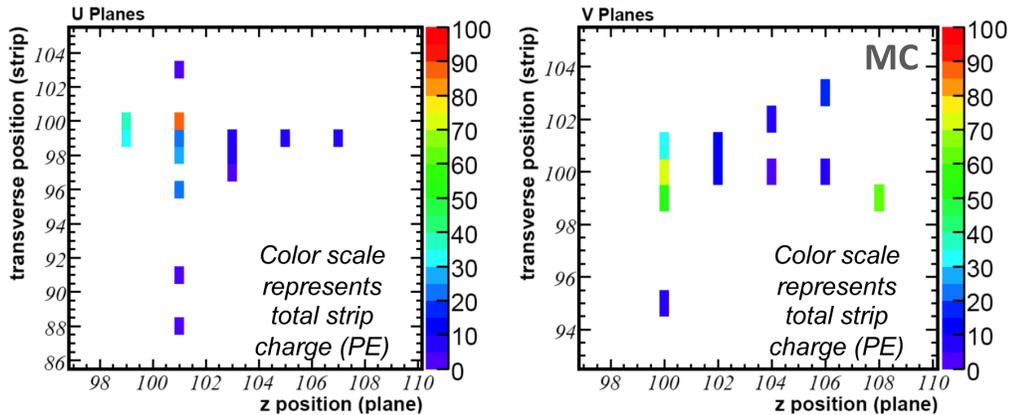
$$\ell^{jk}(n_A^{jk}, n_B^{jk}) = \int_0^\infty P(n_A^{jk}; \lambda) P(n_B^{jk}; \lambda) d\lambda, \quad (4.6)$$

where $P(n; \lambda)$ is a Poisson distribution with mean λ and observation n . We then define the event likelihood as the product of the single-strip likelihoods and, for convenience,

Raw MINOS Event:



After PE Cut and Repositioning:



After Strip Grouping:

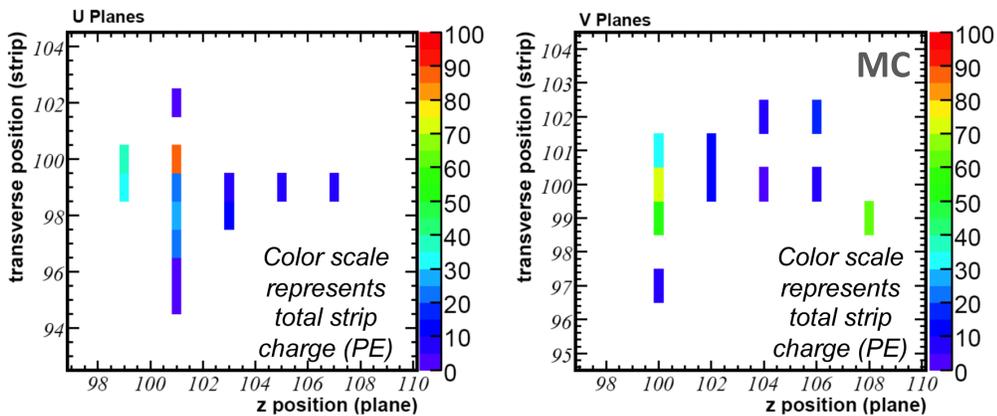


Figure 4.7: A sample Monte Carlo event undergoes compacting. The top image shows the raw event before any compacting. The second image shows the event following the 3-photoelectron cut, attenuation corrections, and repositioning. The third image has the strip grouping applied; note the effect this has on the topology of the event.

calculate the negative log-likelihood:

$$-\mathcal{L} = -\sum_{j,k} \ln \left(\int_0^\infty P(n_A^{jk}; \lambda) P(n_B^{jk}; \lambda) d\lambda \right), \quad (4.7)$$

where the sum is over all strips that have nonzero pulse height in at least one of A and B .

Before the input event is compared to the library events, it is first compared to itself (i.e., the input event is both event A and event B) to calculate the self-likelihood $-\mathcal{L}^{\text{self}}$. Then, as the matching proceeds (i.e., the input event is event A and each library event takes a turn being event B), the resulting likelihood $-\mathcal{L}^{\text{lib}}$ is decremented by $-\mathcal{L}^{\text{self}}$ to obtain

$$-\Delta\mathcal{L} = (-\mathcal{L}^{\text{lib}}) - (-\mathcal{L}^{\text{self}}) = \mathcal{L}^{\text{self}} - \mathcal{L}^{\text{lib}}. \quad (4.8)$$

This quantity is positive, and a smaller value implies a better match. Also, because $-\mathcal{L}^{\text{self}}$ has been subtracted out, values of $-\Delta\mathcal{L}$ can be compared among different input events to inspect the overall quality of the matching process. The values of $-\Delta\mathcal{L}$ are used to rank library events in a given input event's list of its 200 best matches.

Possible imperfections in the event centroid calculation may adversely affect the quality of a match. To mitigate this, a given input event is compared with each library event three times – with the input event as-is, shifted one plane forward, and shifted one plane backward. The smallest resulting value of $-\Delta\mathcal{L}$ is used to rank that library event in the input event's list of best matches. We find that about two-thirds of ν_e CC input events find their best match without any longitudinal shift; the other one-third find their best matches after being shifted ± 1 plane.

Though much effort has been made to make the matching process efficient, we can economize it further. An input event is required to pass a relaxed version of the preselection described in Section 4.2, as follows:

- $0.5 \text{ GeV} < E_{\text{reco}} < 12 \text{ GeV}$, and

- $N_{\text{planes}}^{\text{track}} < 30$,

where E_{reco} is the reconstructed energy of the event and $N_{\text{planes}}^{\text{track}}$ is the number of track planes. Any input event failing this preselection is assigned a LEM PID value of -1 and does not undergo matching. As the matching proceeds on an input event passing this preselection, each library event (event B) is checked for “compatibility” with the input event (event A), as defined by the following criteria:

- $|N_{\text{planes}}^A - N_{\text{planes}}^B| \leq x$ where $x = 4$ if $\langle N_{\text{planes}} \rangle = \frac{N_{\text{planes}}^A + N_{\text{planes}}^B + 1}{2} > 20$, $x = 3$ if $\langle N_{\text{planes}} \rangle > 15$, and $x = 2$ otherwise (the number of planes N_{planes} must not differ drastically between events),
- $2 \left| \frac{N_{\text{strips}}^A - N_{\text{strips}}^B}{N_{\text{strips}}^A + N_{\text{strips}}^B} \right| \leq 0.2$ (the number of strips N_{strips} must not differ by more than 20% between events), and
- $2 \left| \frac{Q_{\text{total}}^A - Q_{\text{total}}^B}{Q_{\text{total}}^A + Q_{\text{total}}^B} \right| \leq 0.2$ (the total attenuation-corrected charge Q_{total} must not differ by more than 20% between events).

If the library event is found to be incompatible with the input event, it is skipped, and the matching proceeds to the next library event. These sets of conditions virtually guarantee that events with long muon tracks or high energy are ignored by the matching process and assigned a LEM PID value of -1 . We also find that of those signal ν_e CC events in the Far Detector Monte Carlo that pass the preselection defined in Section 4.2, 99.8% successfully complete the matching process [99]. One final reduction in processing time is achieved by immediately aborting any comparison in which $-\Delta\mathcal{L}$ exceeds the value of $-\Delta\mathcal{L}$ for that input event’s 200th best match.

Figure 4.8 shows distributions of $-\Delta\mathcal{L}$ for the 50 best library matches to Far Detector Monte Carlo events of various classes, including signal ν_e CC events assuming θ_{13} at the 90% C.L. CHOOZ limit. Note that each distribution shows the values of $-\Delta\mathcal{L}$ for only one class of library event (ν_e CC or NC). We see that as compared to NC events, signal ν_e CC events match ν_e CC library events more frequently and with a sharper distribution of $-\Delta\mathcal{L}$ that peaks at a slightly lower value. In contrast, NC

events match NC library events more frequently. We can conclude that the matching process appears to have worked as intended.

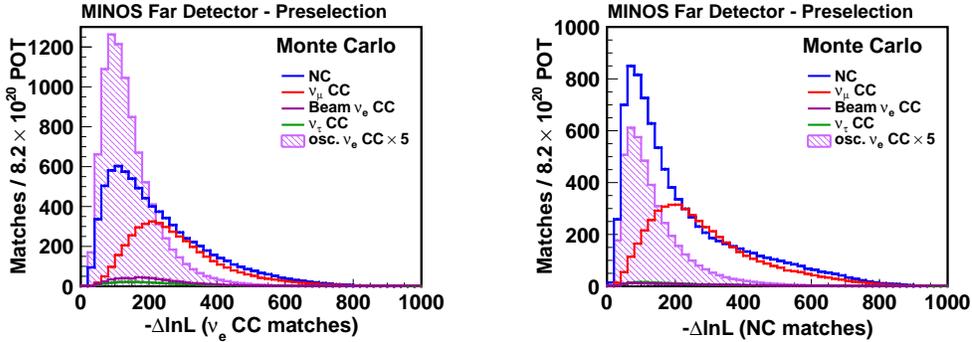


Figure 4.8: $-\Delta\mathcal{L}$ distributions for the 50 best library matches to Far Detector Monte Carlo events of various classes. Note that each distribution is for only one class (ν_e CC or NC) of library event. The size of the signal, labeled “osc. ν_e CC,” assumes θ_{13} to be at the 90% C.L. CHOOZ limit and has been scaled by a factor of 5 for clarity.

4.3.4 Gauging the Signal-Like Nature of the Best-Matched Events

With the list of the 200 best library matches for each input event in hand, we now require one or more variables that summarize the best matches and thus characterize the input event. As our goal is to determine how likely it is that a given input event is a signal event, the variables chosen should each have a different distribution for signal versus background, and they should be sufficiently different from each other in order to form the basis of a powerful discriminant when used jointly. Before proceeding, we note that in the original implementation of LEM, it was determined that using the information from the 50 best matches maximized the *FOM* [92]. Some optimal number of matches is expected – too few will not contain enough information, and too many will include low-quality matches as a result of the finiteness of the library. We will use this result in the present analysis as well; henceforth, any reference to “the best matches” is specifically to the 50 best matches.

We start by asking: what fraction of the best matches are ν_e CC (signal) events? One would expect this quantity, f_{50} , to be larger overall for signal events than for

background events and thus provide discrimination power. Figure 4.9 shows that a trend of this sort is, in fact, observed.

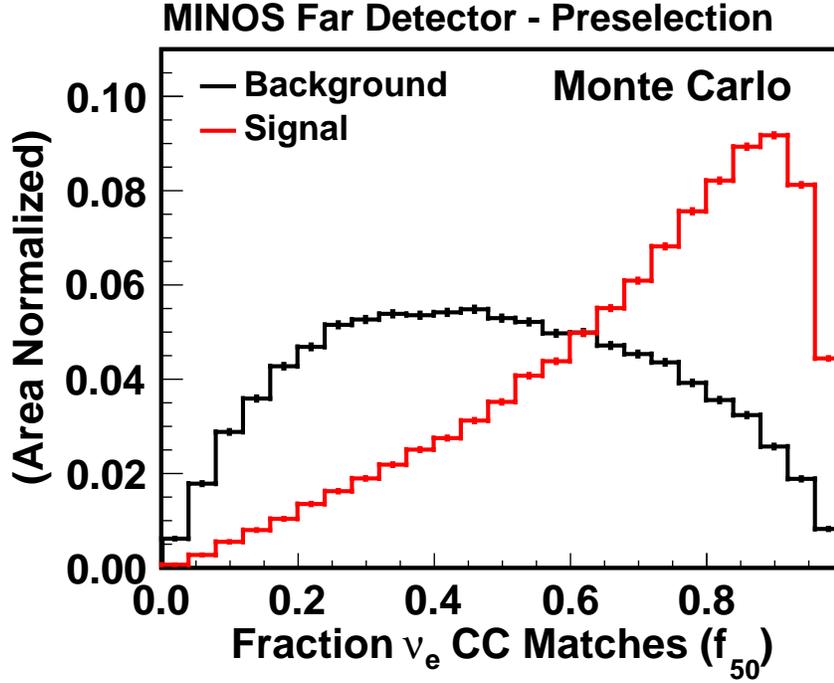


Figure 4.9: Distributions of f_{50} for signal and background events in the Far Detector Monte Carlo sample that pass the preselection defined in Section 4.2. Each distribution has been normalized to unity.

It is possible for a NC event to have little or no electromagnetic activity but still have a large number of ν_e CC best matches, albeit also with low electromagnetic activity.⁷ The resulting large value of f_{50} will suggest that this NC event is likely signal, a conclusion we would like to avoid. We define the inelasticity, or hadronic y , of a CC event as

$$y = \frac{E_\nu - E_\ell}{E_\nu}, \quad (4.9)$$

where E_ν is the energy of the incoming neutrino and E_ℓ is the energy of the outgoing charged lepton. Thus, a ν_e CC event in which a large fraction of the event energy is in the electromagnetic shower (generated by the electron) would have y close to 0, but

⁷Recall that a ν_e CC event consists of overlapping electromagnetic and hadronic showers, and it is possible for only a small fraction of the event's energy to be contained in the electromagnetic shower.

those ν_e CC events matching the hypothetical NC event described above would tend to have larger values of y . We define the summary variable y_{50} for a given input event as the mean y of the ν_e CC events found among its 50 best matches. Figure 4.10 shows distributions of y_{50} for signal and background events in the Far Detector Monte Carlo sample.

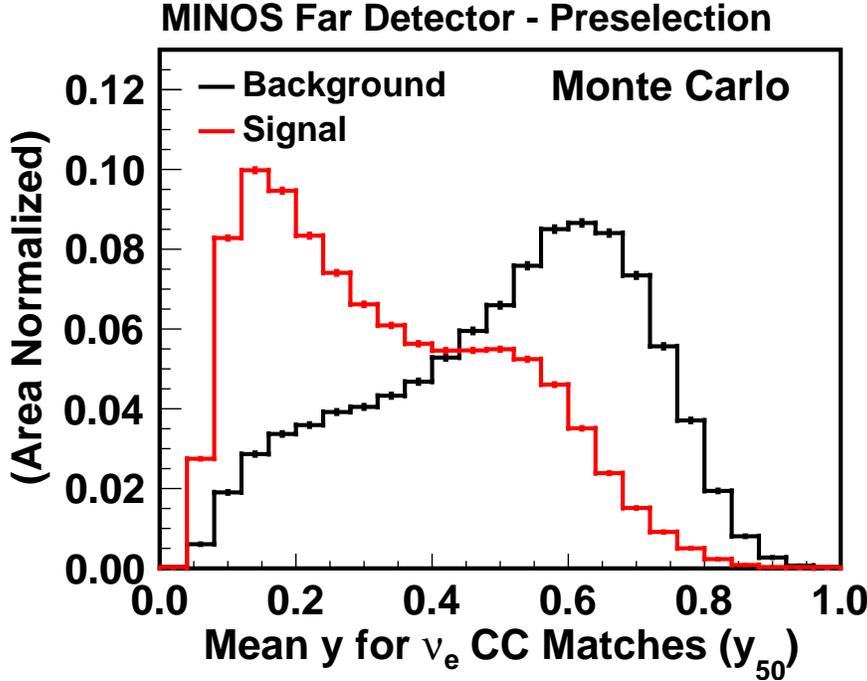


Figure 4.10: Distributions of y_{50} for signal and background events in the Far Detector Monte Carlo sample that pass the preselection defined in Section 4.2. Each distribution has been normalized to unity.

We also include a variable that summarizes the quality of the match between the input event and its ν_e CC best matches. First, we define the fraction of the total pulse height of two events that overlaps between the events:

$$Q_{\text{frac}} = \frac{Q_{\text{matched}}}{Q_{\text{matched}} + Q_{\text{unmatched}}}. \quad (4.10)$$

A few examples that clarify the calculation of each quantity in Equation 4.10 follow.

- Suppose that in strip 101 of plane 105, event A has a pulse height of 5 photoelectrons, and event B has a pulse height of 8 photoelectrons. The value of

Q_{matched} will be incremented by $13(= 5 + 8)$.

- Suppose that in strip 102 of plane 107, event A has a pulse height of 4 photoelectrons, but event B has no activity. The value of $Q_{\text{unmatched}}$ will be incremented by 4.

Upon completion of the matching process between a given input event and the current library event, Q_{frac} is calculated according to Equation 4.10 using the final values of Q_{matched} and $Q_{\text{unmatched}}$ and is stored. The summary variable constructed from this information is q_{50} , the mean value of Q_{frac} for the ν_e CC events among the given input event's 50 best matches. Figure 4.11 provides distributions of q_{50} for signal and background events in the Far Detector Monte Carlo sample.

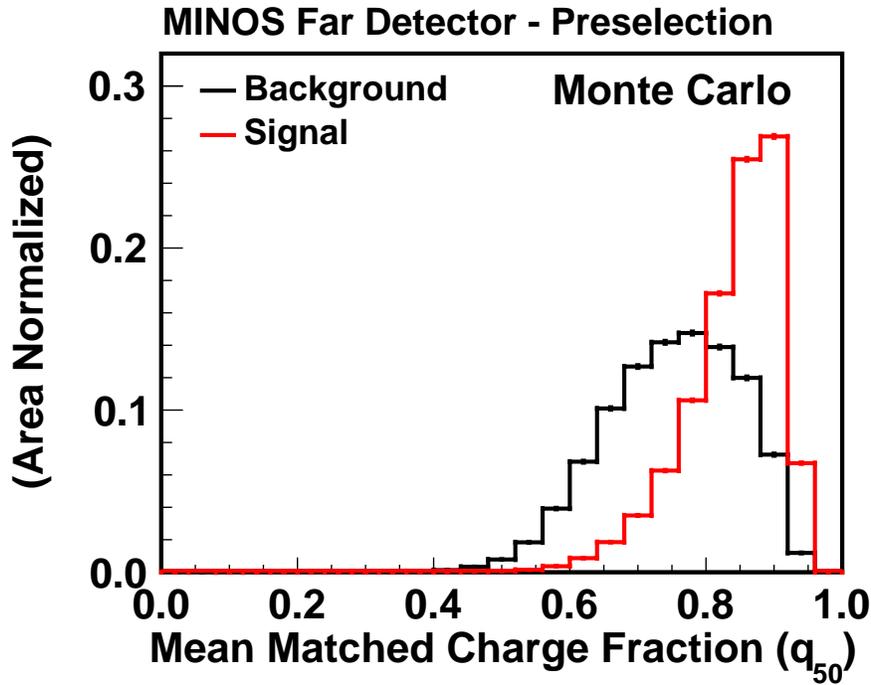


Figure 4.11: Distributions of q_{50} for signal and background events in the Far Detector Monte Carlo sample that pass the preselection defined in Section 4.2. Each distribution has been normalized to unity.

Finally, the signal and background reconstructed energy spectra differ. Signal ν_e CC events are the result of $\nu_\mu \rightarrow \nu_e$ oscillations, which in turn have explicit energy dependence (see, for example, Equation 3.1). In fact, the topology of any CC event

depends, as we have seen, on the flavor of the neutrino producing it, which in turn is susceptible to change via oscillation, so some related energy dependence is expected in the ν_μ CC, beam ν_e CC, and ν_τ CC backgrounds as well. The NC event rate, on the other hand, is totally independent of oscillation effects⁸ and thus lacks any associated energy dependence. As NC events are the primary component of the background, their contribution to the background energy spectrum will dominate. Thus, reconstructed energy is also included as a discriminant variable.

4.3.5 The LEM PID

The final step of the procedure is to distill the information contained in the discriminant variables defined in Section 4.3.4 into a single variable that characterizes how signal-like a given input event is. Such a multivariate approach typically achieves better signal–background separation than do simple cuts on individual variables. The original implementation of LEM used a simple likelihood:

$$\text{LEM PID} = \frac{\prod_{i=f_{50},y_{50},q_{50}} P_{\text{sig}}^i(x_i)}{\prod_{i=f_{50},y_{50},q_{50}} P_{\text{sig}}^i(x_i) + \prod_{i=f_{50},y_{50},q_{50}} P_{\text{bg}}^i(x_i)}, \quad (4.11)$$

where x_i is the value of variable i and $P_{\text{sig}}^i(x_i)$ and $P_{\text{bg}}^i(x_i)$ are the values of the signal and background probability density functions, respectively, for variable i evaluated at x_i [92]. These probability density functions were binned in energy and populated with Far Detector Monte Carlo events. As we see in Figure 4.12, however, the discriminant variables are highly correlated, and the likelihood in Equation 4.11 does not take these correlations into account.

A four-variable probability density function would be a straightforward solution to the problem posed by the correlations, but the available number of Far Detector Monte Carlo events would not sufficiently populate such a function. An artificial neural network, on the other hand, handles the correlations well despite the limited statistics, and this was the approach used to compute the LEM PID in this analysis.

⁸This statement assumes that there are no sterile neutrinos – that is, all neutrino flavors are active and participate in the weak interaction.

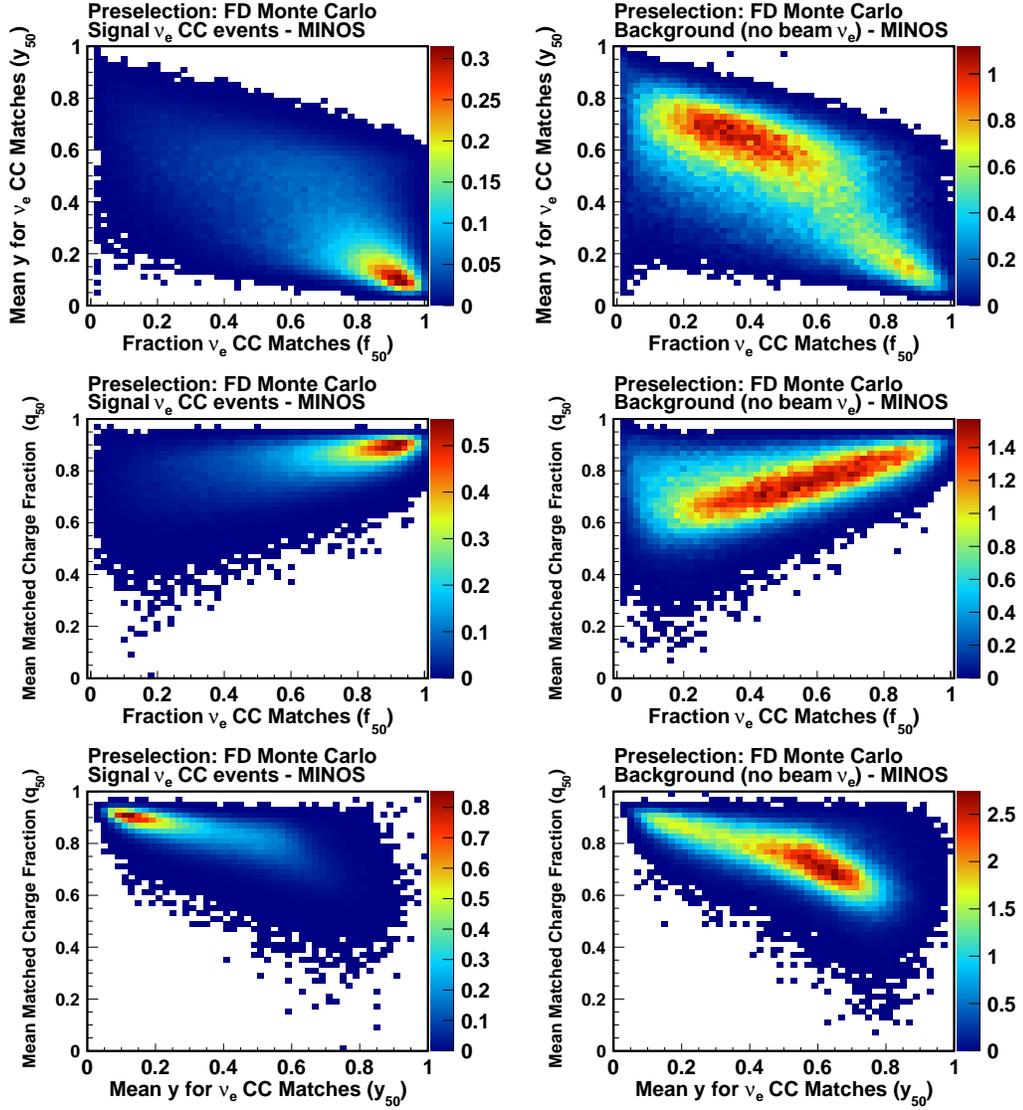


Figure 4.12: Relationships between the discriminant variables f_{50} , y_{50} , and q_{50} taken two at a time. The plots in the left column are for signal events, and those in the right column are for background events, excluding intrinsic ν_e CC events. The color scale represents Far Detector Monte Carlo events scaled to the 8.2×10^{20} -PoT exposure of this analysis; the signal is calculated assuming θ_{13} at the 90% C.L. CHOOZ limit. We can clearly see, as discussed, that these three discriminant variables are highly correlated. There is a particularly interesting feature to note in the top-right plot. While most background events have low f_{50} (not many ν_e CC events among their best matches) and high y_{50} (the ν_e CC events among their best matches tend to have greater hadronic character), a significant number have high f_{50} and low y_{50} and thus mimic the signal; compare the corresponding region in the top-left plot.

In a broad sense, an artificial neural network consists of a set of artificial neurons.⁹ These neurons are given a set of external inputs to which they respond in a characteristic way. These neurons also communicate with each other, via a network of weighted connections. This network of neurons can be taught to perform a specific task – in our case, to discriminate between signal and background events – by manipulating the weights associated with these connections, as we will see shortly. The neural network implementation used for LEM was the ROOT [101] class `TMultiLayerPerceptron` in the multivariate analysis package `TMVA` [102]. Dubbed LEMNN (LEM Neural Network), the network itself consists of three layers of neurons:

1. the input layer, consisting of four neurons corresponding to the four input variables f_{50} , y_{50} , q_{50} , and reconstructed energy,
2. a hidden layer, consisting of nine neurons with which the input layer neurons communicate directly, and
3. the output layer, consisting of one neuron with which the hidden layer neurons communicate directly.

For a given set of inputs to the input layer, corresponding to a single input event, the output neuron produces a number between 0 and 1 that quantifies the network’s determination of how backgroundlike (0) or signal-like (1) the event was. We say that the “architecture” of the system is 4:9:1 – simply a list of the number of neurons in each layer.

LEMNN was taught how to discriminate between signal and background using training sample of Far Detector Monte Carlo events (all passing the preselection of Section 4.2), divided into a signal set (286,676 ν_e CC events) and a background set (251,935 NC, ν_μ CC, and ν_τ CC events). Each event consisted of a set of values of the four discriminant variables, the event’s true identity (i.e., signal or background), and information about its relative weight (associated with the Monte Carlo tuning described in Section 2.6.1 as well as with the oscillation probability of the event’s associated neutrino). The training process used a supervised learning technique known as

⁹A relatively thorough treatment of the topic is provided in [100].

“backpropagation,” which, assuming some initial set of values for the aforementioned neural network connection weights, proceeded as follows:

1. Provide values of the discriminant variables for an input event to the input neurons and allow the network to produce an output (between 0 and 1).
2. Reveal what the output *should* have been – 0 for a background event or 1 for a signal event (this is the “supervised” part).
3. Calculate the difference between the actual and desired outputs (this is the “backpropagation”¹⁰ part) and adjust the connection weights so as to reduce the square of this difference.

These steps were repeated over the full training sample a maximum of 600 times as the training algorithm sought to minimize the squared difference of the actual and desired outputs. Finally, the neural network could be considered trained. It was tested on a separate sample of Far Detector Monte Carlo events passing preselection; the resulting distribution of LEM PID – the output of LEMNN – is shown in Figure 4.13.

With the LEM PID ready, we would like to study its behavior and decide how to use it. Does it behave similarly for data and Monte Carlo in the Near Detector, where no signal is expected? What cut on the LEM PID will maximize the *FOM* of Equation 4.1 or some other figure of merit? How does the performance of LEM compare with that of ANN? Until we have the means to predict what we may observe in the Far Detector, we can only answer the first of these questions, which we do in Section 4.3.6. As for the remaining questions, we will discuss the necessary techniques in the next few chapters and address these issues immediately thereafter.

4.3.6 LEM in the Near Detector

In Chapter 3, we noted that this is a blind analysis, which precludes us from comparing the behavior of LEM in the Far Detector data with its behavior in the Far Detector

¹⁰This is an abbreviation for “backward propagation of errors.”

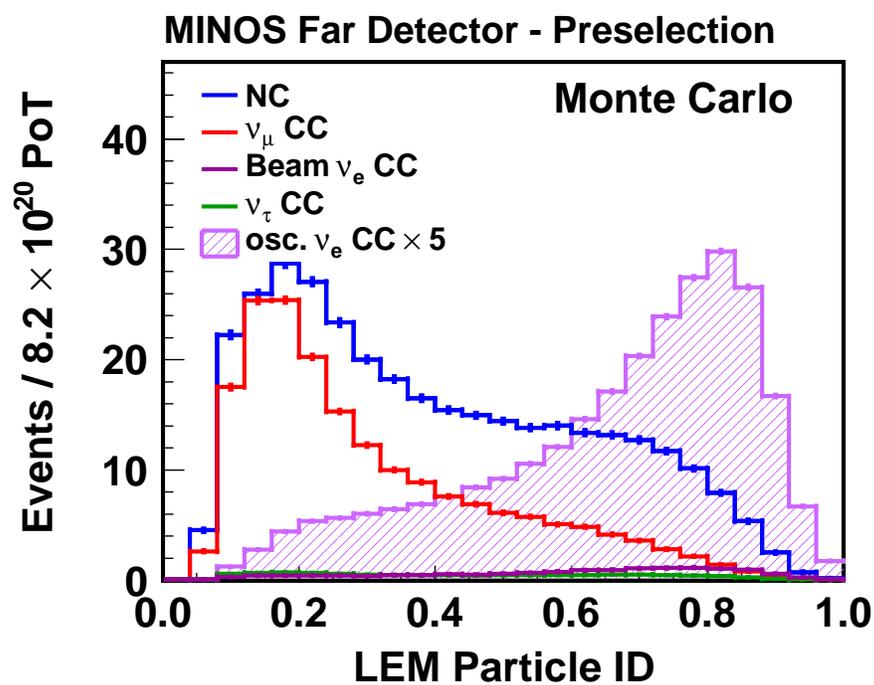


Figure 4.13: The LEM PID distributions for all event classes in the Far Detector Monte Carlo. The signal distribution has been scaled by a factor of 5 for clarity and is computed assuming θ_{13} to be at the 90% C.L. CHOOZ limit.

Monte Carlo. Such a comparison in the Near Detector, however, is completely acceptable; at this stage, it is actually preferable, as we expect no oscillation effects in the Near Detector. We calculate the LEM PID for all events in the preselected Near Detector data and Monte Carlo samples and normalize each of the resulting LEM PID distributions to an exposure of 1×10^{19} PoT (by convention). The results are shown in Figure 4.14, with a systematic error band on the Monte Carlo distribution resulting from systematic effects discussed in Chapter 7.

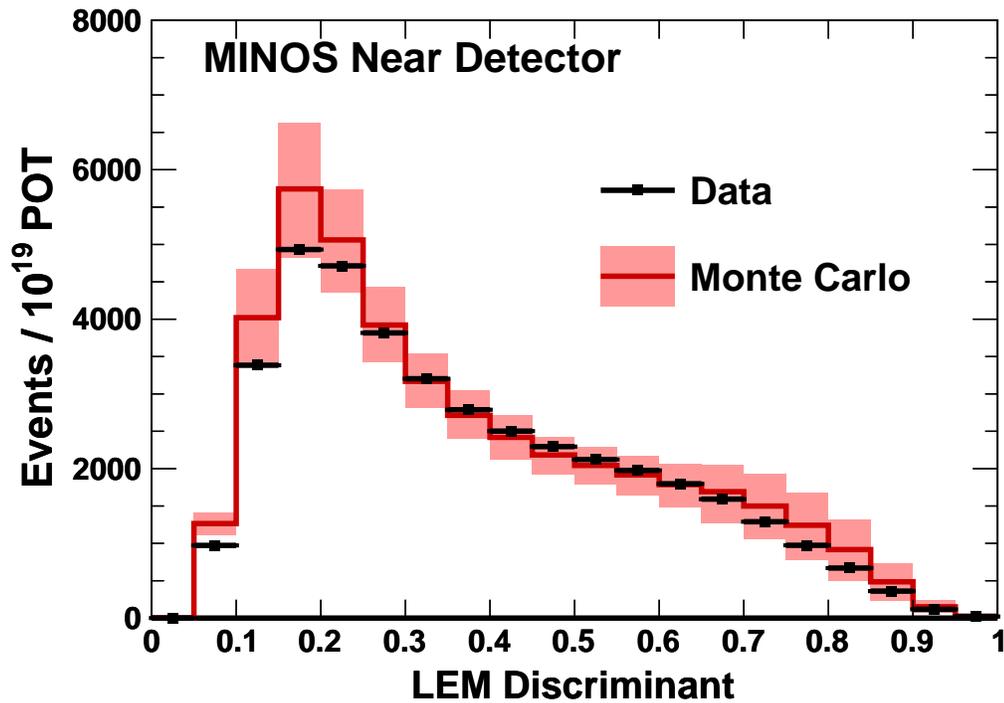


Figure 4.14: Distribution of LEM PID for Near Detector data and Monte Carlo events passing the preselection, normalized to 1×10^{19} PoT. The shaded band on the Monte Carlo distribution represents the total 1σ systematic uncertainty resulting from a variety of sources described in Chapter 7.

Note that while the data lie within the 1σ systematic error band and are thus consistent with the Monte Carlo, the data–MC difference is still somewhat disconcerting. Recall from Chapter 3, however, that as long as the underlying cause of any systematic effect is common to both Near and Far Detectors, the resulting systematic errors will largely cancel in the analysis. It is instructive to see how the details of the simulation can affect the magnitude of the data–MC difference. As an example, we

consider one aspect of the simulation that we found to be most suspect in producing the observed difference: the AGKY hadronization model discussed in Section 2.6.2. In particular, we discovered [103] a correlation between the hadronic character of a simulated event (y of Equation 4.9 for a ν_e CC event, Y_{NC} ¹¹ for a NC event) and its value of y_{50} ; see Figure 4.15.

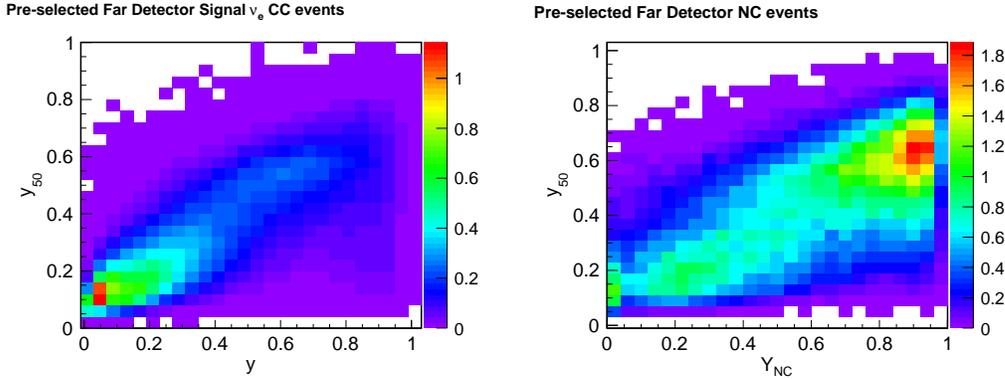


Figure 4.15: Correlation between y_{50} and true hadronic character. Together, these plots support the idea that y_{50} can act as an effective measure of a simulated event’s level of hadronization, or hadronic character. (Left) y_{50} vs. true y for preselected FD signal ν_e CC events. A strong correlation exists between the two quantities. (Right) y_{50} vs. Y_{NC} for preselected FD NC events. Though not as strong as in the plot to the left, a correlation between the two quantities is discernible nonetheless.

In a way, one can imagine that the hadronization model produces the energy deposition pattern of a Monte Carlo event from the y/Y_{NC} (and, of course, other details) of the event’s underlying interaction. LEM, in turn, establishes a link between that pattern and a library of other patterns – patterns that, we should note, are of Monte Carlo events generated using the *same* hadronization model – in a way that tends to accurately recover the y/Y_{NC} of the underlying interaction, in the form of y_{50} . If the hadronization model is doing a poor job of simulating reality, however, the value of y_{50} determined by LEM for a *real* event will *not* accurately reflect that event’s true y/Y_{NC} , and the resulting distribution of LEM PID for real events will disagree with the corresponding distribution for Monte Carlo events. Thus, we hypothesize that at least some of the data–MC difference seen in Figure 4.14 is due to the mismodeling

¹¹This quantity – defined as $Y_{NC} = 1 - \frac{E^{\pi^0}}{E_{vis}}$, where E^{π^0} is the energy carried by π^0 s in the event and E_{vis} is the total visible energy of the event – is a measure of the hadronic character of the event.

of hadronic showers.

A series of variations on features of the AGKY hadronization model allows us to simulate the associated uncertainties and to determine their effect on the overall systematic error in the analysis [104]. As a test of the hypothesis developed above, each of these variations was applied to the Monte Carlo through a reweighting scheme, and each resulting “reweighted” LEM PID distribution was compared to the original distribution. Most of these variations had effects of only a few percent, but one, “Baryon x_f Selection” (which we will describe shortly), had an effect on the order of the required 20%–30%. The LEM PID distributions for data and Monte Carlo, before and after this reweighting, are shown in Figure 4.16. The associated y_{50} distributions are shown in Figure 4.17.

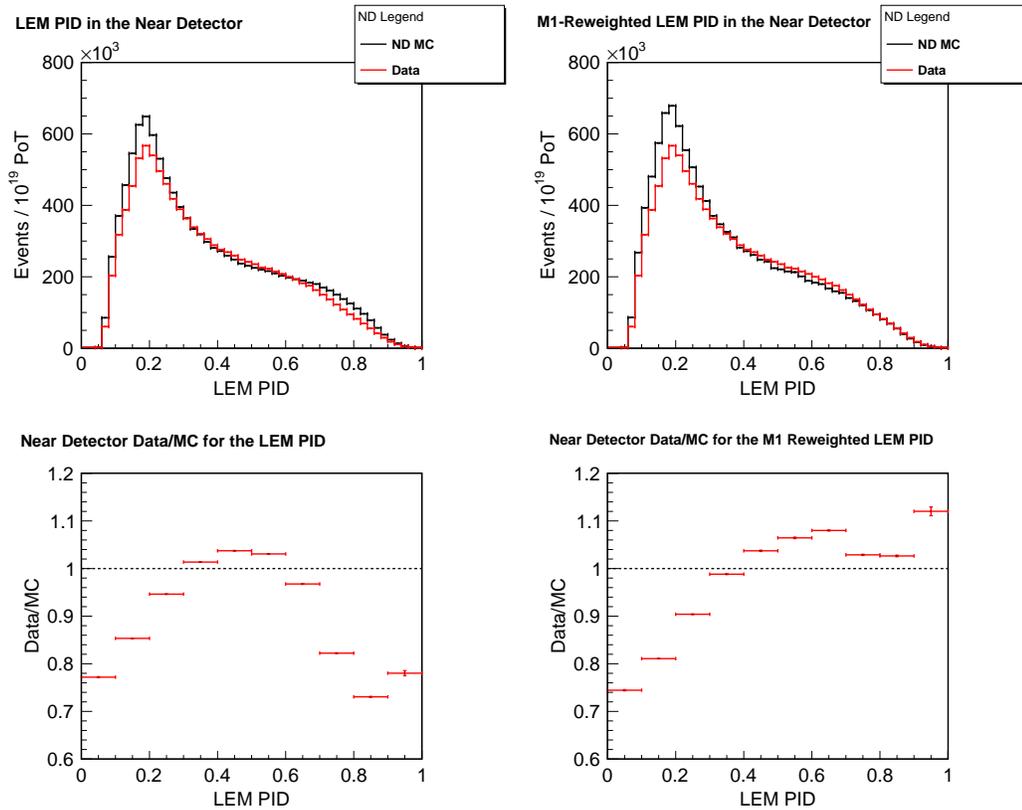


Figure 4.16: LEM PID in the Near Detector before (left) and after (right) the “Baryon x_f Selection” (code-named “M1”) reweighting. The data distribution is the same in both plots. The lower plots show the corresponding ratio of data to Monte Carlo. Note that while the disagreement remains large in the low-PID (backgroundlike) region, there is a marked improvement in the high-PID (signal-like) region.

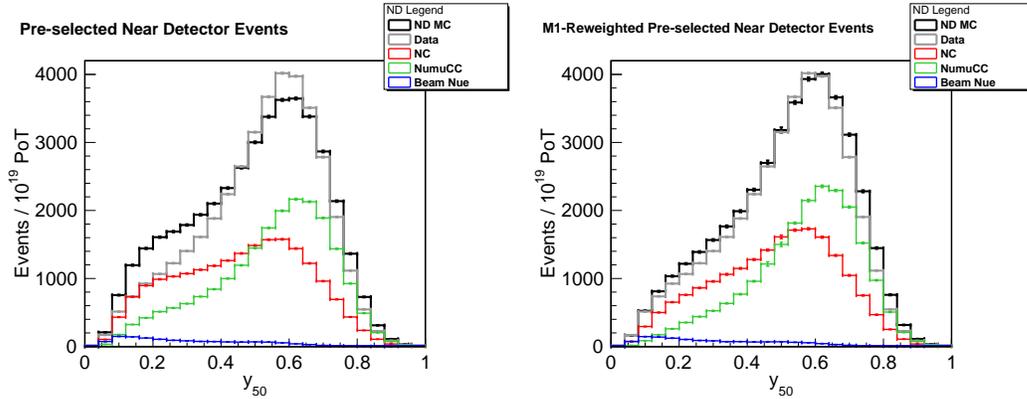


Figure 4.17: y_{50} in the Near Detector before (left) and after (right) the “Baryon x_f Selection” (code-named “M1”) reweighting. The Monte Carlo components are shown individually (NC, ν_μ CC, beam ν_e CC) and in total (ND MC). Note the dramatic improvement in agreement between the data and total Monte Carlo distributions for $0 < y_{50} \lesssim 0.6$, due mostly to a shift in the NC distribution.

The AGKY hadronization model as employed in the MINOS Monte Carlo implements parton model arguments regarding the four-momenta of baryons produced in hadronic showers resulting from neutrino interactions; in particular, these arguments predict that the baryon is more likely to be found in the backward hemisphere. The “Baryon x_f Selection” variation in the model undoes this feature to favor an isotropic distribution, generating the baryon four-momentum in the center of mass using a phase-space decay [104]. Figure 4.16 suggests that such features of the hadronization model are indeed capable of producing at least part of the observed difference between data and Monte Carlo. Fortunately, as noted before, the same hadronization model is used in simulating both Near Detector and Far Detector events, and as we shall see in Chapter 7, the systematic uncertainty on the Far Detector prediction induced by uncertainties in the hadronization model are actually quite small.

Chapter 5

The Near Detector Background Decomposition

The data–MC difference discussed in Section 4.3.6 enhances the motivation for the data-driven extrapolation technique introduced in Chapter 3, as we now clearly see that Monte Carlo alone cannot serve as a reliable means of making a Far Detector prediction. We also noted that in order to perform the extrapolation, the Near Detector data must be broken up, or decomposed, into its various components. This decomposition will inevitably rely on the Monte Carlo in some way, but in light of the aforementioned data–MC difference, we would prefer that this reliance not be too sensitive to the physics modeling details of the Monte Carlo. As introduced in Chapter 3 and alluded to in Section 4.3.6, this can be done by using *ratios* of Monte Carlo event counts, in which systematic uncertainties due to a common source cancel to first order.

We now describe two different techniques for performing the Near Detector background decomposition. The first, “PORP,” is a simple, computationally versatile technique used in PID cut optimization studies, analysis binning studies, and any other investigations requiring rapid turnaround. “HOOHE” is a more rigorous but less versatile technique that relies on a proper fit incorporating statistical and systematic errors to compute the background components; it is used in later stages of the analysis.

5.1 The PORP Method

The PORP¹ method relies on two simple assumptions regarding the Near Detector data and Monte Carlo:

- The intrinsic (“beam”) ν_e CC component is simulated well in the Monte Carlo, so the difference between data and Monte Carlo arises solely from the NC and ν_μ CC components.
- The energy- and PID-dependent ratio between data and Monte Carlo (i.e., the disagreement) is the same for both NC and ν_μ CC components.

Several findings support the first assumption. First, the data–MC ratio of electron selection efficiencies in CalDet was found [105] to be within 5% of unity and, in fact, consistent with unity given the statistics. The PID selection was also applied to a special sample of hybrid events known as “MRE” (muon-removed electron-added), discussed in Chapter 6, to estimate the data–MC ratio of ν_e CC selection efficiencies; this ratio was found to be within 2% of unity for reconstructed energies between 1 and 4 GeV and no more than 12% for higher energies. Thus, at least in the high-PID region, ν_e CC events are well modeled by the Monte Carlo. The second assumption is justified by the fact that a ν_μ CC event surviving the preselection consists mostly of hadronic shower activity and is thus likely to be just as mismodeled as a NC event is.

Under these assumptions, the Near Detector data in each analysis bin are broken up as follows:

$$N_{\text{beam } \nu_e}^{\text{data}} = N_{\text{beam } \nu_e}^{\text{MC}}, \quad (5.1)$$

$$N_{\text{NC}}^{\text{data}} = N_{\text{NC}}^{\text{MC}} \times \frac{N_{\text{total}}^{\text{data}} - N_{\text{beam } \nu_e}^{\text{MC}}}{N_{\text{total}}^{\text{MC}} - N_{\text{beam } \nu_e}^{\text{MC}}}, \quad (5.2)$$

$$N_{\nu_\mu \text{ CC}}^{\text{data}} = N_{\nu_\mu \text{ CC}}^{\text{MC}} \times \frac{N_{\text{total}}^{\text{data}} - N_{\text{beam } \nu_e}^{\text{MC}}}{N_{\text{total}}^{\text{MC}} - N_{\text{beam } \nu_e}^{\text{MC}}}, \quad (5.3)$$

¹This is an acronym formed from the initials of the method’s creators: Pedro Ochoa and Ryan Patterson.

where N_i^j is the number of Near Detector events of type i in sample j , normalized to an exposure of 1×10^{19} PoT. Despite our confidence in the assumptions made above, an explicit estimate of the systematic error induced in the total Far Detector background prediction as a result of using PORP would provide additional reassurance. Such an estimate was made² using a fake data approach for several assumptions about the error on the NC component, as follows:

1. randomly generate the beam ν_e CC contribution assuming that the true value in the data is within $\pm 30\%$ of the Monte Carlo value,³
2. randomly generate the NC contribution assuming that the true value in the data is within $\pm x\%$ of the Monte Carlo value,
3. compute the ν_μ CC contribution by requiring that the three contributions add up to the total observed number of Near Detector events,
4. extrapolate each contribution to the Far Detector, and
5. compare the resulting total predicted background in the Far Detector with the prediction made using the PORP-computed values of the Near Detector beam components.

For $x = 20, 50$, we find the variation in the predicted Far Detector background to be 2.4% and 5.0%, respectively; if we take x to be very large – i.e., we assume virtually no knowledge of the systematic error on the Monte Carlo modeling of the NC contribution to the Near Detector data – the error on the predicted background is 6.0%. Ultimately, the relatively small systematic error on the PORP-based Far Detector background prediction can be seen as being due primarily to the constraint that the Near Detector data components must add up to the total observed number of Near Detector events:

$$N_{\text{beam } \nu_e}^{\text{data}} + N_{\text{NC}}^{\text{data}} + N_{\nu_\mu \text{ CC}}^{\text{data}} = N_{\text{total}}^{\text{data}}, \quad (5.4)$$

²In this study, reported in [106], the background decomposition and extrapolation were performed in bins of energy on all events passing an optimized cut on PID value.

³This error is primarily the result of flux uncertainties.

a condition that is automatically satisfied by Equations 5.1–5.3. Thus, PORP is perfectly adequate for the purposes indicated.⁴ For the more delicate tasks of evaluating the sensitivity of the analysis to the null hypothesis (i.e., that $\theta_{13} = 0$) and computing confidence intervals from the Far Detector data, however, we employ the more sophisticated “HOOHE” approach described below.

5.2 The HOOHE Method

In Section 5.1, Equation 5.4 is presented as a constraint on the Near Detector data components, but we may also view it as a linear equation in three variables. This, in turn, leads us to ask if additional equations in these variables may be written, for the resulting system could then be solved for these variables. Note that Equation 5.4 contains external input, in the form of $N_{\text{total}}^{\text{data}}$, so additional equations will require additional input. As discussed in Section 2.1, the NuMI beam can be configured to produce different neutrino spectra, which in turn will produce Near Detector event spectra with differing fractions of NC, ν_{μ} CC, and beam ν_e CC events. In fact, statistically significant data samples have been collected in the “Horn Off” (NuMI target in the LE position, focusing horns turned off) and pHE configurations, and these can provide the additional input to write two additional equations, for a total of three equations in three variables. Figure 5.1 shows the simulated true neutrino energy spectra in each configuration, and Figure 5.2 shows, for each configuration, the simulated fraction of each event type as a function of reconstructed energy for events in the analysis region of LEM PID. The method’s name, HOOHE, is an acronym for Horn On/Off High Energy – the three beam configurations⁵ upon whose data the method is based.

⁴The PID cut optimization and analysis binning studies using PORP were performed by the author and are detailed in Appendix B.

⁵The LE beam configuration is alternatively referred to as the “Standard” or “Horn On” configuration.

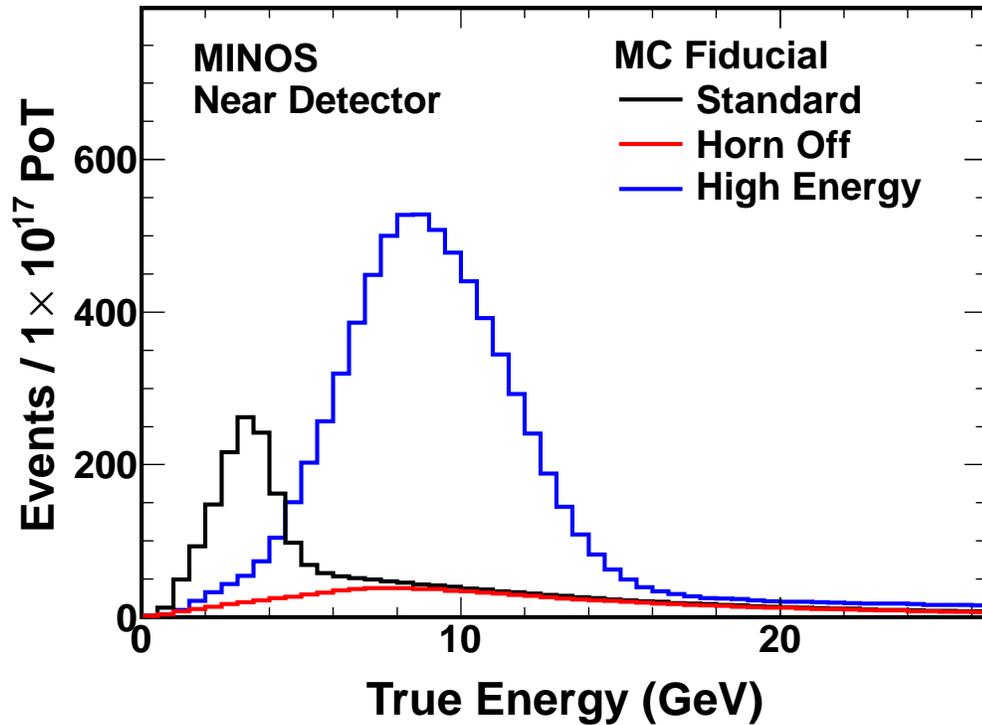


Figure 5.1: True energy spectra of neutrinos producing events in the Near Detector fiducial volume, according to the Monte Carlo simulation for each beam configuration. Note the shift of the focusing peak from the Standard configuration to the High Energy configuration as well as the significantly higher event rate in the High Energy configuration. In the Horn Off configuration, the spectrum lacks a distinctive focusing peak, as expected; also, the event rate is significantly lower than in the other configurations, reflecting the proportion and energy distribution of neutrino parents that are produced nearly parallel to the beam axis.

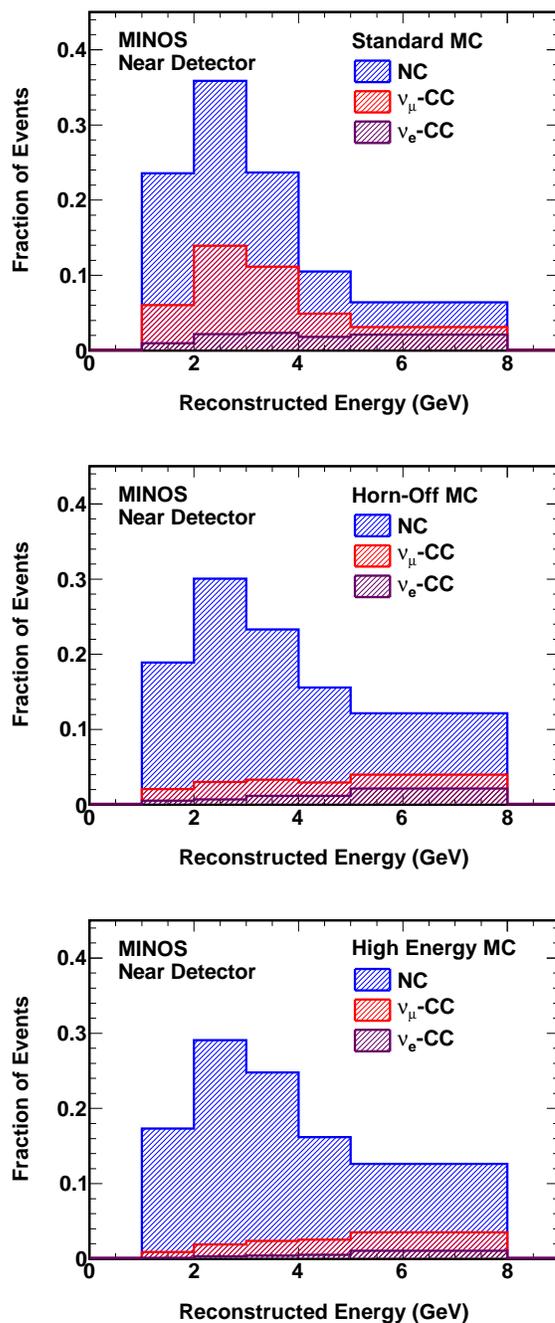


Figure 5.2: Stacked histograms showing, for each beam configuration, the fraction of each event type in the Near Detector as a function of reconstructed energy for preselected events in the analysis region of LEM PID (LEMPID > 0.6). To clarify, the bin contents of a given histogram sum to unity, so the “fraction” indicated is relative to the total number of events with reconstructed energy between 1 and 8 GeV and LEM PID in the analysis region. Note the similarity between the Horn Off and High Energy distributions; it motivates the use of additional constraints in the HOOHE method.

The desired system of equations in each analysis bin is

$$N_{\text{NC}}^{\text{data On}+} + N_{\nu_{\mu} \text{CC}}^{\text{data On}+} + N_{\text{beam } \nu_e}^{\text{data On}} = N_{\text{total}}^{\text{data On}}, \quad (5.5)$$

$$\frac{N_{\text{NC}}^{\text{MC Off}}}{N_{\text{NC}}^{\text{MC On}}} \times N_{\text{NC}}^{\text{data On}+} + \frac{N_{\nu_{\mu} \text{CC}}^{\text{MC Off}}}{N_{\nu_{\mu} \text{CC}}^{\text{MC On}}} \times N_{\nu_{\mu} \text{CC}}^{\text{data On}+} + \frac{N_{\text{beam } \nu_e}^{\text{MC Off}}}{N_{\text{beam } \nu_e}^{\text{MC On}}} \times N_{\text{beam } \nu_e}^{\text{data On}} = N_{\text{total}}^{\text{data Off}}, \quad (5.6)$$

$$\frac{N_{\text{NC}}^{\text{MC pHE}}}{N_{\text{NC}}^{\text{MC On}}} \times N_{\text{NC}}^{\text{data On}+} + \frac{N_{\nu_{\mu} \text{CC}}^{\text{MC pHE}}}{N_{\nu_{\mu} \text{CC}}^{\text{MC On}}} \times N_{\nu_{\mu} \text{CC}}^{\text{data On}+} + \frac{N_{\text{beam } \nu_e}^{\text{MC pHE}}}{N_{\text{beam } \nu_e}^{\text{MC On}}} \times N_{\text{beam } \nu_e}^{\text{data On}} = N_{\text{total}}^{\text{data pHE}}, \quad (5.7)$$

where the notation of Equations 5.1–5.4 has been modified to include the applicable beam configuration. Equation 5.5 is simply Equation 5.4 rewritten with the beam configuration specified. In Equations 5.6 and 5.7, the Monte Carlo ratios may be thought of as conversion factors between the Horn On configuration and the Horn Off or pHE configuration. While the individual quantities in a given ratio are subject to large systematic errors due to a variety of sources, the error on the ratio itself is small due to cancellation to first order of errors due to common sources. These small errors ultimately contribute to the uncertainty on the background decomposition.

For the sake of simplifying the notation, we define the following variables:

$$r_j^k \equiv \frac{N_j^{\text{MC } k}}{N_j^{\text{MC On}}}, \quad X_j \equiv N_j^{\text{data On}}, \quad N^k \equiv N_{\text{total}}^{\text{data } k}. \quad (5.8)$$

We can then rewrite Equations 5.5–5.7 as a matrix equation:

$$\begin{bmatrix} 1 & 1 & 1 \\ r_{\text{NC}}^{\text{Off}} & r_{\nu_{\mu} \text{CC}}^{\text{Off}} & r_{\text{beam } \nu_e}^{\text{Off}} \\ r_{\text{NC}}^{\text{pHE}} & r_{\nu_{\mu} \text{CC}}^{\text{pHE}} & r_{\text{beam } \nu_e}^{\text{pHE}} \end{bmatrix} \begin{bmatrix} X_{\text{NC}} \\ X_{\nu_{\mu} \text{CC}} \\ X_{\text{beam } \nu_e} \end{bmatrix} = \begin{bmatrix} N^{\text{On}} \\ N^{\text{Off}} \\ N^{\text{pHE}} \end{bmatrix}, \quad (5.9)$$

$$R \quad X \quad = \quad N.$$

We wish to solve for X , and it would appear that the solution is simply $R^{-1}N$. Indeed, this would trivially minimize $\chi^2 = (RX - N)^T V^{-1} (RX - N)$, where V is the covariance matrix of $RX - N$. Unfortunately, the similarity of the background compositions in the Horn Off and High Energy configurations (see Figure 5.2) renders

the second and third rows of R almost linearly dependent. This leaves great freedom to one of the three components of X ; that freedom, in turn, is manifested as large uncertainties in the components of $R^{-1}N$. We mitigate this effect by overconstraining the system with additional information from the Monte Carlo. In particular, we compute a purely Monte Carlo-based prediction of X , denoted X_{MC} , and the associated covariance matrix V_{MC} . We then define a new χ^2 :

$$\chi^2 = \begin{bmatrix} (X - X_{\text{DD}})^T & (X - X_{\text{MC}})^T \end{bmatrix} \begin{bmatrix} V_{\text{DD}} & V_{\text{cor}} \\ V_{\text{cor}} & V_{\text{MC}} \end{bmatrix}^{-1} \begin{bmatrix} (X - X_{\text{DD}}) \\ (X - X_{\text{MC}}) \end{bmatrix}, \quad (5.10)$$

where $X_{\text{DD}} = R^{-1}N$ is the data-driven (“DD”) prediction, V_{DD} is its associated covariance matrix, and V_{cor} is a covariance matrix for the correlations between X_{DD} and X_{MC} . If we define

$$Y = \begin{bmatrix} X \\ X \end{bmatrix}, P = \begin{bmatrix} X_{\text{DD}} \\ X_{\text{MC}} \end{bmatrix}, V_P = \begin{bmatrix} V_{\text{DD}} & V_{\text{cor}} \\ V_{\text{cor}} & V_{\text{MC}} \end{bmatrix}, \quad (5.11)$$

we can rewrite Equation 5.10 as

$$\chi^2 = (Y - P)^T V_P^{-1} (Y - P). \quad (5.12)$$

Thus, the decomposition is achieved by χ^2 minimization rather than by the simple solution of a linear system. Statistical and systematic errors are included in V_P , with systematic errors evaluated in a manner similar to that described in Chapter 7; further details on the procedure used can be found in [107]. The equations written thus far have been for a single analysis bin at a time. However, one can easily enlarge the vectors and matrices in the equations above to include the applicable quantities from *all* analysis bins. While one bin’s contribution to the enlarged P can be computed independently from that of another bin, the systematic error evaluation involved in computing the enlarged V_P will account for correlations among all bins. Thus, the final χ^2 that we minimize in order to perform the decomposition will account for

all statistical and systematic errors as well as for correlations among the analysis bins. Additional constraints, discussed in [107], are included in the χ^2 to ensure that (1) the data components in each bin exactly sum to the total data in that bin and (2) no bins contain a negative number of events. We note that the minimization procedure computes not only the number of events of each Near Detector background component in each analysis bin but also an error matrix V^{HOOHE} whose off-diagonal elements describe correlations among components and analysis bins. The results of the decomposition are shown in Figures 5.3 and 5.4; averaging over the full analysis region (LEM PID > 0.6 , reconstructed energy between 1 and 8 GeV), we find that the Near Detector data sample is $(59.8 \pm 0.3)\%$ NC, $(28.8 \pm 0.2)\%$ ν_μ CC, and $(11.4 \pm 0.4)\%$ beam ν_e CC. Note the relatively small size of the errors, a consequence of cancellations to first order of systematic errors in ratios of Monte Carlo event counts and of the constraint that the components of the Near Detector data sample must add up to the total observed number of Near Detector events.

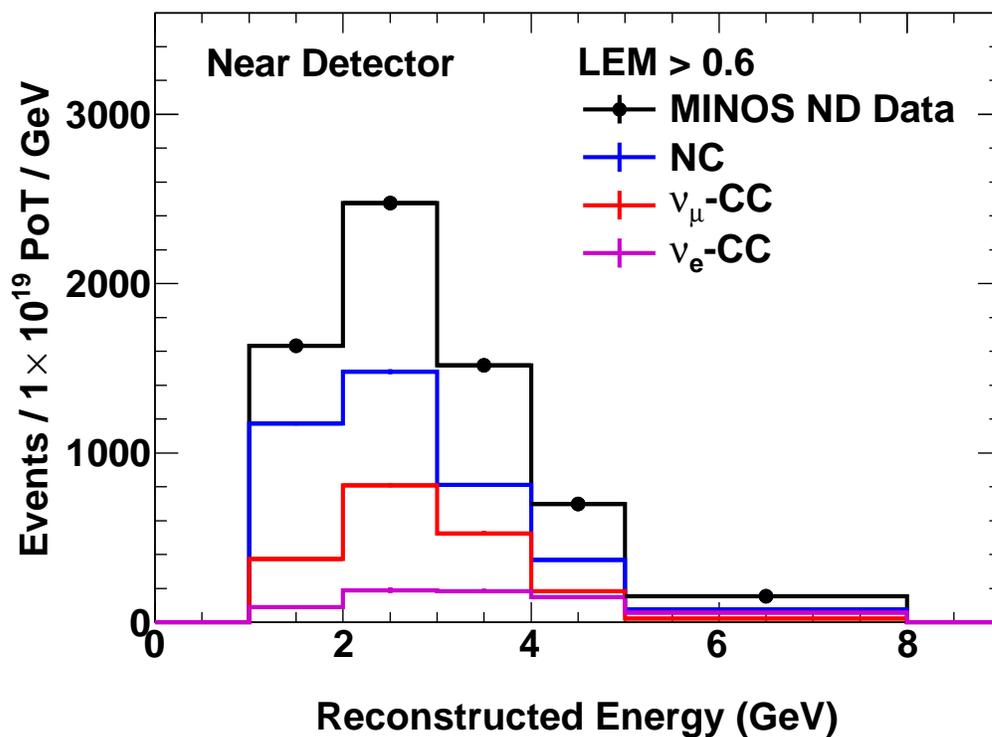


Figure 5.3: The components of the Near Detector data as determined by the HOOHE method, as a function of reconstructed energy. While the full decomposition is performed (and subsequently used in the analysis) in bins of LEM PID (in the analysis region) and reconstructed energy, this plot shows the decomposition projected onto the reconstructed energy axis. Uncertainties on the data are statistical and too small to be visible on this scale. Uncertainties on the individual components are combined statistical and systematic and are just barely visible for NC and ν_e CC in the 2–3 GeV bin.

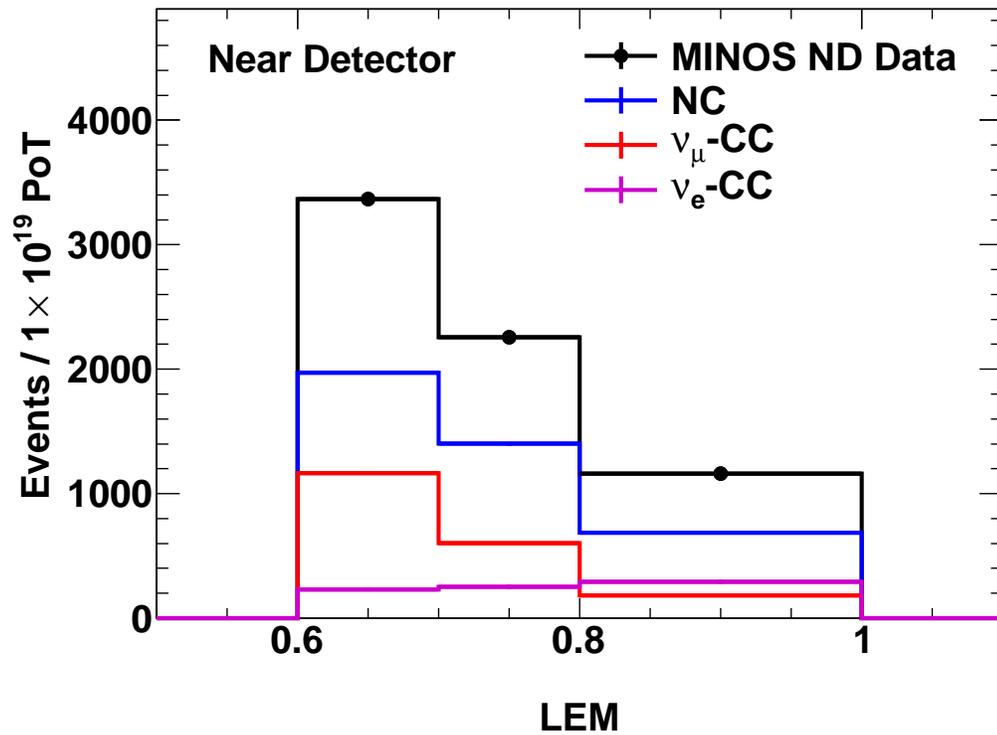


Figure 5.4: The components of the Near Detector data as determined by the HOOHE method, as a function of LEM PID in the analysis region. While the full decomposition is performed (and subsequently used in the analysis) in bins of LEM PID (in the analysis region) and reconstructed energy, this plot shows the decomposition projected onto the LEM PID axis. Uncertainties on the data are statistical and too small to be visible on this scale. Uncertainties on the individual components are combined statistical and systematic and are not visible on this scale.

Chapter 6

The Far Detector Prediction

With the Near Detector background components computed, we may proceed with the extrapolation procedure outlined in Chapter 3. Recall, however, that predicting the signal and the ν_τ CC background will require a different approach, which we describe in Section 6.2 below.

6.1 Predicting the ND-Type Background Components

The formula for the NC, ν_μ CC, and beam ν_e CC was presented as Equation 3.2. We change the notation slightly and rewrite the formula here:

$$F_i^{\alpha \text{ pred}} = N_i^{\alpha \text{ data}} \times \frac{F_i^{\alpha \text{ MC}}}{N_i^{\alpha \text{ MC}}}, \quad (6.1)$$

where $\alpha \in \{\text{NC}, \nu_\mu \text{ CC}, \text{beam } \nu_e \text{ CC}\}$ and i is the analysis bin (reconstructed energy and LEM PID) index – that is, the prediction is performed component-by-component *and* bin-by-bin. The “far-to-near” ratio $\frac{F_i^{\alpha \text{ MC}}}{N_i^{\alpha \text{ MC}}}$ accounts for the difference in neutrino flux at the two detectors (dominated by the usual $1/R^2$ fall-off, but also due to beam-line geometry, focusing, and decay kinematics), size difference between the detectors (fiducial volume and acceptance), and other interdetector differences that can lead to differences in selection efficiency. The $F_i^{\alpha \text{ MC}}$ term also includes oscillation proba-

bilities for CC events.¹ In this way, Equation 6.1 “extrapolates” the activity in the Near Detector to the Far Detector. Systematic uncertainties on the prediction due to various sources are discussed in Chapter 7.

6.2 Predicting the Signal ν_e CC and Background ν_τ CC Components

Unlike the event types in Section 6.1, the signal ν_e CC and background ν_τ CC components arise from oscillations of ν_μ s and thus do not occur in the Near Detector. Consequently, we cannot use Equation 6.1 to form predictions for these event types, but we can devise a similarly inspired extrapolation that accounts for the different nature of these events.

Naturally, we want this approach to be data-driven as well. Considering the source of the event types in question, a sample of clearly identified ν_μ CC events in the Near Detector data is the right place to start. These “ ν_μ CC-like” events are in contrast to the ν_μ CC events in the standard analysis sample (i.e., those passing the ν_e -like preselection and ultimately having a LEM PID value calculated) in that they typically have long muon tracks that distinguish them as ν_μ CC interactions. Beyond the usual fiducial volume cut, we select ν_μ CC-like events by requiring that

- The event contains a reconstructed track.
- One of the following is satisfied:²
 1. The reconstructed track passed the track fitting algorithm’s internal checks.
 2. There are no more than 5 planes separating the first U-view and first V-view planes encountered by the track; there are no more than 40 planes separating the last U-view and last V-view planes encountered by the track; and neither of these last planes is beyond the 270th plane of the detector.

¹NC events are flavor blind, so oscillations between flavors do not affect their rate.

²These requirements simply check the quality of the fit performed by the track fitting algorithm.

- The event’s value of the ν_μ -CC/NC discriminant used in the ν_μ disappearance analysis reported in [108] exceeds 0.3.³

The resulting event distribution is denoted $N_k^{\nu_\mu \text{ CC-like data}}$, with k indexing the muon-calibrated reconstructed energy. We can apply the same selection to the Near Detector and Far Detector Monte Carlo samples, with the latter assuming no oscillations away from ν_μ , to obtain $N_k^{\nu_\mu \text{ CC-like MC}}$ and $F_k^{\nu_\mu \text{ CC-like MC}}$, respectively. The usual extrapolation formula provides us with a data-driven prediction of the ν_μ CC-like reconstructed energy spectrum in the Far Detector:

$$N_k^{\nu_\mu \text{ CC-like data}} \times \frac{F_k^{\nu_\mu \text{ CC-like MC}}}{N_k^{\nu_\mu \text{ CC-like MC}}} \quad (6.2)$$

Note the care we have taken to define the current sample as ν_μ CC-*like*, as the sample is not 100% pure and the selection is not 100% efficient. We can correct for these effects using the purity and efficiency as estimated by the Monte Carlo, but we must first “migrate” from reconstructed event energy to true neutrino energy; that is, we transform the distribution from one of reconstructed event energy to one of true neutrino energy, as the purity and efficiency are properly defined in terms of the latter. This is achieved using a “reconstructed-to-true energy migration matrix” $RT_{kj}^{\nu_\mu \text{ CC-like MC}}$, populated with ν_μ CC-like events in the Far Detector and with each reconstructed energy bin’s contents normalized to unity.⁴ The formula:

$$\sum_k N_k^{\nu_\mu \text{ CC-like data}} \times \frac{F_k^{\nu_\mu \text{ CC-like MC}}}{N_k^{\nu_\mu \text{ CC-like MC}}} \times RT_{kj}^{\nu_\mu \text{ CC-like MC}} \quad (6.3)$$

thus provides a data-driven prediction of the ν_μ CC-like *true* energy spectrum in the Far Detector, where true neutrino energy is indexed by j . Note that the migration

³This discriminant, described in [109], is the output of a multivariate likelihood based on four variables that characterize a muon track. It selects ν_μ CC events with high efficiency while providing very high background rejection against NC events. The cut value of 0.3 was determined in a sensitivity optimization study, reported in [110], performed in preparation for the cited ν_μ disappearance analysis.

⁴This normalization turns each reconstructed energy bin into a probability distribution in true neutrino energy, reflecting the fact that an event with a certain reconstructed energy could be produced, with some probability, by a neutrino with any true energy.

matrix effectively “divides up” a single event of a certain reconstructed energy over a range of true neutrino energies; the event count in a given true neutrino energy bin j , then, is the sum of the fractional contributions to the bin from events of a variety of reconstructed energies. As indicated earlier, we multiply by the purity P_j to correct for the fact that the sample defined in Equation 6.3 is not purely ν_μ CC. We also divide by the efficiency E_j of the ν_μ CC-like selection defined above, thus “restoring” the true ν_μ CC events that were cut by the selection. Finally, we obtain

$$F_j^{\nu_\mu \text{ CC fid}} = \left[\sum_k N_k^{\nu_\mu \text{ CC-like data}} \times \frac{F_k^{\nu_\mu \text{ CC-like MC}}}{N_k^{\nu_\mu \text{ CC-like MC}}} \times RT_{kj}^{\nu_\mu \text{ CC-like MC}} \right] \frac{P_j}{E_j}. \quad (6.4)$$

Equation 6.4 thus provides the predicted ν_μ CC true neutrino energy spectrum in the fiducial volume (“fid”) of the Far Detector, from which we may obtain the predicted signal ν_e CC and background ν_τ CC spectra in the Far Detector. In fact, the formulas for these components are identical, so we will write the formula once, for neutrino flavor $\ell \in \{e, \tau\}$.

We start by correcting for the difference between the ν_μ and ν_ℓ cross sections, as a function of true neutrino energy. Next, we multiply by the oscillation probability $P_j(\nu_\mu \rightarrow \nu_\ell)$ using the desired oscillation parameters. To obtain the electron-calibrated reconstructed energy distribution, we multiply by a true-to-reconstructed energy migration matrix $TR_{ji}^{\nu_\ell \text{ CC MC}}$ – populated with true ν_ℓ CC events in the Far Detector fiducial volume and with each true neutrino energy bin’s contents normalized to unity⁵ – and sum over the true neutrino energy indexed by j . At this point, we have the reconstructed energy spectrum for ν_ℓ CC events in the Far Detector fiducial volume. Multiplying by the Monte Carlo–based cumulative selection efficiency $\epsilon_{im}^{\nu_\ell \text{ CC}}$ of the preselection and bin m of the LEM PID for ν_ℓ CC events in the Far Detector,

⁵This matrix is conceptually similar to $RT_{kj}^{\nu_\mu \text{ CC-like MC}}$ in the preceding equations, with the roles of reconstructed energy and true energy reversed and the event type used to populate the matrix changed.

we obtain

$$F_{im}^{\nu_\ell \text{ CC pred}} = \left[\sum_j F_j^{\nu_\mu \text{ CC fid}} \times \frac{\sigma_j^{\nu_\ell}}{\sigma_j^{\nu_\mu}} \times P_j(\nu_\mu \rightarrow \nu_\ell) \times TR_{ji}^{\nu_\ell \text{ CC MC}} \right] \epsilon_{im}^{\nu_\ell \text{ CC}}. \quad (6.5)$$

Note that in the case of the optimal LEM PID cut of 0.7, m will have only one possible value.

6.2.1 Selection Efficiency Considerations

Recall that Equation 6.1, by design, accounts for changes in the NC, ν_μ CC, and beam ν_e CC selection efficiencies between the Near and Far Detectors via the Monte Carlo ratio $\frac{F_i^{\alpha \text{ MC}}}{N_i^{\alpha \text{ MC}}}$ (which, as noted before, is largely independent of the Monte Carlo modeling details). In addition, the selection efficiency in data is known, via the $N_i^{\alpha \text{ data}}$ term, so the Far Detector predictions we obtain for these three event types are faithful representations of the real (i.e., data) selection efficiencies for their respective event types. Equation 6.5, on the other hand, relies on $\epsilon_{im}^{\nu_\ell \text{ CC}}$ – a purely Monte Carlo–based quantity – for the ν_τ CC and signal ν_e CC selection efficiencies. In light of the LEM PID data–MC disagreement discussed in Section 4.3.6, this approach requires further consideration.

As we will see in Section 6.3, the expected ν_τ CC background is quite small ($\sim 4\%$) and relatively constant for several assumptions of θ_{13} and δ_{CP} . In Chapter 7, we will find that the very large uncertainty on the ν_τ CC cross section translates into a systematic error on the ν_τ CC background prediction of approximately 50%. Hadronic shower mismodeling may produce a significant data–MC difference in the selection efficiency $\epsilon_{im}^{\nu_\tau \text{ CC}}$, but in light of these observations, this possible difference is expected to be largely inconsequential to the analysis. The signal ν_e CC prediction, on the other hand, is obviously dependent on θ_{13} and δ_{CP} , as well as the neutrino mass hierarchy, and the systematic error on its prediction is not as large as it is for the ν_τ CC prediction. Taking these observations as well as the signal prediction’s central role in this analysis into account, it would be wise to determine whether or not a data–MC difference exists in the signal ν_e CC selection efficiency and, if it does

exist, to account for it.

One particularly straightforward approach to this investigation would be to evaluate the LEM PID for data and Monte Carlo ν_e CC events and compare the selection efficiencies in each sample. Unfortunately, one cannot know if an event in data is actually ν_e CC – this is, after all, the main challenge in this analysis. Our solution is to perform the study on a set of specialized events dubbed “MRE” (muon-removed electron-added) events, which are constructed as follows [111] from Near Detector data and Monte Carlo events passing the ν_μ CC-like selection defined in Section 6.2:

1. Isolate and remove the reconstructed muon track from the event, deleting all hits associated with the track and leaving behind the shower remnant.⁶
2. Record the removed track’s position and momentum.
3. Generate a Monte Carlo electron with position and momentum identical to those of the removed track.
4. Convert the simulated electron to a list of hits.
5. Merge the electron hits with the shower remnant from the original event.
6. Reconstruct the resulting MRE event.

In Section 4.3.6, we discussed the role played by hadronic shower mismodeling in producing the data–MC disagreement observed in the LEM PID distribution; naturally, we suspect this mismodeling to be primarily responsible for any data–MC difference in signal ν_e CC selection efficiency. MRE events retain the source of this difference (the original hadronic showers) but also incorporate the effect of electron showers. Analysis of the MRE samples thus allows us to determine – as directly as possible – the impact of hadronic shower mismodeling on signal ν_e CC selection efficiency. In particular, we can use the data–MC ratio of the MRE event selection efficiencies in each analysis bin as a multiplicative correction factor to $\epsilon_{im}^{\nu_e CC}$,

⁶This step, followed by re-reconstruction of the shower remnant, is referred to as the MRCC (muon-removed charged-current) process [112], and the resulting event is an MRCC event. These events constitute the analysis sideband that is signal-free by construction, referred to in Chapter 3.

thus obtaining an improved signal prediction that better reflects the signal selection efficiency in data. Figure 6.1 shows the MRE data and Monte Carlo selection efficiencies, as well as their associated correction factor, as a function of the reconstructed energy for events with $\text{LEMPID} > 0.7$. For the full analysis as well as for the predictions shown in Section 6.3, the correction factor computed in each of the 15 analysis bins – that is, as a function of the reconstructed energy for $0.6 < \text{LEMPID} < 0.7$, $0.7 < \text{LEMPID} < 0.8$, and $0.8 < \text{LEMPID} < 1.0$ – is used.

6.3 Far Detector Predictions

We may now compute the number of background events of each type and the number of signal events predicted in the Far Detector, collectively referred to as the “Far Detector predictions,” for any desired set of oscillation parameters. As mentioned in Chapter 3, we will ultimately do this over a grid of θ_{13} and δ_{CP} values for both neutrino mass hierarchies in order to determine the sensitivity of the analysis and to compute the confidence intervals on the oscillation parameters implied by the Far Detector data. For now, however, we will compute predictions assuming the following representative parameter values, in order to demonstrate the method and to obtain a sense of what we may expect in the Far Detector data:

- $\theta_{23} = \pi/4$ – That is, we assume maximal atmospheric mixing, which is consistent with the best available result which finds that $\sin^2 2\theta_{23} > 0.92$ at 90% C.L. [10].
- $\theta_{12} = 0.60$ [51].
- $\theta_{13} = 0, \theta_{13} = 0.16$ – These are sample values of θ_{13} ; corresponding to $\sin^2 2\theta_{13} = 0$ and $\sin^2 2\theta_{13} = 0.1$, respectively; for which we will compute Far Detector predictions.
- $\Delta m_{21}^2 = 7.59 \times 10^{-5} \text{ eV}^2$ [10].
- $\Delta m_{32}^2 = 2.32 \times 10^{-3} \text{ eV}^2$ [45] – The absolute value of Δm_{32}^2 is from [45]. We are assuming that it is positive – i.e., the normal neutrino mass hierarchy – for

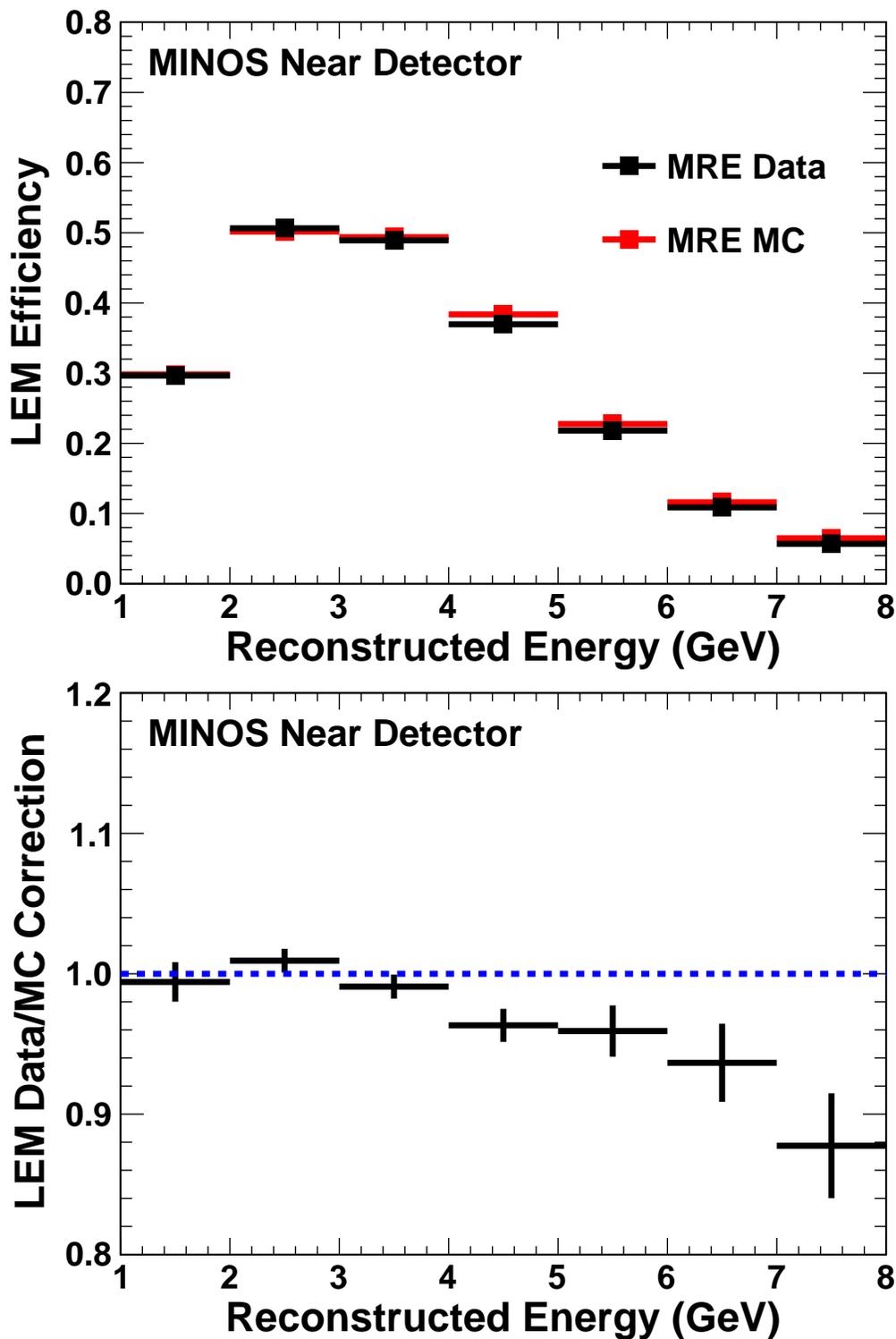


Figure 6.1: (Top) The LEMPID > 0.7 selection efficiencies for MRE data and Monte Carlo. The errors are statistical and are not visible at this scale. (Bottom) The ratio of the efficiencies above. This serves as a correction factor to the signal ν_e CC selection efficiency. The errors are statistical.

the present predictions.

- $\delta_{\text{CP}} = 0$ – There is no definitive experimental data providing any information about the value of δ_{CP} . We assume it to be zero for our present purposes.

Table 6.1 shows the predicted event counts in each LEM PID bin of the analysis region and for the single optimal cut on LEM PID, summed over the energy range of the analysis (1–8 GeV), for the oscillation assumptions listed above. A representative set of figures follows, for the $\sin^2 2\theta_{13} = 0.1$ assumption only.⁷ These include the predicted LEM PID distribution (Figure 6.2), the predicted reconstructed energy distribution for LEM PID > 0.7 (Figure 6.3), and the predicted reconstructed energy distribution in each LEM PID bin of the analysis region (Figure 6.4).

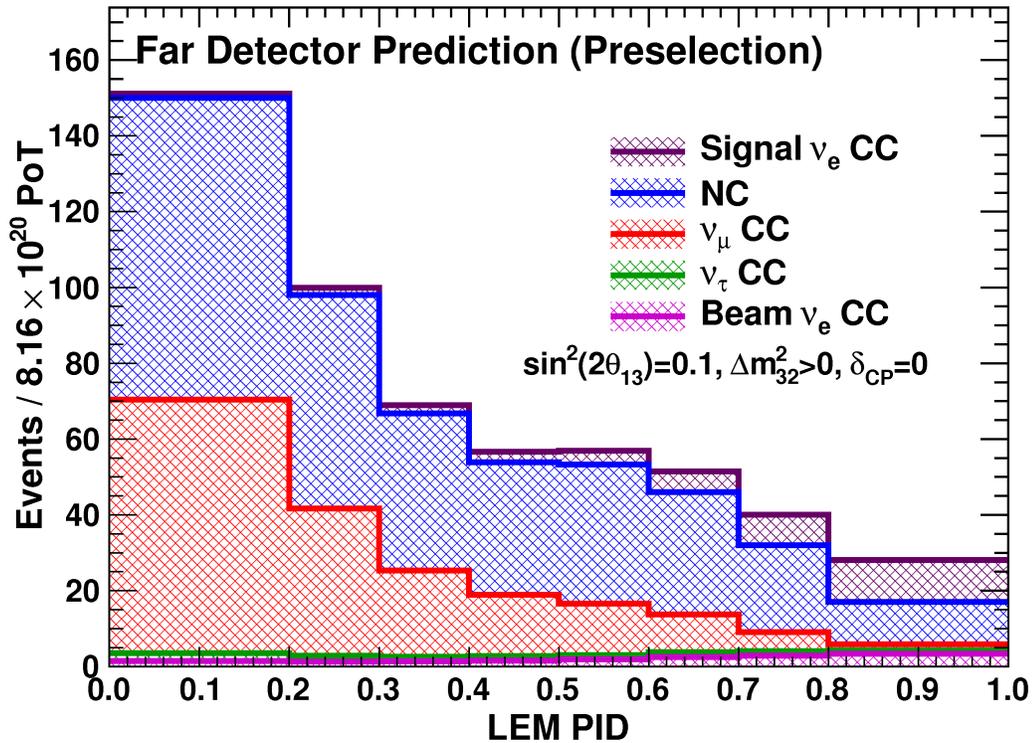


Figure 6.2: Stacked histograms showing the predicted LEM PID distribution for all events passing the ν_e -like preselection. Note that with increasing LEM PID, the total background decreases while the signal increases.

⁷A close look at Table 6.1 reveals that the four background components are affected negligibly by a change in θ_{13} . Thus, one can visualize the $\sin^2 2\theta_{13} = 0$ version of these figures simply by ignoring the signal ν_e CC contributions.

	Range of LEM PID			
	0.6–0.7	0.7–0.8	> 0.8	> 0.7
$\sin^2 2\theta_{13} = 0$				
NC	32.4	22.9	11.2	34.1
ν_μ CC	10.0	5.1	1.6	6.7
Beam ν_e CC	2.6	2.9	3.5	6.4
ν_τ CC	1.3	1.3	0.9	2.2
Total Background	46.3	32.2	17.2	49.4
Signal ν_e CC	0.1	0.1	0.1	0.2
$\sin^2 2\theta_{13} = 0.1$				
NC	32.3	22.9	11.2	34.1
ν_μ CC	9.9	5.1	1.6	6.7
Beam ν_e CC	2.5	2.8	3.4	6.2
ν_τ CC	1.3	1.2	0.9	2.1
Total Background	46.1	32.0	17.0	49.0
Signal ν_e CC	5.4	8.1	11.0	19.1

Table 6.1: Far Detector predictions for the oscillation assumptions listed above assuming a total exposure of 8.16×10^{20} PoT. Note that event counts are shown for each LEM PID bin of the analysis region and for the single optimal cut on LEM PID. The nonzero signal in the $\sin^2 2\theta_{13} = 0$ case is due to Δm_{21}^2 -driven (solar) oscillations.

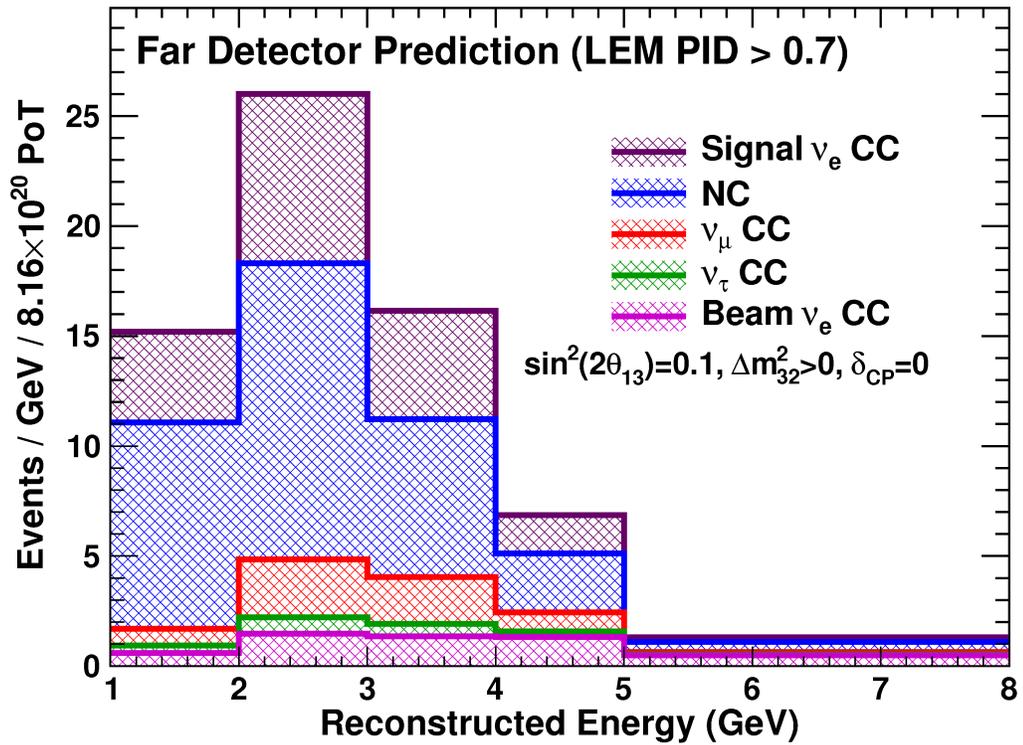


Figure 6.3: Stacked histograms showing the predicted reconstructed energy distribution for the single optimal cut LEMPID > 0.7. Note that the signal is largest in the 2–3 GeV bin, which coincides with the Δm_{32}^2 -driven (“atmospheric”) oscillation maximum for the MINOS baseline.

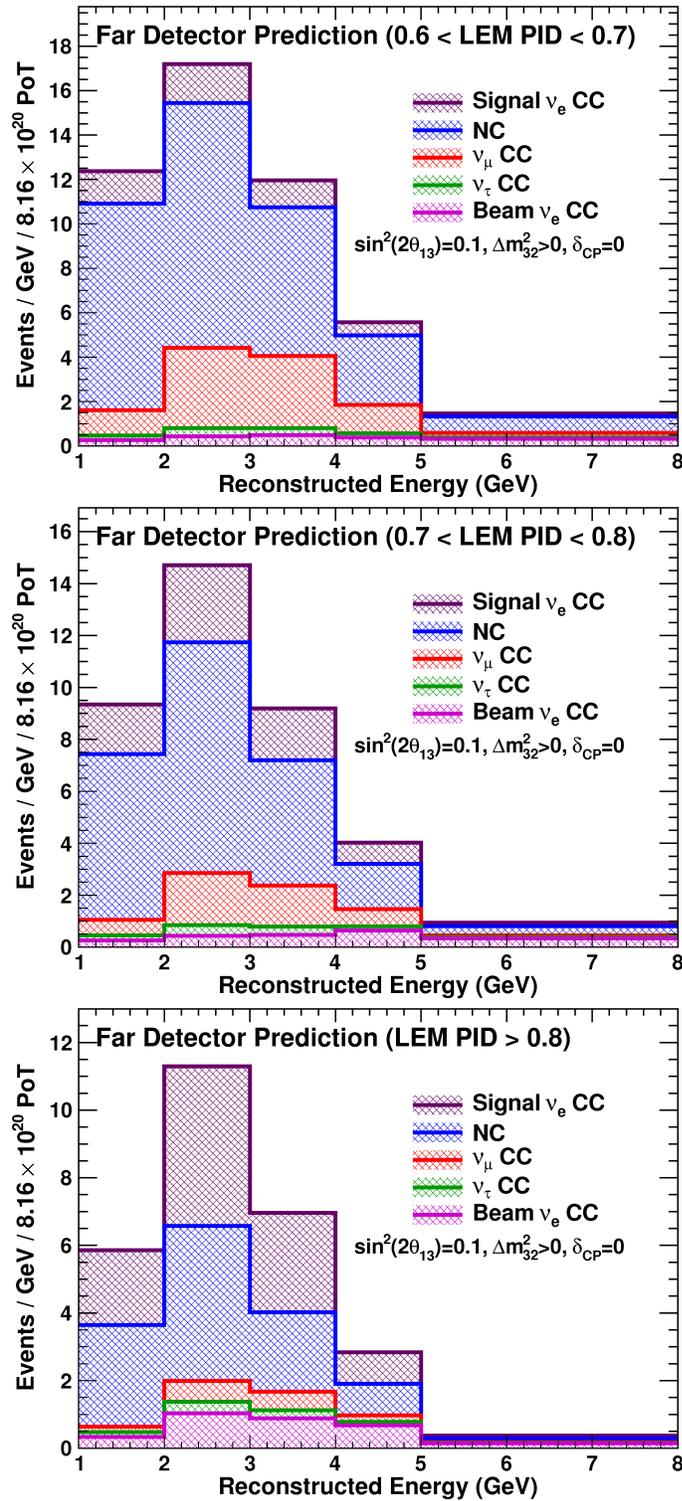


Figure 6.4: Stacked histograms showing the predicted reconstructed energy distributions for the LEM PID bins in the analysis region. Note that with increasing LEM PID (from the top plot to the bottom plot), both the number of signal events and the fraction of the events that are signal increase.

Chapter 7

Systematic Errors

In our analysis, a variety of systematic effects produce associated errors on the Far Detector predictions described in Chapter 6. We wish to quantify these errors and ultimately combine them in a covariance matrix V for the binned total Far Detector prediction (signal plus background for a given set of oscillation parameters). The systematic effects considered in our analysis are, by their nature, Gaussian and independent. Some of these effects are on the order of a few percent and can thus be approximated as being linear in their impact on the prediction; others are considerably larger but are known to affect the prediction in a linear fashion. We can state these properties more precisely: a given systematic effect k is governed by some parameter x_k (e.g., a physical model parameter or a parameter characterizing the detector performance) that is Gaussian-distributed with mean μ_k and variance σ_k^2 and is independent of all other x_l for $l \neq k$, and the total prediction $F_i^{\text{total pred}}$ in analysis bin i either varies approximately linearly with $(x_k - \mu_k)$ because σ_k/μ_k is small or varies linearly with $(x_k - \mu_k)$ because we know this to be the case. As a result, if we define $\delta_k F_i^{\text{total pred}}$ to be the 1σ shift in $F_i^{\text{total pred}}$, induced by varying x_k from μ_k to $\mu_k + \sigma_k$, we can approximate the covariance matrix elements as a sum of products of these shifts over the various systematic effects:

$$V_{ij} = \sum_k \delta_k F_i^{\text{total pred}} \times \delta_k F_j^{\text{total pred}}, \quad (7.1)$$

where i, j are analysis bin indices.¹ We actually compute these shifts on the total prediction by summing the shifts $\delta_k F_i^{\alpha \text{ pred}}$ on the individual components $F_i^{\alpha \text{ pred}}$ of the prediction:

$$\begin{aligned} \delta_k F_i^{\text{total pred}} = & \delta_k F_i^{\text{NC pred}} + \delta_k F_i^{\nu_\mu \text{ CC pred}} + \delta_k F_i^{\text{beam } \nu_e \text{ CC pred}} \\ & + \delta_k F_i^{\nu_\tau \text{ CC pred}} + \delta_k F_i^{\text{signal } \nu_e \text{ CC pred}}. \end{aligned} \quad (7.2)$$

Our primary task in this chapter, then, is to evaluate the terms on the right side of Equation 7.2.² Note that in principle, we would need to compute these shifts and, subsequently, the full covariance matrix for *every* desired set of oscillation parameters, as the individual event types are sensitive to them. Instead, we compute the *relative* errors $\frac{\delta_k F_i^{\alpha \text{ pred}}}{F_i^{\alpha \text{ pred}}}$ assuming oscillation parameters as listed in Section 6.3 – specifically, with $\theta_{13} \neq 0$ – and make the approximation that these relative errors do not vary with subsequent changes in the assumed oscillation parameters. We may then compute the prediction $F_i^{\alpha \text{ pred}}$ for each event type α in each analysis bin i under some particular choice of oscillation parameters and multiply it by the appropriate relative error to obtain the corresponding shift in the prediction. Finally, we obtain the full covariance matrix for that particular set of oscillation parameters via Equations 7.2 and 7.1.

The following sections provide a brief overview of the various systematic effects that affect this analysis and the basic means by which their induced systematic errors are quantified. More detailed discussions can be found in the references cited herein and in [113]. We save a summary of the approximate sizes of these systematic errors to the end of the chapter.

¹We explicitly note that in this approximation, the expectation value of the product of deviations from means normally found in the definition of a covariance matrix is replaced by the product of the described shifts. This implies 100% bin-to-bin correlations for any single systematic effect, which is appropriate for the systematic effects relevant to our analysis.

²Appendix C indicates how this is done for the Near Detector-type background components.

7.1 Effects Common to Both Detectors

A number of systematic effects are common to the Near and Far Detectors, and while the associated error in an individual detector may be large, these errors cancel to first order³ and thus produce comparatively much smaller errors on the Far Detector prediction. These effects include uncertainties in the neutrino flux, uncertainties in the neutrino interaction modeling, and absolute energy calibration uncertainties.

7.1.1 Neutrino Flux Uncertainties

Recall from Section 2.6.1 that our nominal neutrino flux is the product of a simulation by FLUKA05 and GEANT3 and that the nominal Monte Carlo produced using that flux is tuned via the SKZP technique so that the simulated Near Detector ν_μ CC spectrum better matches the observed spectrum. In particular, SKZP is implemented as an event weighting scheme, and a 1σ error on these weights – resulting from uncertainties in hadronic production in the target as well as uncertainties in various beam and target parameters – is computed as well. These errors on the weights are used to reweight the events in order to evaluate the SKZP systematic error.

The flux simulation used in MINOS was recently changed to FLUGG [114, 115], and the validation of the new simulation [116] showed an increase in the beam ν_e flux relative to FLUKA05 on the order of 20% in the Near Detector and 10% in the Far Detector. Time constraints prevented us from adopting the FLUGG-based Monte Carlo for our analysis, so we account for this difference with a systematic error. Specifically, we reweight the events initiated by beam ν_e s (primarily beam ν_e CC events, plus a small number of NC events) by the following double ratio as a function of true neutrino energy:

$$\frac{\frac{\text{number of } \nu_e \text{ generated by FLUGG}}{\text{number of } \nu_\mu \text{ generated by FLUGG}}}{\frac{\text{number of } \nu_e \text{ generated by FLUKA05}}{\text{number of } \nu_\mu \text{ generated by FLUKA05}}}, \quad (7.3)$$

and compute the resulting fractional shift in the prediction for each event type to

³See Appendix C, particularly Equation C.9, for a demonstration of this property.

determine the associated systematic error.

Finally, the total flux in Runs II and III was found to be decaying as a function of time due to degradation of the NuMI target – specifically, target NT-02, which had been exposed to 5.80×10^{20} PoT by the end of Run III (see Table 2.2). Studies showed [117] that the target degradation is well modeled by assuming that the two segments of the target at the shower maximum were missing.⁴ The associated systematic error was evaluated by forming a weighting function from the neutrino fluxes simulated with and without those two segments in place and applying it only to the simulated events in Runs II–VI.⁵

7.1.2 Neutrino Interaction Modeling Uncertainties

Uncertainties in the various cross section parameters used in the NEUGEN simulation produce uncertainties in event counts that are expected to cancel to first order in the far-to-near ratio. Nonetheless, small residual effects could be important. To evaluate the impact of this systematic effect, we modify the cross section by varying each underlying parameter within its recommended [118] $\pm 1\sigma$ error; then, the ratio of the systematically modified cross section to the nominal cross section is applied as a weight on the simulated events. While this approach is appropriate for ν_e CC and ν_μ CC interactions, there is significantly greater uncertainty in the ν_τ CC cross section. In particular, it was shown [119] that an overall $\pm 50\%$ uncertainty should be applied to quasi-elastic and resonance ν_τ CC events, and an overall $\pm 10\%$ uncertainty should be applied to deep inelastic ν_τ CC events.⁶ These two uncertainties average to an uncertainty of approximately $\pm 45\%$ in the region of interest to this analysis.

Some of the largest uncertainties in our simulation come from our modeling of the hadronization process associated with the quark inside a nucleon that is scattered by

⁴Such studies were necessary because the target is too hot to be inspected visually.

⁵Note that this represents a conservative estimate of the associated error, as the target used in Runs IV–VI was new and did not show signs of degradation.

⁶The rather large uncertainty for quasi-elastic and resonance events is a consequence of our lack of knowledge about the pseudoscalar form factor, which is proportional to the square of the ratio of the lepton mass to the nucleon mass and thus makes a significant contribution to the ν_τ cross section (and very small contributions to the ν_e and ν_μ cross sections).

the incident neutrino. The model is tuned to data from previous bubble chamber neutrino experiments and experiments with pion beams scattering on iron, but these data do not fully cover the kinematic range of interest in this analysis. Several aspects of the hadronization model, determined to be most important to this analysis and poorly constrained by external data, were identified; their effects on the analysis were then evaluated via a reweighting scheme [120]. Systematic errors introduced by uncertainties in the INTRANUKE intranuclear cascade model were computed by yet another reweighting scheme [121].

7.1.3 Energy Scale Uncertainties

Two energy scale uncertainties affect both detectors in a similar fashion and are thus expected to cancel to first order in the far-to-near ratio, but we check them for any residual effects. Errors in the absolute energy scale in the MINOS detectors could affect the analysis through the preselection energy cuts, which constrain the selected events to be within the 1–8 GeV window, and through possible variations in the LEM PID input variables. Calibration studies [122] have found the error on the absolute energy scale in the MINOS detectors to be 5.7%. We evaluate the associated systematic error by reconstructing raw nominal Monte Carlo events using shifted values of the relevant calibration constant and computing the resulting relative errors in the predictions.

As we have seen, this analysis makes use of events that are composed of a wide range of possible ratios of electromagnetic energy to hadronic energy, and LEM attempts to distinguish among signal and background events based on the topological differences between electromagnetic and hadronic showers. Thus, the possibility that the hadronic and electromagnetic energy scales may vary relative to each other in an uncorrelated fashion constitutes another systematic effect. To compute the associated systematic error, we begin by determining the dependence of the LEM selection efficiency on true electromagnetic energy and true hadronic energy. While holding the electromagnetic portion of the total energy constant, we shift the hadronic energy

by $\pm 5\%$ (as suggested by CalDet hadronic data) and compute a new effective selection efficiency. This allows the calculation of appropriate event weights. However, a change in the hadronic energy changes the total reconstructed energy of the event, so we must also determine the dependence of the reconstructed energy on the hadronic energy and apply it as an energy shift on individual weights. Thus, this systematic error is evaluated by a combination of reweighting and energy shifting.

7.2 Effects Differing Between the Detectors

Certain systematic effects are present in both detectors in such a way that they are not expected to cancel in the far-to-near ratio. These include the relative event rate normalization, relative calibration errors, and PMT crosstalk modeling details.

7.2.1 Normalization Uncertainties

The relative event normalization error is due to uncertainty in relative PoT exposure and fiducial mass of the Near and Far Detectors. The error is estimated [123] to be 2.0% for the NC, ν_μ CC, and beam ν_e CC components and 2.4% for the ν_τ CC background and the ν_e CC signal. The 2.0% error is the sum in quadrature of several errors: 0.32% for PoT counting, 0.2% for steel thickness, 0.2% from scintillator thickness, and 2% from uncertainty in the fiducial volume. The additional uncertainty in the ν_τ CC background and ν_e CC signal arises from a possible difference between the detectors in the efficiency for selecting ν_μ CC-like events (recall that the ν_τ CC and signal ν_e CC predictions are based on the ν_μ CC-like sample).

7.2.2 Calibration Uncertainties

Uncertainties in the calibration chain lead not only to differences in the reconstructed energy of events but also to changes in the event topology. As was done for the absolute energy scale systematic error, various calibration-related systematic errors were evaluated by reconstructing raw nominal Monte Carlo events using modified

values of the relevant calibration constants and computing the resulting relative error in the prediction. In particular, the following effects were investigated.

Relative energy scale. This effect was studied in the same manner as the study of the absolute energy scale uncertainty, except only one detector's energy scale was shifted and the other's was held constant. In particular, the Near Detector energy scale was shifted by $\pm 1.9\%$ while the Far Detector energy scale was held constant; then, the Far Detector energy scale was shifted by $\pm 0.9\%$ while the Near Detector energy scale was held constant. The relative energy scale uncertainties quoted for each detector were determined in [124].

PMT gain. The PMT gains in each detector are known to within a systematic shift of 5% and to within random channel-to-channel variation of 7%. Raw nominal Monte Carlo events were reconstructed with the gain shifted by $\pm 5\%$, with an additional 7% random Gaussian variation independently applied to each channel. Systematic errors were then computed in the same fashion as for the relative energy scale – i.e., shift the gains in one detector while keeping the other detector's gains unchanged.

Attenuation. Following the attenuation correction step of the calibration, residual differences on the order of 1% remain in the mean response along the strip. To assess any potential impact of these differences on the analysis, the response as a function of position along a strip in Monte Carlo is scaled to match the corresponding data response, and the resulting fractional shift in the prediction relative to the nominal Monte Carlo is calculated.

Strip-to-strip. After the strip-to-strip calibration is performed, variations on the order of 0.5% among strips may persist. We simulate this effect by varying the number of SigCor (see Section 2.3) in each strip, independently of other strips, according to a Gaussian distribution of width 0.5%.

Linearity. In order to determine how well we correct our data for PMT nonlinearity, we vary the number of SigLin in each strip, independently of other strips,

according to a Gaussian distribution whose width is set equal to the error on the SigLin (see Section 2.3) value calculated by the calibration computer code when performing the linearity fit. The resulting relative error in the prediction is then calculated.

7.2.3 PMT Crosstalk Modeling Uncertainties

The multianode PMTs exhibit two types of crosstalk. Electrical crosstalk results in a peak at 0.3 photoelectrons in the strip pulse height distribution and is primarily the result of secondary electrons leaking into neighboring pixels downstream in the dynode chain. Optical crosstalk produces a peak at 1 photoelectron and is the result of a secondary electron cascade in a nonilluminated pixel (e.g., due to light leaking from an optical fiber illuminating a given pixel to a neighboring pixel). Variations in the detectors' light readout systems produce differences in their crosstalk characteristics. Thus, a deficiency in the crosstalk simulation can result in an inaccurate far-to-near ratio; fortunately, since the reconstruction discards hits below 2 photoelectrons [125] and LEM discards hits below 3 photoelectrons (Section 4.3.2.3), much of this potential inaccuracy is mitigated. The remaining effect is small and can be treated as a systematic error.

We estimate the electrical crosstalk systematic error by comparing the number of selected events obtained using the standard simulation with that obtained using a simulation with electrical crosstalk turned off. The effect turned out to be negligible [113] and therefore is ignored. The optical crosstalk systematic was evaluated by comparing the number of events selected in the standard simulation with the number of events selected in a simulation using the more accurate crosstalk map described in [97]. The resulting fractional change in the far-to-near ratio was taken as the crosstalk systematic error.

7.3 Effects Unique to ν_τ CC and Signal ν_e CC Events

The ν_τ CC and signal ν_e CC component predictions are susceptible to a unique set of systematic effects due to their reliance on the ν_μ CC-like sample and to the fact that they occur only in the Far Detector. We begin by noting that these components are sensitive to the SKZP, FLUGG, target degradation, and NEUGEN cross section systematics through their dependence on the ν_μ CC-like sample. They are also sensitive to the other neutrino interaction systematics, calibration systematics, and PMT crosstalk systematic in the same way as are the NC, ν_μ CC, and beam ν_e CC components, save for any first-order cancellations in a far-to-near ratio.

Certain systematic effects unique to ν_τ CC and signal ν_e CC events are the result of these components' reliance on the ν_μ CC-like sample. These include a 50% error on the number of NC events selected by the ν_μ CC-like selection and a 10% error on the ν_μ CC energy scale. These shifts are applied to the Monte Carlo samples used in predicting the ν_τ CC and signal ν_e CC components, and the resulting shifts in the predictions are used to compute the associated systematic errors.

Finally, the signal ν_e CC component entails two additional systematic errors: one associated with the MRE correction discussed in Section 6.2.1 and another associated with data–MC discrepancies in the modeling of the interactions of electrons in CalDet. Essentially, these errors account for the impact of uncertainties in the modeling of the hadronic and electromagnetic components, respectively, of ν_e CC events on the selection efficiency for these events. Further details on the systematic error associated with the MRE correction can be found in [126].

7.4 Systematics and the HOOHE Decomposition

In Section 5.2, we noted that the covariance matrix V_P (Equation 5.11) used in the HOOHE χ^2 (Equation 5.12) includes systematic errors. This covariance matrix is computed in a manner conceptually similar to that discussed thus far in this chapter, i.e., via an appropriately modified form of Equation 7.1. While the details of the

method are covered in [107], we note that the key differences lie in the absence of any Far Detector Monte Carlo and in the use of three different Near Detector Monte Carlo samples – Horn On, Horn Off, and High Energy. Systematically modified versions of each of these samples are prepared for each of the applicable systematic effects discussed in this chapter, and the resulting shifts in the components of the prediction P (Equation 5.11) are used to compute the elements of V_P .

7.5 Summary

To summarize the various systematic effects and indicate the approximate size of the associated errors, we present the relative error on each component, due to the systematic effects described, in Tables 7.1 and 7.2 for the LEM PID > 0.7 selection averaged over 1–8 GeV. To keep these systematic errors in perspective, we use the Far Detector predictions for $\sin^2 2\theta_{13} = 0.1$ given in Table 6.1 to compute the statistical errors on the total ND-type background, ν_τ CC background, and the signal and find them to be 15%, 69%, and 23%, respectively.⁷

In conclusion, we make a series of notes and observations regarding the errors:

- The total systematic error in each case (ND-type background, ν_τ CC background, signal ν_e CC) is considerably less than the statistical error. Thus, the sensitivity of this analysis is largely limited by statistics.
- The systematic errors induced by effects discussed in Section 7.1 are indeed dramatically reduced in the far-to-near ratio. For example, the absolute energy scale shifts in the Near and Far Detectors⁸ were 11.5% and 9.7% [113], respectively (note that the signs are the same), but the shift in the far-to-near ratio is only -1.2% . Yet another example: the total systematic error associated with uncertainties in the hadronic model is 31.8% and 32.3% in the Near and Far Detectors, respectively, but the corresponding total error on the far-to-near

⁷While the backgrounds (and thus their statistical errors) do not change appreciably with θ_{13} , the signal decreases (and thus its statistical error *increases*) with decreasing θ_{13} .

⁸See Appendix C for an indication of how individual-detector systematic shifts are computed.

Systematic Effect	NC	ν_μ CC	beam ν_e CC	Total
SKZP	-0.36%	-0.16%	4.50%	0.27%
FLUGG	-0.10%	0.00%	-5.04%	-0.69%
Target Degradation	-0.14%	-0.11%	0.14%	-0.10%
Cross Section	0.24%	0.79%	0.46%	0.23%
Hadronic Model	0.89%	4.70%	2.14%	0.84%
INTRANUKE	0.62%	2.60%	1.69%	0.67%
Absolute Energy Scale	-1.52%	0.00%	-0.90%	-1.21%
EM/Hadronic Energy Scale	0.33%	0.25%	0.08%	0.28%
Normalization	2.00%	2.00%	2.00%	2.00%
Relative Energy Scale (FD)	1.32%	2.77%	1.84%	1.60%
Relative Energy Scale (ND)	-3.59%	-5.53%	-2.99%	-3.81%
PMT Gain (FD)	0.93%	1.07%	-1.55%	0.65%
PMT Gain (ND)	-0.42%	-0.92%	-0.19%	-0.46%
Attenuation	-0.12%	1.04%	-0.11%	0.05%
Strip-to-Strip	0.17%	0.55%	-0.65%	0.12%
Linearity	-0.21%	-0.35%	-0.05%	-0.21%
Crosstalk	-2.58%	3.15%	-5.23%	-2.04%
Total	—	—	—	5.21%

Table 7.1: The systematic shifts in the NC, ν_μ CC, and beam ν_e CC component predictions and their sum, calculated for a single analysis bin ($\text{LEMPID} > 0.7$ and $1 \text{ GeV} < E < 8 \text{ GeV}$). The shifts in the predictions listed are for $+1\sigma$ shifts in the assumption, parameter, etc., underlying the indicated systematic effect. The resulting systematic error is obtained by summing the shifts in the right-hand column in quadrature. Note that the cross section, hadronic model, and INTRANUKE shifts listed here are actually the sums in quadrature of the shifts induced by the uncertainties in each of several parameters governing each model.

Systematic Effect	ν_τ CC	signal ν_e CC
ν_μ SKZP	0.08%	-0.52%
ν_μ FLUGG	-0.01%	-0.01%
ν_μ Target Degradation	-0.30%	0.01%
ν_μ Cross Section	0.78%	0.44%
ν_τ Cross Section	45.2%	—
Hadronic Model	-0.51%	2.94%
INTRANUKE	—	3.03%
Absolute Energy Scale	9.99%	4.41%
EM/Hadronic Energy Scale	-0.82%	-0.48%
Normalization	2.40%	2.40%
PMT Gain (FD)	-1.03%	0.21%
Attenuation	-0.02%	-0.02%
Strip-to-Strip	1.11%	-0.18%
Linearity	-0.01%	0.00%
Crosstalk	0.71%	0.63%
NC in ν_μ CC-like sample	-0.59%	-0.55%
ν_μ CC Energy Scale	-0.92%	3.90%
MRE	—	0.52%
CalDet	—	1.40%
Total	46.6%	7.88%

Table 7.2: The systematic shifts in the ν_τ CC background and signal ν_e CC predictions for a single analysis bin ($\text{LEMPID} > 0.7$ and $1 \text{ GeV} < E < 8 \text{ GeV}$). The shifts in the predictions listed are for $+1\sigma$ shifts in the assumption, parameter, etc., underlying the indicated systematic effect. The resulting systematic errors are obtained by summing the shifts in each column in quadrature. The cross section, hadronic model, and INTRANUKE shifts listed here are actually the sums in quadrature of errors induced by uncertainties in each of several parameters governing each model – except in the case of the ν_τ CC hadronic model shift, which represents the shift due to uncertainty in only one aspect of the model. An error in the computer code prevented calculation of the INTRANUKE error for the ν_τ CC component, but the total error for that component is already dominated by the uncertainty in the ν_τ CC cross section.

ratio is only 0.84%. We also note that the size of the Near Detector systematic errors are reflected in the size of the systematic error band in Figure 4.14.

- The systematic errors associated with gain and other calibration-related errors are dramatically reduced in the current implementation of LEM. In the original implementation, for example, the FD-associated gain and relative energy scale errors were 5.4% and 4.9% [127], respectively; Table 7.1 indicates that these errors are now 0.7% and 1.6%.

Chapter 8

Sensitivity, Sidebands, and Results

As pointed out in Chapter 3, we are now in a position to determine directly what our data imply about the θ_{13} mixing angle. We will do this in several stages, beginning with a calculation of the sensitivity of this analysis to the null hypothesis $\theta_{13} = 0$.

8.1 Sensitivity to θ_{13}

Recall from Chapter 3 our (simplified) definition of sensitivity as the 90% confidence level upper limit we can set on the value of θ_{13} , as a function of δ_{CP} and the neutrino mass hierarchy, should we observe a number of events equal to the expected background. More precisely, the upper limit we present will be on $2 \sin^2 \theta_{23} \sin^2 2\theta_{13}$ – again, as a function of δ_{CP} and the mass hierarchy.¹ Note that this quantity (actually, half of it) is the amplitude of the leading-order term P_{atm} of $P(\nu_\mu \rightarrow \nu_e)$. It explicitly depends on θ_{23} and thus has the desirable feature of incorporating the uncertainty on that mixing angle into the parameter we ultimately limit or measure. In fact, the final result we present at the end of this chapter will also be two continua (one for each mass hierarchy) of confidence intervals for $2 \sin^2 \theta_{23} \sin^2 2\theta_{13}$, each a function of δ_{CP} ; see, for example, Figure 1.15. Note also that in the case of maximal atmospheric mixing ($\theta_{23} = \pi/4$), $2 \sin^2 \theta_{23} \sin^2 2\theta_{13} = \sin^2 2\theta_{13}$.

We begin by dividing the δ_{CP} vs. $2 \sin^2 \theta_{23} \sin^2 2\theta_{13}$ plane into a grid. Taking

¹To clarify further, we do not make any claims regarding δ_{CP} and the mass hierarchy. We simply compute and present the limit on $2 \sin^2 \theta_{23} \sin^2 2\theta_{13}$ for every possible combination of δ_{CP} and mass hierarchy.

$\theta_{23} = \pi/4$, each grid point represents a combination of values for δ_{CP} and θ_{13} . We use those values along with the techniques of Chapter 6 and the best-fit values of the measured oscillation parameters ($\theta_{23}, \theta_{12}, |\Delta m_{32}^2|, \Delta m_{21}^2$) to generate two predictions at that point – one for the normal mass hierarchy ($\Delta m_{32}^2 > 0$) and another for the inverted mass hierarchy ($\Delta m_{32}^2 < 0$). Repeating this for every grid point, we obtain two grids of Far Detector predictions, one for each mass hierarchy.

Next, we use a likelihood ratio L to quantify how well a certain prediction matches a given observation (i.e., observed events classified into analysis bins). In particular, suppose we observe $F_i^{\text{total data}}$, where i indexes the 15 analysis bins as usual. For a prediction $F_i^{\text{total pred}}$ at some grid point and in the case of statistical error only, the likelihood ratio satisfies

$$-2 \ln L = 2 \sum_i \left(F_i^{\text{total pred}} - F_i^{\text{total data}} + F_i^{\text{total data}} \ln \frac{F_i^{\text{total data}}}{F_i^{\text{total pred}}} \right). \quad (8.1)$$

In order to incorporate the systematic error, we return to the two grids of Far Detector predictions and, following the prescription in the introduction to Chapter 7, compute the covariance matrix V on the total prediction at each grid point; thus, wherever V appears, it is understood to be the covariance matrix computed for $F_i^{\text{total pred}}$ assuming some set of oscillation parameters. We then define a set of nuisance parameters f_i – one parameter for each analysis bin – and make the substitution $F_i^{\text{total pred}} \rightarrow F_i^{\prime \text{total pred}} = F_i^{\text{total pred}} + f_i$ in Equation 8.1. Defining \mathbf{f} to be a column vector of the f_i , we augment Equation 8.1:

$$-2 \ln L = 2 \sum_i \left(F_i^{\prime \text{total pred}} - F_i^{\text{total data}} + F_i^{\text{total data}} \ln \frac{F_i^{\text{total data}}}{F_i^{\prime \text{total pred}}} \right) + \mathbf{f}^T V^{-1} \mathbf{f}, \quad (8.2)$$

and, at each grid point, use the MINUIT [128] optimization code to minimize $-2 \ln L$ over the f_i . Finally, because the confidence intervals we present will be functions of δ_{CP} and mass hierarchy, we find the minimum value $-2 \ln L_{\text{min}}$ of $-2 \ln L$ for each value of δ_{CP} on each of the two grids and subtract it from all values of $-2 \ln L$ sharing that value of δ_{CP} on that same grid. The resulting value, which we write as $-2\Delta \ln L$,

is the sole likelihood quantity with which we concern ourselves from this point forward.

We now possess all the tools necessary to compute the desired sensitivity. We set the $F_i^{\text{total data}}$ equal to the $F_i^{\text{total pred}}$ computed for $\theta_{13} = 0$ (δ_{CP} and the mass hierarchy do not affect the prediction in this case) and the best-fit values of θ_{23} , θ_{12} , $|\Delta m_{32}^2|$, and Δm_{21}^2 provided in Section 6.3, and we follow the procedure described above to compute $-2\Delta \ln L$ for each grid point on each grid. The 90% C.L. upper limits on $2\sin^2 \theta_{23} \sin^2 2\theta_{13} = \sin^2 2\theta_{13}$, shown in Figure 8.1, are given by the points on these grids where $-2\Delta \ln L = 2.71$ – the cutoff value on $\Delta\chi^2$ or $-2\Delta \ln L$ that produces a 90% C.L. Gaussian confidence interval.

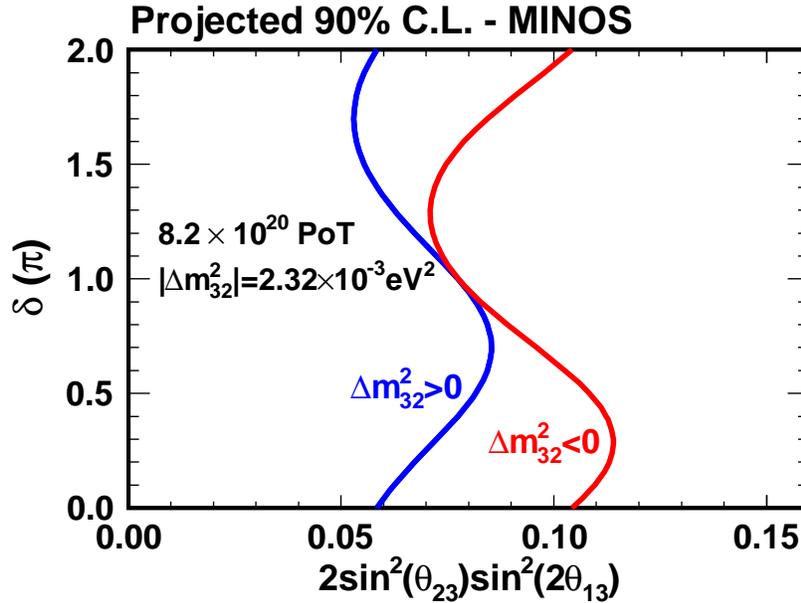


Figure 8.1: Expected 90% C.L. upper limits on $2\sin^2 \theta_{23} \sin^2 2\theta_{13}$ corresponding to a true value of $\theta_{13} = 0$, for each mass hierarchy.

Note that in addition to the improvements in the LEM technique itself, discussed in previous chapters, the analysis presented in this thesis constitutes two key improvements over previous MINOS analyses: (1) use of the LEM PID instead of the ANN PID mentioned in Section 4.3 and (2) fitting to the binned PID distribution instead of performing a counting experiment (i.e., count the number of events above the optimal cut on the PID). We now quantify the impact of these improvements by comparing the sensitivities of three analyses: (1) the ANN PID > 0.7 counting exper-

iment,² (2) the LEM PID > 0.7 counting experiment, and (3) the present analysis. Figure 8.2 shows these sensitivities for the normal mass hierarchy. Using the sensitivity at $\delta_{\text{CP}} = 0$ as a figure of merit, we find that the replacement of ANN with LEM provides a 13% improvement in sensitivity, and fitting to the LEM PID distribution provides an additional 13% improvement. The total improvement in sensitivity due to these analysis changes is approximately 25%.

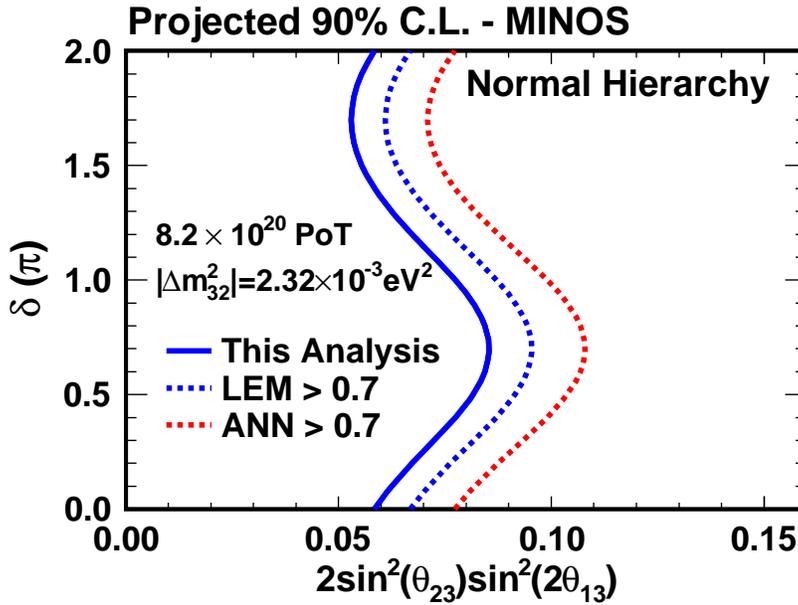


Figure 8.2: 90% C.L. sensitivities for three different selections assuming the normal mass hierarchy, illustrating the improvement in sensitivity achieved by replacing ANN with LEM (dashed red curve to dashed blue curve) and by fitting to the PID distribution instead of performing a counting experiment (dashed blue curve to solid blue curve).

8.2 Validating the Analysis with Sidebands

Prior to looking at the Far Detector data in the analysis region of reconstructed energy and LEM PID, we run the analysis on two auxiliary data samples, or “sidebands,” as noted in Chapter 3.

The first sideband, known as the “anti-LEM” sideband, consists of all events passing the ν_e -like preselection and having LEM PID < 0.5 . Note that this in the

²The optimal cut on the ANN PID is 0.7.

backgroundlike region of the LEM PID. In particular, the cut value of 0.5 is chosen so that a large number of events are selected while the statistical significance of a signal driven by a large value of θ_{13} (e.g., $\sin^2 2\theta_{13} = 0.1$) in this region is less than 1σ , all while maintaining a “buffer” region between the sideband and the analysis region.³ Assuming $\sin^2 2\theta_{13} = 0.1$, the full analysis chain (fiducial and preselection cuts, LEM selection, HOOHE decomposition, extrapolation) yields 369 ± 19 (stat.) predicted background events and 8 predicted signal events; the requirement that the statistical significance of a signal driven by large θ_{13} be less than 1σ is thus satisfied by the LEM PID < 0.5 cut.

When we change the oscillation parameter assumption so that $\sin^2 2\theta_{13} = 0$, the prediction changes slightly, to 370 ± 19 (stat.). For the anti-LEM sideband check itself, we require that the observed number of Far Detector events with LEM PID < 0.5 differ from this prediction by no more than 2σ . We find 377 Far Detector events in the anti-LEM sideband – an excess of 7 events, well within 2σ of the background prediction. This strongly suggests that our full analysis chain functions as intended. Figure 8.3 shows the predicted and observed LEM PID distributions for the anti-LEM sideband.

In Footnote 6 of Section 6.2.1, we described the construction of so-called MRCC (muon-removed charged current) events. This particular set of events, which can be constructed from data and Monte Carlo in both detectors, offers several benefits as a sideband. They are the shower remnants of clearly identified ν_μ CC events, so they constitute a showerlike (specifically, NC-like) sample that is signal-free by construction. Also, this sample is totally separate from the analysis sample, so we are free to investigate *any* range of LEM PID. In particular, the MRCC sideband allows for an independent probe of the data–MC disagreement in the analysis region of LEM PID (see Section 4.3.6), specifically the impact of the disagreement on the analysis as a result of the disagreement being different in the two detectors.

We must point out that the unique nature of the MRCC sideband necessitates

³This is to confound inadvertent “guessing” of the number of events in the analysis region based on the total number of preselected events in previous analyses and the number of events in this sideband.

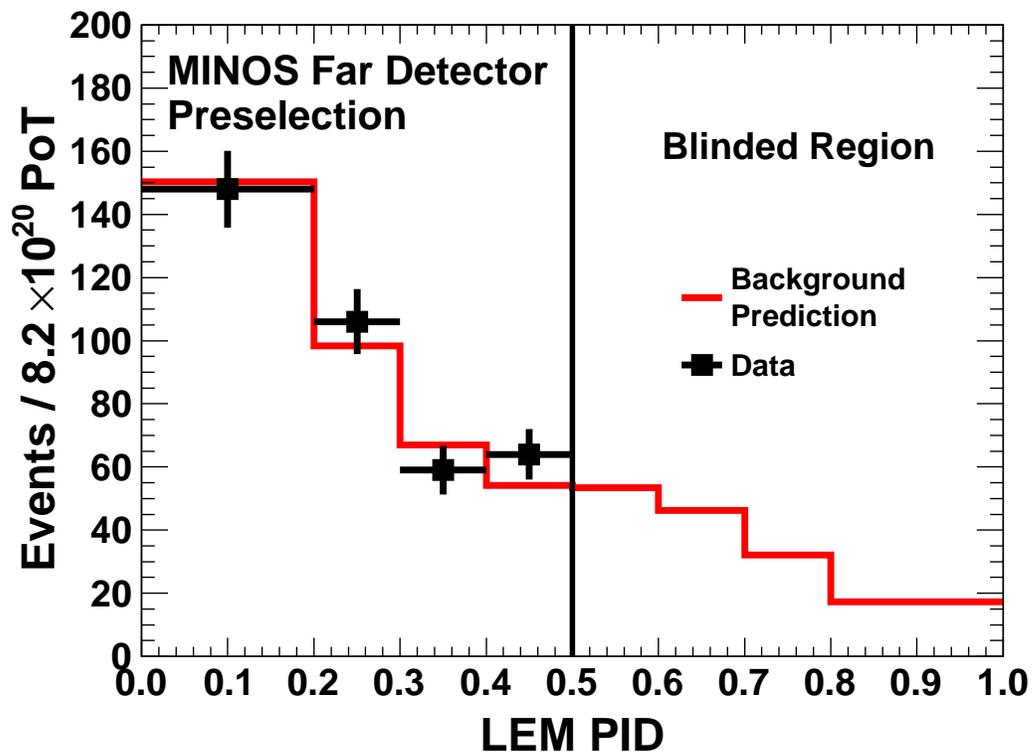


Figure 8.3: Distribution of the LEM PID for preselected events in the Far Detector, with the data distribution for LEMPID > 0.5 blinded. Errors on the data are statistical only. The background prediction is computed assuming $\theta_{13} = 0$. In particular, the *shape* of the background prediction shown is taken directly from the Monte Carlo, but the normalization of the *total* prediction in each extrapolated region (the anti-LEM region, the buffer region $0.5 < \text{LEMPID} < 0.6$, and each bin of the analysis region) is set to match the actual prediction obtained via the full analysis chain.

slight modifications to the details of our analysis technique. For example, the original (pre-muon removal) event is required to pass a separate set of quality and event property cuts, similar to the set of cuts used for the ν_μ CC-like sample that serves as the basis for the ν_τ CC and signal ν_e CC Far Detector predictions. Also, since the MRCC sample consists of only one type of event, no Near Detector decomposition is required, and the extrapolation formula is modified slightly. Details of these differences, along with a documentation of the MRCC analysis itself, are provided in [129]; the key is that the principle of using a simulated far-to-near ratio to extrapolate from the Near Detector data to a Far Detector prediction remains intact.

Our requirement for the MRCC sideband analysis is that the data and prediction agree within 2σ ($p > 0.0455$). The full LEM distributions of the Far Detector MRCC prediction and data are shown in Figure 8.4; they agree, with $\chi^2/N_{\text{d.o.f.}} = 9.7/8$ ($p = 0.287$) using statistical errors only. Performing the prediction in 40 bins – the 8 LEM PID bins shown in Figure 8.4 by the 5 reconstructed energy bins of the official analysis binning – the data and prediction agree, with $\chi^2/N_{\text{d.o.f.}} = 47.5/40$ ($p = 0.194$). In the analysis region only, the data and prediction still agree but with a slightly larger $\chi^2/N_{\text{d.o.f.}}$ of 20.4/15 ($p = 0.157$). Our analysis technique is thus validated despite the data–MC disagreement, both in the analysis region and over the full range of LEM PID.

8.3 The Far Detector Data

We now proceed to look at the full Far Detector data set, starting with some basic distributions for the preselected sample. Figure 8.5 shows the ratio of the PoT-normalized preselected event rate in the Far Detector data to the PoT-normalized preselected event rate in the Near Detector data, as a function of time. We fit a constant (i.e., a line of zero slope) to the points in order to assess whether the observed variation in the Far-to-Near ratio is consistent with statistical fluctuation within 2σ , and this appears to be the case, with $\chi^2/N_{\text{d.o.f.}} = 34.4/24$ ($p = 0.077$) for the fit. We also check the timing distribution of preselected Far Detector events –

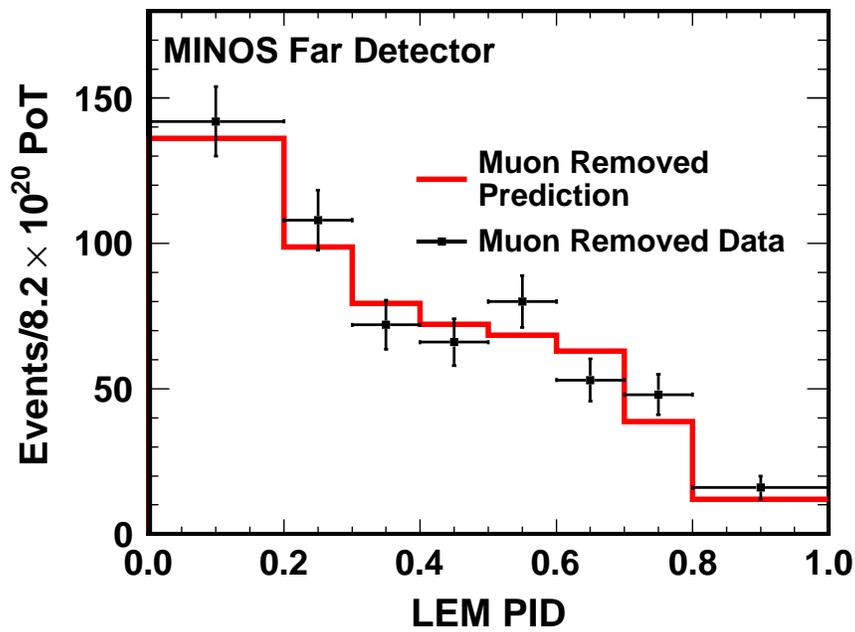


Figure 8.4: Distribution of LEM PID for MRCC events in the Far Detector. Errors on the data are statistical only. The observed and predicted distributions agree, with $\chi^2/N_{\text{d.o.f.}} = 9.7/8$.

specifically, the time from the event to the nearest NuMI beam spill – in Figure 8.6 and find it to be free of any obvious anomalies (e.g., no pileup of events at either end of the time window that might suggest the presence of non-beam-induced events).

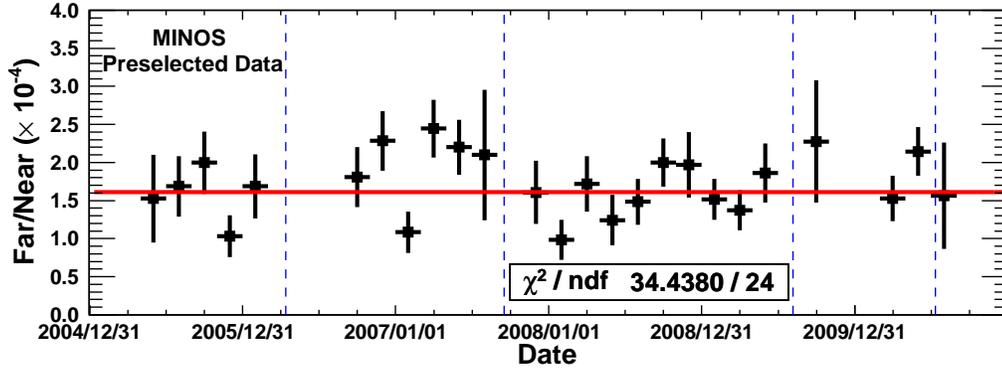


Figure 8.5: Ratio of PoT-normalized preselected data event rate in the Far Detector to PoT-normalized preselected data event rate in the Near Detector, as a function of time. Errors are statistical only. The red line is a constant fit to the data. The dashed vertical blue lines separate (from left to right) Runs I, II, III, IV+V, and VI.

Figure 8.7 shows distributions of event vertex coordinates for Far Detector events passing the preselection. The background prediction included with each distribution is the predicted distribution for $\theta_{13} = 0$. Our goal in making these distributions is qualitative, checking for any pathologies in the distribution of events in the Far Detector. Any possible signal would, of course, appear as an excess above this background; the purpose of the background distributions, however, is not to assist in locating event excesses but simply to serve as a visual guide to the data distributions, approximating their expected shapes. A qualitative inspection of the distributions reveals no obvious pathologies.

Satisfied with the distributions of these low-level quantities, we proceed to look at distributions of higher-level variables. Figure 8.8 shows the full LEM PID distribution, with the buffer and analysis regions unblinded. Again, the background prediction shown is computed assuming $\theta_{13} = 0$, but keeping Figure 6.2 in mind, we expect any possible signal to be a modest contribution to the total distribution, even for larger values of LEM PID. Thus, we may reasonably expect the observed distribution to follow the background distribution somewhat closely, and we see that it does. To

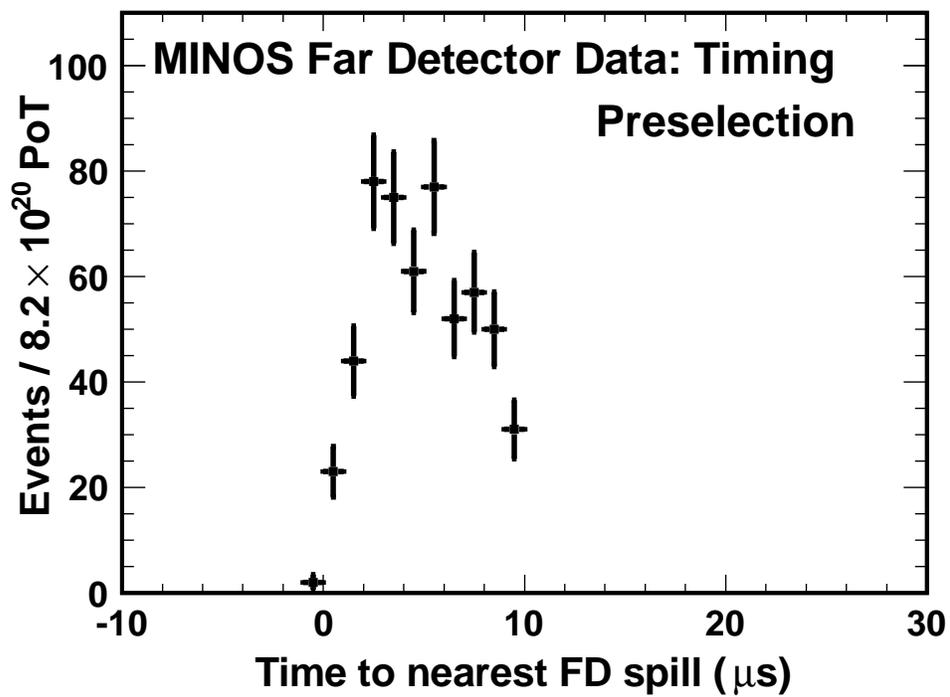


Figure 8.6: Distribution of time to nearest NuMI beam spill for preselected Far Detector events. Errors are statistical only. The data are consistent with the 10- μ s spill length.

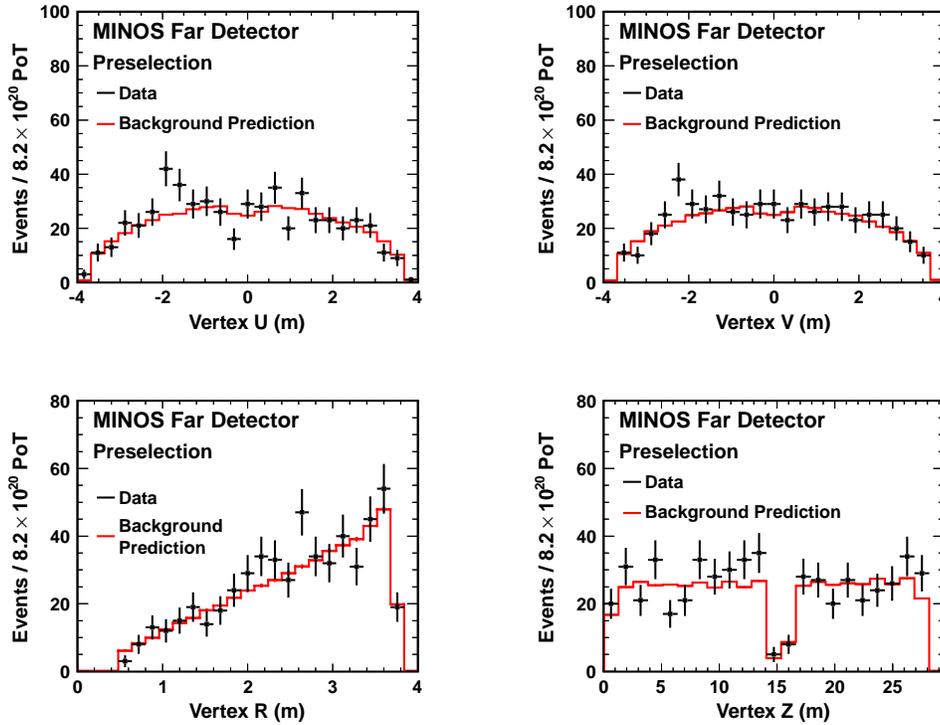


Figure 8.7: Distributions of vertex coordinates for preselected Far Detector events. The background predictions shown are computed assuming $\theta_{13} = 0$. Errors are statistical only. U (top left) and V (top right) are orthogonal coordinates in each detector plane, so named because the axes along which they are measured coincide with the orientations of scintillator strips in U-view and V-view planes. The origin for each axis is the coil hole. R (bottom left) is the distance from the coil hole to the event vertex (i.e., radius), and Z (bottom right) is the longitudinal coordinate, in the beam direction, measured downstream from the first detector plane. The dip in the middle of the Z distribution coincides with the gap between supermodules 1 and 2.

check for any pathologies associated with event location (e.g., unusual clustering of events, excessive number of events near the detector edge), we plot the vertex of each event in the x - y plane – specifically, one plot for all preselected events and a second only for those events in the analysis region $\text{LEM PID} > 0.6$. These plots, shown in Figure 8.9, do not suggest any such problems.

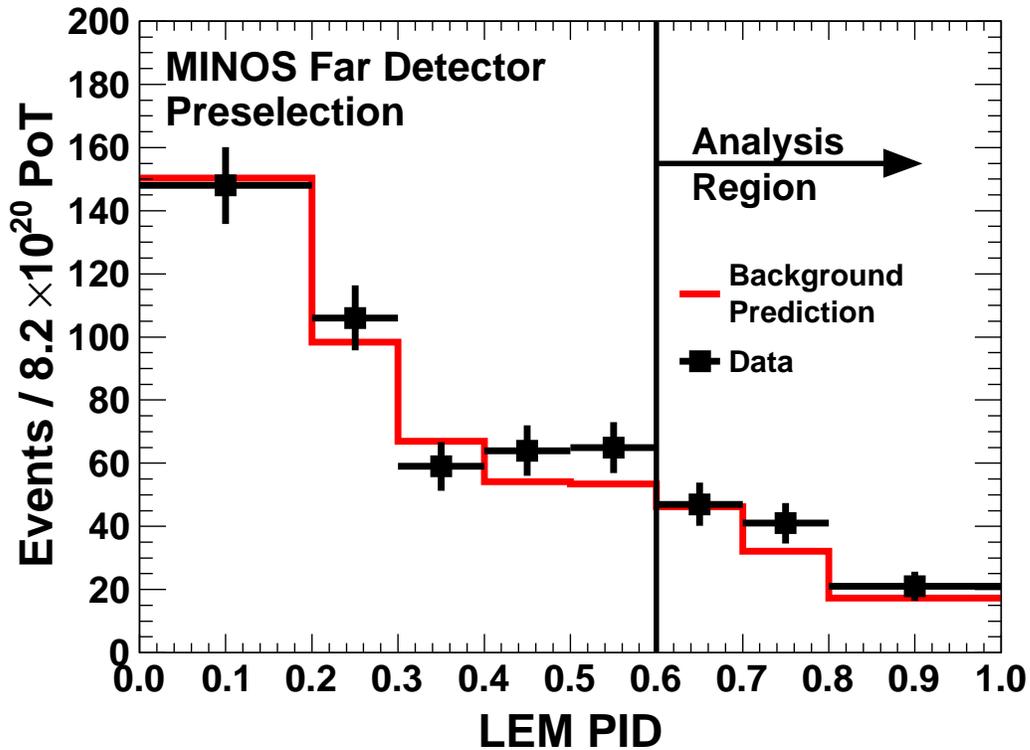


Figure 8.8: LEM PID distribution of all Far Detector events passing the preselection. Errors on the data are statistical only. The background prediction is computed assuming $\theta_{13} = 0$.

Next, we look at the reconstructed energy distribution for Far Detector events passing the preselection, shown in Figure 8.10. Note that the 5–8 GeV bin has been divided into three 1-GeV bins, solely for the purpose of making this plot; the shape of the background prediction in that energy range is taken directly from the Monte Carlo, but each 1-GeV bin’s content is scaled so that the sum of their contents equals the total prediction in the 5–8 GeV range. We immediately notice the excess (found to be at the level of 2.6σ) in the 5–6 GeV bin, especially as it contrasts with the relatively good agreement between the data and the prediction in the other bins. Figures 6.3 and 6.4 indicate that this excess is not in the right energy range to be

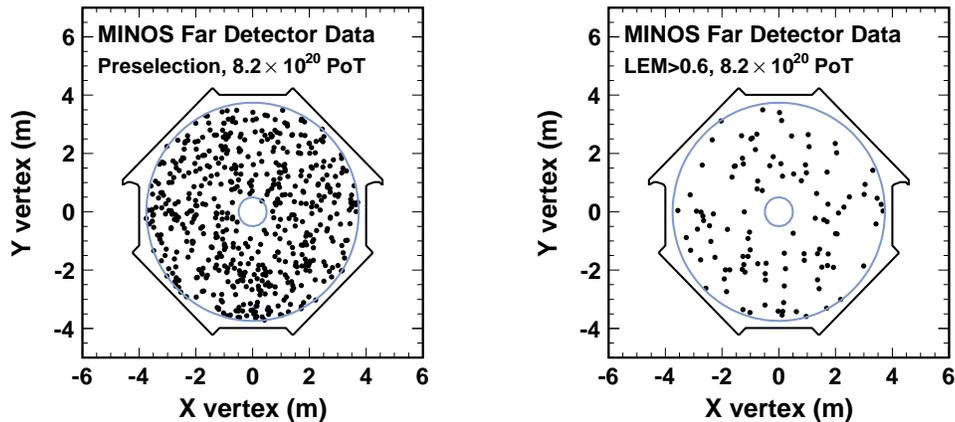


Figure 8.9: Vertex distributions in the x - y plane for Far Detector events passing the preselection (left) and for Far Detector events with LEMPID > 0.6 (right). The black lines indicate the physical extent of the Far Detector; the blue circles delimit the fiducial volume.

associated with a signal.

A variety of checks were performed to determine whether the excess was pathological or simply a statistical fluctuation. Using statistical errors only, we find that $\chi^2/N_{\text{d.o.f.}} = 12.0/7$ ($p = 0.10$) for the full reconstructed energy distribution, so the overall variation from the background prediction is consistent with a statistical fluctuation. One possibility we considered was that our background predictions at higher energies were somehow faulty. This was ruled out by extending our prediction beyond the 8-GeV preselection cut – to 12 GeV – as shown in Figure 8.11; any potential signal is negligible in the 8–12 GeV range, and we see that the data agree reasonably well with the background prediction. The vertex distributions in Figure 8.9 were remade for events with reconstructed energy between 5 and 6 GeV and are shown in Figure 8.12; these reveal no pathologies associated with event location (e.g., clustering of events in some region of the detector, at the edge or elsewhere). Further checks revealed no pathologies, and we concluded that the excess was a statistical fluctuation.⁴

⁴Originally, the analysis was to be performed in one bin of reconstructed energy. Binning optimization studies showed that only binning in LEM PID provided appreciable improvement in sensitivity; no such improvement was achieved by binning in reconstructed energy. Upon discovery of the 5–6 GeV excess, however, we opted for the current binning. While such a change is not wholly consistent with a blind analysis, it happens to be sensitivity-neutral, and we would have been

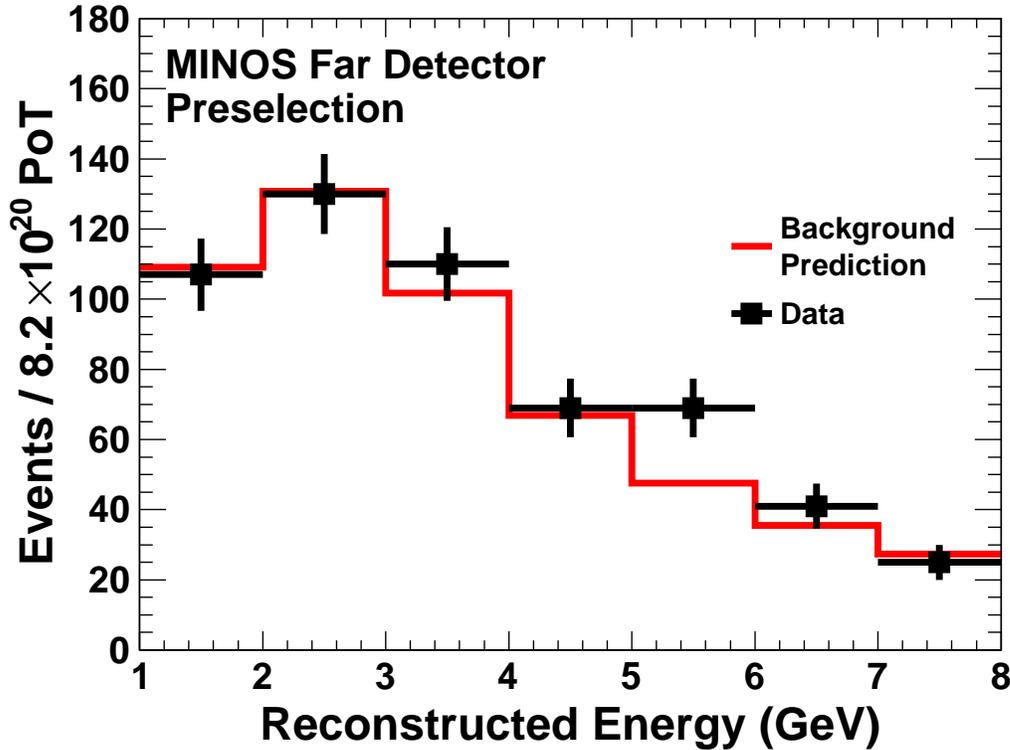


Figure 8.10: Reconstructed energy distribution for Far Detector data. Errors on the data are statistical only. The background prediction is computed assuming $\theta_{13} = 0$.

8.4 Fitting the Data

With the data checked for any obvious problems, we proceed to the final stage of the analysis: determining confidence intervals on $2 \sin^2 \theta_{23} \sin^2 2\theta_{13}$. In Section 8.1, we set the $F_i^{\text{total data}}$ equal to the background prediction, computed $-2\Delta \ln L$ over the two grids of predictions, and took the points where $-2\Delta \ln L = 2.71$ as our 90% C.L. upper limits. One might think that the same approach can be taken in fitting the actual data, but there is a subtlety here. In computing the sensitivity, we knew that the best-fit value of $2 \sin^2 \theta_{23} \sin^2 2\theta_{13}$ was necessarily zero; thus, the confidence interval we obtained would be one sided – an upper limit – and that upper limit was found to be some distance away from $2 \sin^2 \theta_{23} \sin^2 2\theta_{13} = 0$. For the data, however, we must be prepared for any possibility, including an excess of events that lifts the best-fit value away from zero, which in turn gives rise to the possibility of obtaining

wise to choose this binning in the first place in order to accommodate possible energy-dependent fluctuations like the one seen.

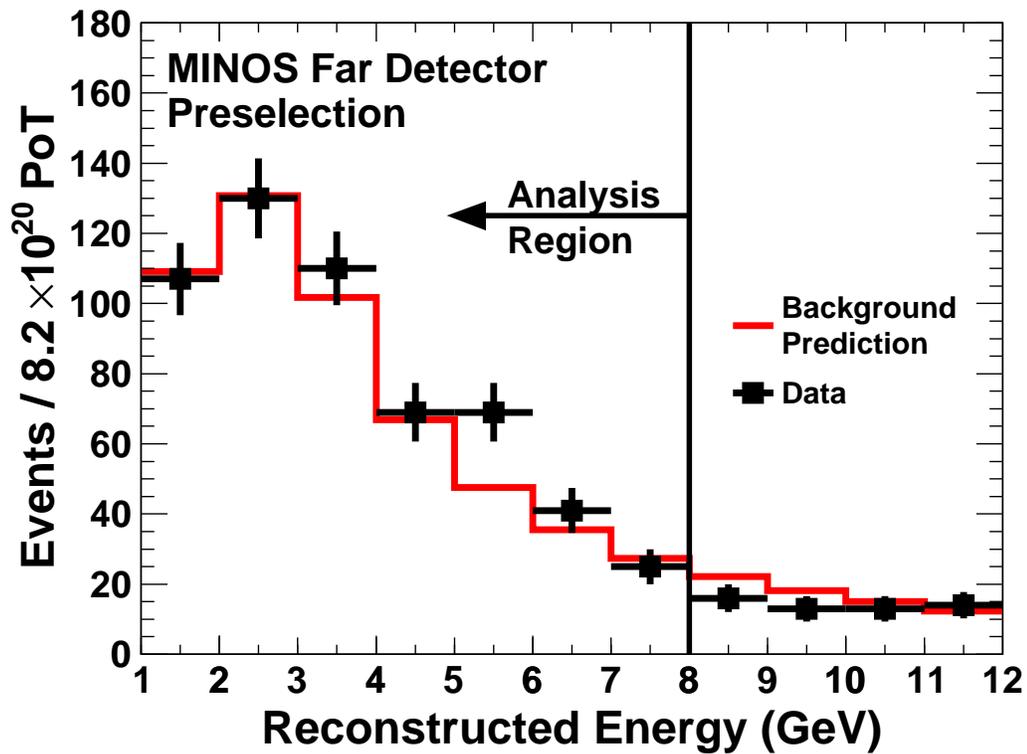


Figure 8.11: Reconstructed energy distribution for Far Detector data, extended beyond the 8-GeV preselection cut – to 12 GeV. Errors on the data are statistical only. The background prediction is computed assuming $\theta_{13} = 0$.

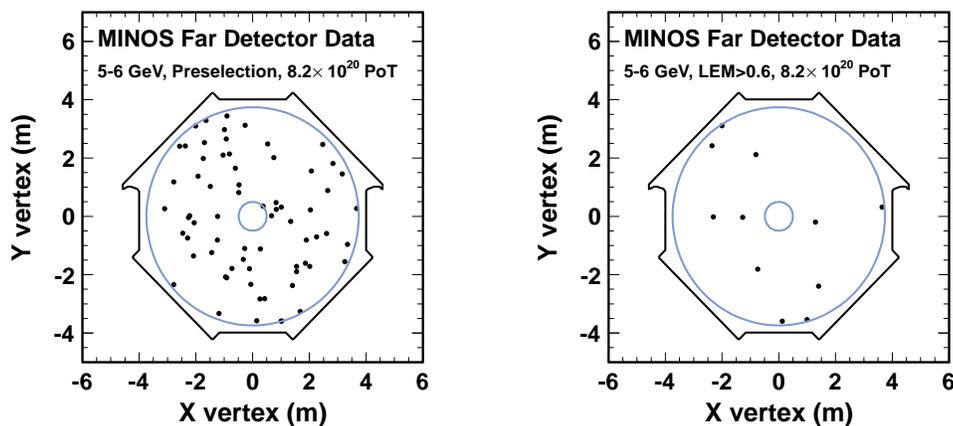


Figure 8.12: Vertex distributions in the x - y plane for Far Detector events passing the preselection (left) and for Far Detector events with $LEM > 0.6$ (right); in both distributions, only events with reconstructed energy between 5 and 6 GeV are shown. The black lines indicate the physical extent of the Far Detector; the blue circles delimit the fiducial volume.

a two-sided confidence interval with a lower limit close to $2 \sin^2 \theta_{23} \sin^2 2\theta_{13} = 0$. Unfortunately, Gaussian confidence intervals near physical boundaries, such as the boundary at $2 \sin^2 \theta_{23} \sin^2 2\theta_{13} = 0$, are known to provide the wrong coverage of the parameter space. In other words, a 90% C.L. Gaussian confidence interval on $2 \sin^2 \theta_{23} \sin^2 2\theta_{13}$, obtained by a cutoff value of 2.71 on $-2\Delta \ln L$, does *not* have a 90% chance of containing the true value of $2 \sin^2 \theta_{23} \sin^2 2\theta_{13}$ if the confidence interval is close to the physical boundary.

Our solution is to follow the prescription of Feldman and Cousins [130], which provides the means to compute confidence intervals that give the correct coverage of the parameter space. In this approach, we require an ordering principle for values of $2 \sin^2 \theta_{23} \sin^2 2\theta_{13}$ – i.e., some way of ranking values of $2 \sin^2 \theta_{23} \sin^2 2\theta_{13}$ according to how well their resulting predictions match a given observation. We already have such a metric: $-2\Delta \ln L$. We may thus proceed to set the $F_i^{\text{total data}}$ equal to our observations and compute $-2\Delta \ln L$ over the two grids of predictions, in precisely the same fashion as in Section 8.1. The difference in the present approach arises in the choice of cutoff value on $-2\Delta \ln L$. In particular, we compute the cutoff value for each combination of $2 \sin^2 \theta_{23} \sin^2 2\theta_{13}$, δ_{CP} , and the mass hierarchy – i.e., for each point in the two parameter grids – by means of conducting “pseudoexperiments,” as follows:

1. Generate 10,000 15-tuples of correlated Gaussian random variables with mean equal to the 15-bin prediction at this grid point and covariance given by the systematic error covariance matrix at this grid point.⁵
2. Generate 10,000 new 15-tuples of Poisson-distributed random variables, taking the i th element of the j th 15-tuple from the previous step as the mean for the i th element of the j th new 15-tuple. These new 15-tuples constitute 10,000 pseudoexperiments – 10,000 possible outcomes, taking systematic (previous step) and statistical (this step) errors into account, of running our experiment in a

⁵Systematic errors due to uncertainties in the oscillation parameters have been added to this covariance matrix via methods described in Appendix D.

universe where the true values of $2 \sin^2 \theta_{23} \sin^2 2\theta_{13}$ and δ_{CP} and the true mass hierarchy are those defined by the current grid point.

3. Compute $-2\Delta \ln L$ at the current grid point for each pseudoexperiment,⁶ and construct a histogram of these values. The value of $-2\Delta \ln L$ below which $x\%$ of the pseudoexperiments lie is the Feldman–Cousins $x\%$ C.L. cutoff value.

Intuitively, the 10,000 pseudoexperiments at a given grid point constitute a reasonably large sample of the possible outcomes of running our experiment in a universe where $2 \sin^2 \theta_{23} \sin^2 2\theta_{13}$, δ_{CP} , and the mass hierarchy are given by that grid point. An experiment run in such a universe has a $x\%$ probability of having a value of $-2\Delta \ln L$ less than or equal to the Feldman–Cousins $x\%$ C.L. cutoff value determined via the procedure outlined above. From a frequentist point of view, this cutoff value thus manifestly provides the desired coverage probability. We illustrate the variation of the 90% C.L. cutoff value with $2 \sin^2 \theta_{23} \sin^2 2\theta_{13}$ for the normal mass hierarchy and $\delta_{\text{CP}} = 0$ in Figure 8.13, with the Gaussian cutoff value of 2.71 superimposed for comparison. This figure also shows $-2\Delta \ln L$ for the Far Detector data, assuming normal mass hierarchy and $\delta_{\text{CP}} = 0$. In particular, note that $-2\Delta \ln L \approx 1.3$ at $2 \sin^2 \theta_{23} \sin^2 2\theta_{13} = 0$ – just below the Feldman–Cousins 90% cutoff but (as one may compute) near the Gaussian 75% C.L. cutoff. This explicitly illustrates the fact that Gaussian confidence intervals provide the wrong coverage probability near physical boundaries.

We proceed to compute the Feldman–Cousins 68% C.L. and 90% C.L. cutoff curves for all values of δ_{CP} and both hierarchies. Imposing these cutoffs on the grids of $-2\Delta \ln L$ for the Far Detector data, we obtain the confidence intervals illustrated in Figure 8.14. We find that $2 \sin^2 \theta_{23} \sin^2 2\theta_{13} < 0.12(0.20)$ at 90% C.L. for $\delta_{\text{CP}} = 0$ and the normal (inverted) mass hierarchy. The $\theta_{13} = 0$ hypothesis is excluded at 89% C.L. The best-fit values are $2 \sin^2 \theta_{23} \sin^2 2\theta_{13} = 0.041_{-0.031}^{+0.047}(0.079_{-0.053}^{+0.071})$ for $\delta_{\text{CP}} = 0$ and

⁶More explicitly, set the $F_i^{\text{total data}}$ equal to the pseudoexperiment “observation” and, keeping δ_{CP} and the mass hierarchy the same, compute $-2 \ln L$ for the full interval of $2 \sin^2 \theta_{23} \sin^2 2\theta_{13}$ being considered. Subtract the resulting minimum value of $-2 \ln L$ from the value of $-2 \ln L$ at the current grid point to obtain $-2\Delta \ln L$.

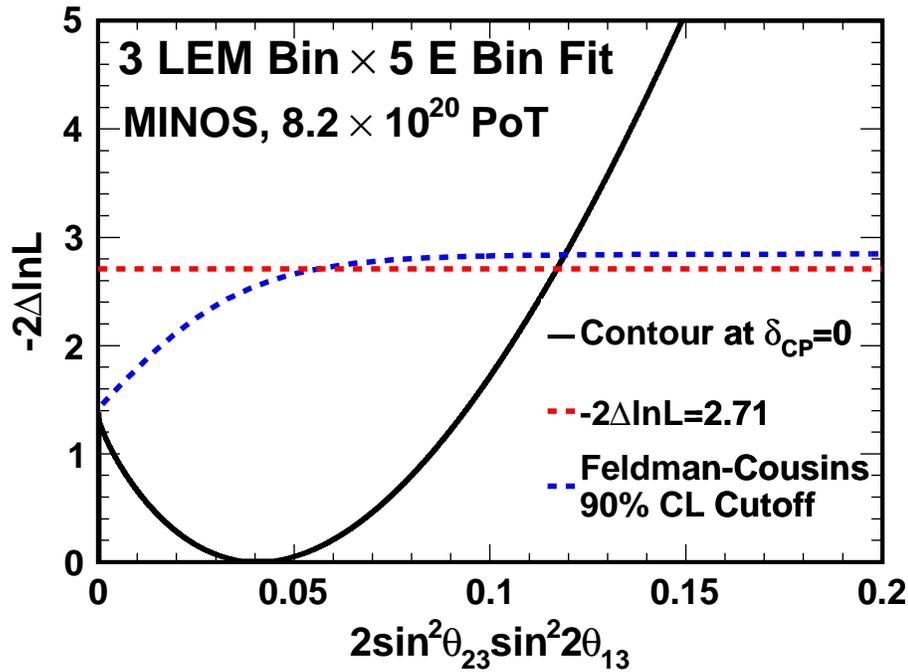


Figure 8.13: Values of $-2\Delta\ln L$ for the Far Detector data, assuming $\delta_{CP} = 0$ and normal mass hierarchy, as a function of $2\sin^2\theta_{23}\sin^22\theta_{13}$ (solid black curve). The dashed red line shows the Gaussian 90% C.L. cutoff value (2.71) on $-2\Delta\ln L$, and the dashed blue curve shows the Feldman–Cousins 90% C.L. cutoff value, which varies with $2\sin^2\theta_{23}\sin^22\theta_{13}$. Note that far from $2\sin^2\theta_{23}\sin^22\theta_{13} = 0$, the Feldman–Cousins cutoff is roughly constant and close to the Gaussian cutoff. As we approach $2\sin^2\theta_{23}\sin^22\theta_{13} = 0$, however, the Feldman–Cousins cutoff drops below the Gaussian cutoff.

the normal (inverted) mass hierarchy; the best-fit signal for the normal hierarchy is shown in Figure 8.15 stacked atop the background prediction as a function of the reconstructed energy in each bin of LEM PID.

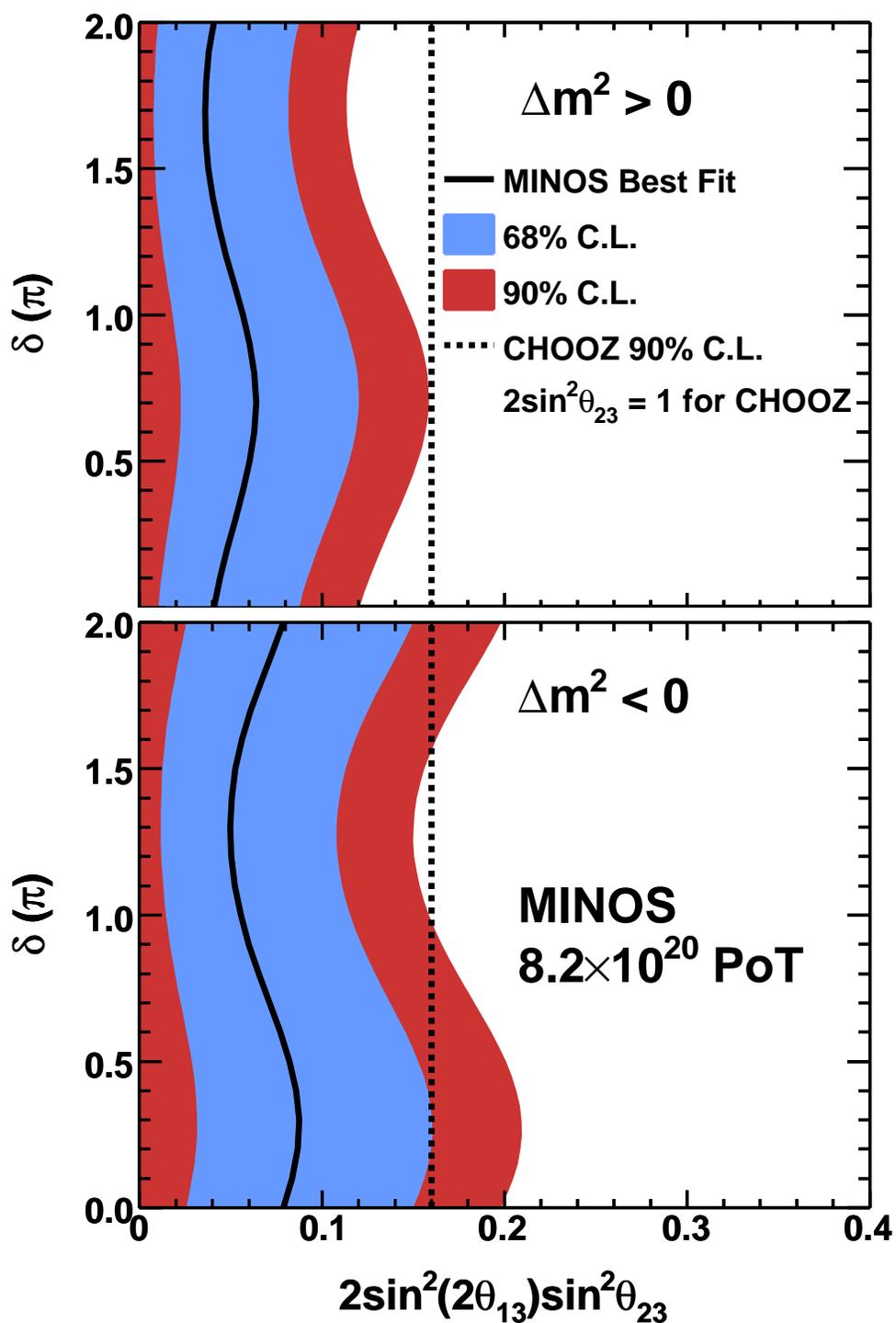


Figure 8.14: Best-fit values and 68% and 90% Feldman–Cousins confidence intervals for $2\sin^2\theta_{23}\sin^2 2\theta_{13}$, as a function of δ_{CP} for the normal (top) and inverted (bottom) mass hierarchies. The 90% C.L. CHOOZ limit, assuming maximal atmospheric mixing and the best-fit value for Δm_{32}^2 , is shown for comparison.

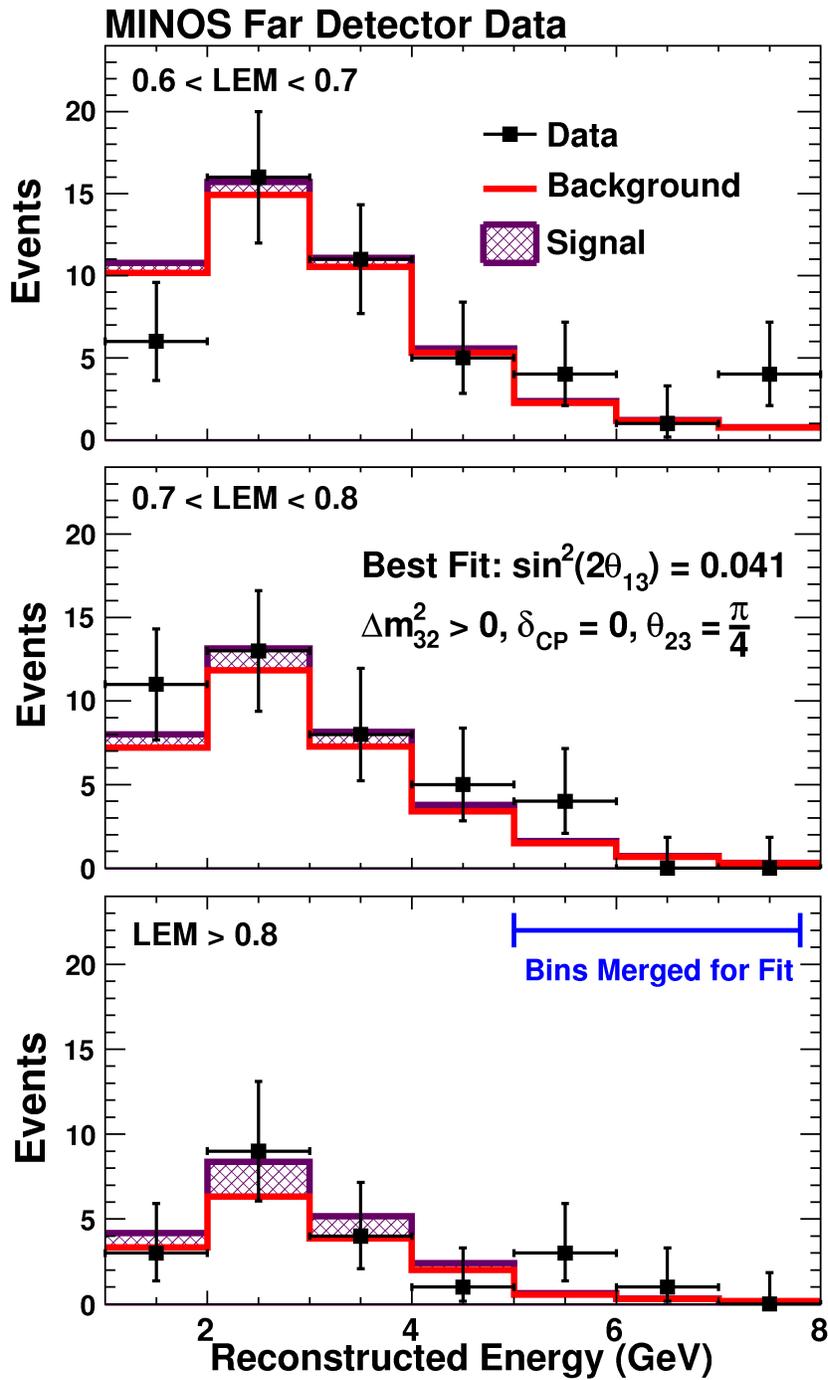


Figure 8.15: Best-fit signal for the normal hierarchy and $\delta_{CP} = 0$ stacked atop the background prediction, as a function of reconstructed energy, for $0.6 < \text{LEMPID} < 0.7$ (top), $0.7 < \text{LEMPID} < 0.8$ (middle), and $\text{LEMPID} > 0.8$ (bottom). The 5–8 GeV bin is divided into three 1-GeV bins as described earlier for display purposes only. The data are shown with statistical errors only.

Chapter 9

Conclusions and a Look at the Future of Neutrino Physics

Nearly a century after Chadwick's observation of the continuous beta spectrum [2], we are at a turning point in the story of the neutrino. Its existence has been hypothesized and confirmed. The solar and atmospheric neutrino problems have been discovered and solved, with oscillations of massive neutrinos as the only experimentally supported explanation for the observed behavior. Several of the parameters governing neutrino oscillations have been measured with great precision. Yet, an equally impressive list of questions remains, including

- What is the value of the third mixing angle θ_{13} ?
- What is the neutrino mass hierarchy?
- To what extent, if any, does CP violation occur in the leptonic sector? That is, what is δ_{CP} ?
- Is atmospheric mixing truly maximal ($\theta_{23} = \pi/4$), or only nearly so?
- What are the absolute values of the neutrino masses?
- Are the neutrinos of definite mass Dirac or Majorana particles? In other words, are antineutrinos distinct from or identical to neutrinos?

The last two questions are fundamentally beyond the abilities of neutrino oscillation experiments. The KATRIN [131] experiment is designed to measure the electron

neutrino mass with subelectronvolt sensitivity via ultrahigh precision measurements of the kinematics of tritium beta decay. As for the question of the Dirac or Majorana nature of neutrinos, a variety of experiments, either under construction or already taking data, are observing samples of certain radioactive isotopes that are capable of undergoing double beta decay, whereby an isotope of atomic number Z decays to an isotope of atomic number $Z + 2$ ($Z - 2$) with the release of two electrons (positrons) and two electron antineutrinos (neutrinos).¹ Such current or upcoming experiments include COBRA [132], observing cadmium and tellurium; EXO [133], observing ^{136}Xe ; and Majorana [134] and GERDA [135], observing ^{76}Ge . If neutrinos happen to be Majorana particles, the antineutrino is identical to the neutrino, and it is possible for the two electron antineutrinos (neutrinos) to annihilate. Such “neutrinoless double beta decay,” in which the isotope in question decays by the emission of two electrons (positrons) and no neutrinos, would be a clear sign that neutrinos are Majorana particles.

The question of the value of θ_{13} has recently become the topic of great interest. In 2010, a global analysis [136] of all available data (solar, atmospheric, reactor, and accelerator experiments), up to and including the second MINOS ν_e appearance result [57], found a best-fit value of $\sin^2 2\theta_{13} = 0.038_{-0.028}^{+0.050}$. The analysis described in this thesis found that $2 \sin^2 \theta_{23} \sin^2 2\theta_{13} = 0.041_{-0.031}^{+0.047} (0.079_{-0.053}^{+0.071})$ for $\delta_{\text{CP}} = 0$ and the normal (inverted) mass hierarchy – wholly consistent with the global best-fit value assuming maximal atmospheric mixing. In fact, our result is in excellent agreement with the global best-fit value for both mass hierarchies and virtually all values of δ_{CP} , assuming maximal atmospheric mixing.

The T2K experiment, located in Japan, uses an approach similar to MINOS, using a two-detector setup to look for ν_e appearance in an accelerator-produced ν_μ beam. A key enhancement is a so-called off-axis arrangement of the detectors, in which the detectors are placed slightly off the neutrino beam axis (2.5° in the case of T2K) in order to take advantage of a relativistic narrowing of the energy spectrum; in partic-

¹Other decay modes include double electron capture and single electron capture followed by positron emission. Two electron neutrinos are released in both cases.

ular, the design parameters are chosen so as to produce a spectrum peak coinciding with the first oscillation maximum, enhancing the flux of neutrinos with maximum oscillation probability while reducing backgrounds resulting from interactions involving higher-energy neutrinos. Almost simultaneously with the announcement of the results described in this thesis, the T2K experiment released its own findings regarding θ_{13} [137] based on an exposure of 1.43×10^{20} PoT, shown overlaid with our results in Figure 9.1. Note that our results significantly constrain the range of $2 \sin^2 \theta_{23} \sin^2 2\theta_{13}$ allowed by T2K.

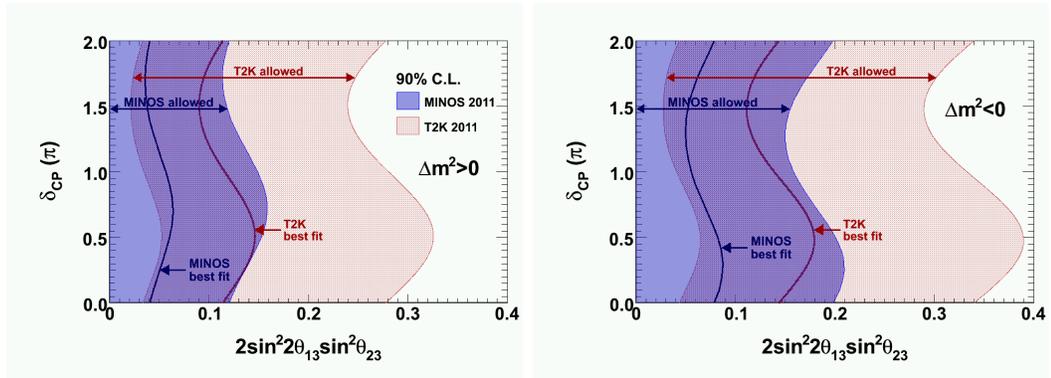


Figure 9.1: Best-fit values and 90% C.L. confidence intervals for $2 \sin^2 \theta_{23} \sin^2 2\theta_{13}$ as determined by T2K [137], overlaid with the results reported in this thesis, as a function of δ_{CP} for normal (left) and inverted (right) hierarchy. The T2K allowed regions are significantly constrained by our results.

From late December 2011 to early April 2012, three next-generation reactor experiments looking for Δm_{32}^2 -driven electron antineutrino disappearance released measurements of θ_{13} . While similar in the oscillation mode being investigated, these experiments represent significant improvements over the CHOOZ experiment, most notably their use of near and far detectors. The Double Chooz experiment released the first of the three results, based on 101 days of data collected with only their far detector, which is located approximately 1050 m from the two 4.25-GW_{th} (gigawatts of thermal power) reactors of the Chooz Nuclear Power Plant in France. Using both rate and shape information from the prompt positron spectrum, they found $\sin^2 2\theta_{13} = 0.086 \pm 0.041$ (stat.) ± 0.030 (syst.); at 90% C.L., $0.017 < \sin^2 2\theta_{13} < 0.16$ [138], and $\theta_{13} = 0$ is ruled out at 94.6% C.L. The addition of their near detector

will lead to reduced systematic uncertainties and an expected precision on $\sin^2 2\theta_{13}$ of approximately 0.02.

The next result came from the Daya Bay experiment, in which antineutrinos from six 2.9-GW_{th} reactors located along the southern coast of China are detected by a total of six detectors – two near detectors at a flux-weighted baseline of 470 m, another near detector at a flux-weighted baseline of 576 m, and three far detectors at a flux-weighted baseline of 1648 m. Based on a rate-only analysis of a 43 kton-GW_{th}-day livetime exposure (collected over 55 days), they found $\sin^2 2\theta_{13} = 0.092 \pm 0.016$ (stat.) ± 0.005 (syst.), ruling out $\theta_{13} = 0$ at 5.2σ [139]. This was shortly followed by a measurement from the RENO experiment, in which antineutrinos from six 2.8-GW_{th} reactors in Korea are detected by two identical detectors located at flux-weighted baselines of 409 and 1444 m. A rate-only analysis of data collected over 229 days found $\sin^2 2\theta_{13} = 0.113 \pm 0.013$ (stat.) ± 0.019 (syst.), ruling out $\theta_{13} = 0$ at 4.9σ [140].

In light of these recent results, particularly those of Daya Bay and RENO, we can safely conclude that θ_{13} is nonzero. While a proper global fit of all available accelerator and reactor analyses is made nontrivial by the involvement of θ_{23} in the accelerator experiment results, one can obtain a relatively good approximation by assuming $\theta_{23} = \pi/4$. Such a combined result, performed following the prescriptions of [10], is shown in Figure 9.2 for each hierarchy assuming $\delta_{\text{CP}} = 0$.

Under the stated assumptions, the combined result is $\sin^2 2\theta_{13} = 0.091 \pm 0.012$ ($\sin^2 2\theta_{13} = 0.095 \pm 0.012$) for the normal (inverted) hierarchy – consistent with our result, given in Section 8.4, under the same assumptions. Figure 9.3 shows the signal implied by $\sin^2 2\theta_{13} = 0.091$ (the best-fit combined value for the normal hierarchy) in MINOS, stacked atop the background prediction as a function of reconstructed energy in each bin of LEM PID; compare Figure 8.15.

Further data taking by the experiments discussed will serve to improve the precision of our measurement of the value of θ_{13} . What is more, in light of the discussion in Section 1.3.3.2 regarding J_{CP} along with the fact that θ_{13} is decidedly nonzero, there is a great opportunity to address the question of leptonic CP violation. CP violation

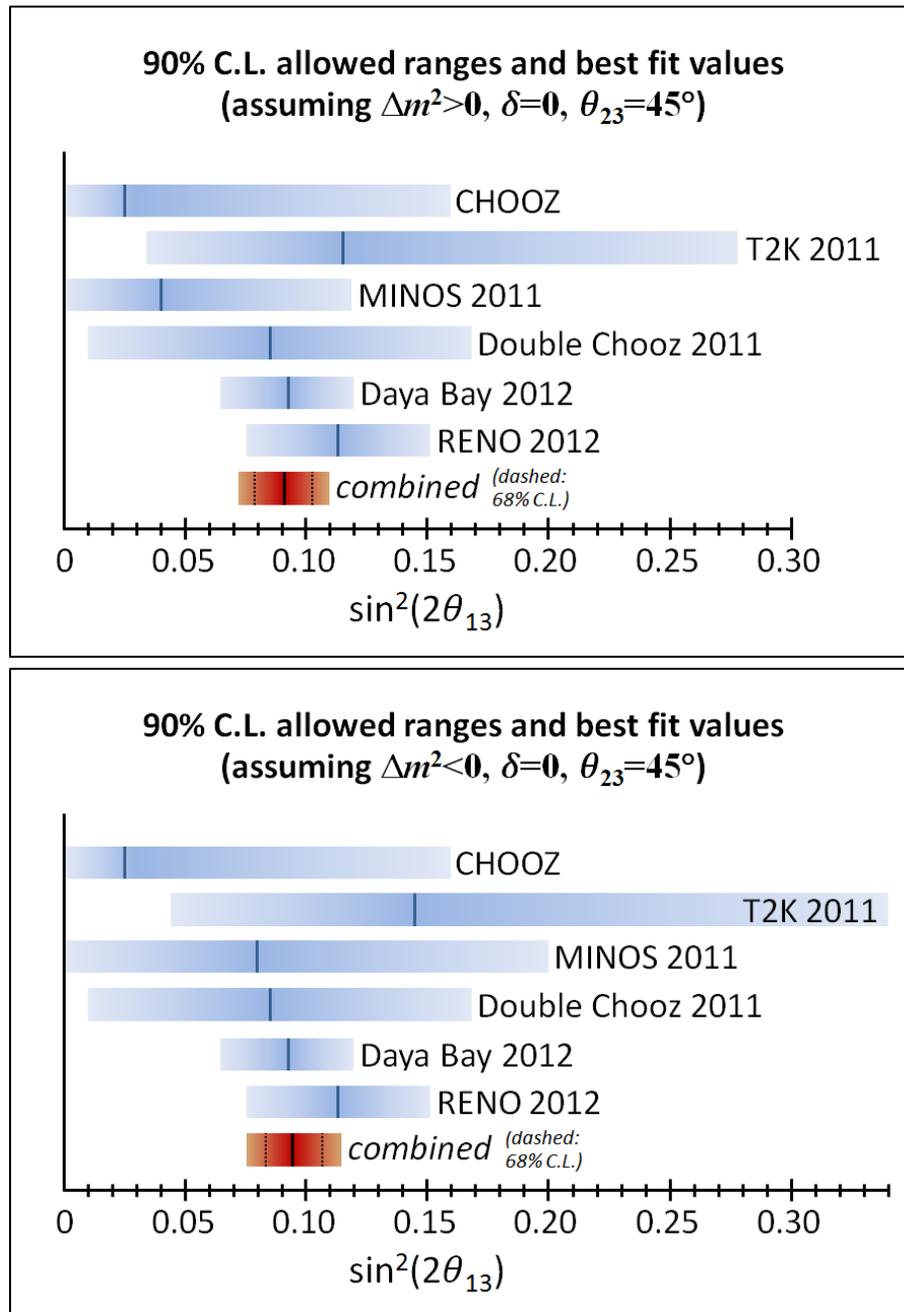


Figure 9.2: Summary of $\sin^2 2\theta_{13}$ measurements and limits from various experiments. The blue bands indicate the 90% C.L. confidence intervals on $\sin^2 2\theta_{13}$, with $\theta_{23} = \pi/4$, $\delta_{\text{CP}} = 0$, and normal (inverted) hierarchy assumed for the accelerator experiments (MINOS and T2K) in the left (right) plot. The red band indicates the 90% C.L. confidence interval on $\sin^2 2\theta_{13}$ obtained by combining the given intervals from the individual experiments, with the dashed lines indicating the 68% C.L. (1σ) confidence interval. On each band, the heavy vertical line indicates the best-fit value. Courtesy of Ryan Patterson.

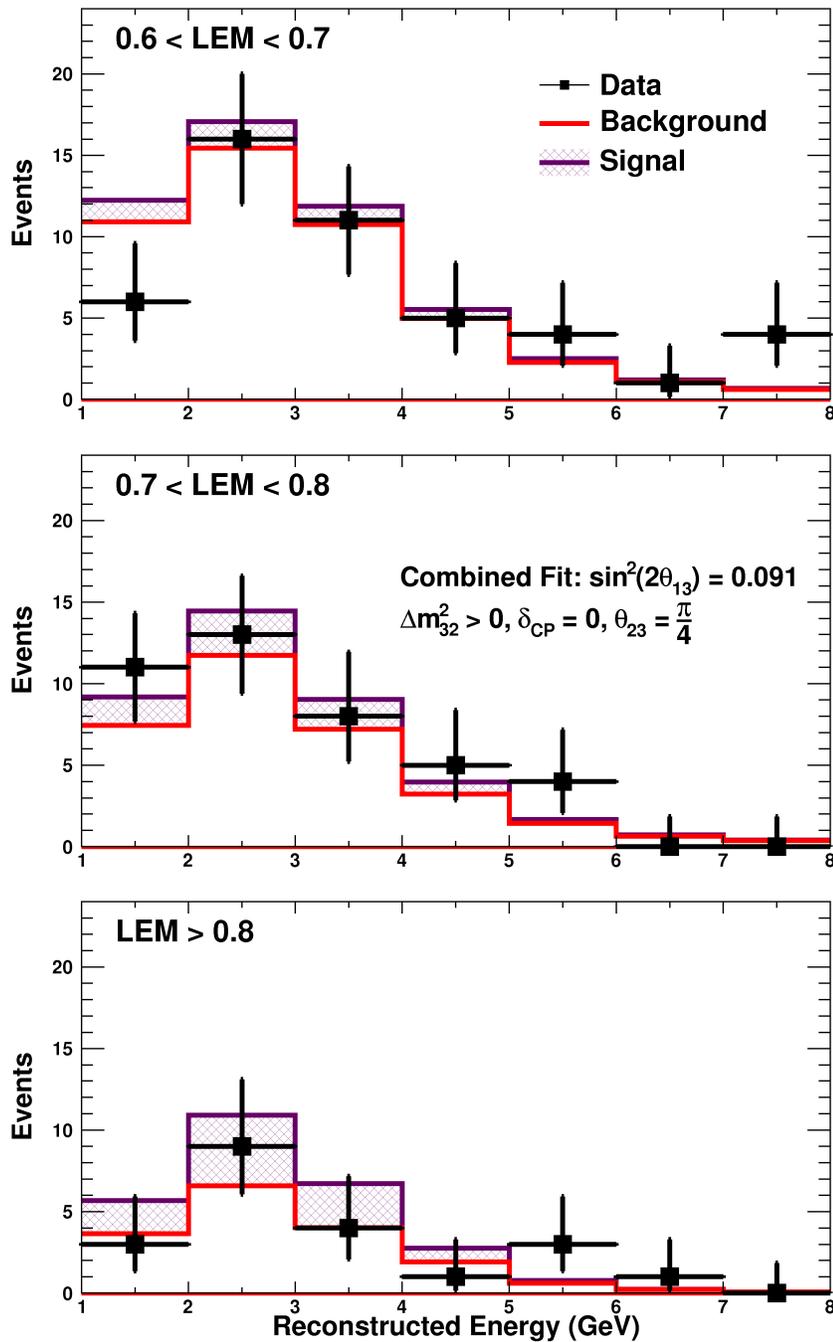


Figure 9.3: Signal implied by the best-fit combined value of $\sin^2 2\theta_{13}$ (assuming the normal hierarchy) given in the text, stacked atop the background prediction as a function of reconstructed energy, for $0.6 < \text{LEM PID} < 0.7$ (top), $0.7 < \text{LEM PID} < 0.8$ (middle), and $\text{LEM PID} > 0.8$ (bottom). The data are shown with statistical errors only. Compare to Figure 8.15; note, however, that the background predictions do not agree exactly between the two figures. This is because the spectra in Figure 8.15 were the result of the fitting procedure (including nuisance parameters) described in Chapter 8 while the spectra in the present figure were computed directly using the methods described in Chapter 6.

in the lepton sector plays an important role in leptogenesis – a mechanism for the production of a lepton–antilepton asymmetry in the early Universe. So-called sphaleron processes, in turn, could partially transform that lepton–antilepton asymmetry into a baryon–antibaryon asymmetry (i.e., baryogenesis via leptogenesis), thereby producing the matter–antimatter asymmetry observed in the Universe today [24]. The observation of leptonic CP violation could thus help shed light on one of the greatest open questions in modern cosmology.²

The questions of leptonic CP violation and the neutrino mass hierarchy will be the territory of the NO ν A experiment for the foreseeable future. Considered in some ways to be the successor to MINOS, NO ν A [142, 143] is an off-axis two-detector 810-km-baseline neutrino oscillation experiment in the NuMI beamline, optimized to study the Δm_{32}^2 -driven $\nu_\mu \rightarrow \nu_e$ oscillation mode. Its excellent energy resolution on ν_μ CC events will allow NO ν A to reduce the uncertainty on the value of $\sin^2 2\theta_{23}$ significantly, with the requisite sensitivity to answer the question of whether or not the atmospheric mixing is maximal. The near detector at Fermilab is already taking data; the far detector at Ash River in northern Minnesota is currently under construction, and the accelerator complex at Fermilab will soon be upgraded to allow for a nominal 75% increase in the neutrino beam intensity.

The NO ν A run plan includes significant running in both neutrino (π^+ , K^+ focused) and antineutrino (π^- , K^- focused) modes, taking advantage of the relatively large matter effect resulting from the long baseline to help determine the mass hierarchy. Because the matter effect produces apparent CP violation, however, there is a chance – depending on the value of δ_{CP} – of ambiguity in this determination. In that case, comparison of observed oscillation probabilities in NO ν A and T2K can help resolve the mass hierarchy. Then, depending on the hierarchy and again on the value of δ_{CP} , NO ν A, either alone or in concert with T2K, may be able to actually measure δ_{CP} .

²A recent paper [141] by Boris Kayser explains that leptogenesis makes it quite interesting to determine not only whether neutrino oscillations violate CP but also whether neutrinos are their own antiparticles.

Appendix A

Reducing the Sensitivity of LEM to Calibration Systematics

In the first electron neutrino appearance search performed by MINOS, reported in [56], the original implementation of LEM described in [92] was used as a secondary event selection. As we discussed in Chapter 4, that implementation of LEM relied on pulse heights measured in photoelectrons, as computed from ADC counts using PMT gain information obtained from light injection data. In Chapter 7, we noted the impact that this dependence had on the systematic error for the original LEM selection. Though still smaller than the expected statistical error, this systematic error was unsettlingly large, particularly when compared to that of ANN; despite being more sensitive, LEM was not used as the primary event selection method.

The solution was to abandon photoelectron pulse heights as computed in the photoelectron branch of the calibration and switch to pulse heights calibrated in the energy branch, where the constants were better known and more reliable. Deciding which pulse height measurement to use and how to incorporate it into LEM, on the other hand, was not so obvious. Certainly, the principle of LEM – matching input events to a library, ranking matches, calculating discriminant variables from the best matches – was to be maintained, and any change in the pulse height measurement had to fit within that constraint.

In this Appendix, we describe the options we considered, as well as any necessary machinery that had to be developed in the process. We conclude with a discussion

of the relative performance of LEM with each option.¹

A.1 Options for Pulse Height Measurement in LEM

The following discussion relies heavily on concepts discussed in Sections 2.3 and 4.3.2. Please refer to those sections as necessary.

Independently-computed SigMaps

Approach. Start with pulse heights in SigCor. Use the official attenuation correction code, as opposed to the approach described in Section 4.3.2.2, to correct the pulse heights for attenuation; for the hit position, use the centroid position determined by the LEM event-compacting procedure. The resulting pulse heights are thus on an equal footing with SigMaps, except that the hits have not yet been grouped into tracks and showers (a task performed by the reconstruction prior to pulse heights being corrected to SigMaps – see Footnote 7 in Section 2.3). From two histograms of individual simulated hit pulse heights calibrated in SigMaps, one in the Near Detector and one in the Far Detector, compute the ratio of the mean pulse height in the Far Detector to the mean pulse height in the Near Detector (result: 0.8417); multiply the attenuation-corrected Near Detector pulse heights by this ratio. This equalizes light levels in the Near and Far Detectors.

Motivation. Our primary aim was to switch to the energy branch of the calibration while maintaining LEM’s independence from the reconstruction. SigCors fulfilled this need while also providing corrections for drift, nonlinearity, and strip-to-strip response variations. We had a slight suspicion that LEM’s native attenuation correction could have also had a hand in the high systematic error observed, so we switched to the official calibration code and constants. Since hit position was calculated by the reconstruction, we continued to use

¹The development work we summarize here was quite extensive. As a result, LEM was not ready for use in the second electron neutrino analysis, which was performed using ANN for event selection and was reported in [57].

the LEM-computed event centroid position as the hit position to maintain our independence from the reconstruction.

Reconstructed Energy

Approach. Use pulse heights in gigaelectronvolts, as computed by the calibration. All necessary calibration steps have been performed, and light levels in the Near and Far Detectors are already equalized.

Motivation. After initial work on the “Independently-computed SigMaps” option, we found that approach to be unnecessarily complicated. The reconstruction had undergone thorough validation and was found to be reliable, so we felt comfortable in giving up LEM’s independence from the reconstruction. We thus freed ourselves to proceed to the end of the calibration chain and use pulse heights in gigaelectronvolts. Further work on the “Independently-computed SigMaps” option was abandoned.

PE-scaled SigCors

Approach. Start with pulse heights in SigCor. Find the means of two histograms of the ratio of pulse height in photoelectrons to pulse height in SigCors, one for individual simulated hits in the Near Detector (result: 0.00802) and another for individual simulated hits in the Far Detector (result: 0.01286); scale hit pulse heights using the ratio for the detector in which the hit occurred. With pulse heights now in PE-sized units, proceed exactly as in the original implementation of LEM.

Motivation. The “Reconstructed Energy” option required a substitute for the Poisson distribution in Equations 4.6 and 4.7, as we were no longer dealing with pulse heights in photoelectrons; we describe this substitute in Section A.2. The present option was crafted as a simple alternative; it required minimal modifications to the original implementation of LEM while achieving the primary aim of abandoning the photoelectron branch of the calibration.

Thus, two options – “Reconstructed Energy” and “PE-scaled SigCors” – were pursued to completion. Results are discussed in Section A.3.

A.2 Poisson Distribution Substitute for Reconstructed Energy Option

In Section 4.3.3, we used the Poisson distribution $P(n; \lambda)$ to calculate the probability that n photoelectrons will be observed by a PMT for a mean deposition of λ photoelectrons. The Reconstructed Energy implementation uses pulse heights measured in energy units, and the Poisson nature of the photon statistics underlying the measurement of those pulse heights is obscured. We thus require a different distribution $F(n; \lambda)$ to calculate the probability that a certain reconstructed energy n will be observed when a true energy λ is deposited in a scintillator strip. The Monte Carlo simulation naturally provides both of these quantities for any given hit, and our approach is to obtain the desired distributions empirically from the Monte Carlo.

We start by populating a two-dimensional histogram of true deposited energy versus reconstructed hit energy with individual hits from a large sample of Far Detector Monte Carlo events. The resulting histogram is shown in Figure A.1. We then take thin (0.2-MeV) true-energy slices of this histogram; three examples are shown in Figure A.2. These slices, when fitted (as indicated in Figure A.2) and normalized, are the desired $F(n; \lambda)$, where we take λ to be the midpoint of the true-energy interval constituting the slice.

It is interesting to note that while the Poisson distribution naturally provides the correct probability of a null observation for a nonzero mean, the asymmetric Gaussian² used in fitting these distributions – a practical choice, with no immediate theoretical motivation – does not. This must, however, be remedied; the compacting

²This is an approximate description of the appearance of the function. Known as the Novosibirsk function, it is defined as $f(x) = A \exp\left(-0.5 \left[\frac{\ln^2\{1+\Lambda\tau(x-x_0)\}}{\tau^2} + \tau^2 \right]\right)$, where A is a normalization constant, $\Lambda = \frac{\sinh(\tau\sqrt{\ln 4})}{\sigma\tau\sqrt{\ln 4}}$, x_0 is the peak position, σ is the width, and τ is a parameter that governs the asymmetry of the function’s tails. As $\tau \rightarrow 0$, $f(x)$ approaches a Gaussian with mean x_0 and variance σ^2 .

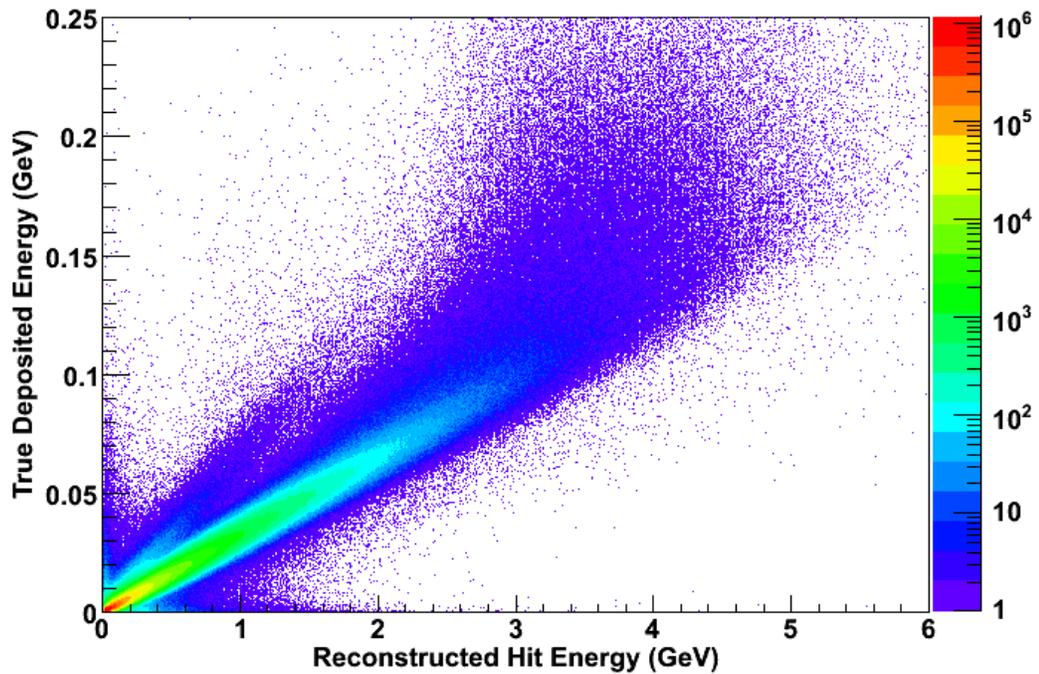


Figure A.1: Two-dimensional histogram of individual hits (color axis) from a large sample of Far Detector Monte Carlo events. The true deposited energy (vertical axis) is the actual energy deposited by a particle in a given scintillator strip (in GeV), while the reconstructed hit energy (horizontal axis) is the reconstruction's determination of the energy (in GeV) lost by the particle in traversing the *plane* – scintillator strip *and* steel. This explains the difference in scale between the two axes.

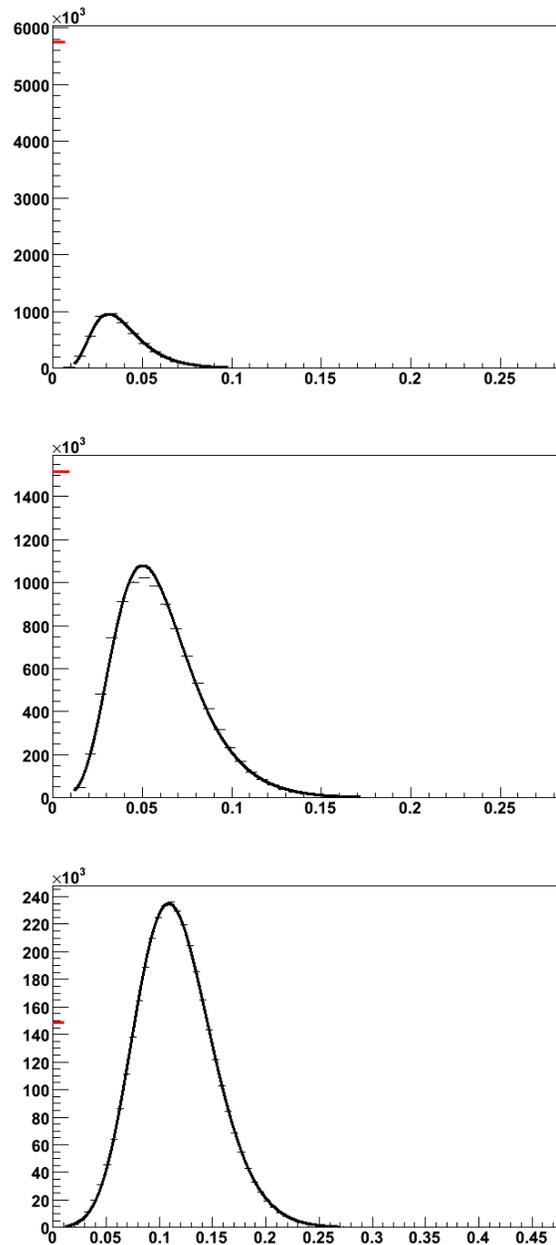


Figure A.2: True-energy slice of the histogram in Figure A.1 – i.e., projections of (short) true-energy intervals onto the reconstructed hit energy axis. Thus, the horizontal axis is the reconstructed hit energy, and the vertical axis is the number of hits with a given reconstructed hit energy. The distributions shown are for hits with true deposited energies between 0.8 and 1.0 MeV (top), 1.8 and 2.0 MeV (middle), and 3.8 and 4.0 MeV (bottom). The fit curves (solid black) are the asymmetric Gaussians discussed in the text. Note that they start at the *second* reconstructed energy bin, as the spike at zero reconstructed energy – highlighted in red in each plot at a height of approximately 5.8×10^6 (top), 1.5×10^6 (middle), and 1.5×10^5 (bottom) – is fitted with a separate step function.

procedure in LEM discards hits with pulse height less than 3 photoelectrons, so a strip with zero reconstructed energy could have had a nonzero true deposited energy. In addition, real hits can be discarded during the reconstruction process itself, so a hit with nonzero true deposited energy can ultimately be reconstructed as having zero reconstructed energy. Indeed, the distributions obtained by slicing the two-dimensional histogram in Figure A.1 have spikes at zero reconstructed energy; each example in Figure A.2 has such a spike, indicated in red.³ We address this issue by fitting with the asymmetric Gaussian starting at the distribution’s *second* bin; defining $F(n; \lambda)$ in a piecewise fashion, as a step function of width one bin for n in the first bin’s range and an asymmetric Gaussian for n beyond that bin; and preserving the relative normalization between the two pieces when normalizing the combination to obtain $F(n; \lambda)$.

A.3 Comparison of the Two Options

Both the Reconstructed Energy and the PE-scaled SigCors options achieved the desired reductions in LEM’s sensitivity to calibration systematics, as reported in [144] and [113], respectively. Recall from Section 7.5, for example, that the original implementation of LEM had Far Detector–associated gain and relative energy scale systematic errors of 5.4% and 4.9%, respectively. These dropped to 2.5% and 0.21% for the Reconstructed Energy implementation and to 0.65% and 1.6% for the PE-scaled SigCors implementation. As a result, the choice of which implementation to adopt was based entirely on maximizing sensitivity to θ_{13} . In particular, we used the full machinery of the extrapolation, along with the versatile PORP decomposition, to optimize a cut on LEM PID by maximizing a figure of merit defined by

$$FOM = \frac{N_{\text{sig}}}{\sqrt{N_{\text{bg}} + (0.01xN_{\text{bg}})^2}}, \quad (\text{A.1})$$

³Note also that the height of this spike relative to the remainder of the distribution declines with increasing true deposited energy, as a larger true deposited energy has a lower probability of producing a hit that is ultimately reconstructed as having zero energy.

which compares the significance of a signal N_{sig} relative to statistical and $x\%$ systematic fluctuations in the background N_{bg} .⁴ For $x = 5$ (i.e., 5% systematic error, a good estimate in retrospect), the PE-scaled SigCors implementation achieved a *FOM* approximately 3% greater than that of the Reconstructed Energy implementation [145] and was thus adopted as the official implementation of LEM for this analysis.

⁴This is similar to the *FOM* given in Equation 4.1, except that the present *FOM* includes systematic fluctuations as well.

Appendix B

PID Cut Optimization and Analysis Binning Studies with PORP

With our analysis technique fully laid out over the course of this thesis, we can now describe how we determined the optimal binning in reconstructed energy and LEM PID, including the determination of the optimal cut on LEM PID. As noted in Section 5.1, the background decomposition for these studies was performed using the PORP method.

The first binning study [146] was performed using statistical errors, as well as a conservative 5% background decomposition systematic error and a 50% ν_τ CC background error; none of the other systematic errors described in Chapter 7 were yet ready to be included. The decomposed Near Detector data and the Monte Carlo in both detectors were binned into bins of reconstructed energy 1 GeV wide, from 1 to 8 GeV, and bins of LEM PID 0.05 units wide, from 0 to 1. We considered a variety of binning schemes as follows, with the constraints that the full reconstructed energy range always be covered and the uppermost (or only) LEM PID bin have an upper edge of 1:

- 1 energy bin ($1 < E < 8$), 1 LEM PID bin ($x < \text{LEM PID} < 1$),
- 1 energy bin ($1 < E < 8$), 2 LEM PID bins ($y < \text{LEM PID} < x, x < \text{LEM PID} < 1$),

- 1 energy bin ($1 < E < 8$), 3 LEM PID bins ($z < \text{LEM PID} < y, y < \text{LEM PID} < x, x < \text{LEM PID} < 1$),
- 2 energy bins ($1 < E < w, w < E < 8$), 1 LEM PID bin ($x < \text{LEM PID} < 1$),
- 2 energy bins ($1 < E < w, w < E < 8$), 2 LEM PID bins ($y < \text{LEM PID} < x, x < \text{LEM PID} < 1$),

where E is the reconstructed energy in giga-electronvolts. All possible such binning schemes were generated by allowing w, x, y, z to take on all allowed values. A standard χ^2 , given by

$$\chi^2 = \sum_{i,j} \frac{(F_i^{\text{total pred}} - F_i^{\text{total data}})(F_j^{\text{total pred}} - F_j^{\text{total data}})}{V_{ij}}, \quad (\text{B.1})$$

was used as the likelihood statistic, where $F_i^{\text{total pred}}$ and $F_i^{\text{total data}}$ are the total prediction and total observation, respectively, in bin i and V is the covariance matrix containing the aforementioned errors. The figure of merit used to evaluate a given binning scheme was the sensitivity to θ_{13} at 90% C.L. provided by that binning scheme assuming the normal mass hierarchy and no CP violation.¹ Out of all of the binning schemes tested, the best sensitivity obtained was 0.0506, provided by the 1-energy-bin/3-LEM-PID-bin scheme with $z = 0.6, y = 0.7, x = 0.8$, although shifting these bin edges up or down by 0.05 LEM PID units changed the sensitivity beyond the fourth decimal place. We also found that for a given number of LEM PID bins, increasing the number of energy bins either left the sensitivity unchanged or hurt the sensitivity slightly. The optimal cut on LEM PID (i.e., x in the 1-energy-bin/1-LEM-PID-bin scheme with the best sensitivity) was found to be 0.75, but a cut of 0.7 provided only negligibly less sensitivity.

Soon thereafter, the PMT crosstalk systematic error and the various calibration-related systematic errors were ready to be included, and the study was repeated [147]

¹More precisely, this is the 90% C.L. upper limit on $\sin^2 2\theta_{13}$ one would obtain using that binning scheme – assuming $\delta_{\text{CP}} = 0$, the normal mass hierarchy, and all known oscillation parameters equal to their best-fit values – if the observation $F_i^{\text{total data}}$ were equal to the prediction $F_i^{\text{total pred}}$ for $\theta_{13} = 0$. The minimum value of χ^2 in this case is zero, at the $\theta_{13} = 0$ boundary, so $\Delta\chi^2 = \chi^2 - \chi_{\text{min}}^2 = \chi^2$ and the 90% C.L. upper limit is given by $\chi^2 = 2.71$.

with these errors as well as the relative normalization systematic error included. We also tested the 1-energy-bin/4-LEM-PID-bin scheme to see if any tangible gain in sensitivity could be achieved with an additional bin of LEM PID. The gain over the next best scheme, with one energy bin and three LEM PID bins, was about 0.5%; also, using four LEM PID bins inevitably violated our desire to maintain a buffer region between the anti-LEM sideband and the analysis region. Thus, we proceeded to consider no more than three LEM PID bins. This second study found that the 1-energy-bin/3-LEM-PID-bin scheme was still the best but with bin edges shifted slightly relative to the first study – $z = 0.65, y = 0.75, x = 0.8$. However, this scheme provided a sensitivity of 0.0529 while the optimal scheme from the first study provided a sensitivity of 0.0530. Considering the insignificance of this 0.2% difference, we saw no reason to give up on the optimal binning from the first study.

In a third and final study, we included the full, official set of systematic errors and used the official likelihood calculation method described in Section 8.1 instead of Equation B.1. Scanning over the same set of binning schemes, we found that the 1-energy-bin/3-LEM-PID-bin scheme with $z = 0.6, y = 0.7, x = 0.8$ still provided the best sensitivity, now 0.0582. Finally, we also found that the optimal cut on LEM PID had shifted, to 0.7.

As mentioned in Footnote 4 of Section 8.3, we ultimately decided to use five energy bins in the extrapolation and fitting. An additional study [148] showed that this changed the sensitivity to 0.0584 – just 0.3% worse than the optimal binning scheme.

Appendix C

Mathematical Details of the Systematic Error Evaluation

In Chapter 7, we noted that our primary task is to evaluate the $\delta_k F_i^{\alpha \text{ pred}}$ for various systematic effects k . To compute $\delta_k F_i^{\alpha \text{ pred}}$ for $\alpha \in \{\text{NC}, \nu_\mu \text{ CC}, \text{beam } \nu_e \text{ CC}\}$, we begin by taking differentials on both sides of Equation 6.1:

$$\delta_k F_i^{\alpha \text{ pred}} = \left[\delta_k N_i^{\alpha \text{ data}} \times \frac{F_i^{\alpha \text{ MC}}}{N_i^{\alpha \text{ MC}}} \right] + \left[N_i^{\alpha \text{ data}} \times \delta_k \left(\frac{F_i^{\alpha \text{ MC}}}{N_i^{\alpha \text{ MC}}} \right) \right], \quad (\text{C.1})$$

where $\delta_k N_i^{\alpha \text{ data}}$ and $\delta_k \left(\frac{F_i^{\alpha \text{ MC}}}{N_i^{\alpha \text{ MC}}} \right)$ are the 1σ shifts in $N_i^{\alpha \text{ data}}$ and $\frac{F_i^{\alpha \text{ MC}}}{N_i^{\alpha \text{ MC}}}$, respectively, induced by varying x_k from μ_k to $\mu_k + \sigma_k$. We note that among the systematic effects we consider, only one – the errors due to the HOOHE decomposition, given by V^{HOOHE} (see Section 5.2) – affects the Near Detector data, and the Monte Carlo far-to-near ratio is completely independent of this particular systematic effect. Let us index the systematic effects so that $k = 0$ is the effect of the HOOHE decomposition errors and $k > 0$ are the various systematic effects discussed in Section 7.1 and 7.2 that affect the Monte Carlo far-to-near ratio. Then

$$\delta_k N_i^{\alpha \text{ data}} = 0 \text{ if } k > 0, \quad (\text{C.2})$$

$$\delta_0 \left(\frac{F_i^{\alpha \text{ MC}}}{N_i^{\alpha \text{ MC}}} \right) = 0, \quad (\text{C.3})$$

and the only remaining nonzero terms will be $\delta_0 N_i^{\alpha \text{ data}}$ and $\delta_k \left(\frac{F_i^{\alpha \text{ MC}}}{N_i^{\alpha \text{ MC}}} \right)$ for $k > 0$. The shifts in the Far Detector predictions associated with these terms will be

$$\delta_0 F_i^{\alpha \text{ pred}} = \delta_0 N_i^{\alpha \text{ data}} \times \frac{F_i^{\alpha \text{ MC}}}{N_i^{\alpha \text{ MC}}}, \quad (\text{C.4})$$

$$\delta_k F_i^{\alpha \text{ pred}} = N_i^{\alpha \text{ data}} \times \delta_k \left(\frac{F_i^{\alpha \text{ MC}}}{N_i^{\alpha \text{ MC}}} \right). \quad (\text{C.5})$$

Let us consider the $k = 0$ case (the HOOHE decomposition errors). First, we note that $\delta_0 F_i^{\nu_\tau \text{ CC pred}} = \delta_0 F_i^{\text{signal } \nu_e \text{ CC pred}} = 0$, as the ν_τ CC and signal ν_e CC components are not involved in the HOOHE decomposition and thus have no decomposition-related errors. Defining $\alpha_1, \alpha_2 \in \{\text{NC}, \nu_\mu \text{ CC}, \text{beam } \nu_e \text{ CC}\}$, we can write

$$\begin{aligned} \delta_0 F_i^{\text{total pred}} \times \delta_0 F_j^{\text{total pred}} &= \sum_{\alpha_1, \alpha_2} \delta_0 F_i^{\alpha_1 \text{ pred}} \times \delta_0 F_j^{\alpha_2 \text{ pred}} \\ &= \sum_{\alpha_1, \alpha_2} \delta_0 N_i^{\alpha_1 \text{ data}} \times \frac{F_i^{\alpha_1 \text{ MC}}}{N_i^{\alpha_1 \text{ MC}}} \times \delta_0 N_j^{\alpha_2 \text{ data}} \times \frac{F_j^{\alpha_2 \text{ MC}}}{N_j^{\alpha_2 \text{ MC}}} \quad (\text{C.6}) \\ &= \sum_{\alpha_1, \alpha_2} \frac{F_i^{\alpha_1 \text{ MC}}}{N_i^{\alpha_1 \text{ MC}}} \times V_{\alpha_1 i, \alpha_2 j}^{\text{HOOHE}} \times \frac{F_j^{\alpha_2 \text{ MC}}}{N_j^{\alpha_2 \text{ MC}}}, \end{aligned}$$

where we have substituted Equation C.4 in the second line. In the third line, we have identified $\delta N_i^{\alpha_1 \text{ data}} \delta N_j^{\alpha_2 \text{ data}}$ as the $\alpha_1 i, \alpha_2 j$ th element of V^{HOOHE} .¹ Notational complexities aside, Equation C.6 tells us precisely how to incorporate the errors from the HOOHE fit into the total covariance matrix.

For $k > 0$, we need to compute the $\delta_k \left(\frac{F_i^{\alpha \text{ MC}}}{N_i^{\alpha \text{ MC}}} \right)$ term in Equation C.5, which we recall is the 1σ shift induced in the Monte Carlo far-to-near ratio by varying the underlying parameter x_k from its mean to one standard deviation above its mean. Since the Monte Carlo event count in a given bin and a given detector is a function of x_k , the ratio of these event counts is also a function of x_k . Thus, we can write

$$\delta_k \left(\frac{F_i^{\alpha \text{ MC}}}{N_i^{\alpha \text{ MC}}} \right) = \frac{F_i^{\alpha \text{ MC}}}{N_i^{\alpha \text{ MC}}}(x_k = \mu_k + \sigma_k) - \frac{F_i^{\alpha \text{ MC}}}{N_i^{\alpha \text{ MC}}}(x_k = \mu_k). \quad (\text{C.7})$$

¹Recall that V^{HOOHE} is indexed by analysis bin and Near Detector background component type.

This equation precisely describes our general approach to systematic shift evaluation in this analysis. In particular, we modify the standard Far Detector and Near Detector Monte Carlo samples, shifting the value of x_k in order to simulate the effect of the uncertainty under study. In practice, this is done either by reweighting simulated events under a suitable reweighting scheme (when the systematic effect happens to affect only the frequency with which certain events occur) or by regenerating the simulated events with x_k shifted up from μ_k by σ_k (when the systematic effect changes the intrinsic character of events). We then substitute our result from Equation C.7 into Equation C.5 and divide by $F_i^{\alpha \text{ pred}}$ (i.e., the nominal prediction, computed assuming $x_k = \mu_k$) to obtain the desired relative shift (see introduction to Chapter 7). Thus, using Equation 6.1 to simplify, we find that for a given systematic effect k , the relative shift in the prediction for a given component is equal to the relative shift in the far-to-near ratio for that component:

$$\frac{\delta_k F_i^{\alpha \text{ pred}}}{F_i^{\alpha \text{ pred}}} = \frac{\delta_k \left(\frac{F_i^{\alpha \text{ MC}}}{N_i^{\alpha \text{ MC}}} \right)}{\frac{F_i^{\alpha \text{ MC}}}{N_i^{\alpha \text{ MC}}}} = \frac{\frac{F_i^{\alpha \text{ MC}}}{N_i^{\alpha \text{ MC}}}(x_k = \mu_k + \sigma_k) - \frac{F_i^{\alpha \text{ MC}}}{N_i^{\alpha \text{ MC}}}(x_k = \mu_k)}{\frac{F_i^{\alpha \text{ MC}}}{N_i^{\alpha \text{ MC}}}(x_k = \mu_k)}. \quad (\text{C.8})$$

One can use essentially the same approach to compute relative shifts in a single detector. In fact, the sum in quadrature of such shifts was employed to obtain the error band in Figure 4.14. Finally, the errors on the ν_τ CC and signal ν_e CC components are computed based on Equation 6.5 and, through the $F_j^{\nu_\mu \text{ CC fid}}$ term, Equation 6.4; systematic effects affect these components through the various terms in those equations. The computation of these components is conceptually similar to that for the NC, ν_μ CC, and beam ν_e CC components but technically more challenging.

We close by noting that in Chapter 3 (and on several occasions afterward), we have stated that a number of systematic effects are “common to both detectors” and “cancel to first order” in the far-to-near ratio. We are now in a position to better understand that statement. Recall our interpretation of the individual-detector Monte Carlo event counts and their ratio as functions of x_k . If we expand these functions to first order in x_k about μ_k and use these expressions to obtain an expansion for the

fraction on the right side of Equation C.8, we find

$$\begin{aligned} \frac{\delta_k \left(\frac{F_i^{\alpha \text{ MC}}}{N_i^{\alpha \text{ MC}}} \right)}{\frac{F_i^{\alpha \text{ MC}}}{N_i^{\alpha \text{ MC}}}} &= \frac{\frac{F_i^{\alpha \text{ MC}}}{N_i^{\alpha \text{ MC}}}(x_k = \mu_k + \sigma_k) - \frac{F_i^{\alpha \text{ MC}}}{N_i^{\alpha \text{ MC}}}(x_k = \mu_k)}{\frac{F_i^{\alpha \text{ MC}}}{N_i^{\alpha \text{ MC}}}(x_k = \mu_k)} \\ &= \left(\frac{\left. \frac{\partial F_i^{\alpha \text{ MC}}}{\partial x_k} \right|_{x_k = \mu_k}}{F_i^{\alpha \text{ MC}}(\mu_k)} - \frac{\left. \frac{\partial N_i^{\alpha \text{ MC}}}{\partial x_k} \right|_{x_k = \mu_k}}{N_i^{\alpha \text{ MC}}(\mu_k)} \right) \sigma_k + O[\sigma_k^2]. \end{aligned} \quad (\text{C.9})$$

Thus, if systematic effect k affects the Far Detector and Near Detector event counts – to first order in x_k – by the same *relative* amount, the term in parentheses vanishes, and the associated systematic shift in the far-to-near ratio is second order or higher in σ_k . This is the case for the systematic effects discussed in Section 7.1. The effects discussed in Section 7.2, on the other hand, are associated with far–near differences, so their induced errors do not cancel in this manner.

Appendix D

Oscillation Parameter Uncertainties and Pseudoexperiment Generation

Throughout the analysis, we have used the best-fit values of θ_{23} , θ_{12} , $|\Delta m_{32}^2|$, and Δm_{21}^2 listed in Section 6.3 to form Far Detector predictions. Errors on these values (see Table D.1) induce errors on the predictions, and we must account for this effect in computing our Feldman–Cousins confidence intervals for $2 \sin^2 \theta_{23} \sin^2 2\theta_{13}$. We do this via an additional systematic error covariance matrix V^{osc} that is computed by generating a special set of pseudoexperiments (distinct from those generated for the Feldman–Cousins $-2\Delta \ln L$ cutoff value calculation) and is added on to the systematic error covariance matrix obtained in Chapter 7. Note that the latter step is done only for the purposes of the Feldman–Cousins pseudoexperiment generation; in computing $-2\Delta \ln L$, only the systematic error covariance matrix obtained in Chapter 7 is ever used.

Parameter	Best-fit Value	Error
θ_{23}	$\pi/4$	± 0.10
θ_{12}	0.60	± 0.02
$ \Delta m_{32}^2 $	$2.32 \times 10^{-3} \text{ eV}^2$	${}^{+0.12}_{-0.08} \times 10^{-3} \text{ eV}^2$
Δm_{21}^2	$7.59 \times 10^{-5} \text{ eV}^2$	${}^{+0.19}_{-0.21} \times 10^{-5} \text{ eV}^2$

Table D.1: Best-fit values of and errors on the known oscillation parameters used in this analysis.

At each grid point on each of the two parameter grids, perform 2000 pseudoexperiments as follows:

1. Use a Gaussian random number generator to compute values of θ_{23} , θ_{12} , $|\Delta m_{32}^2|$, and Δm_{21}^2 fluctuated within their errors.
2. Use the value of θ_{23} obtained above and the value of $2 \sin^2 \theta_{23} \sin^2 2\theta_{13}$ at the current grid point to compute the implied value of θ_{13} .
3. Compute a Far Detector prediction as prescribed in Chapter 6 using the oscillation parameter values just computed, the value of δ_{CP} at the current grid point, and the mass hierarchy represented by the current grid.
4. Sum all five components (NC, ν_μ CC, beam ν_e CC, ν_τ CC, and signal ν_e CC) to obtain the total “fluctuated prediction” F_i^m , where i indexes the analysis bin (as usual) and m indexes the pseudoexperiment.

Let F_i^{nom} represent the nominal prediction at this grid point, i.e., the prediction computed using the best-fit values in Table D.1 and the corresponding implied value of θ_{13} . We compute V^{osc} in the usual manner:

$$V_{ij}^{\text{osc}} = \langle (F_i^m - F_i^{\text{nom}}) (F_j^m - F_j^{\text{nom}}) \rangle_m, \quad (\text{D.1})$$

where the subscript m implies that the expectation value is computed over pseudoexperiments. Proceeding in this manner, we compute V^{osc} for each grid point in the two parameter grids.

Bibliography

- [1] W. Pauli. Letter to the physical society of Tübingen. 1930.
- [2] J. Chadwick. The intensity distribution in the magnetic spectrum of beta particles from radium (B + C). *Verh. Phys. Gesell.*, 16:383–391, 1914.
- [3] C. D. Ellis and W. A. Wooster. The average energy of disintegration of radium E. *Proceedings of the Royal Society of London. Series A*, 117(776):109–123, 1927.
- [4] E. Fermi. An attempt of a theory of beta radiation. 1. *Z. Phys.*, 88:161–177, 1934.
- [5] H. Bethe and R. Peierls. The “neutrino”. *Nature*, 133:532, 1934.
- [6] F. Reines and C. L. Cowan. Detection of the free neutrino. *Phys. Rev.*, 92:830–831, 1953.
- [7] C. L. Cowan, F. Reines, F. B. Harrison, H. W. Kruse, and A. D. McGuire. Detection of the free neutrino: A confirmation. *Science*, 124:103–104, 1956.
- [8] H. B. Newman and T. Ypsilantis (eds.). *History of Original Ideas and Basic Discoveries in Particle Physics (NATO Science Series B: Physics Vol. 352)*. Springer, 1996.
- [9] D. Griffiths. *Introduction to Elementary Particles*. Wiley-VCH, 2008.
- [10] K. Nakamura et al. Review of particle physics. *J. Phys. G*, 37:075021, 2010.

-
- [11] K. Zuber. *Neutrino Physics, Second Edition (Series in High Energy Physics, Cosmology and Gravitation)*. Taylor & Francis, 2011.
- [12] S. L. Glashow. Partial symmetries of weak interactions. *Nucl. Phys.*, 22:579–588, 1961.
- [13] S. Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, 1967.
- [14] A. Salam. Elementary particle theory. *N. Svartholm, Stockholm, Almquist and Wiksell*, 367, 1968.
- [15] G. 't Hooft. Renormalization of massless Yang-Mills fields. *Nucl. Phys. B*, 33:173–199, 1971.
- [16] G. 't Hooft. Renormalizable Lagrangians for massive Yang-Mills fields. *Nucl. Phys. B*, 35:167–188, 1971.
- [17] P. W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, 1964.
- [18] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, 1964.
- [19] N. Cabibbo. Unitary symmetry and leptonic decays. *Phys. Rev. Lett.*, 10:531–533, 1963.
- [20] M. Kobayashi and T. Maskawa. CP violation in the renormalizable theory of weak interaction. *Prog. Theor. Phys.*, 49:652–657, 1973.
- [21] M. Goldhaber, L. Grodzins, and A. W. Sunyar. Helicity of neutrinos. *Phys. Rev.*, 109:1015–1017, 1958.
- [22] G. Backenstoss, B. D. Hyams, G. Knop, P. C. Marin, and U. Stierlin. Helicity of mu-mesons from pi-meson decay. *Phys. Rev. Lett.*, 6:415–416, 1961.
- [23] M. Bardon, P. Franzini, and J. Lee. Helicity of μ^- mesons: Mott scattering of polarized muons. *Phys. Rev. Lett.*, 7:23–25, 1961.

-
- [24] S. Bilenky. *Introduction to the Physics of Massive and Mixed Neutrinos (Lecture Notes in Physics)*. Springer, 2010.
- [25] J. N. Bahcall, A. M. Serenelli, and S. Basu. New solar opacities, abundances, helioseismology, and neutrino fluxes. *Astrophys. J.*, 621:L85–L88, 2005.
- [26] J. N. Bahcall and A. M. Serenelli. How do uncertainties in the surface chemical abundances of the Sun affect the predicted solar neutrino fluxes? *Astrophys. J.*, 626:530, 2005.
- [27] J. N. Bahcall and R. Davis Jr. On the problem of detecting solar neutrinos. 1964. BNL-8217.
- [28] R. Davis. A review of measurements of the solar neutrino flux and their variation. *Nucl. Phys. Proc. Suppl.*, 48:284–298, 1996.
- [29] B. T. Cleveland, T. Daily, R. Davis Jr., J. R. Distel, K. Lande, et al. Measurement of the solar electron neutrino flux with the Homestake chlorine detector. *Astrophys. J.*, 496:505–526, 1998.
- [30] K. S. Hirata et al. Real time, directional measurement of B-8 solar neutrinos in the Kamiokande-II detector. *Phys. Rev. D*, 44:2241, 1991.
- [31] Y. Fukuda et al. Solar neutrino data covering solar cycle 22. *Phys. Rev. Lett.*, 77:1683–1686, 1996.
- [32] J. N. Abdurashitov et al. Measurement of the solar neutrino capture rate with gallium metal. III: Results for the 2002–2007 data-taking period. *Phys. Rev. C*, 80:015807, 2009.
- [33] M. Altmann et al. Complete results for five years of GNO solar neutrino observations. *Phys. Lett. B*, 616:174–190, 2005.
- [34] J. Hosaka et al. Solar neutrino measurements in Super-Kamiokande-I. *Phys. Rev. D*, 73:112001, 2006.

-
- [35] Z. Maki, M. Nakagawa, and S. Sakata. Remarks on the unified model of elementary particles. *Prog. Theor. Phys.*, 28:870–880, 1962.
- [36] B. Pontecorvo. Neutrino experiments and the problem of conservation of leptonic charge. *Sov. Phys. JETP*, 26:984–988, 1968.
- [37] 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, 2002.
- [38] J. N. Bahcall. Solar models and solar neutrinos: Current status. *Phys. Scripta*, T121:46–50, 2005.
- [39] B. Aharmim et al. Electron energy spectra, fluxes, and day-night asymmetries of B-8 solar neutrinos from measurements with NaCl dissolved in the heavy-water detector at the Sudbury Neutrino Observatory. *Phys. Rev. C*, 72:055502, 2005.
- [40] C. W. Kim and A. Pevsner. *Neutrinos in Physics and Astrophysics (Contemporary Concepts in Physics)*. Harwood Academic Publishers, 1993.
- [41] Y. Fukuda et al. Atmospheric muon-neutrino/electron-neutrino ratio in the multiGeV energy range. *Phys. Lett. B*, 335:237–245, 1994.
- [42] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81:1562–1567, 1998.
- [43] Y. Ashie et al. Evidence for an oscillatory signature in atmospheric neutrino oscillation. *Phys. Rev. Lett.*, 93:101801, 2004.
- [44] M. H. Ahn et al. Measurement of neutrino oscillation by the K2K experiment. *Phys. Rev. D*, 74:072003, 2006.
- [45] P. Adamson et al. Measurement of the neutrino mass splitting and flavor mixing by MINOS. *Phys. Rev. Lett.*, 106:181801, 2011.

- [46] M. Apollonio et al. Search for neutrino oscillations on a long baseline at the CHOOZ nuclear power station. *Eur. Phys. J. C*, 27:331–374, 2003.
- [47] F. Boehm, J. Busenitz, B. Cook, G. Gratta, H. Henrikson, et al. Final results from the Palo Verde neutrino oscillation experiment. *Phys. Rev. D*, 64:112001, 2001.
- [48] K. Eguchi et al. First results from KamLAND: Evidence for reactor anti-neutrino disappearance. *Phys. Rev. Lett.*, 90:021802, 2003.
- [49] B. Kayser. On the quantum mechanics of neutrino oscillation. *Phys. Rev. D*, 24:110, 1981.
- [50] R. W. L. Jones, G. Borissov, M. W. Grunewald, and M. Smizanska. Precision electroweak measurements on the Z resonance. *Physics Reports*, 427(5-6), 2006.
- [51] B. Aharmim et al. An independent measurement of the total active B-8 solar neutrino flux using an array of He-3 proportional counters at the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 101:111301, 2008.
- [52] S. Abe et al. Precision measurement of neutrino oscillation parameters with KamLAND. *Phys. Rev. Lett.*, 100:221803, 2008.
- [53] L. Wolfenstein. Neutrino oscillations in matter. *Phys. Rev. D*, 17:2369–2374, 1978.
- [54] S. P. Mikheyev and A. Y. Smirnov. Resonance amplification of oscillations in matter and spectroscopy of solar neutrinos. *Sov. J. Nucl. Phys.*, 42:913–917, 1985.
- [55] E. K. Akhmedov, R. Johansson, M. Lindner, T. Ohlsson, and T. Schwetz. Series expansions for three flavor neutrino oscillation probabilities in matter. *JHEP*, 0404:078, 2004.
- [56] P. Adamson et al. Search for muon-neutrino to electron-neutrino transitions in MINOS. *Phys. Rev. Lett.*, 103:261802, 2009.

- [57] P. Adamson et al. New constraints on muon-neutrino to electron-neutrino transitions in MINOS. *Phys. Rev. D*, 82:051102, 2010.
- [58] J. Hylen, W. M. Smart, A. Wehmann, S. R. Menary, A. J. Malensek, et al. Conceptual design for the technical components of the neutrino beam for the Main Injector (NuMI). FERMILAB-TM-2018, 1997.
- [59] P. Adamson et al. A study of muon neutrino disappearance using the Fermilab Main Injector neutrino beam. *Phys. Rev. D*, 77:072002, 2008.
- [60] R. R. Wilson. The Tevatron. *Physics Today*, 30N10:23–30, 1977.
- [61] J. Hylen et al. *NuMI Technical Design Handbook*. http://www-numi.fnal.gov/numwork/tdh/tdh_index.html.
- [62] A. G. Abramov, N. A. Galyaev, V. I. Garkusha, J. Hylen, F. N. Novoskoltsev, A. D. Ryabov, and V. G. Zarucheisky. Beam optics and target conceptual designs for the NuMI project. *Nucl. Instr. and Meth. A*, 485(3):209–227, 2002.
- [63] D. G. Michael et al. The magnetized steel and scintillator calorimeters of the MINOS experiment. *Nucl. Instr. and Meth. A*, 596:190–228, 2008.
- [64] P. Adamson, G. Crone, L. Jenner, R. Nichol, R. Saakian, et al. The MINOS calibration detector. *Nucl. Instr. and Meth. A*, 556:119–133, 2006.
- [65] ANSYS Inc. Engineering commercial software. Canonsburg, PA 15317.
- [66] P. Adamson, J. Alner, B. Anderson, T. Chase, P. J. Dervan, et al. The MINOS light injection calibration system. *Nucl. Instr. and Meth. A*, 492:325–343, 2002.
- [67] P. Adamson, L. Barrett, A. Belias, A. Blake, A. Cabrera, et al. On the linearity of the MINOS light-injection calibration system. *Nucl. Instr. and Meth. A*, 521:361–366, 2004.
- [68] B. J. Rebel. *Neutrino-induced muons in the MINOS Far Detector*. PhD thesis, Indiana University, 2004. Advisor: S. Mufson.

- [69] T. Cundiff et al. The MINOS Near Detector front end electronics. *IEEE Trans. Nucl. Sci.*, 53:1347–1355, 2006.
- [70] R. J. Yarema, G. W. Foster, J. Hoff, M. Sarraj, and T. Zimmerman. A fast, wide range charge integrator and encoder ASIC for photomultiplier tubes. *IEEE Trans. Nucl. Sci.*, 40:750–752, 1993.
- [71] J. Oliver et al. Design and performance of the readout system of the MINOS Far Detector. *IEEE Trans. Nucl. Sci.*, 51:2193–2195, 2004.
- [72] A. Belias, S. Madani, T. C. Nicholls, G. F. Pearce, G. J. Crone, et al. The MINOS data acquisition system. *IEEE Trans. Nucl. Sci.*, 51:451–455, 2004.
- [73] J. J. Hartnell. *Measurement of the calorimetric energy scale in MINOS*. PhD thesis, University of Oxford, 2005. Advisors: A. Weber and G. Pearce.
- [74] D. E. Groom, N. V. Mokhov, and S. I. Striganov. Muon stopping power and range tables 10-MeV to 100-TeV. *Atom. Data Nucl. Data Tabl.*, 78:183–356, 2001.
- [75] R. Brun et al. GEANT detector description and simulation tool. CERN Program Library Long Writeup W5013, 1994.
- [76] P. L. Vahle. *Electromagnetic interactions in the MINOS detectors*. PhD thesis, University of Texas at Austin, 2004. Advisor: K. Lang.
- [77] C. Zeitnitz and T. A. Gabriel. The GEANT–CALOR interface and benchmark calculations of ZEUS test calorimeters. *Nucl. Instr. and Meth. A*, 349:106–111, 1994.
- [78] M. A. Kordosky. *Hadronic interactions in the MINOS detectors*. PhD thesis, University of Texas at Austin, 2004. Advisor: K. Lang.
- [79] MINOS collaboration. Calibration position paper for pre-shutdown data. minos-doc-3941, 2008.

- [80] P. Adamson et al. First direct observation of muon antineutrino disappearance. *Phys. Rev. Lett.*, 107:021801, 2011.
- [81] P. Adamson et al. Search for the disappearance of muon antineutrinos in the NuMI neutrino beam. *Phys. Rev. D*, 84:071103, 2011.
- [82] A. I. Himmel. *Antineutrino oscillations in the atmospheric sector*. PhD thesis, California Institute of Technology, 2011. Advisor: H. B. Newman.
- [83] P. Ballester. Hough transform for robust regression and automated detection. *Astron. Astrophys.*, 286:1011, 1994.
- [84] R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME – Journal of Basic Engineering, Series D*, 82:35–45, 1960.
- [85] A. Ferrari, P. R. Sala, A. Fasso, and J. Ranft. FLUKA: A multi-particle transport code (program version 2005). 2005.
- [86] S. Kopp, Ž. Pavlović, P. Vahle, and R. Ospanov. Constraining the beam neutrino MC flux using the MINOS ND data. *minos-doc-2965*, 2007.
- [87] H. Gallagher. The NEUGEN neutrino event generator. *Nucl. Phys. Proc. Suppl.*, 112:188–194, 2002.
- [88] T. Yang, C. Andreopoulos, H. Gallagher, K. Hoffmann, and P. Kehayias. A hadronization model for few-GeV neutrino interactions. *Eur. Phys. J. C*, 63:1–10, 2009.
- [89] T. Sjostrand, S. Mrenna, and P. Z. Skands. PYTHIA 6.4 physics and manual. *JHEP*, 0605:026, 2006.
- [90] Z. Koba, H. B. Nielsen, and P. Olesen. Scaling of multiplicity distributions in high-energy hadron collisions. *Nucl. Phys. B*, 40:317–334, 1972.

- [91] R. Merenyi, W. A. Mann, T. Kafka, W. Leeson, B. Saitta, et al. Determination of pion intranuclear rescattering rates in muon-neutrino Ne versus muon-neutrino D interactions for the atmospheric neutrino flux. *Phys. Rev. D*, 45:743–751, 1992.
- [92] J. P. Ochoa-Ricoux. *A search for muon neutrino to electron neutrino oscillations in the MINOS experiment*. PhD thesis, California Institute of Technology, 2009. Advisor: H. B. Newman.
- [93] I. Bird. Computing for the Large Hadron Collider. *Ann. Rev. Nucl. Part. Sci.*, 61:99–118, 2011.
- [94] D. R. Yocum, E. Berman, P. Canal, K. Chadwick, T. Hesselroth, et al. Fermi-Grid. 2007. FERMILAB-CONF-07-125-CD.
- [95] A. Holin. Nue muon removal background estimation. minos-doc-4018, 2008.
- [96] J. Ling. MINOS detector WLS fiber attenuation calibration documentation. minos-doc-4850, 2008.
- [97] T. Yang and G. Pawloski. Crosstalk tuning. minos-doc-5319, 2008.
- [98] J. P. Ochoa-Ricoux. The status of the new MCNN. minos-doc-4888, 2008.
- [99] J. P. Ochoa-Ricoux. The Monte Carlo Nearest Neighbors selection for ν_e CC events. minos-doc-5031, 2008.
- [100] D. J. C. MacKay. *Information Theory, Inference and Learning Algorithms*. Cambridge University Press, 2003.
- [101] R. Brun and F. Rademakers. ROOT: An object oriented data analysis framework. *Nucl. Instr. and Meth. A*, 389:81–86, 1997.
- [102] A. Hocker, J. Stelzer, F. Tegenfeldt, H. Voss, K. Voss, et al. TMVA – Toolkit for multivariate data analysis. *PoS*, ACAT:040, 2007.

-
- [103] J. P. Ochoa-Ricoux, M. H. Orchanian, R. B. Patterson, and R. B. Toner. The Library Event Matching (LEM) particle ID. minos-doc-7991, 2011.
- [104] J. Boehm, H. R. Gallagher, and T. Yang. Hadronization model uncertainties for the electron neutrino appearance analysis. minos-doc-5392, 2008.
- [105] P. Vahle. Nue PID on CalDet electrons. minos-doc-5546, 2009.
- [106] J. P. Ochoa and R. B. Patterson. The impact of background composition on ν_e sensitivity. minos-doc-3760, 2007.
- [107] J. A. B. Coelho. Near Detector beam decomposition for the 3rd nue analysis. minos-doc-7986, 2011.
- [108] P. Adamson et al. Measurement of neutrino oscillations with the MINOS detectors in the NuMI beam. *Phys. Rev. Lett.*, 101:131802, 2008.
- [109] R. Ospanov. *A measurement of muon neutrino disappearance with the MINOS detectors and NuMI beam*. PhD thesis, University of Texas at Austin, 2008. Advisor: K. Lang.
- [110] D. A. Petyt. PID sensitivity study status. minos-doc-2743, 2007.
- [111] J. Boehm. Measuring ν_e CC selection efficiency and systematic error using the Muon Removal with MC Electron Addition process. minos-doc-5385, 2009.
- [112] A. Holin. Muon removal background estimation method. minos-doc-4141, 2008.
- [113] M. Betancourt et al. Systematic uncertainties for the third nue box opening. minos-doc-7976, 2011.
- [114] M. Campanella, A. Ferrari, P. R. Sala, and S. Vanini. First calorimeter simulation with the FLUGG prototype. ATL-SOFT-99-004, 10 December 1999.
- [115] M. Campanella, A. Ferrari, P. R. Sala, and S. Vanini. Reusing code from FLUKA and GEANT4 geometry. ATL-SOFT-98-039, 13 October 1998.

-
- [116] A. I. Himmel. The NuMI beam simulation with Flugg. minos-doc-6316, 2010.
- [117] M. Dorman. Target decay weighting function in MCReweight. minos-doc-6477, 2009.
- [118] C. Andreopoulos et al. Updated cross section model uncertainties for the charged current analysis. minos-doc-2989, 2007.
- [119] H. Gallagher. Uncertainty on the CC tau rate in the Far Detector. minos-doc-5442, 2008.
- [120] J. Boehm, H. Gallagher, and T. Yang. Hadronization model uncertainties for the electron neutrino appearance analysis. minos-doc-5392, 2008.
- [121] M. Kordosky. A procedure to reweight events to account for uncertainties in final state interactions. minos-doc-3449, 2007.
- [122] S. Budd. Dogwood calibration errors, runs 1-6. minos-doc-7962, 2011.
- [123] P. Vahle. Updated normalization systematic. minos-doc-6854, 2010.
- [124] MINOS Collaboration. 2009 calibration position paper on runs I-II-III. minos-doc-6717, 2010.
- [125] P. A. Rodrigues. 2PE cut implementation. minos-doc-5141, 2008.
- [126] A. P. Schreckenberger. MRE 3rd analysis note. minos-doc-7980, 2011.
- [127] J. Boehm et al. Systematic uncertainties on the measurement of ν_e CC appearance. minos-doc-5387, 2008.
- [128] TMinuit package. <http://root.cern.ch/root/html/TMinuit.html>.
- [129] M. H. Orchanian and R. B. Toner. The Far Detector sidebands for the third nue analysis. minos-doc-8000, 2011.
- [130] 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.

-
- [131] A. Osipowicz et al. KATRIN: A next generation tritium beta decay experiment with sub-eV sensitivity for the electron neutrino mass. Letter of intent. *hep-ex/0109033*, 2001.
- [132] K. Zuber. COBRA: Double beta decay searches using CdTe detectors. *Phys. Lett. B*, 519:1–7, 2001.
- [133] M. Danilov, R. DeVoe, A. Dolgolenko, G. Giannini, G. Gratta, et al. Detection of very small neutrino masses in double beta decay using laser tagging. *Phys. Lett. B*, 480:12–18, 2000.
- [134] D. G. Phillips II, E. Aguayo, F. T. Avignone III, H. O. Back, A. S. Barabash, et al. The Majorana experiment: An ultra-low background search for neutrinoless double-beta decay. *arXiv:1111.5578*, 2011.
- [135] S. Schonert et al. The GERmanium Detector Array (GERDA) for the search of neutrinoless beta beta decays of Ge-76 at LNGS. *Nucl. Phys. Proc. Suppl.*, 145:242–245, 2005.
- [136] M. C. Gonzalez-Garcia, M. Maltoni, and J. Salvado. Updated global fit to three neutrino mixing: Status of the hints of $\theta_{13} > 0$. *JHEP*, 1004:056, 2010.
- [137] K. Abe et al. Indication of electron neutrino appearance from an accelerator-produced off-axis muon neutrino beam. *Phys. Rev. Lett.*, 107:041801, 2011.
- [138] Y. Abe et al. Indication for the disappearance of reactor electron antineutrinos in the Double Chooz experiment. *Phys. Rev. Lett.*, 108:131801, 2012.
- [139] F. P. An et al. Observation of electron-antineutrino disappearance at Daya Bay. *Phys. Rev. Lett.*, 108:171803, 2012.
- [140] J. K. Ahn et al. Observation of reactor electron antineutrino disappearance in the RENO experiment. *Phys. Rev. Lett.*, 108:191802, 2012.
- [141] B. Kayser. Two questions about neutrinos. *arXiv:1012.4469*, 2010.

-
- [142] D. S. Ayres et al. NOvA: Proposal to build a 30 kiloton off-axis detector to study $\nu_\mu \rightarrow \nu_e$ oscillations in the NuMI beamline. *hep-ex/0503053*, 2004.
- [143] B. Rebel. Status of NOvA. *AIP Conf. Proc.*, 1382:88–90, 2011.
- [144] M. H. Orchanian. Initial systematics studies for LEM and antiPID update. *minos-doc-7272*, 2010.
- [145] R. B. Toner. LEM status update. *minos-doc-7089*, 2010.
- [146] M. H. Orchanian. First look at multibin sensitivities for LEM and binning optimization. *minos-doc-7699*, 2010.
- [147] M. H. Orchanian. Next round of multibin sensitivities for LEM. *minos-doc-7746*, 2010.
- [148] R. B. Toner. Predictions of a 3x5 fit. *minos-doc-8212*, 2011.