

Searching for Light Sterile Neutrinos with NOvA
Through Neutral-Current Disappearance

Dissertation

Presented to the Faculty of the Graduate School

of the University of Cincinnati

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

in the Physics Department

of College of Arts and Sciences

by

Shaokai Yang, M.S.

UNIVERSITY OF CINCINNATI

July 2019

Committee Chair: Alexandre Sousa, Ph.D.

Abstract

Most appearance and disappearance neutrino oscillation experiments have observed evidence for neutrino oscillations that agree with the three-flavor neutrino paradigm. Explaining unexpected signal from short-baseline appearance experiments, such as LSND and MiniBooNE, in terms of neutrino oscillations requires the existence of light sterile neutrino(s). The study of light sterile neutrino induced flavor mixing in NOvA uses a long-baseline of 810 km between the Near Detector (ND) at Fermilab and the Far Detector (FD) in Minnesota. The signal for light sterile neutrino-driven oscillations is a deficit of Neutral-Current (NC) neutrino interactions at the FD concerning the ND prediction. In our analysis, we are employing current innovations in machine learning technologies for NC sterile neutrino searches with NOvA. I will detail the state-of-art algorithm which we employed and the improved results obtained in the NOvA NC Disappearance analysis.

To My Parents, Yachun and Xiuju

Acknowledgements

I want to take this opportunity to express my appreciation to all who have played a role in supporting me throughout these years and enabled the completion of this thesis. I am forever grateful.

I owe all this work to my supervisor, Alexandre Sousa. Alex, you have been not only a never-ending source of inspiration for me during these five years but also the reason I joined neutrino physics. Thank you for the opportunity of becoming your advisee, and for your supervision, support, and patience.

I cannot fully express my thanks to Dr. Deshan Yang and Dr. Zhizhong Xing, I would not be a physicist had it not been for your encouragement.

Thanks to the other members of my thesis committee: Randy Johnson, Michael Ma, and Philip Argyres, for reading this document. Thanks also to the other faculty of the UC NOvA group, Adam Aurisano, for your support.

Thank you, Gavin Davies, Louise Suter, Sijith Edayath, Gareth Kafka, Siva Prasad Kasetti, and Rijeesh Keloth, for helping me learn the ropes early on and teaching me how to stand on my own two feet.

Thank you, Jose Andres Sepulveda-Quiroz, Xin Tian, Enhao Song, Prabhjot Singh, Jianming Bian, Martin Frank, Nitin Yadav, Biao Wang, Leo Aliaga, Zhuting Sun, Saleh Qutub, Onur Rauf, Joshua Armstrong Eby, and Navaneeth Poonthottathil, I would not choose different friends if I had to do it all over again.

x

Thank you to my parents, sisters, brothers, and friends for your love and support. Thank you for believing in me, and encouraging me when times got complicated.

Moreover, last but not least, thank you Yan Sha, for keeping me sane through all of this. I love you!

Contents

1	Analysis Motivation and Research Strategy	2
1.1	Analysis Motivation	3
1.1.1	Neutrino. A Brief History of a Unique Particle	4
	The Conundrum of the Continuous β Decay Spectrum	4
	Pauli's Proposal	5
	The First Theory of β Decay	6
	Happy to See You, Little Neutral One!	7
	Different Types of Neutrino	8
	Neutrino and Discovery of the Neutral Current Weak Interaction	10
	What They Knew!	12
1.1.2	Neutrinos in the Standard Model and Beyond	13
	The Standard Model of Particle Physics	14
	Standard Model Particles and Their Interactions	15
	The Only Massless Matter Particles in the Standard Model	17
	Neutrino Oscillation as a Window to New Physics	18
1.1.3	The $\bar{\nu}_e$ Appearance Anomaly	18
	The LSND Signal	18
	The KARMEN Constraint	20
	The MiniBooNE ν_e and $\bar{\nu}_e$ Appearance Searches	20
1.1.4	What We Know?	21

1.2	Research Strategy	21
1.2.1	Accelerator neutrinos	21
1.2.2	Neutrino Detection	21
1.2.3	General Strategy	22
1.2.4	Machine-Learning Based Improvements	23
2	The NOvA Experiment	24
2.1	The NuMI Beam	26
2.1.1	Neutrino-Generating Process	29
2.1.2	Horn Configurations	32
2.1.3	NuMI Upgrade	33
2.2	NOvA Detectors	34
2.2.1	Off-Axis Location of Detectors	36
2.2.2	Detector Technology	37
	Liquid Scintillator	39
	Optical Fiber	40
	Photodetector and Electronics	42
	PVC Modules	44
2.2.3	The Near Detector	46
2.2.4	The Far Detector	48
3	Analysis Files Production	49
3.1	Data Taking	52
3.1.1	DAQ Overview	52
3.1.2	Triggering Systems	54
3.1.3	Timing System	57
3.1.4	Live Data-Taking	58
3.1.5	Accumulated NOvA Data Set	59

3.2	Energy Calibration	59
3.2.1	Relative Calibration	61
3.2.2	Absolute Calibration	62
3.3	Event Reconstruction	63
3.3.1	Reconstruction Philosophy	63
3.3.2	Reconstruction Steps	63
	Interaction Separation with DBSCAN	63
	Guidelines with Hough Transform	64
	Vertex Identification with Elastic-Arms	65
	Further Reconstruction Objects	65
3.4	NOvA Simulation File Production	65
3.4.1	Beam Flux Simulation	67
3.4.2	Interaction Simulation	68
	GENIE	68
	Cross Section Model	68
	Neutrino-induced Hadron Production	70
	Intranuclear Hadron Transport	70
	Nuclear Physics Model	71
3.4.3	Detector Simulation	71
	Photon Transport	71
	Electronic Readout	72
4	Neutrino Interaction Classification and Signal Selection	73
4.1	Neutral Current Signal Pre-Selection	77
4.1.1	Pre-Selection Cuts	77
4.1.2	Data Quality	77
	Timing cut	77

	Beam Quality Cuts	78
	Data Quality Cuts	78
4.1.3	Analysis Mask	79
4.1.4	Cosmic Veto	79
4.1.5	Event Quality	80
4.1.6	Fiducial Volume and Containment	82
	Far Detector	83
	Near Detector	84
4.1.7	Cosmic Rejection Pre-Selection Cuts	87
	Cosmic Photon Rejection	88
	Combined Transverse Momentum Fraction and Shower Position Cut	89
	Time Gap Cut	91
4.1.8	NC/CC Separation Pre-Selection Cuts	93
	Number of Hits	94
	Visible Energy	94
	ND p_T/p Cut	94
4.2	NC-Dedicated Cosmic Rejection PID	97
4.2.1	Cosmic Rays and NOvA Far Detector	97
4.2.2	Machine Learning Algorithm Selection	99
4.2.3	Training Phase	103
	Combining algorithms	103
	Input Event-Variable Selection	107
	Growing the Decision Tree	122
	Overtraining and Pruning	123
	Final Hyper-Parameter Setting	124
	BDT Output	124

4.2.4	Application Phase	125
4.3	NC/CC Binary Classifier	127
4.3.1	Input Event-Variables Selection	128
4.3.2	Training and Application Results	144
4.4	PID Tuning for Nus17 Official and Re-Analysis Selections	147
4.4.1	2017 Official	147
FD	147
ND	147
4.4.2	2017 Re-Analysis	148
FD	148
ND	148
4.5	Predicted Near and Far Detector Spectra	150
5	Extraction of Sterile Mixing Parameters	153
5.1	3 + 1 Model	157
5.2	Far Detector Spectra Prediction	161
5.2.1	Near Detector Decomposition	162
5.2.2	Near-to-Far Extrapolation	162
Truth Energy Extrapolation	163
Reco Energy Extrapolation	165
Base MC Extrapolation	165
Final Extrapolation Output	166
5.2.3	Far Detector Spectra Prediction	166
5.3	Systematic Uncertainty Analysis	167
5.3.1	Mixing Parameter Values Used in Systematic Studies	167
5.3.2	Beam Systematic Uncertainty	167
Beam Transport Systematic	168

	The beam transport systematic effects on ND spectra . . .	170
	The beam transport systematic effects on FD spectra (the direct selection results)	171
	The beam transport systematic effects on FD spectra (the extrapolated selection result)	172
	Package to Predict the FluX Systematic	173
	The PPFX systematic effects on ND spectra.	174
	The PPFX systematic effects on FD spectra (direct selection results)	175
	The PPFX systematic effects on FD spectra (the extrapo- lated selection results)	176
5.3.3	GENIE Simulation	177
	The GENIE systematic effects on ND spectra.	178
	The GENIE systematic effects on FD spectra (the direct se- lection results)	179
	The GENIE systematic effects on FD spectra (the extrapo- lated selection results)	180
5.3.4	Calibration Systematic	181
	Absolute Calibration	181
	The Absolute Calibration systematic effects on ND spectra	182
	The Absolute Calibration systematic effects on ND spectra (the direct selection results)	183
	The Absolute Calibration systematic effects on ND spectra (the extrapolated selection results)	184
	Relative Calibration	184
	The Relative Calibration systematic effects on ND spectra	186

	The Relative Calibration systematic effects on FD spectra (the direct selection results)	187
	The Relative Calibration systematic effects on FD spectra (the extrapolated selection results)	188
	Cherenkov Systematic	189
	The Cherenkov and Light Level systematic effects on ND spectra	190
	The Cherenkov and Light Level systematic effects on FD spectra (the direct selection results)	191
	The Cherenkov and Light Level systematic effects on FD spectra (the extrapolated selection results)	192
5.3.5	ND MC Decomposition Systematic	193
	The ND MC Decomposition systematic effects on ND spec- tra	194
	The ND MC Decomposition systematic effects on FD spec- tra (the direct selection results)	195
	The ND MC Decomposition systematic effects on FD spec- tra (the extrapolated selection results)	196
5.3.6	Acceptance Study	197
	The Acceptance systematic effects on ND One spectra . . .	198
	The Acceptance systematic effects on ND Two spectra . .	199
	The Acceptance systematic effects on ND Three spectra . .	200
	The Acceptance systematic effects on ND Four spectra . .	201
5.3.7	3-favor oscillation parameters	202
5.3.8	The normalization uncertainty	202
5.3.9	Uncertainty Summary	203
5.4	Fitting Procedure	204

5.4.1	Shape Fitting	204
5.4.2	Treatment of PMNS Mixing Parameters	205
5.4.3	Treatment of Extra Mixing Parameters in 3+1 Model	206
5.4.4	Treatment of Systematic Uncertainties	207
5.5	Analysis Results	207
5.5.1	Official 2017 Results	208
	2017 Official 2D Contour on θ_{34} and θ_{24}	212
	Global Picture on θ_{34} and θ_{24}	212
5.5.2	2017 Re-Analysis vs 2017 Official Analysis	214
	2017 NC Selection Comparison	214
	Re-Analysis Selection Sideband Study	219
	2017 Re-Analysis Selection Results	220
	2017 Re-Analysis Selected NC-like Data Distributions	220
	2017 Re-Analysis R-Ratio	224
	2017 Re-Analysis 2D Contour on θ_{34} and θ_{24}	224
	Global Picture on θ_{34} and θ_{24}	226
6	Conclusion and Future Improvements	229
A	2017 Re-Analysis Based NC Disappearance δ_{24} Dependence Study	232
B	2017 Official and 2017 Re-Analysis Per-Period Study	238
B.1	NOvA Per-Period Data Set Details	238
B.2	2017 Official Vs 2017 Re-Analysis Per-Period Selection Results	239
B.3	2017 Re-Analysis Per-Period Selected Data Distribution	240
B.3.1	Period One	240
B.3.2	Period Two	241
B.3.3	Period Three	242

B.3.4 Period Five 243

References 254

List of Figures

1.1	β Decay Electron Energy Spectrum	5
1.2	The Gargamelle heavy bubble chamber detector	11
1.3	Standard Model Particles	15
1.4	Standard Model Interaction Feynman diagrams	16
1.5	Three-Flavor Neutrino Model Oscillation	18
1.6	A Schematic Drawing of the LSND Detector	19
1.7	The LSND Energy Distribution	19
1.8	The KARMEN Constraint	20
1.9	Minos Beam	22
1.10	General Analysis Chain	23
2.1	NOvA Detectors Location	25
2.2	NuMI Horn	27
2.3	NuMI Beamline Path	28
2.4	Fermilab Accelerator Complex	29
2.5	The Components of NuMI beamline	30
2.6	The Components of the NuMI Target	31
2.7	NuMI Focusing Horn	33
2.8	3D schematic diagram of NOvA detector and neutrino interaction.	35
2.9	The Neutrino Flux Energy	37
2.10	The Simulated Neutrino Flux Energy.	38

2.11	The Expected Neutrino Interaction Rates.	38
2.12	A Schematic of a Single Cell	39
2.13	NOvA Detector with Glued Cells	40
2.14	Scintillator Wavelength.	41
2.15	NOvA WLS fiber absorption and emission spectra	42
2.16	One NOvA APD	43
2.17	The Schematic of the APD and the FEB	43
2.18	Profile of the PVC Cells	44
2.19	Schematic of the NOvA PVC Module	45
2.20	Schematic of the NOvA and MINOS ND hall	46
2.21	Pictures of the NOvA Near Detector	47
2.22	Schematic of the NOvA Far Detector	48
3.1	Diagram of DCM placement on the FD.	52
3.2	A schematic overview of NOvA Far Detector readout and DAQ systems.	53
3.3	DCM layout in FD.	54
3.4	NOvA Data-Driven Trigger System.	54
3.5	NOvA running triggers for 2017 Data Set.	55
3.6	Scheme of SNEWS Notification Propagation.	56
3.7	A Scheme of the SuperNova Triggering System	57
3.8	Timing System.	58
3.9	Live Data Taking	58
3.10	NOvA Data Taking	59
3.11	NOvA Simulation Chain	66
4.1	FD Energy spectra after applied Event Quality cuts.	82
4.2	ND Energy Spectra after Applied Event Quality cuts	82

4.3	Prong Distance to the Top Side of the Detector.	84
4.4	FD Energy spectra after applied containment cuts	85
4.5	ND Energy spectra after applied Fiducial Volume cuts	86
4.6	ND Energy spectra after applied Preselection cuts	87
4.7	Example of a cosmic ray event selected as an NC interaction by the 2016 analysis.	89
4.8	Neutral-Current 2D Distribution	90
4.9	Charged-Current 2D Distribution	90
4.10	Cosmic 2D Distribution	91
4.11	NC 2D distribution of the Distance from Closest Slice	92
4.12	CC 2D Distribution of the Distance from Closest Slice	92
4.13	Cosmic 2D distribution of the Distance from Closest Slice vs Clos- est Distance from Top	93
4.14	ND Energy spectra after applied containment cuts.	95
4.15	NC 2D distribution of the Distance from Closest Slice	96
4.16	A Simulated Air Shower	98
4.17	Secondary Cosmic Ray and the NOvA FD,	99
4.18	FD NC and Cosmic Ray Event Comparison	100
4.19	A simple Decision Tree Example	102
4.20	ND Data Vs MC CVN CosmicID.	109
4.21	FD Signal Vs Background CVN CosmicID.	109
4.22	ND Data Vs MC Simulated FD Number of Showers.	110
4.23	FD Signal vs Background Number of Showers.	110
4.24	ND Data Vs MC Simulated FD Leading Shower Y Direction.	111
4.25	FD Signal Vs Background Leading Shower Y Direction.	111
4.26	ND Data Vs MC Length of Shower.	112
4.27	FD Signal Vs Background Length of Shower	112

4.28	ND Data Vs MC Fraction of Transverse Momentum	113
4.29	FD Signal Vs Background Fraction of Transverse Momentum. . .	113
4.30	ND Data Vs MC Reconstructed Shower Feature	114
4.31	FD Signal Vs Background Reconstructed Shower Feature	114
4.32	ND Data Vs MC Shower Hits	115
4.33	FD Signal Vs Background Number of Hits	115
4.34	ND Data Vs MC XplusY	116
4.35	FD Signal Vs Background XplusY	116
4.36	ND Data Vs MC XminusY	117
4.37	FD Signal Vs Background XminusY	117
4.38	ND Data Vs MC XoverY	118
4.39	FD Signal Vs Background XoverY	118
4.40	ND Data Vs MC MIP	119
4.41	FD Signal Vs Background MIP	119
4.42	ND Data Vs MC Shower of Gap	120
4.43	FD Signal Vs Background Shower Gap	120
4.44	ND Data Vs MC Shower of Energy.	121
4.45	FD Signal Vs Background Shower of Energy.	121
4.46	TMVA-Based NC-dedicated Cosmic Rejection PID Distribution . .	126
4.47	CAF-Based NC-dedicated Cosmic Rejection PID Distribution . . .	126
4.48	ND Ratio CVN CosmicID.	128
4.49	FD Signal Vs Background CVN CosmicID.	128
4.50	ND Ratio Shower Length.	129
4.51	FD Signal Vs Background Leading Shower Length.	129
4.52	ND Ratio MIP Hits.	130
4.53	FD Signal Vs Background MIP Hits.	130
4.54	ND Ratio Number of Calibrated Hits.	131

4.55	FD Signal Vs Background Number of Calibrated Hit.	131
4.56	ND Ratio Number of Hits in the Leading Prong.	132
4.57	FD Signal Vs Background Number of Hits in the Leading Prongs.	132
4.58	ND Ratio X View Number of Hits in the Leading Prong.	133
4.59	FD Signal Vs Background X View Number of Hits in the Leading Prong.	133
4.60	ND Ratio Y View Number of Hits in the Leading Prong.	134
4.61	FD Signal Vs Background Y View Number of Hits in the Leading Prong.	134
4.62	ND Ratio Leading Prong Calibrated Energy.	135
4.63	FD Signal Vs Background Leading Prong Calibrated Energy.	135
4.64	ND Ratio Energy Weighted Summarized X View Hits.	136
4.65	FD Signal Vs Background Energy Weighted Summarized X View Hits.	136
4.66	ND Ratio Energy Weighted Summarized Y View Hits.	137
4.67	FD Signal Vs Background Energy Weighted Summarized Y View Hits.	137
4.68	ND Ratio Summarized X View Hits.	138
4.69	FD Signal Vs Background Summarized X View Hits.	138
4.70	ND Ratio Summarized Y View Hits.	139
4.71	FD Signal Vs Background Summarized Y View Hits.	139
4.72	ND Ratio FuzzyK Prong.	140
4.73	FD Signal Vs Background Number of Showers.	140
4.74	ND Ratio Leading Shower Y Direction.	141
4.75	FD Signal Vs Background Leading Shower Y Direction.	141
4.76	ND Ratio Maximum Y.	142
4.77	FD Signal Vs Background Maximum Y in the box.	142

4.78	ND Ratio Minimum Y.	143
4.79	FD Signal Vs Background Minimum Y in the box.	143
4.80	TMVA-Based NC/CC Classifier Response.	144
4.81	CAF-Based NC/CC Classifier Distribution	146
4.82	FD Standard Energy Spectra	150
4.83	FD 2017 Official Spectra	151
4.84	FD 2017 Re-Analysis Spectra	151
4.85	FD 2017 Re-Analysis Side Band Spectra	152
5.1	NOvA NC Disappearance Analysis flow chart	155
5.2	3+1 Unitary Mixing Matrix	157
5.3	3+1 Model Neutrino Mass Diagram	157
5.4	θ_{23} influence on NC Disappearance	159
5.5	θ_{24} influence on NC Disappearance	159
5.6	θ_{34} influence on NC Disappearance	160
5.7	δ_{24} influence on NC Disappearance	160
5.8	The CAF-based Prediction Chain	161
5.9	ND Data vs MC reconstructed energy spectra	163
5.10	ND 2D plot of comparison of reco energy to the true energy	164
5.11	Beam Transport Uncertainty on ND Signal	170
5.12	Beam Transport Uncertainty on ND Background	170
5.13	Beam Transport Uncertainty on FD Signal (direct selection)	171
5.14	Beam Transport Uncertainty on FD Background (direct selection)	171
5.15	Beam Transport Uncertainty on FD Signal (extrapolated)	172
5.16	Beam Transport Uncertainty on FD Background (extrapolated)	172
5.17	PPFX Uncertainty on ND Signal	174
5.18	PPFX Uncertainty on ND Background	174

5.19	PPFX Uncertainty on FD Signal (direct selection)	175
5.20	PPFX Uncertainty on FD Background (direct selection)	175
5.21	PPFX Uncertainty on FD Signal (extrapolated)	176
5.22	PPFX Uncertainty on FD Background (extrapolated)	176
5.23	GENIE Uncertainty on ND Signal	178
5.24	GENIE Uncertainty on ND Background	178
5.25	GENIE Uncertainty on FD Signal (direct selection)	179
5.26	GENIE Uncertainty on FD Background (direct selection)	179
5.27	GENIE Uncertainty on FD Signal (extrapolated)	180
5.28	GENIE Uncertainty on FD Background (extrapolated)	180
5.29	Absolute Calibration Uncertainty on ND Signal	182
5.30	Absolute Calibration Uncertainty on ND Background	182
5.31	Absolute Calibration Uncertainty on FD Signal (direct selection) .	183
5.32	Absolute Calibration Uncertainty on ND Background (direct selection)	183
5.33	Absolute Calibration Uncertainty on FD Signal (extrapolated) . .	185
5.34	Absolute Calibration Uncertainty on ND Background (extrapolated)	185
5.35	Relative Calibration Uncertainty on ND Signal	186
5.36	Relative Calibration Uncertainty on ND Background	186
5.37	Relative Calibration Uncertainty on FD Signal (direct selection) .	187
5.38	Relative Calibration Uncertainty on FD Background (direct selection)	187
5.39	Relative Calibration Uncertainty on FD Signal (extrapolated) . . .	188
5.40	Relative Calibration Uncertainty on FD Background (extrapolated)	188
5.41	Cherenkov Systematic Uncertainty on ND Signal	190
5.42	Cherenkov Systematic Uncertainty on ND Background	190

5.43 Cherenkov Systematic Uncertainty on FD Signal (direct selection)	191
5.44 Cherenkov Systematic Uncertainty on FD Background (direct selection)	191
5.45 Cherenkov Systematic Uncertainty on FD Signal (extrapolated) .	192
5.46 Cherenkov Systematic Uncertainty on FD Background (extrapolated)	192
5.47 ND MC Decomposition Uncertainty on ND Signal	194
5.48 ND MC Decomposition Uncertainty on ND Background	194
5.49 ND MC Decomposition Uncertainty on FD Signal (direct selection)	195
5.50 ND MC Decomposition Uncertainty on FD Background (direct selection)	195
5.51 ND MC Decomposition Uncertainty on FD Signal (extrapolated) .	196
5.52 ND MC Decomposition Uncertainty on FD Background (extrapolated)	196
5.53 Acceptance Uncertainty on ND Signal (One spectra)	198
5.54 Acceptance Uncertainty on ND Background (One spectra)	198
5.55 Acceptance Uncertainty on ND Signal (Two spectra)	199
5.56 Acceptance Uncertainty on ND Background (Two spectra)	199
5.57 Acceptance Uncertainty on ND Signal (Three spectra)	200
5.58 Acceptance Uncertainty on ND Background (Three spectra)	200
5.59 Acceptance Uncertainty on ND Signal (Four spectra)	201
5.60 Acceptance Uncertainty on ND Background (Four spectra)	201
5.61 Signal Uncertainty Summary	203
5.62 Background Uncertainty Summary	203
5.63 2017 Official FD Vertex XZ and YZ Distributions.	209
5.64 2017 Official FD Vertex XY Distribution.	210
5.65 2017 Official FD Energy Spectrum	210

5.66	NOvA 2017 Official 2D allowed regions	212
5.67	2017 Official vs Re-Analysis Selection Purity and Efficiency	215
5.68	2017 Official vs Re-Analysis ND Energy Spectrum	216
5.69	2017 Official vs Re-Analysis FD Energy Spectrum	217
5.70	2017 Official vs Re-Analysis 2D Sensitivity	218
5.71	Sideband Energy Spectra	219
5.72	2017 Re-Analysis FD Vertex XZ and YZ Distributions.	221
5.73	2017 Re-Analysis FD Vertex XY Distribution.	222
5.74	2017 Official and 2017 Re-Analysis FD Energy Spectrum	223
5.75	NOvA 2017 Re-Analysis 2D allowed regions	225
5.76	Comparison of 2D allowed regions for the 2017 Re-Analysis and the 2017 Official analysis	225
5.77	Comparison of 2D allowed regions for the 2017 Re-Analysis, the 2017 Official analysis, and the 2016 published analysis.	226
5.78	1D χ^2 Profiles for the 2017 Re-Analysis.	227
6.1	GAN Architecture	230
A.1	NC Disappearance δ_{24} Dependence	233
A.2	NC Disappearance δ_{24} Dependence	234
A.3	FD NC Energy Spectrum.	234
A.4	Energy Spectra with fixed δ_{24} ($0, \pi/2, \pi, 3\pi/2$).	235
A.5	Comparison of Energy Spectra with fixed δ_{24} ($0, \pi/2, \pi, 3\pi/2$).	235
A.6	2D Contour while profiling over δ_{24}	236
A.7	2D contours for fixed values of δ_{24} ($0, \pi/2, \pi, 3\pi/2$).	237
A.8	Comparison of 2D contours with fixed δ_{24} ($0, \pi/2, \pi, 3\pi/2$).	237
B.1	2017 Official Per-Period Selection Results	239

B.2	2017 Re-Analysis Per-Period Selection Results	239
B.3	Period One Selected Data Distribution XZ View	240
B.4	Period One Selected Data Distribution YZ View	240
B.5	Period One Selected Data Distribution XY View	240
B.6	Period Two Selected Data Distribution XZ View	241
B.7	Period Two Selected Data Distribution YZ View	241
B.8	Period Two Selected Data Distribution XY View	241
B.9	Period Three Selected Data Distribution XZ View	242
B.10	Period Three Selected Data Distribution YZ View	242
B.11	Period Three Selected Data Distribution XY View	242
B.12	Period Five Selected Data Distribution XZ View	243
B.13	Period Five Selected Data Distribution YZ View	243
B.14	Period Five Selected Data Distribution XY View	243

List of Tables

2.1	Fermilab long-baseline neutrino beams	35
2.2	Active Scintillating Material Components	41
4.1	Beam Quality Cuts.	78
4.2	Data Quality Cuts.	78
4.3	The number of FD events after the Analysis Mask cut is applied. .	79
4.4	FD-Selected Events after Applying Cosmic Veto	80
4.5	Event Quality Cuts.	81
4.6	Number of FD Events that pass the Event Quality Cuts	81
4.7	Number of ND Events that pass the Event Quality Cuts	81
4.8	Containment Cuts Applied to Events in the FD	83
4.9	FD Selected Events before and after applied Containment Cuts . .	84
4.10	Containment and Fiducial Volume Cuts Applied to events in the ND	86
4.11	Number of ND Events before and after applied Containment and Fiducial Volume cuts	87
4.12	FD Cosmic Rejection Pre-Selection Cuts	93
4.13	Number of FD events after applied Transverse Momentum Re- lated Cuts	94
4.14	NC/CC Separation Pre-Selection Cuts.	95
4.15	Number of FD Events after Applied Energy Cut.	95

4.16	Number of ND Events after applying all cuts.	96
4.17	TMVA Unspecific Method Separation Ability	122
4.18	TMVA Method Specific Separation Ability (Cosmic BDT).	125
4.19	TMVA Method Specific Separation Ability (NC/CC)	145
4.20	NC Selection PID Tuning for the 2017 Re-Analysis	149
4.21	FD Direct Selection Results.	150
5.1	Percentage difference between shifted and nominal predictions for number of extrapolated FD events due to beam transport sys- tematic effects.	169
5.2	PPFX Uncertainty On ND	174
5.3	PPFX on FD (direct selection)	175
5.4	PPFX on FD (extrapolated)	176
5.5	Three-Flavor oscillation parameter sets.	206
5.6	The R-Ratio for the 2017 Official Analysis	211
5.7	2017 Official NOvA 1D limits on θ_{24} and θ_{34}	213
5.8	Comparison of NC-selected events for 2017 Official and 2017 Re- Analysis	220
5.9	2017 Re-Analysis R-ratio	224
5.10	Comparison of 1D limits on sterile mixing parameters with results from global experiments	227
A.1	The oscillation parameter settings used in the δ_{24} dependence study.	233
B.1	Period Data Set Details	238

Chapter 1

Analysis Motivation and Research

Strategy

“The scientist only imposes two things, namely truth and sincerity, imposes them upon himself and upon other scientists.”

— Erwin Schrödinger

Our knowledge about neutrinos has changed dramatically in the past two decades. The three-flavor neutrino model has successfully explained a wide range of neutrino oscillation experiment results. On the other hand, the excess signal seen by the LSND and MiniBooNE experiments and the deficit of events observed at the GALLEX and SAGE experiments can be described by the existence of a right-handed chiral neutrino state. However, the above tantalizing results are not entirely convincing, as they are in tension with other results, from both short-baseline and long-baseline neutrino experiments. It has profound meanings for both particle physics and cosmology to resolve the issue of the existence of light sterile neutrinos. The physics reach of the light sterile neutrino searches will be limited by the performance of machine learning algorithms and computational

resources. Our task requires the identification of rare signals in extensive backgrounds. Applying the state-of-the-art machine learning algorithms in particle identification promises to provide improvements in our physics goal. In the following sections of this chapter, I will introduce the analysis motivation and employed research strategies.

Chapter Organization

- Section 1: Analysis Motivation
 1. Neutrino. A brief history of a unique particle
 2. Neutrinos in Standard Model and Beyond
 3. Appearance Anomaly
- Section 2: Research Strategy
 1. Accelerator neutrinos
 2. Neutrino Detection
 3. General Strategy
 4. Machine-Learning Based Improvement

1.1 Analysis Motivation

The story of neutrino physics starts from unexpected results from measurements of the energy spectrum of β decay in the early 20th Century. In the first section, I will go through a short overview of the theoretical background of light sterile neutrinos; why it is called “sterile neutrino” and why it is so often present in many theories describing phenomena beyond the Standard Model; and, finally, I will discuss how NOvA can look for sterile neutrinos.

1.1.1 Neutrino. A Brief History of a Unique Particle

As far as we know, the neutrino is a type of subatomic particle that cannot decay into further smaller constituents, and so is called a fundamental particle. Furthermore, unlike other fundamental particles, leptons and quarks, neutrinos are the only type of fermions that have neither electric nor color charges. Therefore, neutrinos can not feel the electromagnetic and strong interaction, but only gravity and the weak interaction. The weak interaction is aptly named. Its range is about 10^{-18} m, much less than an atomic nucleus diameter, for example 1.6 fm, or about 10^{-15} m for a proton in a hydrogen atom. On the other hand, neutrinos also have comparatively small masses, which are at least 6 orders of magnitude lighter than the masses of leptons and quarks. These observations make neutrinos almost impossible to detect. So, why would anyone have proposed the neutrino existence?

The Conundrum of the Continuous β Decay Spectrum

During the 1920s, nuclear physicists faced the dilemma of trusting energy conservation or not. This was caused by the unexpected continuous spectrum of β decay. β decay is a type of radioactive decay, in which an energetic electron or positron is emitted from an atomic nucleus. Therefore, the nucleus of atomic number Z is transformed from Z to $Z+1$ as a neutron is converted into a proton. Since the daughter atomic nucleus has slightly lower ground state energy than the parent atomic nucleus, the emitted electron or positron was expected to carry off the energy difference in kinetic energy form. However, the electron energy spectrum was measured to be continuous, and the electron to always carry off less energy than expected. Figure 1.1 shows a simple example of this process.

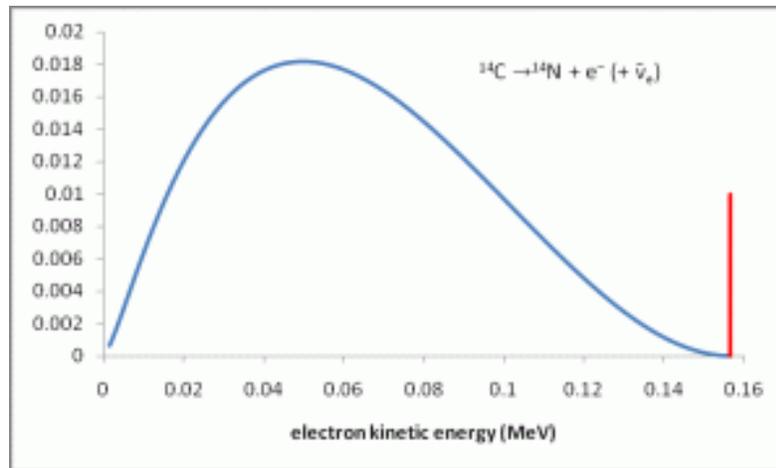


FIGURE 1.1: The decay ${}^6\text{C}^{14} \rightarrow {}^7\text{N}^{14} + e^{-}$ is a good example to illustrate the continuous electron spectrum from β decay [1]. The blue line depicts the measured energy spectrum, which is dramatically different from the red line representing the originally expected energy of the emitted electron.

As it was shown in the above example, the electron energy is a continuous distribution instead of all electrons having a fixed energy value. Therefore, the energy conservation principle seemed not to hold at the atomic level, and this apparent violation of a fundamental law of Physics was not explained until Wolfgang Pauli wrote down his short but famous letter.

Pauli's Proposal

All the stories about neutrino physics, and for sure, the analysis presented in this dissertation started with the famous Pauli letter [2]. In late 1930, to defend the law of energy-momentum conservation, Wolfgang Pauli hypothesized that there is a light neutral particle of spin 1/2 also emitted from the atomic nucleus alongside the electron. Therefore, it can carry off the missing energy. However, this particle could not be detected by any known experimental methods at the time. As one of the greatest minds in the golden age of physics, Pauli's concern about neutrinos being forever undetectable was almost right. Pauli originally

called his undetected neutral particle neutron. However, the name was used for the proton-like neutral hadron, discovered and named by James Chadwick in 1932. Enrico Fermi then renamed Pauli's neutral particle as neutrino, which means little neutral one in Italian [3].

The First Theory of β Decay

There remained the question of how the electron-neutrino pair was produced during a β decay process. By answering this question, Fermi made another fundamental contribution to neutrino physics (his first meaningful contribution being to give the neutrino a beautiful name):

$$n \longrightarrow p + e^{-} + \bar{\nu}_{\mu} \quad (1.1)$$

Let us follow Fermi's idea to write down his beautiful theory of the beta decay of nuclei [4]. The theory was built by an assumption that nuclei are bound states of neutrons and protons. Consider a simple quantum transition:

$$p \longrightarrow p + \gamma \quad (1.2)$$

The corresponding Hamiltonian has the form of the scalar product of the vector electromagnetic field and vector current:

$$\mathcal{H}^{EM}(x) = e\bar{p}(x)\gamma_{\alpha}p(x)A^{\alpha}(x) \quad (1.3)$$

By analogy with the above quantum transition electromagnetic Hamiltonian, Fermi hypothesized that the beta decay Hamiltonian was the scalar product of an electromagnetic vector current and a new vector current which was built by electron and neutrino fields, and their Hermitian conjugates:

$$\mathcal{H}^{\beta}(x) = G_F \bar{p}(x) \gamma^{\alpha} n(x) \bar{e}(x) \gamma_{\alpha} \nu(x) + h.c., \quad (1.4)$$

where G_F was named the Fermi constant.¹

Based on the effective four-fermions Hamiltonian, Fermi calculated the spectrum of emitted electrons.

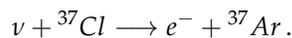
Happy to See You, Little Neutral One!

Soon after Fermi proposed his effective theory, Bethe and Peierls obtained the first estimate of the neutrino-nucleus cross-section [5]. Based on the result they obtained, Bethe and Peierls wrote down "there is no practically possible way of observing the neutrino". Indeed, after their paper, it was widely believed that the neutrino is an undetectable particle. Bruno Pontecorvo was the first physicist to challenge the opinion by proposing a radiochemical method of neutrino detection in 1946 [6]².

Between 1953 and 1959, Frederick Reines and Clyde Cowan performed a series of experiments whose results are the first proof of the existence of the neutrino [8]³. They detected antineutrinos produced in the Savannah River reactor

¹If we further compare the two Hamiltonians, we can find that the electromagnetic one is the Hamiltonian of the interaction of one boson field and two fermion fields, the beta decay one is the interaction of four-fermion fields. Therefore, the two constants, e and G_F , have different dimensions. Taking into account the charge, e , is a dimensionless quantity, the Fermi constant is then not, which means that the four-fermions Hamiltonian is an effective Hamiltonian.

²Pontecorvo suggested a Cl-Ar method, which is based on



This method was employed many years later to observe solar neutrinos in the first solar neutrino experiment [7].

³Reines and Cowan's pioneering experiments not only confirmed the Pauli-Fermi hypothesis but also proved the correctness of the V-A weak interaction theory by measuring the neutrino-nuclei cross-section.

from a β decay chain, by measuring the following process:

$$\bar{\nu} + p \longrightarrow e^+ + n \quad (1.5)$$

Their pioneering experiments also started the era of experiments using reactor neutrinos.

Different Types of Neutrino

Physicists had known four charged leptons, electron, muon, and their antiparticles, when they finally realized neutrinos are detectable.⁴ Pushed by the uncontrollable impulse of finding new particles, two immediate questions spread out among particle physicists following the first experimental measurement of the neutrino, in fact, the antineutrino. They are:

- There is one neutrino that can be produced in association with an electron. Should there be another distinguishable neutrino that can be produced in association with a μ in β decay?
- The positron is the antiparticle for the electron. Should there be another distinguishable neutrino which is the antiparticle of the detected one?

The first question was directly answered by π^+ decay. The two π^+ decay channels are:

$$\pi^+ \longrightarrow \mu^+ + \nu_\mu \quad (1.6)$$

$$\pi^+ \longrightarrow e^+ + \nu_e \quad (1.7)$$

The ratio of the decay widths R can be defined as :

⁴The third generation lepton, τ , was not predicted until 1971.

$$R = \frac{m_e^2(1 - \frac{m_e^2}{m_\pi^2})^2}{m_\mu^2(1 - \frac{m_\mu^2}{m_\pi^2})^2} \quad (1.8)$$

Plugging in values in the above equation, we find that R is about 1.2×10^{-4} , which means the second decay channel (Eq. 1.7) is strongly suppressed with respect to the first decay channel (Eq. 1.6). The Brookhaven neutrino experiment then produced a beam of high-energy neutrinos originating from decays of π^+ which are produced at accelerators⁵. The produced neutrinos then would be detected through one of these processes:

$$\nu + N \longrightarrow \mu^- + X \quad (1.9)$$

or

$$\nu + N \longrightarrow e^- + X \quad (1.10)$$

So, if the ν_e and ν_μ are the same particle, we expect to observe practically equal numbers of e and μ . However, in the Brookhaven experiment [9], 29 μ s and 6 electrons were observed.

The second question, in fact, was answered first through the inverse β decay process first proposed by Pontecorvo as introduced above:

$$\nu + {}^{37}\text{Cl} \longrightarrow e^- + {}^{37}\text{Ar} \quad (1.11)$$

Between 1955 – 1960 (just a little bit later than Reines and Cowan's series of

⁵This is also the experiment that started the era of experiments with accelerator neutrinos.

experiments), Ray Davis studied the above process by employing reactor antineutrinos from the Brookhaven reactor as the source, and carbon tetrachloride as the target. His results showed that the probability of Chlorine-37 to convert to Argon-37 reaction was no more than 10%. The reason is that, based on the law of lepton number conservation⁶, this reaction needs to involve a neutrino (not an antineutrino if they are not identical). Beta decay, on the other hand, needs to involve an antineutrino due to the process requiring an antilepton to balance the electron. Therefore, reactor neutrinos can only convert chlorine-37 to argon-37 when neutrinos and antineutrinos are identical. This great work not only proved an electron neutrino has its own distinguishable antiparticle but also proved that the weak interaction process obeys the law of lepton number conservation.

Neutrino and Discovery of the Neutral Current Weak Interaction

At the early 1970s, the Glashow-Weinberg-Salam theory was just thought as one of many possibilities to explain particle interactions, until the neutral-current (NC) neutrino interaction was discovered in Gargamelle, a liquid bubble chamber detector operating at CERN. The discovery of the NC interaction was the first proof that electromagnetic and weak interactions are a unified theory.

Neutrinos can interact via charged-current (CC) processes, for instance, The neutrino-quark CC interaction, by exchanging a W-boson, is represented by the inclusive process:

$$\nu_{\mu} + N \longrightarrow \mu^{-} + X \quad (1.12)$$

⁶The law of lepton number conservation says that the difference between the total number of leptons and the total number of antileptons is always a constant.

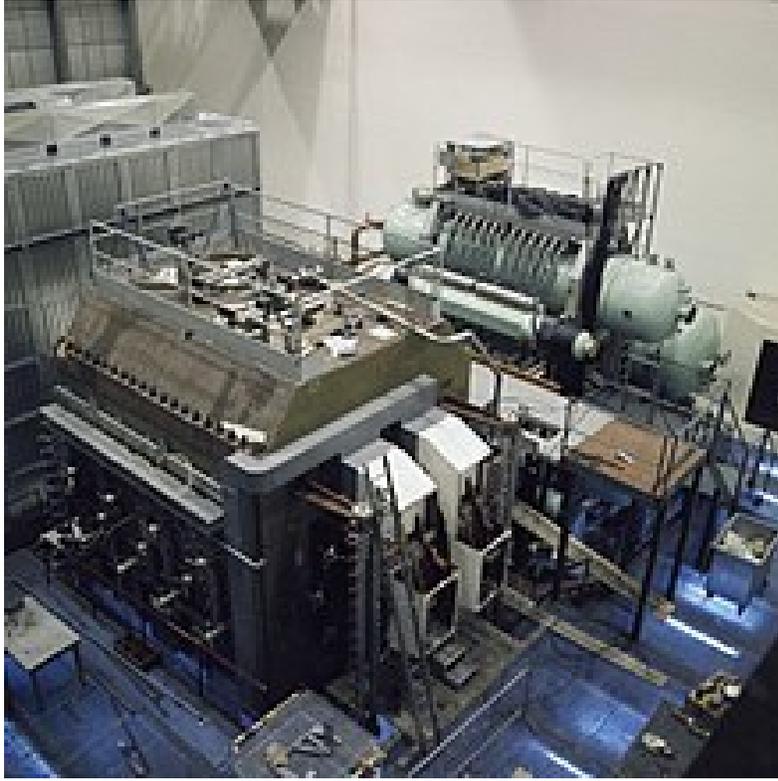


FIGURE 1.2: The Gargamelle heavy bubble chamber detector (4.8 m long, 2 m in diameter) is filled with 18 tonnes of liquid Freon. The Neutral-Current interaction was first observed in it in 1973 at CERN [10].

where X means any possible hadrons. The corresponding effective Hamiltonian is

$$\mathcal{H}^{CC} = \frac{G_F}{\sqrt{2}} 2\bar{\mu}_L \gamma^\alpha \bar{\nu}_{\mu L} j_\alpha^{CC} + h.c. \quad (1.13)$$

If in addition, neutrino and quarks can also interact via NC by exchanging a Z -boson, such as:

$$\nu_\mu + N \longrightarrow \nu_\mu + X \quad (1.14)$$

We then can find only hadronic final states. The corresponding effective Hamiltonian is

$$\mathcal{H}^{\text{NC}} = \frac{G_F}{\sqrt{2}} 2\bar{\nu}_{\mu L} \gamma^\alpha \nu_{\mu L} j_\alpha^{\text{NC}} + h.c. \quad (1.15)$$

Comparing the above two effective Hamiltonians, we see these interactions are both characterized by the Fermi constant. Therefore, the cross-section of these two processes is comparable.

The first NC event was observed at the beginning of 1973, and motivated exhaustive searches for hadronic NC-induced interactions.

What They Knew!

- In 1954, Zhengning Yang and Robert Mills extended the gauge theory for abelian groups to non-abelian groups to explain strong interactions.
- In the late 1950s, the symmetry breaking concept was proposed in superconductivity studies.
- In 1960, Yoichiro Nambu discussed the application of symmetry breaking in particle physics.
- In 1961, Sheldon Glashow ‘merged’ the weak force and electromagnetic interaction.
- In 1964, three independent groups proposed the mass generation theory without "breaking" gauge theory.⁷
- In 1967, Abdus Salam and Steven Weinberg incorporated the Higgs mechanism into the electroweak interaction, to give rise to the masses of electroweak interaction fundamental particles. The Glashow-Weinberg-Salam

⁷The three independent groups are: 1) Peter Higgs; 2) Gerald Guralnik, C. R. Hagen, and Tom Kibble; 3) Robert Brout and François Englert

model was widely accepted from the moment the weak neutral-current interaction mediated by Z boson exchange were observed.

- In 1973, asymptotic freedom was proposed by two independent groups.⁸
- In 1983, the charged and neutral bosons of the weak interaction were discovered experimentally.
- In 1995, the top quark was measured in CDF and DØ.
- In 2000, the tau neutrino was observed by DONUT [11].
- In 2012, the Higgs Boson was finally discovered at the LHC.

After ‘merging’ the above theoretical proposals and experimental discoveries, physicists around the world developed the Standard Model of Particles in stages. The theory can explain three of the four known fundamental interactions and has had great successes in producing experimental predictions.

The Standard Model seemed to be the FINAL THEORY physicists were looking for, BUT ...

1.1.2 Neutrinos in the Standard Model and Beyond

Over the past 45 years, the Standard Model of Particle has been a hugely successful theory. It provides an excellent description of almost all of the phenomena in particle physics. The only sector of this effective model which can not stand up to experimental examination is its assumption of massless neutrinos. The atmospheric neutrino oscillation discovery, SuperK [12] and SNO [13], is one of few

⁸The two groups are: 1) David Politzer; 2) David Gross and Frank Wilczek. They proposed the theory in the same year.

significant recent discoveries in particle physics. The oscillation phenomenon was then also proved by solar experiments (Homestake [7], Gallex [14], SAGE[14]), and the reactor experiment KamLAND [15]. The experimental discovery of the neutrino flavor-changing phenomenon not only proves the neutrinos are not massless particles, but also explained the long-standing solar neutrino problem. The phenomenon caused immediate great experimental and theoretical interest, though Bruno Pontecorvo has first proposed it in 1957. The precise measurement of neutrino oscillation can shed light on several fundamental properties. A couple of neutrino experiments, based on the different type of neutrino sources, have been built to contribute to this field⁹. We begin this section by shortly introducing the Standard Model and its building pieces. We then discuss neutrino properties in the Standard Model circa 1970s and explain how neutrino physics has developed into its current form, driven by both experimental discoveries and theoretical motivation for neutrino mass generation in the last decades.

The Standard Model of Particle Physics

The Standard Model (SM), our current best theory, is a type of periodic table of the particle physics. Instead of listing the chemical elements, the SM lists the two types of the fundamental particles, who can not be broken down into smaller particles, and make up the atoms. The two kinds of matter particles are: 1) hadrons, which can interact by the strong force, 2) leptons, which can not feel the strong force. The most well-known SM particle is the electron, which has a negative electrical charge and can be detected by electromagnetic interaction.

⁹The following experiments dominate the precise measurement of neutrino oscillations in the last ten years. They are: 1) BOREXINO [16], a solar neutrino experiment; 2) MINOS [17], K2K [18], and T2K [19], long-baseline accelerator neutrino experiments; and 3) Daya Bay [20], RENO [21] and Double Chooz [22], reactor neutrino experiments.

On the other hand, neutrinos are the most mysterious piece, due to the fact that they do not carry electrical charge.

Standard Model Particles and Their Interactions

The building pieces of the SM are shown in Fig. 1.3. The SM particles can also be classified into fermions, who are spin-1/2 particles, and bosons which are integer-spin (0 or 1) particles. The total 12 fermions are the fundamental constituents of matter, and they can be further classified into two groups, leptons and quarks. Each fermion group consists of three generations, which can be identified by their masses.

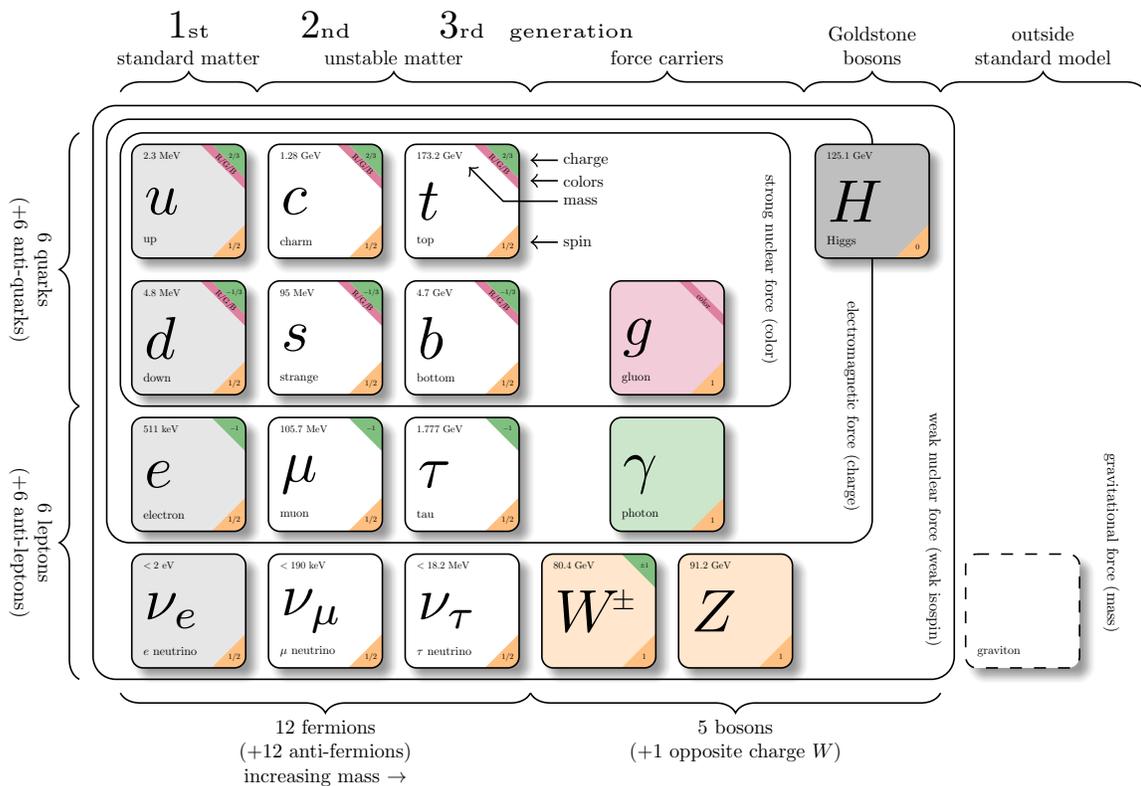


FIGURE 1.3: The Standard Model Particles. The plot was produced by David Galbraith and Carsten Burgard at the CERN Webfest 2012.

Besides fermions, there are four types of bosons, whose spin is 1, which are the mediators of the strong (eight gluons), weak (Z^0 and W^\pm), and electromagnetic (γ) interactions, respectively. The Higgs boson (spin 0) is the last discovered SM particle which is predicted by the electroweak theory, and is responsible for giving massive particles, including itself, their masses.

Standard Model Interactions (Forces Mediated by Gauge Bosons)

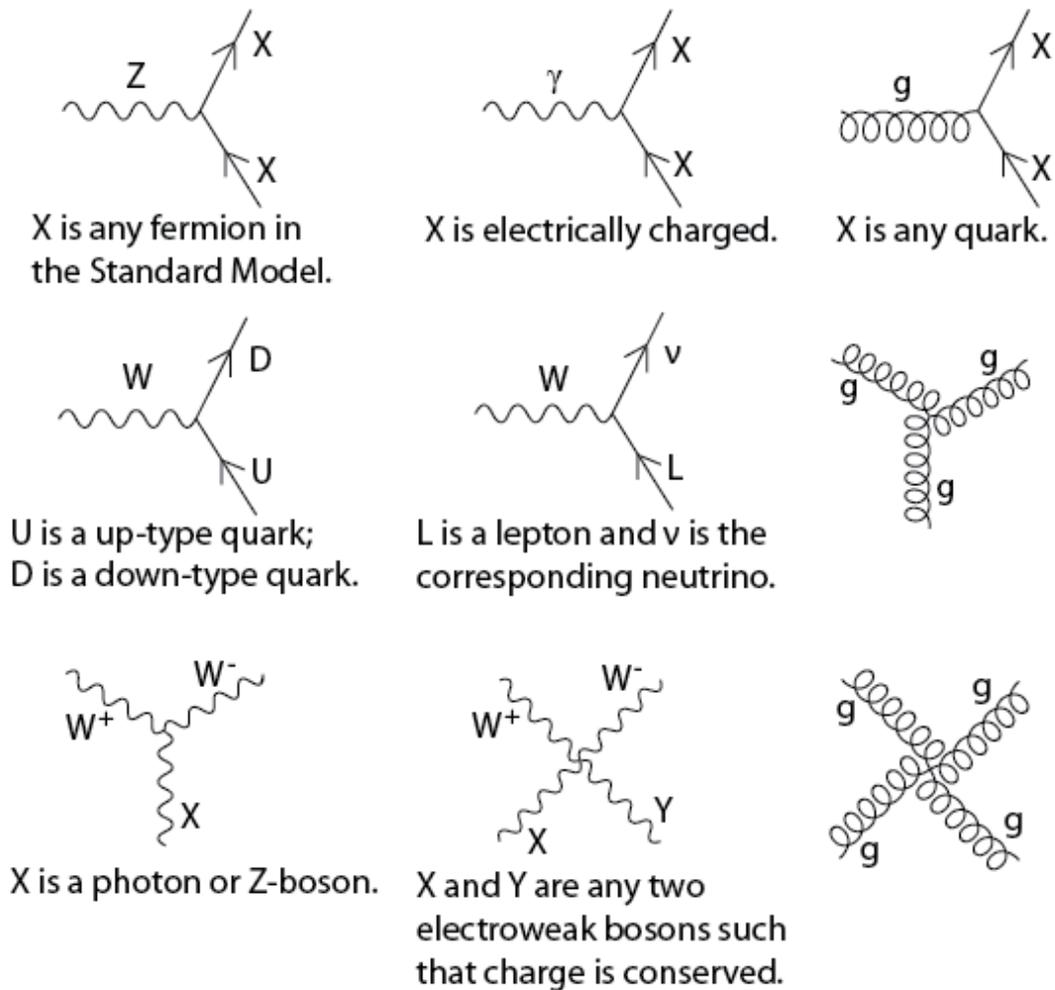


FIGURE 1.4: Standard Model Interaction Feynman diagrams. These interactions form the basis of the Standard Model Interactions [23].

The three types of SM fundamental interactions are listed in Fig.1.4.

- Weak Interaction: almost all SM particles (Z^0 and W^\pm bosons and all fermions) carry the weak charge, and, therefore, experience the weak interaction, whose interaction range is about 10^{-18} meters, which is approximately 0.1% of the diameter of a proton.
- Strong Interaction: only quarks and gluons carry the color charge, and therefore only they can experience the strong interaction, whose interaction range is about 10^{-15} meters, which is approximately the diameter of a medium-sized nucleus.
- Electromagnetic interaction: is experienced by SM particles which have a non-zero electric charge, is a force of infinite range which follows the inverse-square law, and therefore can hold atoms together.

The Only Massless Matter Particles in the Standard Model

Before physicists finally realized the parity-violating nature of weak interaction and its relation with neutrino mass, it was firmly believed that neutrinos must have zero mass due to it being interwoven into the weak interactions. The other massless particles, gluons and photon, are force carriers, not matter particles.

In the SM Lagrangian, the only way to construct a Dirac mass term for fermions is :

$$- \mathcal{L}_{mass} = m(\bar{\psi}_L \psi_R + \psi_L \bar{\psi}_R) \quad (1.16)$$

Due to the absence of the right-hand neutrino and the left-hand antineutrino, the Dirac mass term could not be constructed in the SM framework.

Neutrino Oscillation as a Window to New Physics

Bruno Pontecorvo first introduced the idea of neutrino oscillation in 1957. Neutrino oscillation evidence was accumulated from the observation of solar and atmospheric experiments over the years¹⁰. The flavor-changing phenomenon is a direct proof of massive neutrinos, which are not included in the SM.

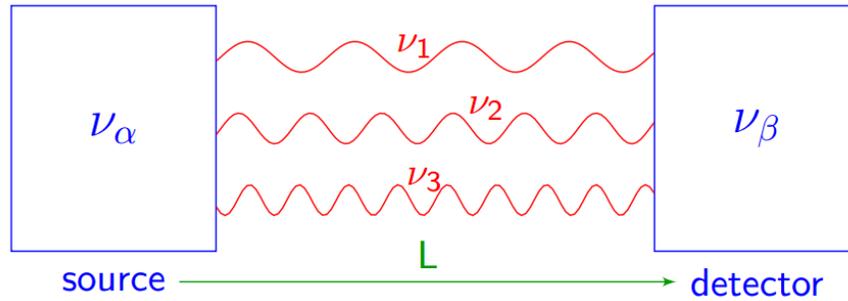


FIGURE 1.5: Three-Flavor Neutrino Model Oscillation. The neutrino flavor changes after it travels for some distance. Plot taken from Stefano Gariazzo's talk.

1.1.3 The $\bar{\nu}_e$ Appearance Anomaly

While the three-flavor neutrino model is well established, there are several oscillation results that cannot be interpreted based on it. However, they could be explained by the introduction of extra neutrinos [24]. The first experimental result supporting the possibility of extra neutrino species came from the Liquid Scintillator Neutrino Detector (LSND) experiment [25].

The LSND Signal

LSND [26], is a single-detector experiment, designed for two purposes: 1) search for $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$ oscillation; 2) measure neutrino-nuclear cross-sections. The $\bar{\nu}_\mu$ travels 30 m to arrive at the detector.

¹⁰The SNO and SuperK results play a decisive role in neutrino oscillation discovery.

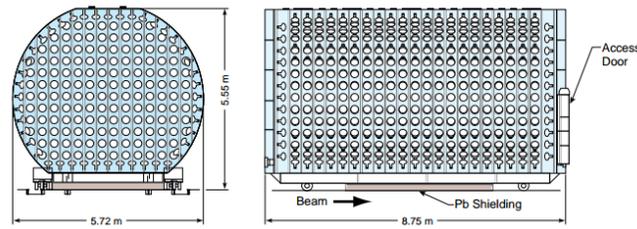


FIGURE 1.6: A schematic drawing of the LSND detector. Plot taken from the LSND public plots webpage [26].

The LSND released result, displayed in Fig. 1.7, shows a total excess of beam events with final-state electrons of $87.9 \pm 22.4 \pm 6.0$ ν_e candidates events with a background estimated at 30 ± 6.0 events. The simplest interpretation of this excess in terms of neutrino oscillations requires a new mass splitting of about 1 eV^2 , three orders of magnitude larger than the atmospheric mass splitting Δm_{atm}^2 , which would be the largest neutrino mass splitting measured to date.

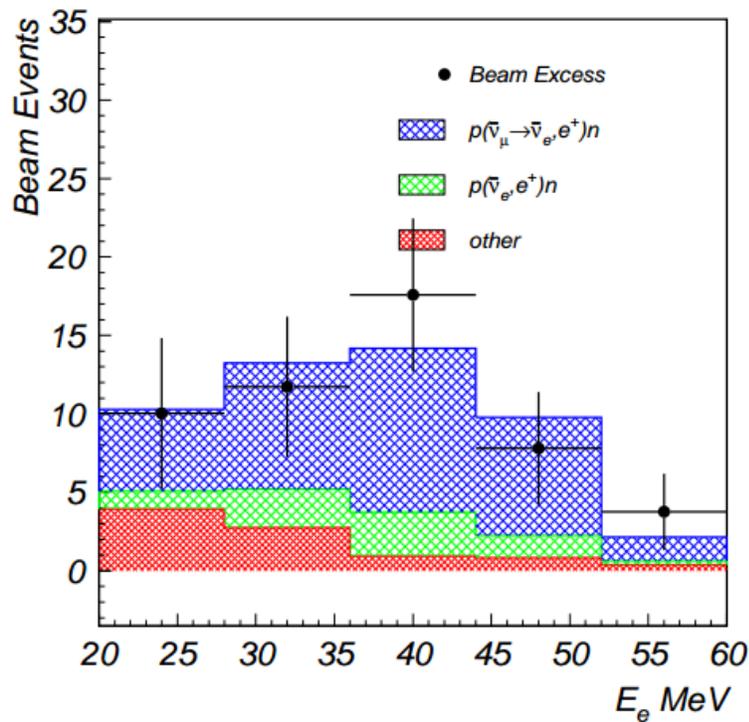


FIGURE 1.7: The LSND energy distribution for events. Plot taken from LSND public plots webpage [26].

The KARMEN Constraint

The KARMEN experiment has a very similar experimental setup to LSND, and was designed to check the puzzling LSND results. KARMEN measured 15 events which survived the analysis selection cuts. The KARMEN results rule out a significant portion of the LSND oscillation allowed region with $\Delta m^2 > 10eV^2$, as shown in Fig. 1.8.

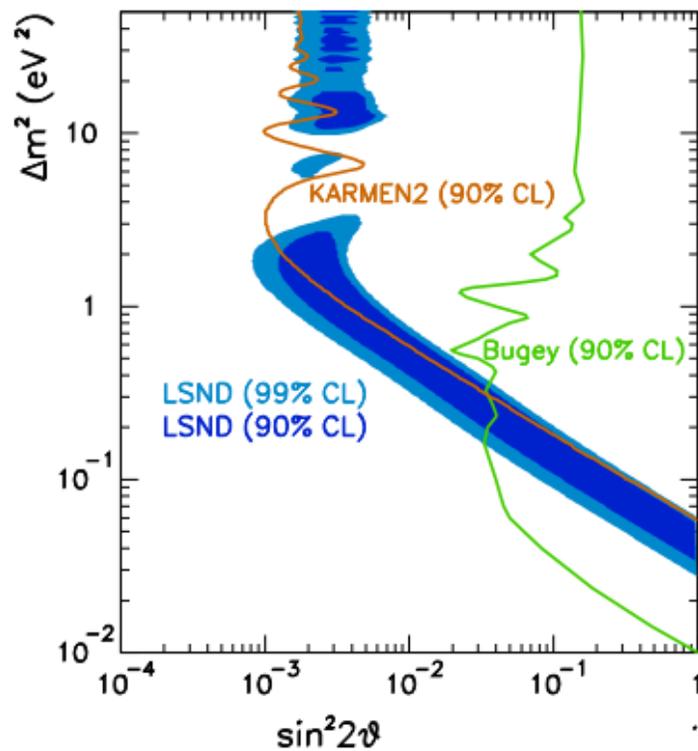


FIGURE 1.8: The $(\sin^2 2\theta, \Delta m^2)$ oscillation parameter fit for the LSND, Bugey, and KARMEN [27].

The MiniBooNE ν_e and $\bar{\nu}_e$ Appearance Searches

The MiniBooNE experiment, which was another experiment designed to test LSND, has reported oscillation results in both neutrino and antineutrino mode. Both results show a low-energy excess in a region not directly compatible with

LSND, but the most recent paper claims to be consistent with and verify the LSND signal [28].

1.1.4 What We Know?

We know there is no mass generation mechanism inside the SM; we know the LSND appearance anomaly cannot be explained within the three-flavor neutrino model, but it can be explained by an additional light sterile neutrino; and we know that a right-hand neutrino can fulfill the SM mass generation mechanism, but it should be much heavier than the one needed to explain the LSND results.

1.2 Research Strategy

In this section, I will go through the detailed ideas and methods which we employed for the NOvA NC Disappearance Analyses.

1.2.1 Accelerator neutrinos

The muon neutrino/anti-neutrino energy spectrum produced through particle accelerators is rather broad, as shown in Fig1.9. The broad energy spectrum can be significantly changed if the neutrino beam is carefully offset slightly from the direction of the Far Detector, as shown later in this thesis.

1.2.2 Neutrino Detection

Neutrinos interact with other fermions, inside the SM, by exchanging a W^\pm boson (charged-current, or CC interaction), or a Z^0 boson (neutral-current, or NC interaction). Moreover, the coupling of the three-generation leptons to the W^\pm

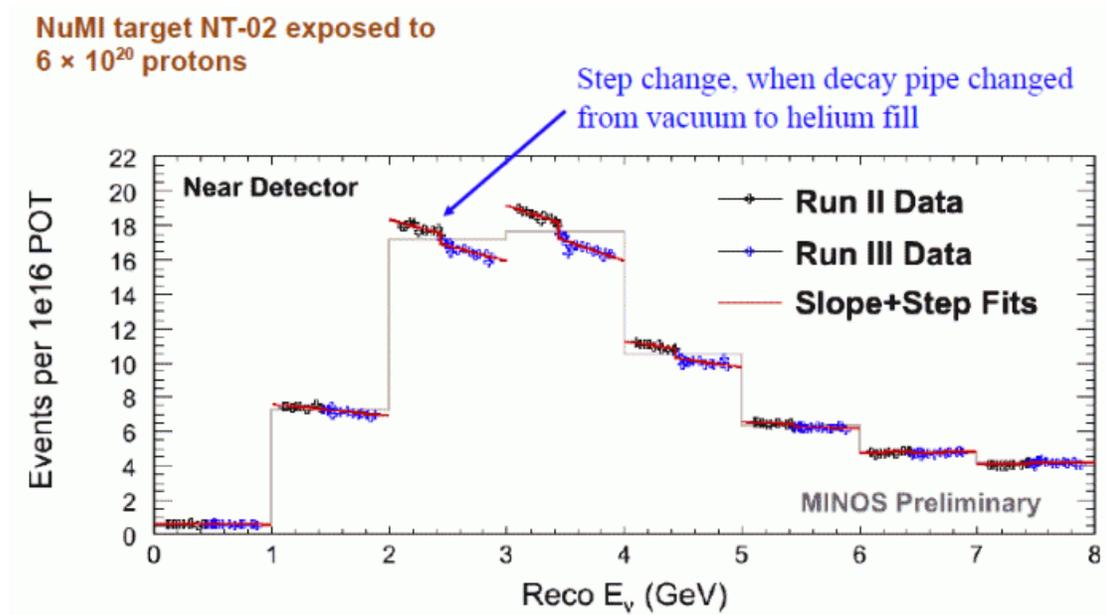


FIGURE 1.9: The reconstructed neutrino energy spectrum for the MINOS beam. Picture taken from the MINOS public plots webpage.

takes place strictly within a particular generation. Therefore, we can only measure the presence of a neutrino in our experiment if it interacts.

Final States of Neutrino Interactions:

- Charged-Current Interaction: the mother neutrino converts into the equivalent charged lepton and the charged lepton is detected in the final state;
- Neutral-Current Interaction : the neutrino remains a neutrino, but transfers energy and momentum to the particles it interacted with.

1.2.3 General Strategy

The following flow-chart explains the general steps in a long-baseline neutrino oscillation analysis.

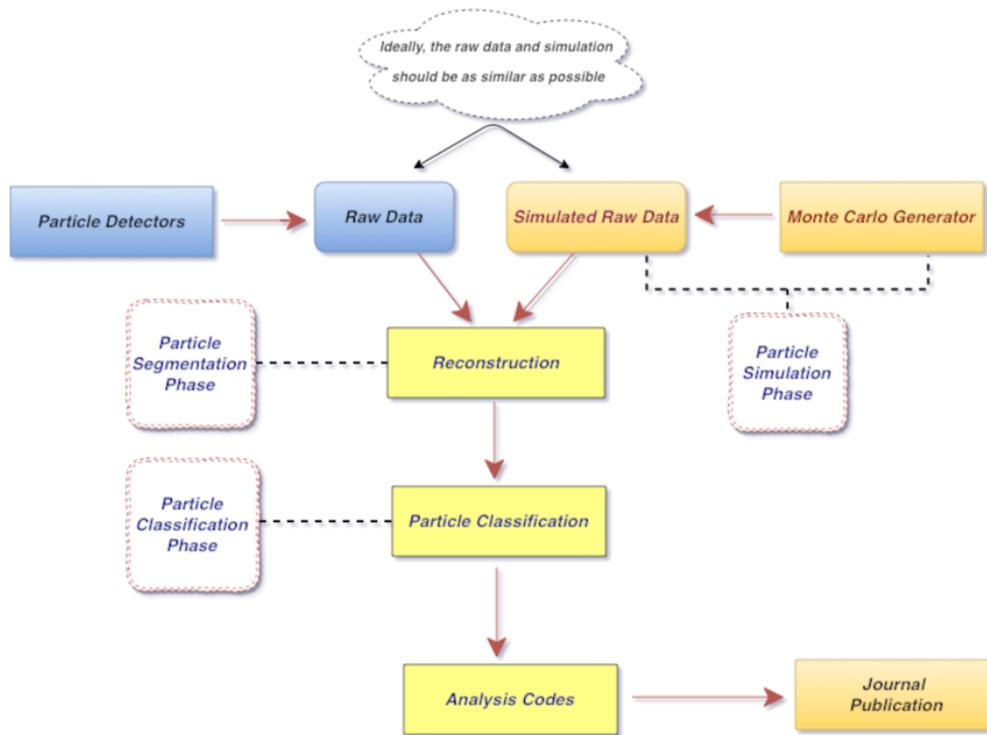


FIGURE 1.10: General Analysis Chain

1.2.4 Machine-Learning Based Improvements

While the analysis chain described above has been successful in producing results [29], it is prone to multiple shortcomings: errors in the reconstruction of high-level features from the raw data may lead to incorrect categorization of the particle interactions; and the features used to characterize the interactions are limited to those which have already been devised and implemented for the experiment. The core problems which physicists are facing share many similarities with the problems confronted in machine learning. Therefore, we decided to employ state-of-the-art machine learning algorithms to help us achieve better analysis results.

Chapter 2

The NOvA Experiment

“I have done something very bad today by proposing a particle that cannot be detected, it is something no theorist should ever do.”

— Wolfgang Pauli

Neutrinos hardly interact with other particles, therefore, we need an intense beam of neutrinos and to deliver it to one or more large enough detector(s) in order to study Neutrino Physics, such as flavor states oscillation and the mass generation mechanism. The Neutrinos at the Main Injector (NuMI) beam provides a high-intensity neutrino beam for long-baseline neutrino oscillation projects. One of the most recent efforts, the NuMI Off-Axis ν_e Appearance (NOvA) experiment, makes use of the NuMI beam to measure ν_μ to ν_e , $\bar{\nu}_\mu$ to $\bar{\nu}_e$, and active-sterile neutrino oscillation by comparing the observed Charged-Current (CC) and Neutral-Current (NC) neutrino-nuclear interactions in two detectors separated by 810 km, as shown in Figure 2.1.

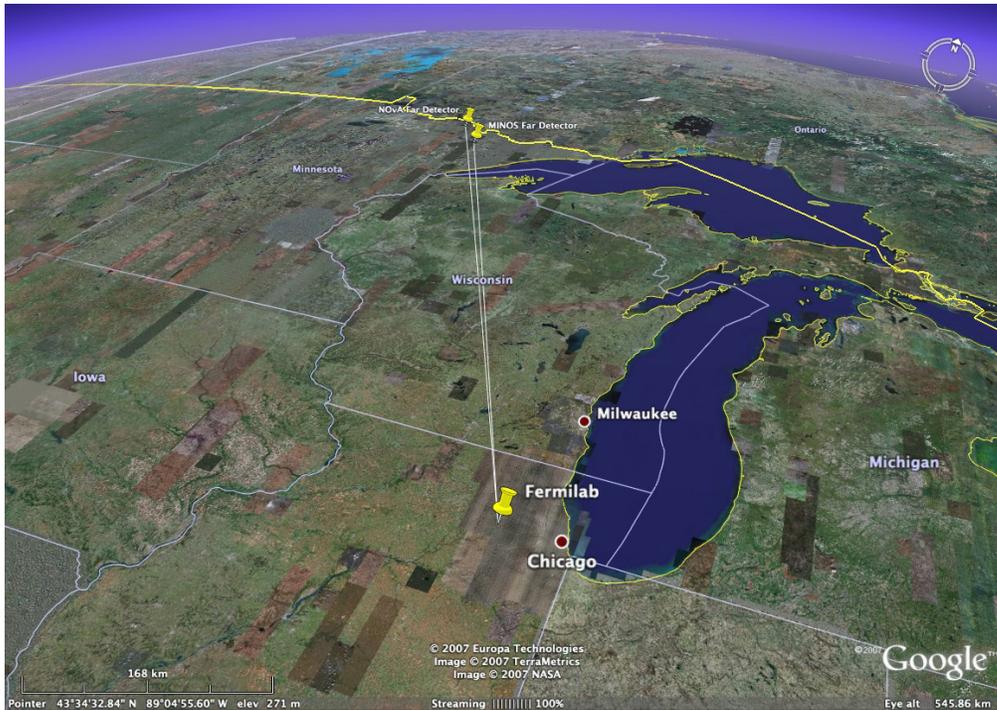


FIGURE 2.1: NOvA detectors location. The MINOS Far detector is on the axis of the NuMI beam, while the NOvA Far Detector is 14 mrad off-axis from the NuMI beam.

The NOvA detectors are functionally equivalent and are both situated at the same off-axis angle from the NuMI beamline to match the beam neutrinos' energy peak, about 2 GeV, to the maximum of the ν_μ to ν_e oscillation probability. The 300-ton Near Detector (ND) is placed approximately 1.05 km downstream of the beam target source, which is located at Fermi National Accelerator Laboratory (Fermilab), and is 100 m beneath the surface. The short distance between the neutrino source and the ND means neutrino oscillation is negligible in the ND, which enables us to understand the initial flavor eigenstates of the neutrinos from the beam and their energy spectra. Meanwhile, the deep underground location helps shield the NOvA Near Detector from cosmic rays, which are the main background noise of our signal in the Far Detector (FD). The 14,000-ton

FD is situated about 810 km away from the neutrino source, at Ash River, Minnesota. Also, it is on the surface, about 300 m above sea level, therefore it is exposed to an abundant rate of secondary cosmic rays. The functionally equivalent detectors enable large cancellation of systematic uncertainties, such as the ones from neutrino beam flux, and neutrino-nuclear cross section uncertainties. Also, neutrino and anti-neutrino beams can be generated by running the NuMI in different modes. Due to matter effects, impacting neutrinos as they go through the earth's crust, the oscillation probabilities are different for neutrinos and anti-neutrinos. Then, by comparing neutrino energy spectra observed in the FD with the prediction from the ND measurement, we can extract or place limits on the mixing parameters.

Chapter Organization

- Section 1: The NuMI Beam
- Section 2: NOvA Detectors

2.1 The NuMI Beam

The NuMI beam is a muon neutrino source, which was constructed in 2005 and designed initially to provide neutrinos for the Main Injector Neutrino Oscillation Search (MINOS) experiment, which was an on-axis long baseline neutrino project. Protons from the Fermilab's Main Injector (MI) are fired into a graphite target shown in Figure 2.2, to create the muon neutrino beam.

The collision of protons with the target produces a variety of fundamental particles, including pions and kaons. These particles are focused into the direction



FIGURE 2.2: One of the NuMI horns.

which we want the neutrinos to travel, by parabolic magnetic horns. After pi-ons and kaons decay into neutrinos and leptons, mostly muon neutrinos and muons, the neutrinos will continue on the same path which their mother particles were traveling. The NuMI beam travels in the direction of the MINOS Far Detector situated in the Soudan mine in Minnesota, which means that neutrinos start to travel from about 50 meters underground and are aimed downward at a 3.3° angle. They will go through the Earth's crust, up to 10 kms below the surface of the earth.

During NuMI's early years of operation, the Fermilab accelerator complex delivered about 350 kW of power to the NuMI beam. The system was then upgraded from 350 kW to 700 kW of power to achieve the physics goals of several latest neutrino projects, such as NOvA, MINOS+, and MINERVA. In the future, the

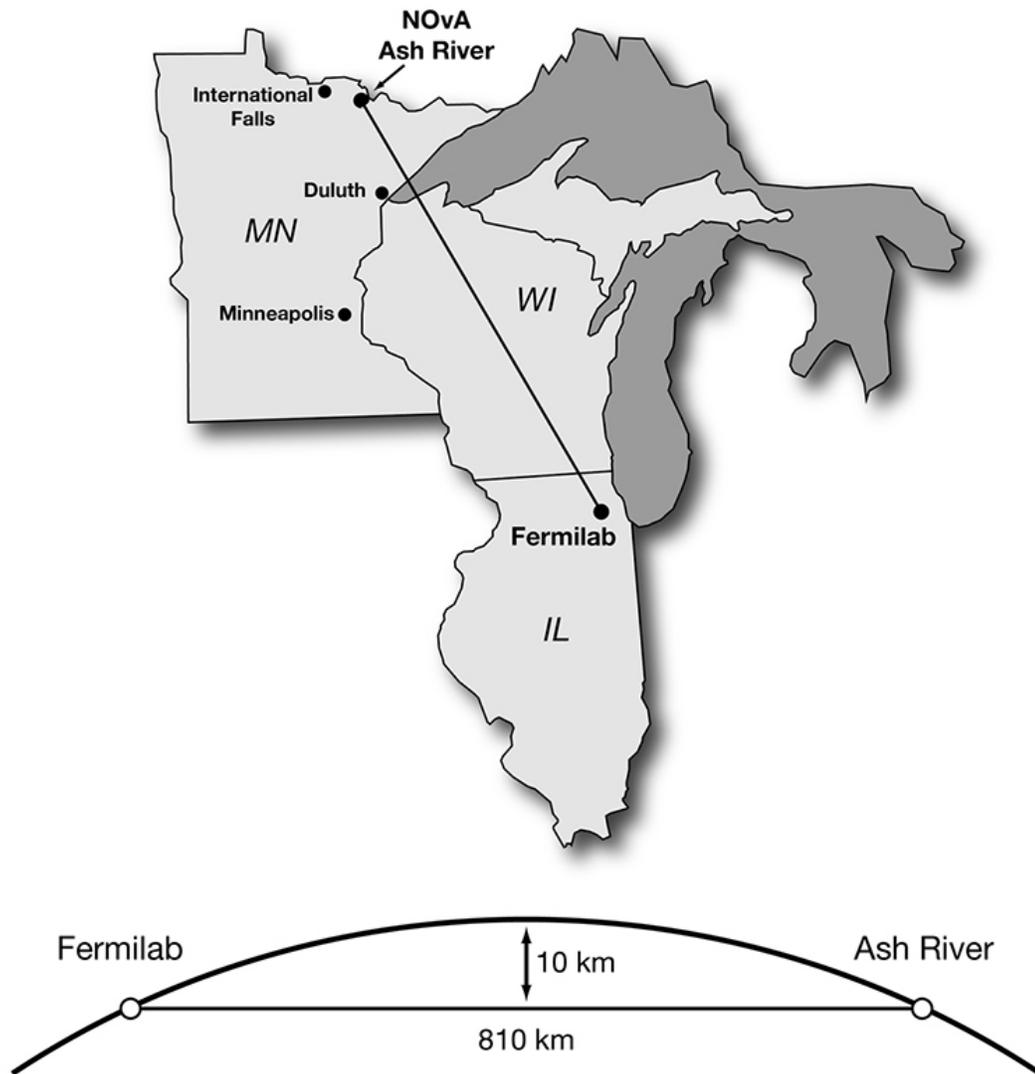


FIGURE 2.3: NuMI Beamline Path

accelerator complex will be capable of delivering 900 kW of power to the NuMI beam for the NOvA project. Furthermore, the NuMI beam can be converted into anti-neutrino beam, basically $\bar{\nu}_\mu$, by reversing the current in the horns, which is called Reverse Horn Current mode (RHC). On the other hand, the neutrino mode of the beam is called Forward Horn Current mode (FHC). The entire data set used in this thesis was taken in FHC mode.

2.1.1 Neutrino-Generating Process

The process of neutrino production starts initially from producing and accelerating negatively charged hydrogen ions (H^-).

Fermilab Accelerator Complex

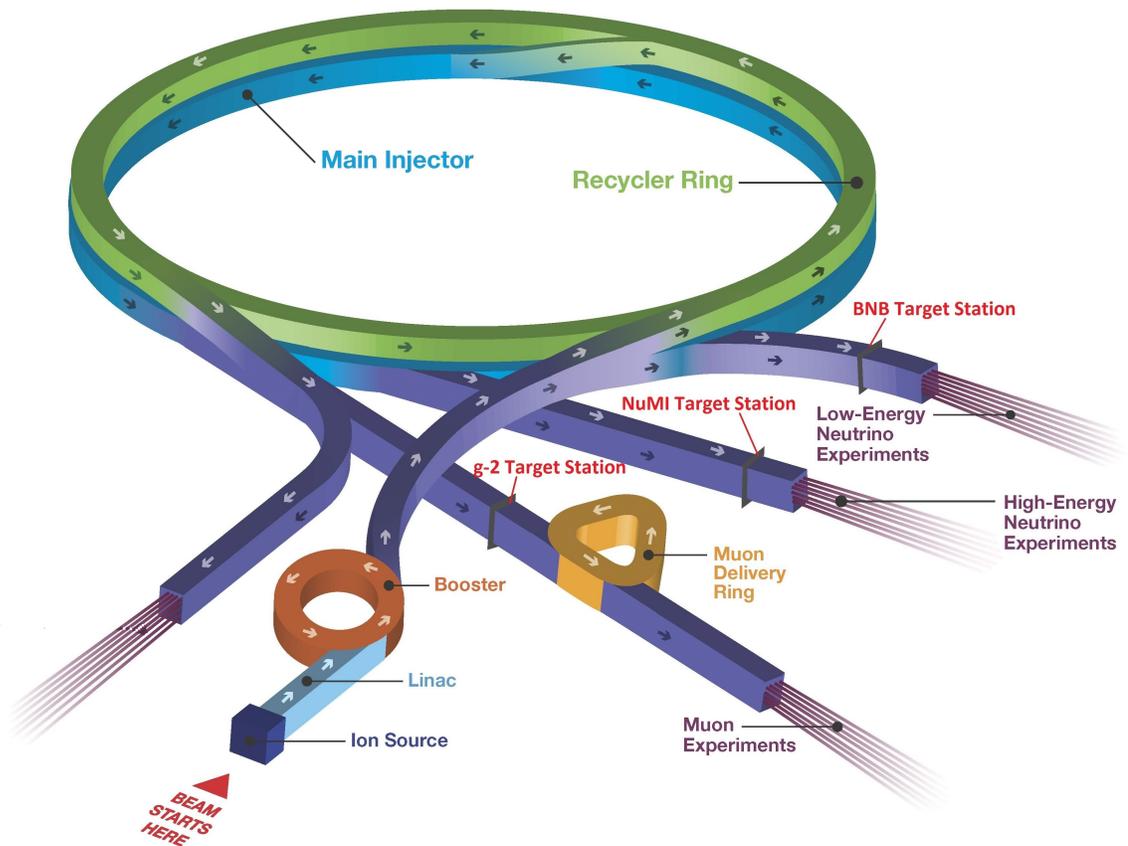


FIGURE 2.4: Fermilab’s accelerator complex delivers the world’s most powerful high-intensity neutrino beam for a broad range of new and existing experiments.

The hydrogen ions from a gaseous hydrogen source are accelerated to 400 MeV from 750 keV in a linear accelerator (Linac) before being fed into a circular synchrotron, the Booster. In the Booster, electrons are stripped. Meanwhile, the surviving protons continue to be accelerated to 8 GeV in the 75.47-meter radius

synchrotron. The beam in the Booster is bunched at 53 MHz to produce batches of protons, 4×10^{12} protons per batch. These batches of protons are then transferred into another larger circular synchrotron with 528.30-meter radius, called the Main Injector (MI), with the injection energy of 8 GeV. The protons are accelerated to a higher energy, 120 GeV, in the Main Injector. When the protons were extracted and accelerated in the Main Injector, the beam was bent by the magnets horizontally downward at an angle of 58 milliradians, about 3.3 degrees, to account for the curvature of the Earth's surface when directed towards the on-axis MINOS Far Detector located at a depth of 716 meters. The proton beam spills from the Main Injector are then transported along a carrier tunnel and collide with a target with a density of about 1.78 g/cm^3 to produce a shower of hadrons via strong interactions. Most of the daughter hadrons are kaons and pions. The target system was designed to maximize the muon neutrino producing rate to provide a high-intensity neutrino beam to the downstream experiments.

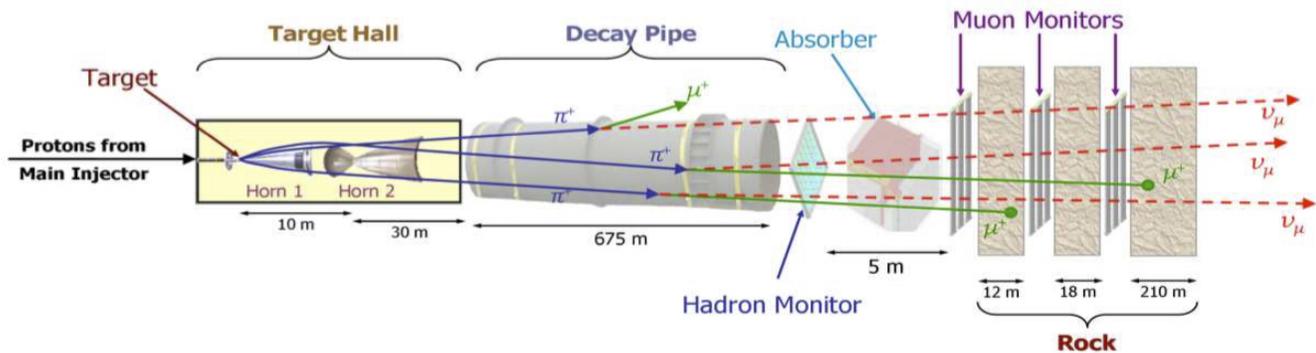


FIGURE 2.5: The Components of the NuMI beamline

This means it should be long enough to allow most of the beam protons to interact with nuclei in the target, on the other hand, it should be thin and narrow in case re-absorption happens so that the daughter mesons can easily escape from the sides. The target, as shown in Figure 2.6, used during the NOvA era consists of twelve 100 mm long and 6.4 mm thick graphite plates. The charged hadronic

daughter particles (primarily π^\pm and K^\pm) then are focused into a collimated beam by the magnetic field produced by two downstream horns.

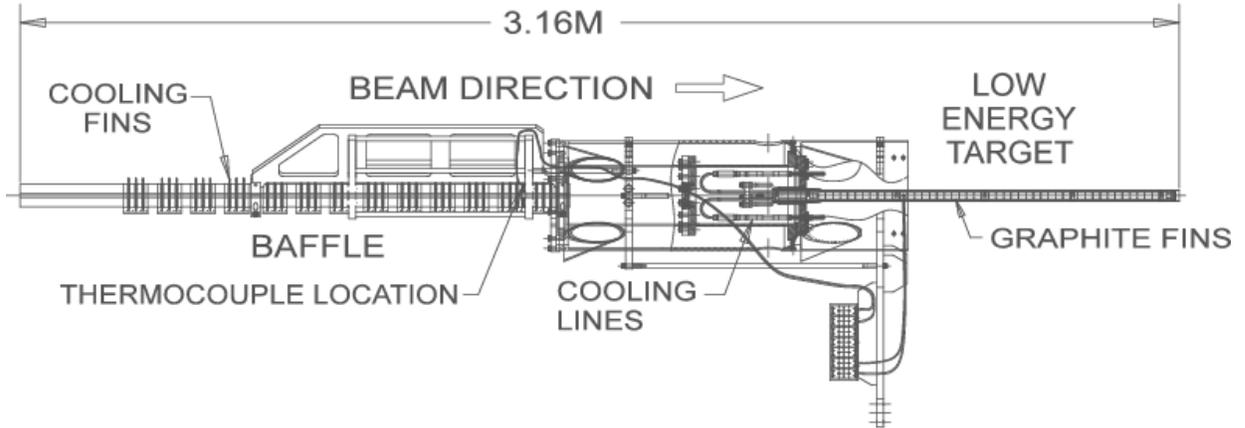
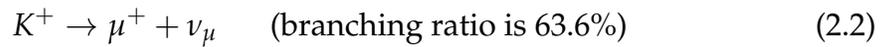
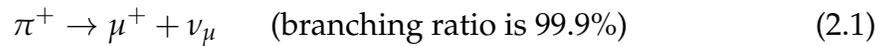
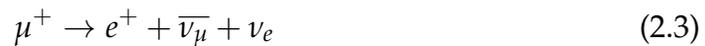


FIGURE 2.6: The Components of the NuMI Target

After being focused by the horns, the charged mesons are directed to a 675 m long decay pipe. The specific pipe length was chosen due to the rough average distance a 10 GeV pion could travel before decaying. Also, the 2 m diameter decay pipe volume was filled with Helium gas at 0.9 atm pressure since December 2007. There, the pions and kaons decay into secondary particles and neutrinos via the primarily channels as below:



Followed by the daughter muon decay:



Based on the above decay modes, we can find that there are intrinsic electron neutrino and muon anti-neutrino components in the NuMI beam, though their

estimated contamination is less than 1% and 2% separately. Also, to acquire a pure neutrino beam, the beam was designed to pass through an absorber, made of water-cooled steel and aluminum, and 240 meters of rock to remove all remaining muons and hadrons. After transferring through this set of apparatuses, the beam heads a little further away to the NOvA Near Detector, and then travel through the earth's crust to the NOvA Far Detector.

2.1.2 Horn Configurations

As mentioned above, the horn configuration has a significant impact on the neutrino energy spectrum as well as the flux component. This section will describe in detail the horn system elements and function. As shown in Figure 2.7, the pair of horns comprise cylindrical-shaped shell conductors and parabolic-shaped inner conductors that generate magnetic fields which function like a lens with the focal length proportional to the hadrons momentum. The two electromagnetic focusing horns, operated in a pulsed mode, can produce a peak strength of 30 kG toroidal magnetic focusing fields with a related nominal current of 200 kA. Therefore, we can set the different polarity modes of the two horns to select positively and negatively charged particles separately.

The magnetic field produced in forward horn current mode, focuses positive mesons, which will later decay into ν_μ . This is also called neutrino mode. On the other hand, the reverse horn current mode (anti-neutrino mode) focuses negative mesons and therefore produces $\bar{\nu}_\mu$. We notice that some opposite sign mesons may go through the center of the horn and therefore cause wrong-sign contamination. When the neutrino mode is running, $\bar{\nu}_\mu$ is expected to be less than 2 percentage of the flux. ν_μ , as the wrong-sign background, is expected to

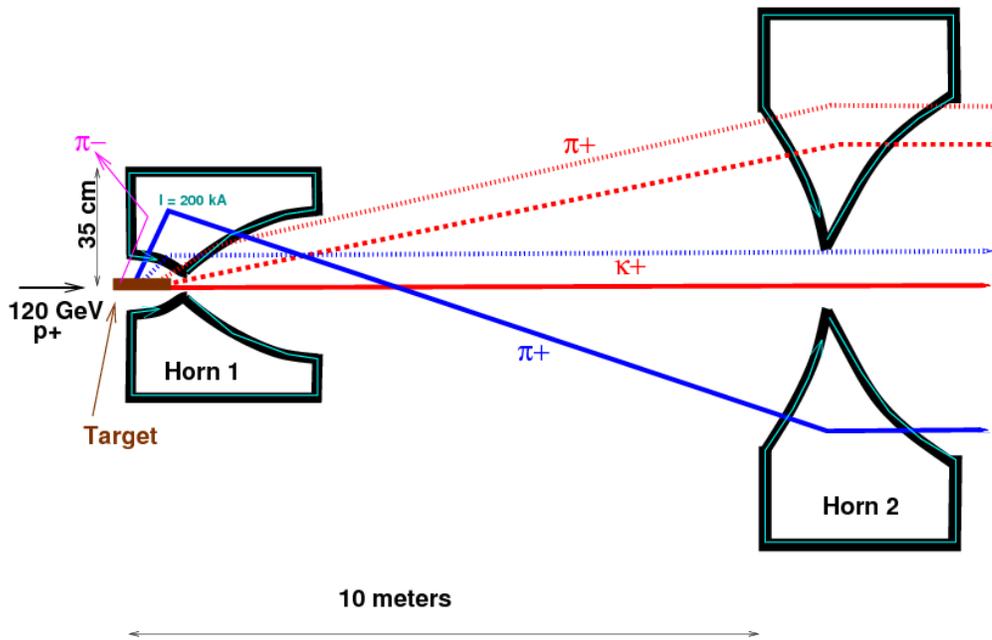


FIGURE 2.7: NuMI focusing horns when pulsed in Forward Horn Current mode (FHC)

be less than 12 percentage when the anti-neutrino mode is running. The wrong-sign particle ratio is higher in reverse horn current mode due to the comparatively smaller interaction cross-sections for anti-neutrinos. Also, the relative position of the pair of the horns can be arranged to change the shape of the neutrino energy spectrum.

2.1.3 NuMI Upgrade

The 400 kW NuMI beam was designed for the MINOS experiment to enable the precision measurement of muon neutrino oscillations. Upgrades to the accelerator complex and NuMI to 700 kW beam power have been accomplished to achieve the goals of the NOvA physics programs. The improvement comes chiefly from two aspects: increased efficiency in proton acceleration and extra protons per batch. Firstly, the Main Injector cycle-time has been increased to 1.33 seconds from 2.2 seconds. This was achieved by adding a Recycler storage ring,

which is available after completion of the Tevatron, into the accelerator complex. Protons from the Booster are now transferred to the Recycler first, which plays the role as a pre-injector to the MI, with resultant reducing the cycle-time since the Recycler can prepare the next batch of protons while the Main Injector is ramping the previous batch. Secondly, a new technology, slip-stacking, has been applied to double the number of protons in each batch before the protons are fed into the MI. The Recycler, which was used to implement slip-stacking, has two radio-frequency (RF) cavities and therefore can store protons with two different cycling frequencies. Six batches are delivered into the RF system one after the other from the Booster, with about 5×10^{12} protons in each batch. Six batches are delivered into the RF system one after the other from the Booster. The protons are then decelerated by progressively reducing the frequency of the RF cavities, falling into a lower momentum orbit. Then, a new batch of protons is fed into the RF system. The new batch was slipped out of the seventh slot and lines up with the earlier six batches due to its slightly higher momentum. By repeating the same process, the first six batches have twice the number of protons. The six slip-stacked batches are then transported into the MI altogether. Finally, each MI spill can include about 4.9×10^{13} protons, and lasts for $10 \mu\text{s}$.

Table 2.1 shows more details about NuMI at different levels of development.

2.2 NOvA Detectors

NOvA employs two functionally-identical detectors, which sit off-axis from the NuMI beam, to measure neutrino interactions. The functionally-identical design was chosen to maximally cancel the effect of systematic errors. The ND is utilized to constrain the neutrino beam flux and the FD is used to detect the

	NuMI/MINOS	NuMI/NOvA	NuMI/DUNE
Proton Beam Power	0.4 MW	0.7 MW	0.7 to 2.3 MW
Proton energy	120 GeV	120 GeV	60 to 120 GeV
Repetition rate	1.87 s (design)	1.33 s	1.33 s
Protons per Spill	4.0×10^{13}	4.9×10^{13}	4.9×10^{13} to 1.6×10^{14}
Baseline	735 km	810 km	1300 km

TABLE 2.1: Fermilab long-baseline neutrino beams

oscillated neutrino spectra. The following sections will illustrate the detector components and how the neutrino interactions were recorded by the detectors. Most of the technical details were drawn from the NOvA Technical Design Report [30], which may also be consulted for more in-depth information.

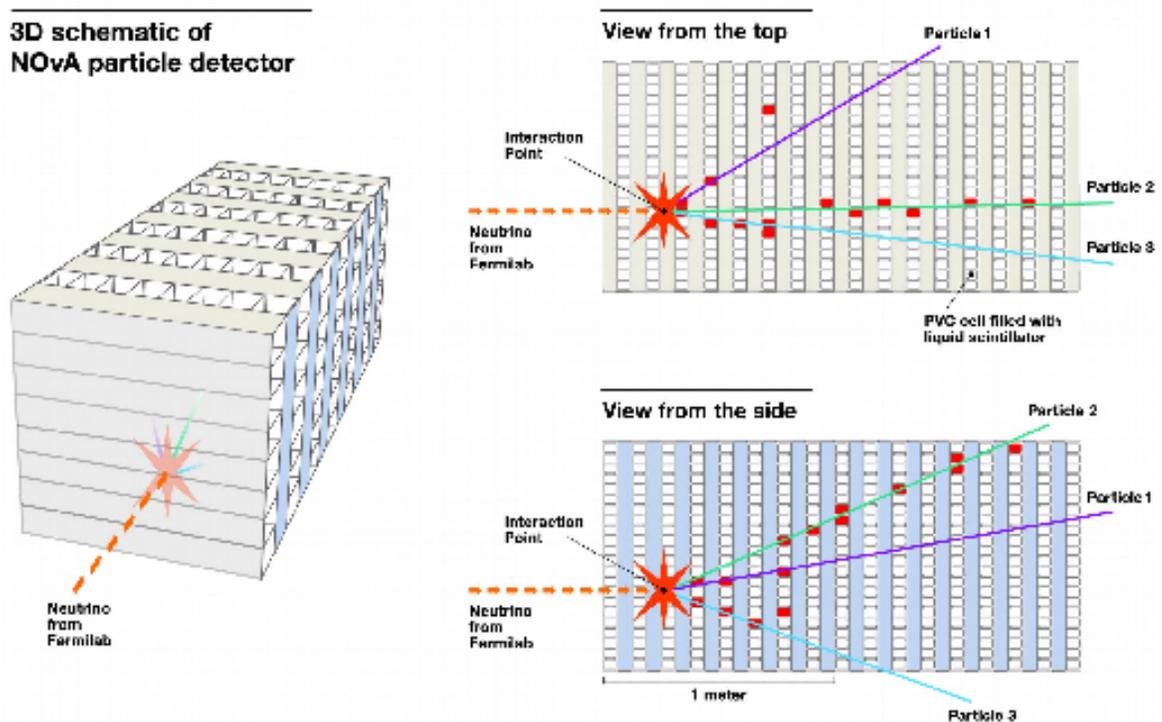


FIGURE 2.8: 3D schematic diagram of NOvA detector and neutrino interactions.

2.2.1 Off-Axis Location of Detectors

The NOvA detectors were designed to sit off-axis to optimize the quantity of detected electron neutrino appearance. In the rest frame of the two-body hadronic decay (mainly π , K) the daughter neutrinos are mono-energetic. Boosting the isotropic decay to the lab frame, we find the relation between the E_h (energy of the mother hadrons), E_ν (energy of the daughter neutrino), the neutrino flux F , the angle θ between the mother and daughter particles, the detector cross-section A , and the the distance to the detector or L are:

$$F = \left(\frac{2\gamma}{1 + \gamma^2\theta^2} \right)^2 \frac{A}{4\pi L^2} \quad (2.4)$$

$$E_\nu = \frac{E_h}{1 + \gamma^2\theta^2} \left(1 - \frac{m_\mu^2}{m_h^2} \right) \quad (2.5)$$

where h denotes the decaying hadron, and γ is the hadron's boost.

The selected angle, 14 milliradians off-axis from the NuMI beam direction, ensures NOvA detectors can receive a narrow-band beam with energy peaked at about 2 GeV, which is close to the first oscillation maximum energy (approximately 1.6 GeV) for muon neutrino to electron neutrino oscillation probability for a baseline of 810 km. The above point is clear to see from Figure 2.9. The narrowness of the spectrum has both pros and cons for detecting active-sterile neutrino mixing via NC disappearance. A reduced neutrino flux decreases feed-down from the neutral-current events.

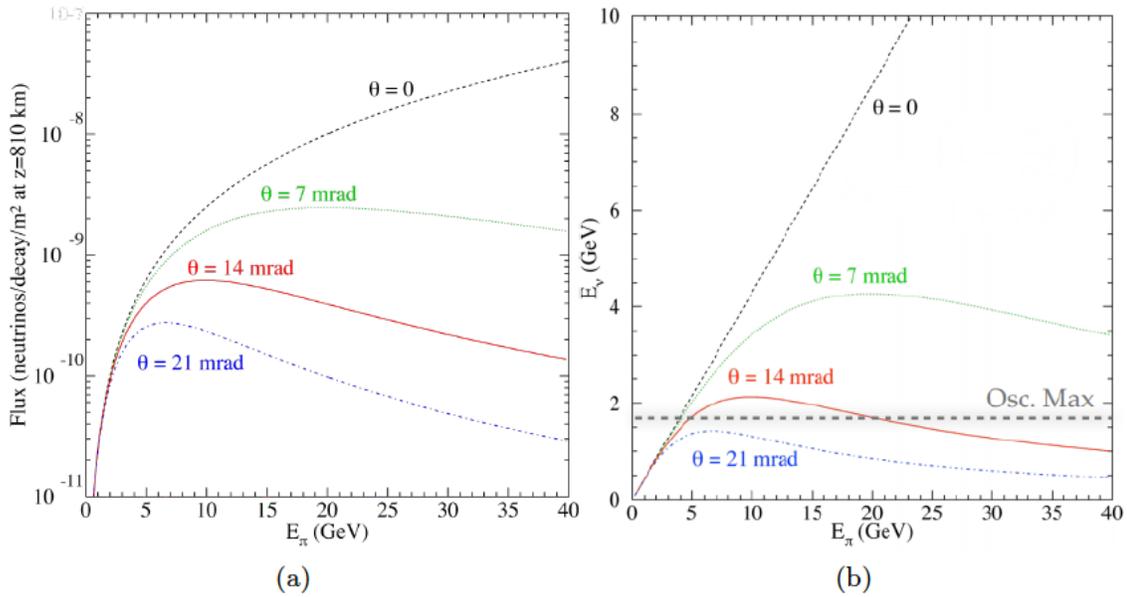


FIGURE 2.9: a) The neutrino flux as a function of pion energy viewed from an angle θ with respect to the NuMI beam. The neutrino flux was normalized to 810 km. b) The energy of daughter neutrinos produced at an angle θ with respect to the mother pion direction as a function of pion energy. To get a $\nu_\mu \rightarrow \nu_e$ maximum oscillation probability, the 14 milliradian angle has been chosen, as the horizontal dashed line shown.

For the NC interaction, the outgoing lepton (the neutrino) carries away an unknown amount of energy, therefore shifting most of the NC events to be below the neutrino energy peak, as shown in Figure 2.11. On the other hand, the CC neutrino interactions are also suppressed due to oscillations, which makes the task to select NC events at the FD easier.

2.2.2 Detector Technology

The two off-axis detectors are finely segmented, liquid scintillator calorimeters. The basic structure of the detectors is a set of extruded PVC plastic cells, filled with a looped optical fiber and liquid scintillator, arranged perpendicularly. The

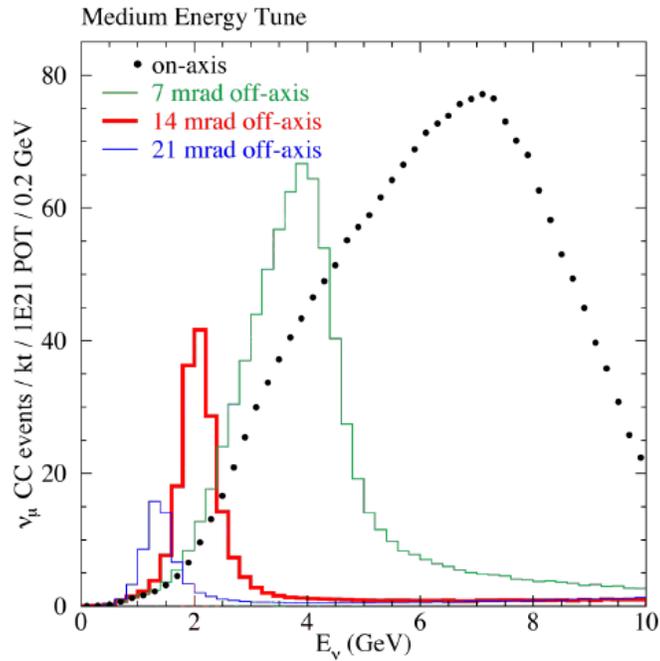


FIGURE 2.10: The NOvA Far Detector simulated neutrino energy spectra. It also shows that the off-axis location design suppresses the high energy neutrino tail.

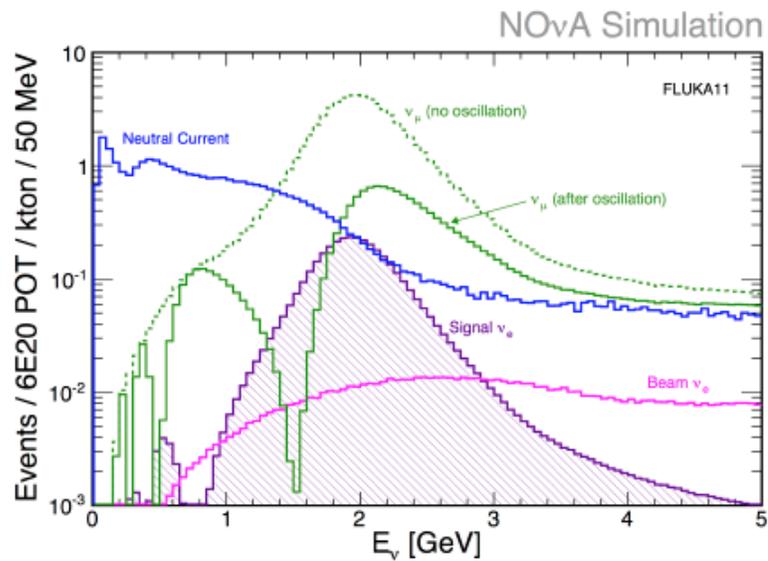


FIGURE 2.11: The expected neutrino interaction rates as a function of the true neutrino energy.

cell length is 4 m in the ND and 15.5 m in the FD, spanning the entire height and width of the corresponding detector. And, each cell has a nearly rectangular

cross section with dimensions of 3.9 cm \times 5.9 cm, as shown in Figure 2.12

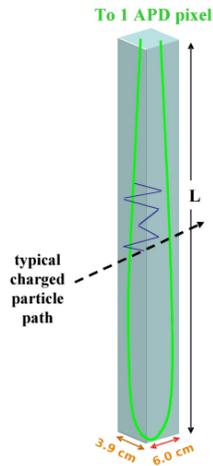


FIGURE 2.12: A schematic of a single cell.

Titanium dioxide was coated to the wall of each cell, which makes the cell highly reflective for scintillator light of 400 to 450 nm. 16 PVC cells are glued together to form one unit. Two units are stuck to create a module. Then, modules are glued together to build a plane.

There are 344,064 cells in the FD, and 18,432 cells in the ND total, which contain and segment the liquid scintillator, the active material in the detector. Figure 2.13 shows a schematic diagram.

Liquid Scintillator

The active scintillating material, 2.7 million gallons, makes about 65% of the NOvA detectors by mass. The liquid scintillator solution consists of several different components, each having a specific function. The primary component is mineral oil, about 95%, which provides the solvent for blending all the other ingredients. Pseudocumene, close to 5% of the solution, is used as the scintillating agent to generate photons in the range 270-320 nm. To convert the UV light, emitted by pseudocumene, into the visible range (380 - 450 nm) of

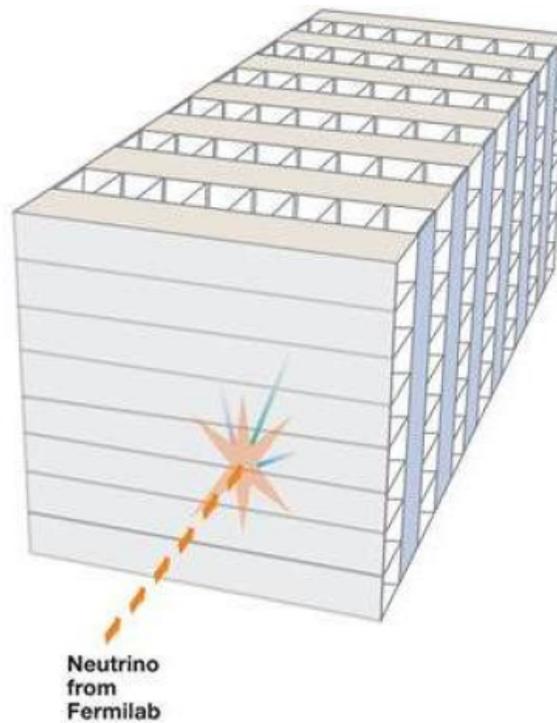


FIGURE 2.13: A schematic of NOvA detector with glued together cells.

the wavelength-shifting fiber, two wavelength shifters, PPO and bis-MSB, have been added to the solution in trace amounts. PPO, excited by the UV light, can emit photons with wavelengths in the range of 340 - 380 nm. Then, the second wavelength shifter, bis-MSB, de-excites through emission in the range 380 - 450 nm.

Table 2.2 shows more details about active scintillating material components.

Optical Fiber

A loop of optical fiber (WLS), as the wavelength shifting agent, rests inside each cell to collect the scintillation light. Titanium dioxide, TiO_2 , as the reflective agent can help to maximize the scintillation light collection. A fluorescent dye,

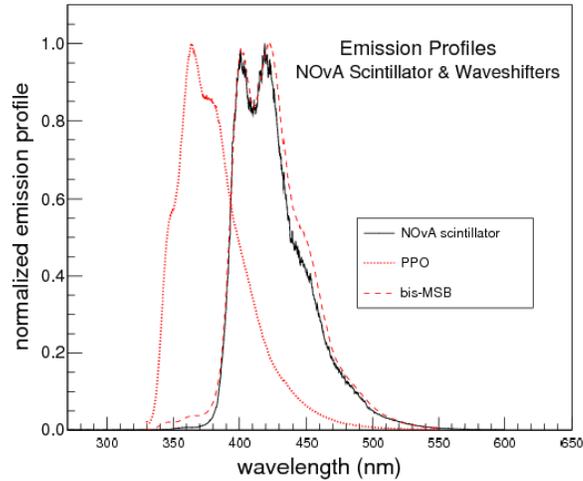


FIGURE 2.14: Emission profiles for NOvA scintillator and waveshifters.

Component	Purpose	Mass Fraction	Mass (kg)
Mineral Oil	solvent	94.91%	691,179
Pseudocumene	scintillant	4.98%	36,2677
PPO	waveshifter	0.11%	801
bis-MSB	waveshifter	0.0016%	11.7
Stadis-425	antistatic	0.001%	7.3
Vitamin E	antioxidant	0.001%	7.1

TABLE 2.2: Active Scintillating Material Components

Y11, in the fiber absorbs this scintillation light (425 nm) and shifts it to the blue-green region (450 - 650 nm). As shown in Figure 2.15, there is a overlap range between the absorption and emission spectra of the Y11, therefore, the light, with wavelength below 520 nm, was heavily attenuated in the fiber.

The photon signal then is transmitted to the fiber then to photodetector and electronics for readout.

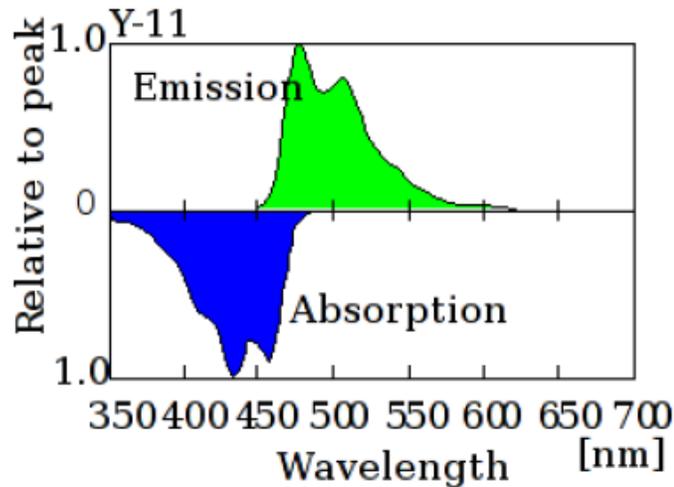


FIGURE 2.15: NOvA WLS fiber absorption and emission spectra.

Photodetector and Electronics

NOvA employs a 32 pixel Hamamatsu avalanche photodiode (APD) as the first component in data readout. The APD, as shown in Figure 2.16 was custom designed for NOvA.

Both ends of the looped fiber can be fitted onto one APD pixel, maximizing the signal collection. NOvA APDs can reach 85% quantum efficiency, which is significantly higher than the traditional photomultiplier tubes (PMT), for the light with 500 - 550 nm wavelengths directed by the fiber. Each APD is cooled to its operating temperature, -15°C , by straight connection to a thermoelectric cooler (TEC). This reduces the noise generated by the thermal current. Then, a water cooling system continuously removes the heat from the TEC, so that the APD operating temperature can star at -15°C . The TEC is a piece of the front-end board (FEB). The FEB and APD are both housed inside one box.

The FEB reads the output from the APD in the same box and then prepares the digital signal for NOvA data storage. The digital signal, from up to 64 FEBs, is

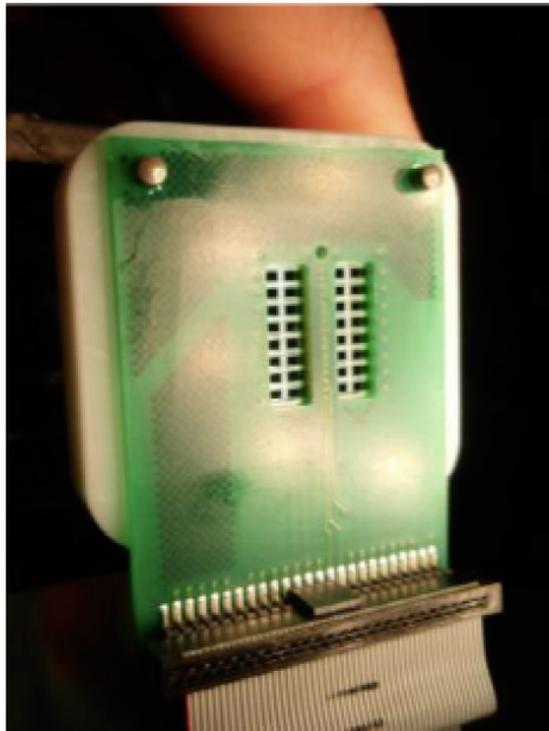


FIGURE 2.16: One NOvA APD

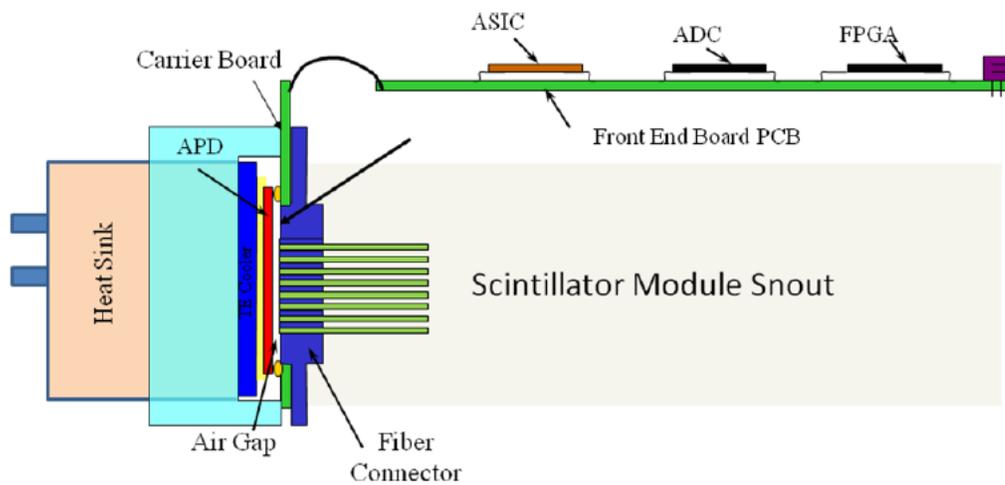


FIGURE 2.17: The Schematic of the APD and the FEB.

aggregated by a Data Concentrator Module (DCM). This signal readout chain is called the data acquisition system. The data acquisition system, DAQ, collects and digitizes the photo signals from all the APD channels and then transforms

them into a data stream. Also, the DAQ includes timing and command distribution systems, whose function is to record the timing information of hits.

PVC Modules

The structural elements of NOvA detectors are PVC modules, which contain the liquid scintillator. The modules make about 30% of the NOvA detectors by mass. The PVC extrusions are designed to optimize minimum module stress, light reflectivity, and a reliable extrusion process. The outer wall of the modules, 4.8 mm thick, is scalloped at the boundaries. The corners of the cells, as shown in Figure 2.18, are scalloped too. This design helps to reduce PVC stress concentration on corners. The inner walls between PVC cells are 3.3 mm thick.

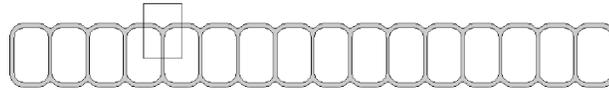


FIGURE 2.18: Profile of the PVC Cells

16 PVC cells were extruded in a group to form one PVC extrusion. Then two extrusions were fastened together with glue to form 32-cell extrusions. The next step is to string the optical fiber in a loop inside the cells of the extrusions. A groove, aligned along the diagonal of the cell cross section, is used to hold the fiber, as shown in Figure 2.19. Then two fiber ends from each of the 32-cell extrusions were embedded into a groove of a fiber raceway, which chains the fiber ends to an optical connector that has 32 holes. Each of the holes connect a single channel on the APD.

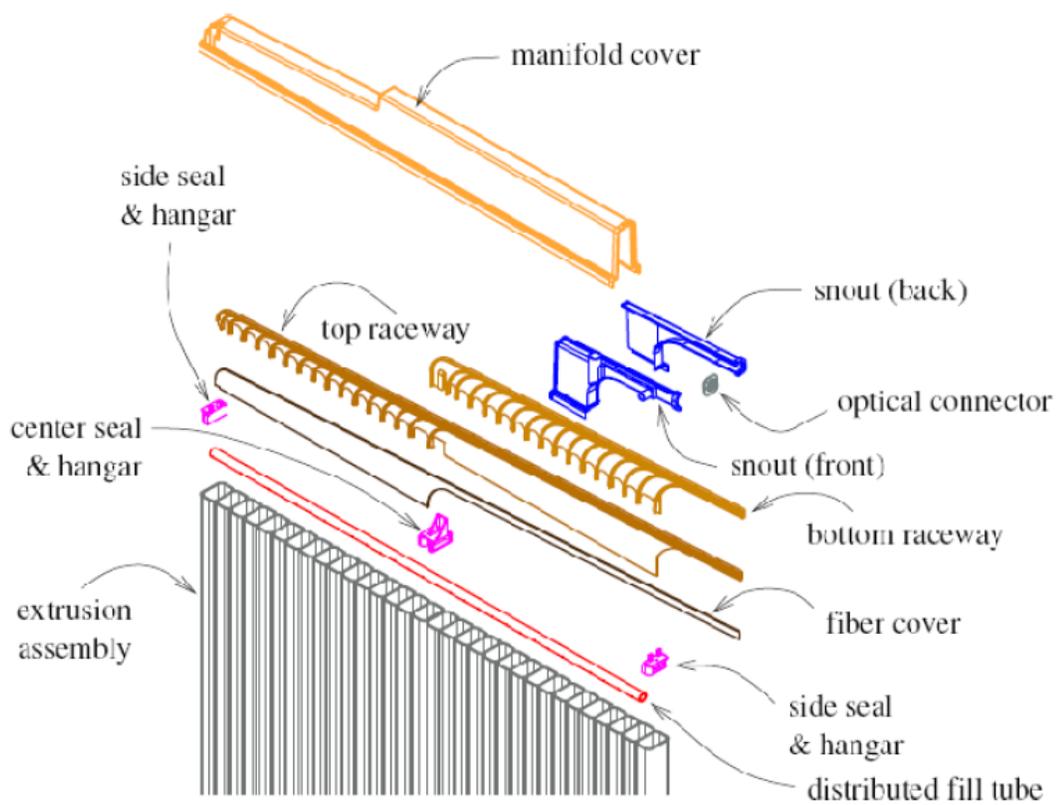


FIGURE 2.19: Schematic of the NOvA PVC module

2.2.3 The Near Detector

The NOvA Near Detector is adjacent to the MINOS ND, about 1015 m from the target system and 105 m beneath the surface, as shown in Figure 2.20

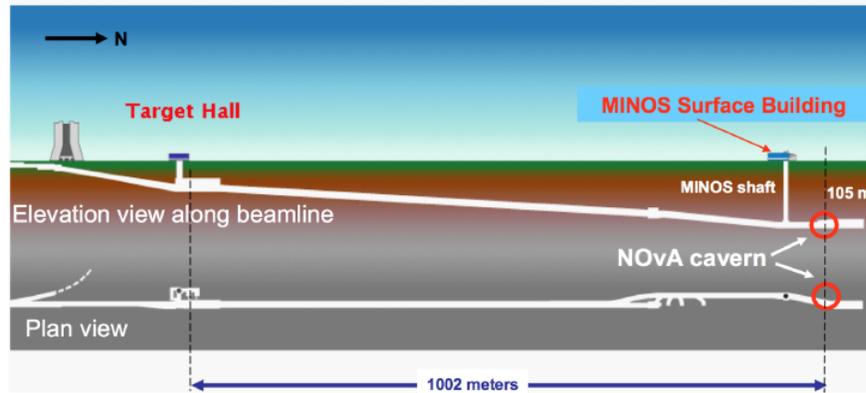


FIGURE 2.20: Schematic of the NOvA and MINOS ND hall.

The ND consists of two different sections, an active region that performs the neutrino-nuclear interaction measurement, followed by a muon catcher region to capture muons that do not stop in the ND active region. The active region of the ND has 8 blocks which are made of 24 planes. Each of the plane has 3 PVC modules. Therefore, totally, the ND consists of 192 planes with dimensions of $4.1 \times 4.1 \times 12.8$ m. The 192 planes have been partitioned into three di-blocks, each consisting of 64 planes. Each of the 64-plane di-blocks has 2 DCMs for the electronics. The beam neutrinos pass through the active region first then go into the downstream muon catcher region where the accurate muon energy measurement can be performed.

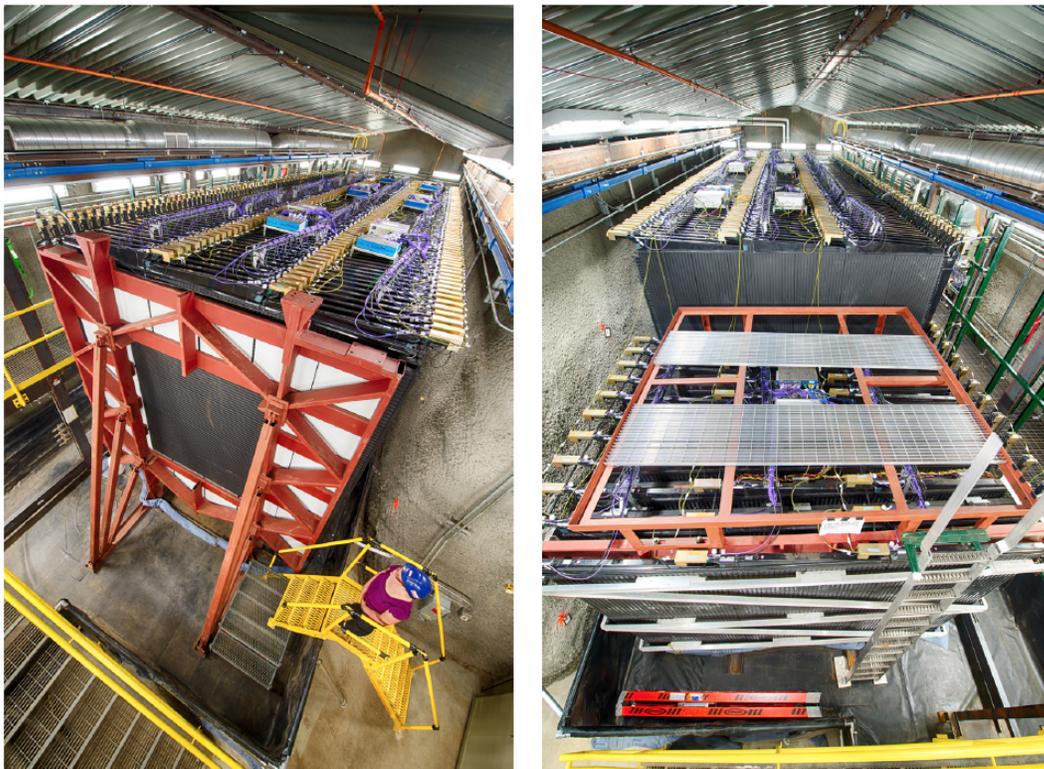


FIGURE 2.21: Pictures of the NOvA Near Detector.

2.2.4 The Far Detector

The NOvA FD is located 810 km away from the target system. The FD is an identical copy of the ND except it is much larger. The fully commissioned FD has 14 diblocks, which consists of 64 planes. Therefore, the FD has dimensions of $15.6 \times 15.6 \times 60$ meters and a mass of 14 kilotons. Unlike the MINOS detector which sits deep underground, the NOvA Far Detector is commissioned on the surface, hence exposed to abundant rate of cosmic rays. The detector was covered in 4 feet concrete and 6 inches of barite, which was designed to shield the NOvA Far Detector from the cosmic rays by about 14 radiation lengths. However, neutron events generated in the rock above detector have a negative influence in the NC selection.



FIGURE 2.22: Schematic of the NOvA Far Detector.

Chapter 3

Analysis Files Production

“It doesn’t matter how beautiful your theory is, it doesn’t matter how smart you are. If it doesn’t agree with experiment, it’s wrong.”

— Richard Feynman

Experimental neutrino physics, and physics in general, is driven by experimental measurements. Therefore, accurately detecting and recording neutrino-like interactions in both NOvA detectors, and processing the data files for subsequent analysis is the starting vital step. The first step in data file production chain is called Data Acquisition (DAQ), which is the process of how NOvA takes data. The following step is the energy calibration process, which translates/interprets the raw data, charge records in the APD pixels, into a physically meaningful energy deposition with the corresponding timing information. An above threshold energy deposition in a detector cell is called (also recorded as) a hit. The third step, event reconstruction, of this process then utilizes a group of hits to reconstruct the corresponding event vertex and daughter particle tracks, and therefore, the whole neutrino nuclear interaction. Considering the structure of a NOvA detector plane, as shown in Fig. 2.13, the recorded hits only show

two-dimensional information for the particle trajectories. The plane expresses the Z coordinate, and the cell expresses the X or Y coordinate corresponding to the vertical or horizontal separately. The event reconstruction process needs both the X view and Y view cells to build a three-dimensional view of a neutrino interaction. The above three steps are part of the file production process for the measured data, also called real data, and are described below in the first three sections. In parallel with the measured data file production, NOvA also produces a variety of simulated data files, also called fake data or Monte Carlo (MC), which is presented in the last section.

Chapter Organization

- Section 1: Data Taking
- Section 2: Energy Calibration
- Section 3: Event Reconstruction
- Section 4: NOvA Simulation File Production

NOvA Analysis Files

- Measured Data Files:
 1. ND Beam File: also called ND File, which records all the (reconstructed) measured neutrino-like interactions in the Near Detector. We employ the ND files for Data Vs Monte Carlo (MC) Agreement Study (introduced in Chapter 4 - **Neutrino Interaction Classification and Signal Selection**), and the ND Spectra Decomposition Analysis presented in Chapter 5-**Extraction of Sterile Mixing Parameters**.

2. FD Beam File: was also called FD File, which records all the (reconstructed) measured neutrino-like interactions in the Far Detector, including all the interactions during the beam window. The FD files are employed for extracting the results as presented in Chapter 5-**Extraction of Sterile Mixing Parameters**.
 3. FD Cosmic File: also called Cosmic File, which records all the (reconstructed) measured Cosmic interactions in the Far Detector. The Cosmic files are employed for training and testing the PIDs as introduced in Chapter 4-**Neutrino Interaction Classification and Signal Selection**.
- Monte Carlo Simulated Files:
 1. ND Monte Carlo Beam File: also called ND MC File, which simulates neutrino-nuclear interactions in the Near Detector. We employ the ND MC files for Data Vs Monte Carlo (MC) Agreement Study (introduced in Chapter 4-**Neutrino Interaction Classification and Signal Selection**), and ND Spectra Decomposition Analysis presented in Chapter 5-**Extraction of Sterile Mixing Parameters**.
 2. FD Monte Carlo Beam File: also called FD MC File, which simulates neutrino-nuclear interactions in the Far Detector. The FD MC files are employed in training and testing the PIDs as introduced in Chapter 4-**Neutrino Interaction Classification and Signal Selection**, and FD Spectra Prediction Analysis presented in Chapter 5-**Extraction of Sterile Mixing Parameters**.
 3. FD Cosmic Monte Carlo File: also called Cosmic MC file, which has been mostly replaced by real Cosmic data in our studies, but is still used in the detector Calibration process.

3.1 Data Taking

The NOvA DAQ system faces unique challenges: 1) Triggering the FD readout hardware in advance due to the long distance between neutrino source and NOvA detector; 2) Optimizing the activity-based trigger systems in the FD to record NuMI beam interactions among the abundant rate of cosmic rays. I will give a brief introduction to the NOvA DAQ system and the way NOvA solves the above challenges, as well as to the DAQ daily operation.

3.1.1 DAQ Overview

The DAQ system operates the following steps after the NOvA readout system takes data from the 20,192 and 344,064 detector channels in the ND and FD, respectively. It starts by aggregating the data from 32 FEBs per module in a Data Concentrator Module (DCM), which is a custom-built computer[31].

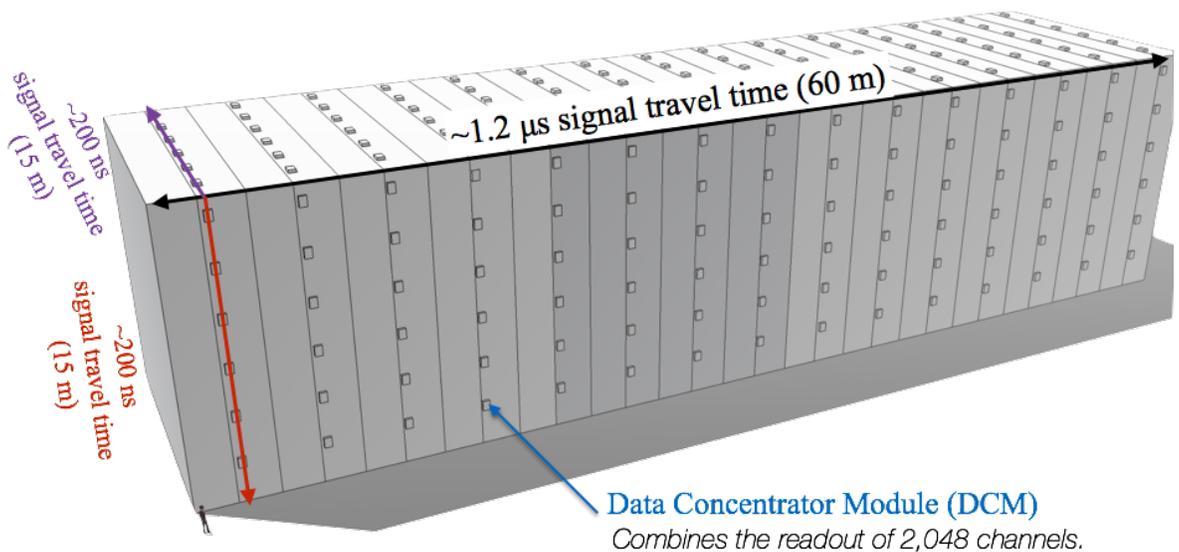


FIGURE 3.1: Diagram of NOvA Far Detector DCM placement [31].

Each of the DCM receives data from 2,048 detector channels and organizes them into a $50 \mu\text{s}$ long unit, which is also called a micro-slice in NOvA. An event

builder module aggregates the micro-slices into mili-slices, which are 5 ms long chunks. Aggregated mili-slices are then fed into a pool of buffer nodes. The duration of these mili-slices is chosen depending on the data transfer requirements. Each buffer is composed of a patch of computers that can save the data for up to 16 minutes so that the NOvA triggering systems have enough time to decide to discard data from the buffer, or to save data to disk as NOvA raw data files, and therefore enter our analysis chain. Fig.3.2 presents the detailed components of the NOvA readout and DAQ systems.

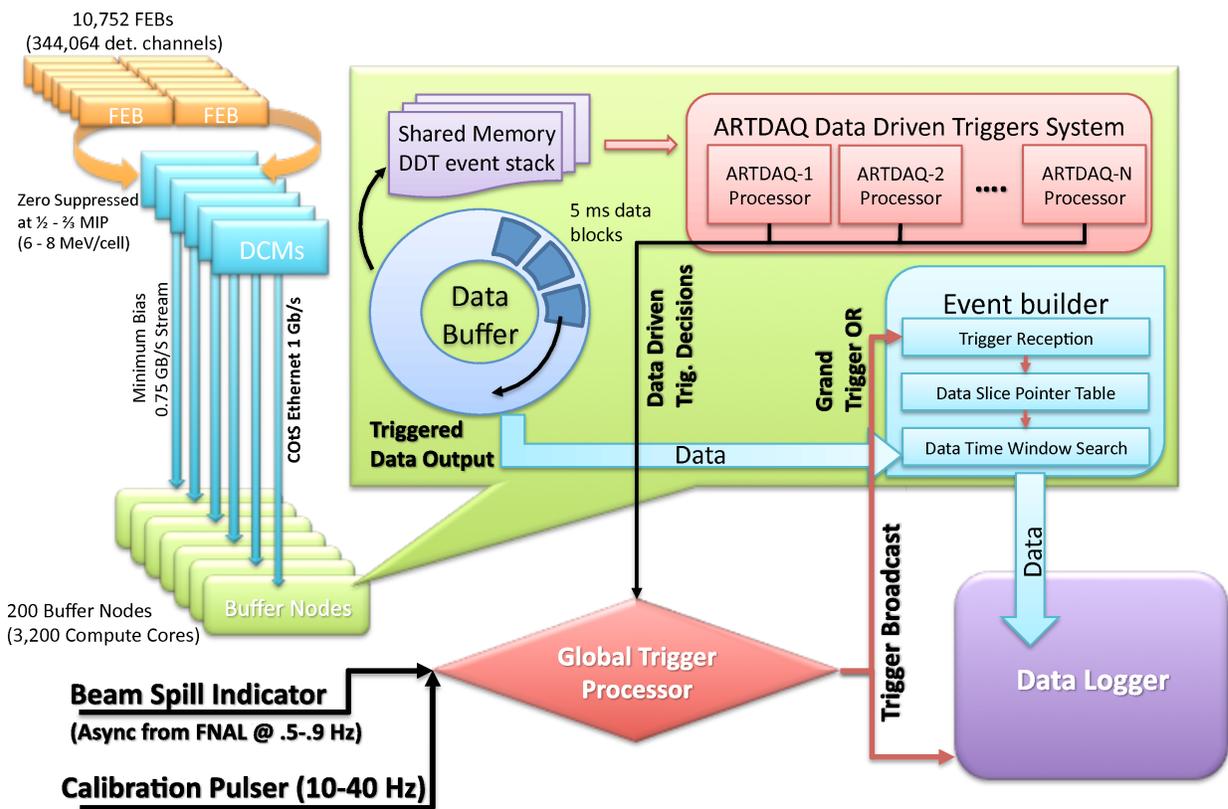


FIGURE 3.2: Overview of the NOvA Far Detector Readout and NOvA DAQ systems [32].

In Figure 3.3, a NOvA Far Detector 2D `EventDisplay` can help to better understand the DCM (in green) distribution.

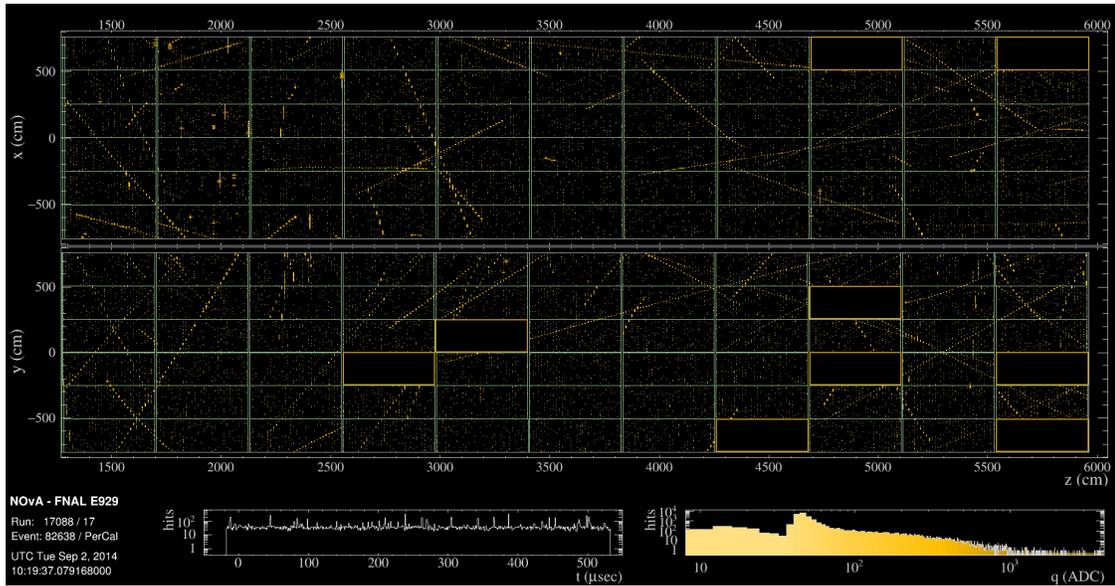


FIGURE 3.3: DCM layout in NOvA Far Detector [32].

3.1.2 Triggering Systems

The NOvA trigger systems play a vital role in data taking process, which can be categorized into the following three groups:

- **Data-Driven Triggers:** also called DDTs. The series of algorithms specifically designed to select data which can fulfill a series of specific conditions, as shown in Fig 3.4. Only when these trigger conditions are met encountered, will the DAQ system save the data into the analysis file.

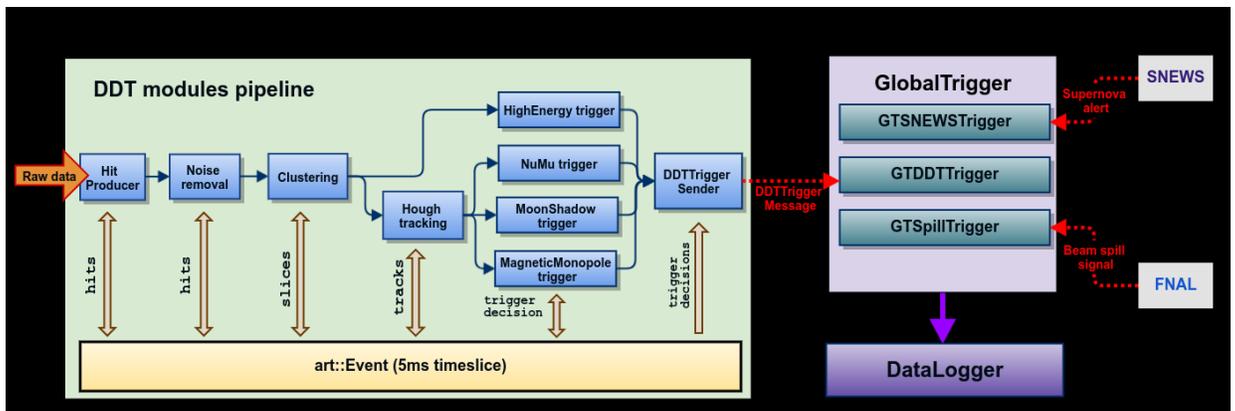


FIGURE 3.4: NOvA Data-Driven Trigger System [33].

- Signal Trigger : an algorithm designed to collect data which is saved for a period of the data-taking process.
- Clock : this trigger system stamps the time interval for all recorded data.

Figure 3.5 outlines the running trigger systems for collecting data for the NOvA 2017 data set, which was employed for the analysis presented in this thesis. Some more recent DDT triggers, such as the new proposed gravitational wave trigger, are not listed here.

Trigger Names	Rate [Hz]	Requested Time Window [μ s]	Data Logger Time Window [μ s]	Fraction of Live Time	Fraction of Recorded Data
NuMI Spill	0.77	500	550	0.04%	5%
Cosmic Pulsar	10	500	550	0.55%	59%
Tricell Calibration	25	1	50	0.13%	13%
High Energy	7	2	50	0.04%	4%
Monopole - Fast	19	1	50	0.10%	10%
Monopole - Slow	0.3	10	50	0.00%	0%
Muon Neutrino	7	1	50	0.04%	4%
Upward Muon	0.8	1	50	0.00%	0%
Upward Muon - Contained	8	1	50	0.04%	4%
Supernova Test - Fast	0.01	50	100	0.00%	0%
Supernova Test - Slow	0	5,000	5,050	0.00%	0%
Supernova	0	2,000,000	2,000,050	0.00%	0%
All Triggers				0.93%	100%

FIGURE 3.5: NOvA running triggers for 2017 Data Set [33].

Among the above listed triggers, there are some specific ones which we will describe in more detail.

- SuperNova Early Warning System Trigger: also called SNEWS. It is a global network, aggregating the supernova trigger signals from various neutrino experiments and giving its subscribers early supernova notifications. If triggers from 3 experiments coincide within a 10s time window, SNEWS would send an alert to subscribed neutrino experiments and a notification to the mailing list [34].

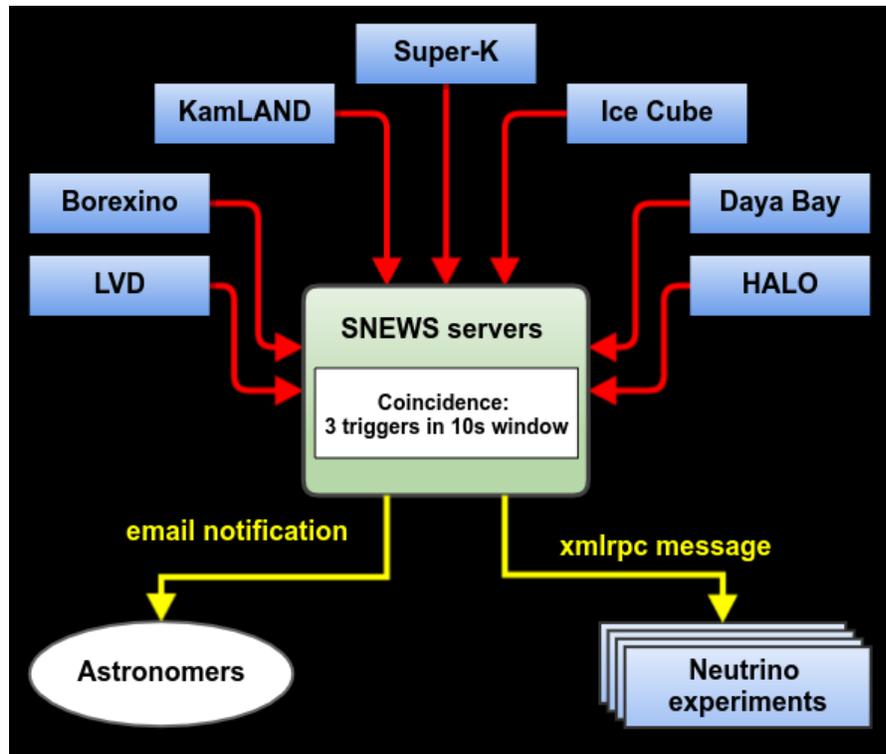


FIGURE 3.6: Scheme of SNEWS notification propagation [33].

- Supernova Trigger : was also called SN trigger system, which removes all the hits from the known sources, reconstruct the interaction candidates (hits groups close in time and space) and sends the rate of such candidates to the GlobalTrigger via DDS message, as shown in Fig. 3.7.
- Cosmic Data Trigger : FD Cosmic Data are saved in a $550 \mu\text{s}$ long time window. This trigger selects cosmic data which is used as a minimum-bias sample for energy calibration as well as to estimate the cosmic backgrounds in the NOvA long-baseline oscillation analyses.
- NuMI Data Trigger : The same $550 \mu\text{s}$ long time window employed as the cosmic data trigger. NOvA neutrino signals are expected to be contained in this data, in addition to the characteristic cosmic events in the FD.

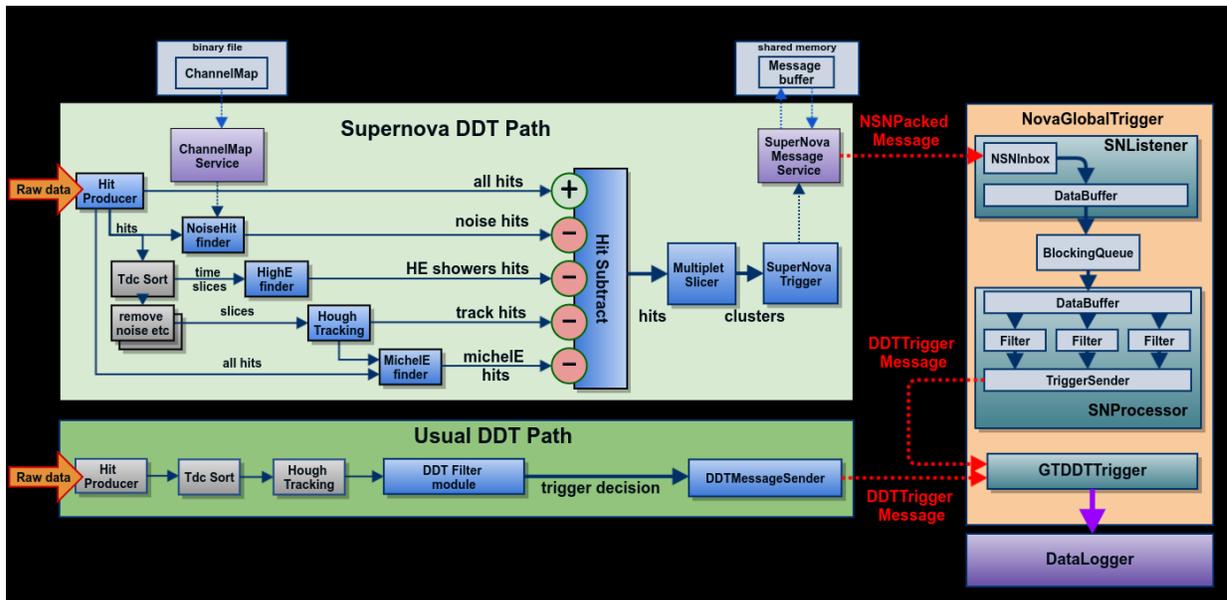


FIGURE 3.7: A scheme of the SuperNova triggering system.[33]

3.1.3 Timing System

NOvA consists of two detectors distanced by 810 km; therefore, it needs to be precisely synchronized in time both within the detectors and externally with the clock of the neutrino beam. This synchronization is critical for our analysis to separate multiple neutrino interactions per beam pulse within the Near Detector, and to select candidates with pulses from the neutrino beam. The technical details of the NOvA timing system, including the system layout, time synchronization, beam triggering, timing resolution, and the calibration technique developed to measure delays between electronics regions of the detector, as shown in Fig 3.8.

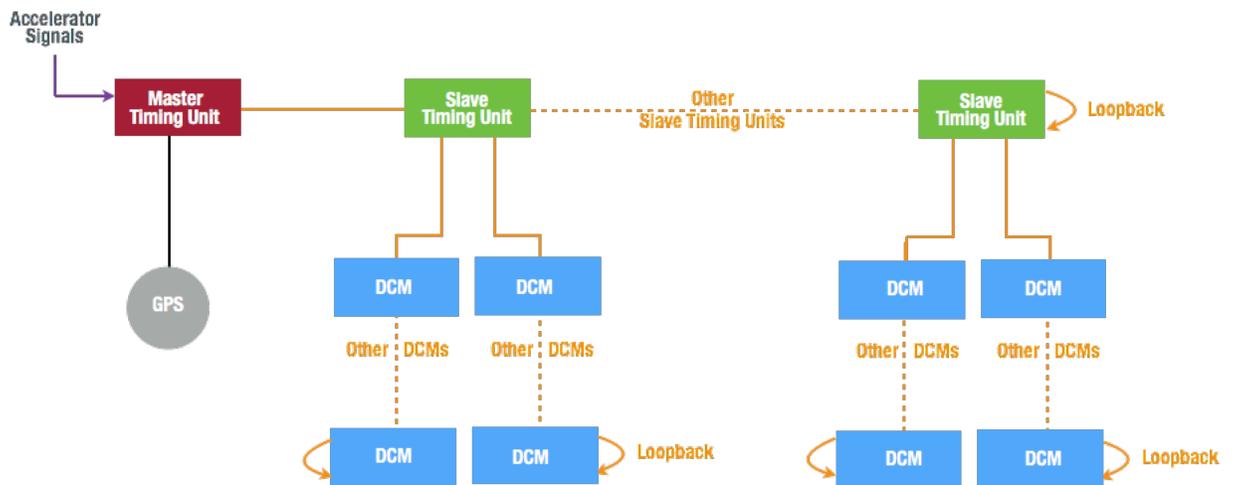


FIGURE 3.8: A simplified schematic detailing how the various elements of the timing system are connected to one another [31].

3.1.4 Live Data-Taking

As is typical for most high-energy physics experiments, the DAQ, and other supporting systems of NOvA, are under monitoring 24/7 by one or several collaborators on rotating shifts. Figure 3.9 shows Cincinnati neutrino group members taking a NOvA shift.



FIGURE 3.9: The NOvA shift operations center at Fermilab.

3.1.5 Accumulated NOvA Data Set

NOvA started to take data since February 2014 when the Far Detector was still under construction. For the 2017 Official and Re-Analysis, the data was taken until March 2017, which corresponds to 8.85×10^{20} POT equivalent beam exposure. Figure 3.10 displays the various periods of data taking by NOvA.

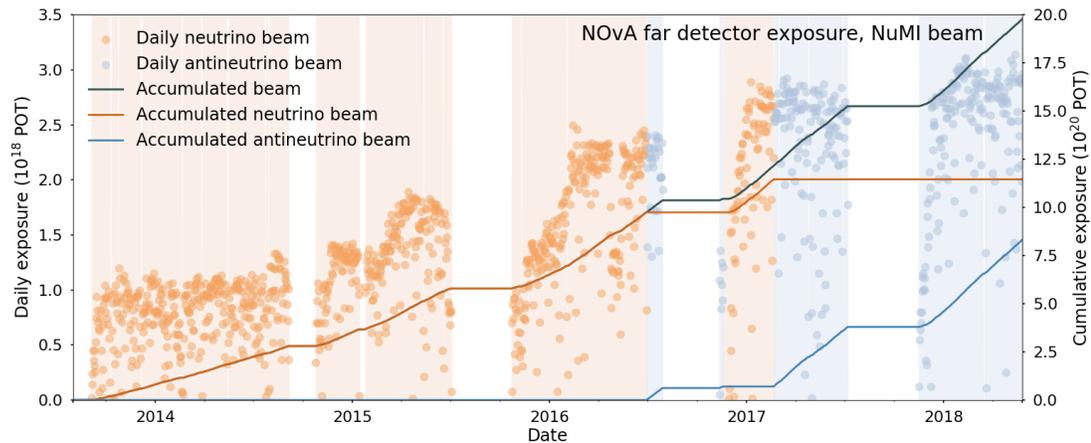


FIGURE 3.10: Time series showing the daily neutrino beam (orange) or antineutrino beam (blue) POT recorded by NOvA, from the start of commissioning to 2018-05-27. Also plotted are lines for the cumulative neutrino beam POT (dark orange), cumulative antineutrino beam POT (dark blue) and total accumulated POT (grey) [35]. Only data collected in neutrino mode up to March 2017 are used in the analysis presented in this thesis.

3.2 Energy Calibration

The energy calibration is the following step in the file producing process, which is done by using measured hits from through-going muons (muons that go into and leave the FD without stopping). Several selection rules have been applied to select the through-going muons from the abundant FD cosmic data. Then,

a NOvA specific tracking algorithm, Window Track, is employed to fit straight lines through measured hits to produce the 3D tracks from the cosmic ray. The selected and reconstructed cosmic muons are used as probes for the energy calibration as they are a source of well-understood energy deposits across the far detector. This process can be further categorized into two sequential stages: attenuation calibration (also known as relative calibration); and absolute calibration. The relative calibration accounts for the threshold effects and attenuation across one single cell by design to produce a uniform response throughout a cell and across the entire detector, ND and FD separately. By using a scale factor on each detector, the absolute calibration converts the calibrated PE scale from the relative calibration into an energy unit.

The units which will be used in the following two calibration processes are defined below:

- ADC: is an unit for "Analog to Digital" Conversion.
- PE: "Photo-Electrons", which is an unit that translates the ADC recorded by an APD to an amount of light incident on the APD.
- PECorr: the first calibrated unit correcting for attenuation and relative cell-to-cell differences.
- MIP: the energy deposited by a minimum-ionizing particle traveling along the z-direction through the depth of one cell.
- GeV: Estimated energy deposited in the scintillator.

3.2.1 Relative Calibration

The full energy calibration process starts from the relative calibration. It is devised to transform the pulse-height recorded in ADC to corrected photoelectron (PE) units. The PE signal from the electronics will be converted into a calibrated unit so that two signals detected in any two sections of one detector have the same true energy deposition. The algorithm of the process accounts for threshold effects and the light attenuation effect. For performing the relative calibration, the measured signal needs to be divided by the path-length in a cell. Path-length of each measured hit corresponds to the distance of the particle travels in the cell. The path-length normalizes the deposited energy in a cell which is belonging to a reconstructed track in a cell, and this channel response is then calibrated. Since path-length on a cell-by-cell basis can be challenging to estimate because of the reconstruction effects, most of the calibration uses tri-cells hits, that is, cell hits where the same cosmic ray also triggered both of the adjacent cells in the plane. The limit guarantees that the cosmic muon entered through the top wall and left through the bottom in that cell, and constrains the path-length to be cell-height, corrected by the direction cosine in that view. The ADC/cm of each selected cell can be expressed as a function of the distance W , which represents the distance along the cell's length to the readout and is determined by 3D track reconstruction. Then the profile of the 2D histogram (ADC/cm vs. W) will be constructed by taking the median value in each W bin. For hits that are around the top and bottom of the cell, a "roll-off" phenomenon has been followed because of the different reflection behavior at the ends of the cells. To correct this effect, an so-called empirical function has been used. Cells remain that have large residuals after the above serial steps, which could be

caused by varying fiber position within the cell or noisy behavior. These residuals do not follow any consistent pattern, thus cannot be fitted by a function. To solve this issue, a LOWESS (Locally Weighted Scatter plot Smoothing) method has been employed to provide a better fit to the curve.

3.2.2 Absolute Calibration

After the above cell-to-cell relative attenuation, the absolute energy calibration method is then applied to the two detectors to convert the corrected PE signal into an energy value, namely, interpret the energy scale from PECorr to GeV. Much like the relative calibration, the tri-cell hits from cosmic ray muons have been applied by the absolute calibration. However, for absolute calibration, we need to know the deposited energy in the cell in order to transform the PE signal into GeV. The cosmic ray muons that stop in the FD are employed to calculate this calibration scale. The deposited energy is learned either from the Bethe-Bloch equation or external sources. With the knowledge of energy deposition and the number of PEs, determining the calibration energy scale is a straightforward procedure. The particle energy used for the absolute calibration is the minimum energy deposition of muons through liquid scintillator. The absolute calibration energy scale is determined from distributions of tri-cell hits based on stopped muons that occur at the end of the muon track. These measured hits are then used to make one-dimensional muon energy unit (MEU) distributions of the relative calibration corrected detector response for both measured and simulated data called MEU-Reco, and the true energy deposition for mC called MEU-truth. The calorimetric energy scaling factor is then taken as the mean of the MEU-truth distribution over the mean of the MEU-REco distribution.

3.3 Event Reconstruction

3.3.1 Reconstruction Philosophy

A variety of event reconstruction algorithms have been developed for producing different purposes in NOvA. The reconstruction chain starts from separating individual neutrino-nuclear interactions from a larger readout window into the object called a slice in NOvA, which is intended to gather all calibrated hits from one neutrino-nuclear interaction and serve as the foundation for all the following reconstruction steps. Then, a modified Hough transform algorithm has been employed to classify straight-line features in each slice. Next, the Hough lines are applied as seeds to a specially designed algorithm to decide the position of the 3-dimensional vertex in each slice. The following step is to produce prongs, which are a collection of calibrated hits with a start point and direction. A fuzzy k-means algorithm performs this process using the 3D vertex as a seed.

3.3.2 Reconstruction Steps

Interaction Separation with DBSCAN

The first step of the reconstruction chain is to separate a variety of interactions into objects known as slices for further processing. The recorded time gap for cell hits in one interaction can range between tens of nanoseconds up to a microsecond, depending on the length and direction of the daughter particles of one interaction as well as the version of the used readout electronics. The average standard deviation of the hit time in a single slice is about 200 ns for single-point timing and about 60 ns for multi-point timing in NOvA Far Detector. On the other hand, in the Near Detector, faster electronic equipment has been applied, therefore, each single slice has a standard deviation of about 10 ns. The

data readout window can be of arbitrary size defined by one trigger algorithm. The one being currently used is $550 \mu\text{s}$ centered on the $10 \mu\text{s}$ neutrino beam spill window. For the Far Detector, there are typically 50-70 cosmic events in one readout window. Meanwhile, in the Near Detector, there are 4 to 5 beam neutrino interactions happening during each beam spill, and these need to be separated by further processing. NOvA reconstruction applies a density-based clustering algorithm, DBSCAN, to develop the slices. This algorithm works by calculating the distance metric between pairs of measured hits in four-dimensional space. The distance metric is computed for all the hits. When the distance of a pair of hits is less than the threshold, they are classified as neighbors. When a hit has at least four neighbors, the hit is considered a core hit. Therefore it can be used as a seed to form a cluster. The core hit and its neighbor hits are then put into a cluster. If any of the neighbor hits are also core hits, the cluster expands to include them and so on until all related core hits and their neighbors are put into one cluster. A limit of at least three hits in each view has been placed on the single slice for reducing extraneous slice noise. To evaluating a single slice purity and efficiency, the completeness is then computed. The cosmic simulation slicing in the Far Detector was established to have the purity and efficiency of 99.3%. The Near Detector simulation slicing has a purity of 98.5% and an efficiency of 94.4%.

Guidelines with Hough Transform

The Hough transform algorithm is used to find the prominent lines in each slice as the next step of the reconstruction chain. The output of the algorithm is a couple of straight trajectory lines in two readout view so that a vertex-finding algorithm can then apply the directions and intersections of the produced straight lines. To be robust against background noise, a modified Hough Transform has

been applied on pairs of points. A line will be drawn through each pair of hits and then transferred into the polar coordinate. This algorithm works in both detector views separately filling a two-dimensional Hough space in the polar coordinate with a Gaussian smeared vote for each hit pair.

Vertex Identification with Elastic-Arms

Then the next phase of the reconstruction chain is to reconstruct the global interaction vertex of each interaction. We assume that all visible energy in the slice is the result of one primary interaction point, which is generally true for neutrinos. For a single slice, there are a set of straight lines which are described as elastic arms. This algorithm is designed to run on neutrino-nuclear interactions and therefore is biased in seeding to prefer vertices at the upstream end of the detector where the beam originates.

Further Reconstruction Objects

The above three sections described the fundamental reconstruction objects and all the others are based on them. There are plenty of reconstructed objects which play vital role in signal selection and mixing parameter extraction. All the corresponding reconstructed objects which have been employed by the NC Disappearance analysis will be detailed in Chapter 4-**Neutrino Interaction Classification and Signal Selection**, and in Chapter 5- **Extraction of Sterile Mixing Parameters**.

3.4 NOvA Simulation File Production

Neutrino based analysis profoundly relies on the Monte Carlo (MC) simulated beam files. Simulating neutrino production, propagation, and interactions in

the NOvA detectors is essential to qualify our understanding of the underlying physics of NOvA. The simulated data can help us build the reconstruction methods, optimize interactions classification machine learning algorithms, tune signal selection cuts, and finally extract values for the mixing parameters by comparing the simulated and detected data. The NOvA MC simulation chain comprises three main stages as illustrated in Figure 3.11 below.

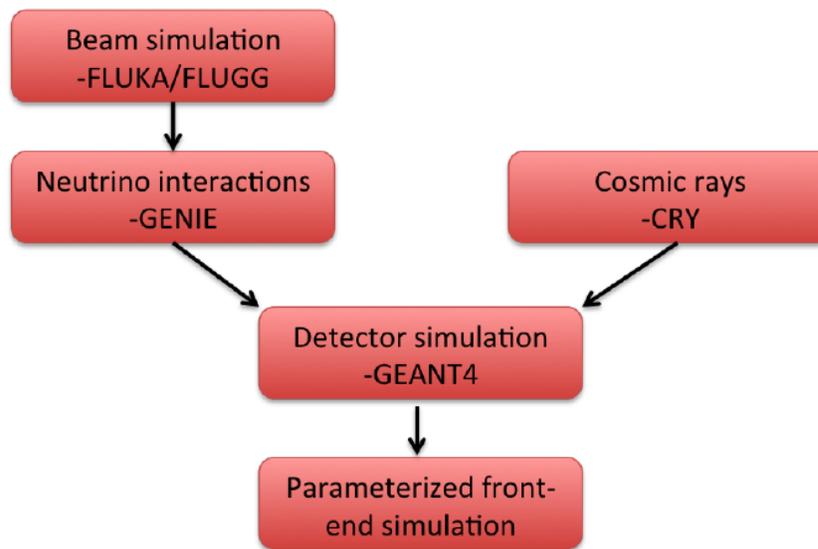


FIGURE 3.11: NOvA Simulation Chain Flowchart. The main steps of the NOvA simulation chain are listed. Credit by Ji Liu.

The simulation main stages are: 1) beam flux simulation; 2) interaction simulation; and 3) detector simulation. The beam flux simulation includes the hadron production process simulation and the corresponding daughter particles propagation through the magnetic horns until the mesons decay. The produced files provide input information for the interactive simulation. The second step consists of simulating neutrinos interactions in NOvA detectors. The final step of the simulation chain is the detector simulation, which takes the previous information to simulate the propagation and interaction with the NOvA detector

elements of the daughter particles from the neutrinos interaction. We then obtain the analysis files which mimic the raw detector data. These three steps are introduced in the following parts.

3.4.1 Beam Flux Simulation

The first link in the whole NOvA simulation chain starts with the simulation of neutrino production in the NuMI beam. Neutrino production simulation begins with the hadron production simulation which is done with the G4NuMI neutrino flux package, based on GEANT4 [36] plus a specific description of NuMI geometry and materials. It models hadron production by protons interacting with the NuMI target and then tracks them into the horns until they decay into neutrinos or are absorbed downstream. The geometry simulation includes the environment where the particles interact in and with, such as the graphite target and relevant elements of the target hall, the horns, and the decaying zone. G4NuMI records parent-particle information of the neutrinos in the NuMI beam along the way so that we can understand where a neutrino comes from, such as from pion or kaon decay. A customized Package, Package to Predict the Flux (PPFX [37, 38]), has been employed to correct the G4NuMI hadron production. It calculates the corrections and the corresponding uncertainties based on the results from hadron production experiments. The output flux files, therefore, include the neutrino flavor, energies, and moving directions as well as the information of their ancestors that produced them.

3.4.2 Interaction Simulation

The Generates Events for Neutrino Interaction Experiments package (GENIE) [39] is a ROOT-based neutrino Monte Carlo Simulation generator that is used to simulate the neutrino-nuclear interactions from the beam in the NOvA detectors. The Cosmic RaY (CRY) generator is employed to simulate the cosmic ray particles interaction in NOvA far detector. CRY is also employed for energy calibration, as mentioned before. The two packages are introduced in more detail below.

GENIE

The theoretical models considered by GENIE can be categorized into three types: cross-section models; hadronization models; and nuclear physics models. We will discuss each of these three category models in the following.

Cross Section Model

In GENIE, cross-section models are applied to calculate the total and differential cross sections with the input flux information. Then, for a specific process, the total cross section will be employed to determine which type of interaction will occur, then utilize the differential cross section for that interaction to calculate the final state kinematics. There are some different targets that a neutrino within different energy ranges can scatter off of, including the whole nucleus, individual nucleons, quarks within the nucleons, and atomic electrons. Below, the three major scattering processes are described in more detail.

Quasi-elastic Scattering One of the primary neutrino interactions happening in the NOvA detectors is quasi-elastic scattering. This type of interaction can be expressed by:

$$\nu + n \longrightarrow l^- + p \quad (3.1)$$

or

$$\bar{\nu} + p \longrightarrow l^+ + n \quad (3.2)$$

GENIE uses the Llewellyn-Smith model, in which the QE events are easily described by a set of Lorentz-invariant form factors.

Baryon Resonance Production Another neutrino interaction commonly happening is baryon resonance, which can be formulated as:

$$\nu + N \longrightarrow l + N^* \quad (3.3)$$

GENIE has employed Rein-Sehgal model that inherits the relativistic quark model of Feynman-Kislinger-Ravndal for baryon resonances. This model obeys $SU(6)$ spin flavor symmetry. Also, it includes the helicity amplitudes of 16 resonances to construct the cross sections for neutrino-production of baryon resonances. The model assumes that the lepton masses are neglected when calculating the differential cross section, but its effect on the phase space boundaries is taken into account.

Deep Inelastic Scattering Deep Inelastic Scattering (DIS) is a process where a neutrino scatters off a nucleon and produces hadron daughters. The Bodek-Yang model has been employed by GENIE which calculates the DIS at low Q zone.

The non-perturbative contributions to the inelastic cross section are calculated. The non-perturbative contributions include kinematic target mass corrections, dynamic higher twist effects, higher order Quantum Chromodynamic (QCD) terms, and nuclear effects on nuclear targets.

Neutrino-induced Hadron Production

The final state and their kinematics are calculated by the hadronization model in GENIE by the nature of neutrino nuclear interaction and event kinematics. This is vital for few-GeV range neutrino interaction simulation. NOvA applied the MINOS developed hadronization model, known as Andreopoulos-Gallagher-Kehayias-Yang (AGKY) model. The AGKY model employs the Koba-Nielsen-Olesen (KNO) model for low invariant mass region, on the other hand, for the higher mass region, the AGKY gradually switches to the PYTHIA/JETSET model to ensure the continuity of all simulated variables as a function of the invariant mass.

Intranuclear Hadron Transport

Neutrino-Nucleon interactions are simulated by AGKY model, then the daughter hadrons may interact with other nucleons which may produce second daughter particles, therefore significantly modify the observable distributions in sampling calorimeters. A subpackage, named INTRANUKE, has been applied to simulate the hadron intranuclear rescattering by a semiclassical model, intranuclear cascade model. In this model, it is assumed that hadrons have a typical mean free path about a few femtometers in the nucleus. They propagate through the nucleus with a reduced interaction probability. Then, the free cross-section

and the nucleons density are both considered to derive the final state interaction rates.

Nuclear Physics Model

The Relativistic Fermi Gas (RFG) nuclear model has been applied for simulating the nuclear physics in NOvA. The Bodek and Ritchie version of RFG has been used and modified to introduce short range nucleon-nucleon correlations.

3.4.3 Detector Simulation

The final step in the simulation chain is the detector simulation which deals with modeling the scintillation light production, transport, and conversion to electrical signals. The process can be classified into two steps: Photon Transport and Electronic Readout simulation.

Photon Transport

GEANT creates energy deposits related to the photons which scatter, reflect, and are absorbed by the fiber. Therefore, a ray-tracing algorithm has been developed. The algorithm calculates the expected light collection rate ahead of time. The input information to the algorithm are the measured scintillator response in NOvA detectors, PVC reflectivity, and the measured absorption spectrum of the fiber. The detector cells are expected to be identical due to the assumption that all cell-to-cell variations in scintillator and fiber response are taken out during calibration. Each side of the loop fiber receives half the collected photon. The lost light in the fiber is modeled based on the average light attenuation measure in the fiber during detector construction.

Electronic Readout

The last step of the detector simulation is the electronic readout. This process starts from the number of PEs saved by the APDs and simulates this to transform the raw signal to corresponding data. We then simulate the APD to have a flat 85% quantum efficiency and a gain of $100 \pm 5\%$ variation.

Chapter 4

Neutrino Interaction Classification and Signal Selection

“All you need is lots and lots of data and lots of information about what the right answer is, and you’ll be able to train a big neural net to do what you want.”

—Geoffrey Hinton

Producing and observing neutrinos require extraordinary resources and detectors. Therefore, how to extract the maximal information from the detected data is a crucial concern for neutrino-nuclear interaction related analyses. Based on the 2017 Summer dataset, two similar but different analyses, 2017 Official and 2017 Re-Analysis, applied different signal selection philosophies, have been performed. Analyses details and results will be presented in the following sections and chapters (This thesis predominantly focus on the 2017 Re-Analysis). In these analyses, we study the use of machine learning algorithms to assist with neutrino mixing parameter measurements. Signal selection is difficult for the NC Disappearance analysis because several types of backgrounds can mimic the signal’s characteristic signature. Furthermore, the vast majority of detected events

are not interactions of interest, so separating the signal from the background is crucial to obtaining results. Hence, particle interaction (event) classification in our analyses is an exciting challenge for machine-learning based data analysis. For the 2017 Re-Analysis, this process occurs in four phases:

1. Pre-selection, in which simple criteria are applied to eliminate detected events that are trivially known to be background or noise.
2. Training of a particle identification algorithm (PID), based on a boosted decision tree (BDT), to remove cosmic-ray related events in the FD. This is the principal original contribution to the NOvA NC/Sterile analysis resulting from the research work in this thesis, along with the Re-Analysis based on the NC/CC classifier described below.
3. Training of a PID algorithm to classify NC and CC neutrino interactions in both the ND and FD using BDT and convolutional neural network (CNN) techniques.
4. Final selection. The final cuts on the output of the Cosmic and NC/CC PIDs are defined by first obtaining a series of selection results with varying values of the PID cuts, and then choosing the final values by assessing which results achieve the highest sensitivity for measuring sterile mixing angles while retaining reasonable selection purity.

The two PIDs described above are based on a supervised learning methods, which means a corresponding label is given to an event during a training phase. This methodology has already been proven successful in other similar analyses in NOvA [40], and other high-energy physics projects [29]. The supervised approach is most effective for binary problems in which the classification accuracy

of the PID algorithm is closely correlated with the quality of the resulting data analysis, and the corresponding systematic uncertainties are minimal.

Chapter Organization

- Section 1: NC Signal Pre-Selection Cuts
- Section 2: NC-Dedicated Cosmic PID Training and Application
- Section 3: NC/CC Binary Classifier Training and Application
- Section 4: PID Tuning for Nus17 Official and Re-Analysis Selection
- Section 5: Predicted Near and Far Detector Spectra

Software Frameworks Used

PID training: For the training phase of the machine learning algorithms, the Toolkit for Multivariate Analysis (TMVA) [41] (version 4.2.0) has been employed, which is implemented in ROOT. Throughout the analysis, version 5.34 of ROOT was used.

Analysis Performance The analyses described in the following were performed in the Common Analysis Format Analysis (CAFAna) software framework, which was designed to support all neutrino analyses in NOvA [42].

Employed Datasets and Normalization Factors

Near Detector ND data and Monte Carlo files were produced by the NOvA production group as introduced in the Data Production Chapter, Chapter 3. An exposure of 8.04×10^{20} POT for ND data and Monte Carlo, which was scaled

to match the data POT, have been employed for the signal selection study. The same files and normalization factor (for Monte Carlo only) are also used for the extrapolated Far Detector prediction.

Far Detector As introduced in the Data Production Chapter, there are three types of FD files: Cosmic Background, Beam Simulation (FD Monte Carlo), and FD data.

- **Cosmic Background:** these are the cosmic trigger data, taken when the NuMI beam is not pulsing. These files were employed for the FD signal selection study (scaled to the equivalent livetime of 440s) and PID training;
- **Beam Simulation:** these are the Monte Carlo files, which were used for the FD signal selection study (normalized to 9.489×10^{20} POT), PID training, and to provide a predicted spectra to carry out fits to the data;
- **FD Data:** these are the detected events in FD, which provide a measured FD spectrum to compare to the predicted ones based on Monte Carlo. Also, the final selected cosmic events estimate is based on the out-of-time window (outside the NuMI beam pulse time window) of the FD data files. Furthermore, for the FD selection study, we employ the direct selection result, as opposed to the extrapolated result, discussed later in Chapter 5.

Signal Selection Standard Given the importance of the FD measured spectrum and the various backgrounds, we decide to apply an FD-Oriented selection standard. It means that FD event selection purity and efficiency have priority when we decide the pre-selection cuts and PIDs tuning results.

4.1 Neutral Current Signal Pre-Selection

4.1.1 Pre-Selection Cuts

Series of pre-selection cuts have been applied with two separate different goals. First, effective pre-selection cuts have been employed to eliminate events that are trivially known to be backgrounds or noise on both Near and Far detectors; Second, the pre-selection is used to find suitable variables with high discrimination power for machine-learning algorithm training. The pre-selection is divided into several stages as described below. The corresponding spectra and cut flow tables show the cuts separation ability. The main references for this section are Refs. [43, 44, 45].

4.1.2 Data Quality

The first type of sliced-based cut is the data quality group cut, which was developed to ensure proper data taking conditions in both detectors and has been applied during the data production process. These cuts are applied per beam spill. The spills which fail these cuts will not be included in POT accounting. We categorize these cuts into three main groups:

Timing cut

A timing cut is employed for FD cosmic data to ensure that the kept data is not too close to the data-taking timing window edge. For cosmic ray interactions within a $500 \mu\text{s}$ timing window, only events between $25 \mu\text{s} < t < 475 \mu\text{s}$ are saved.

Beam Quality Cuts

The details for the series of beam quality cuts were studied and set in Ref. [46].

Spills must meet the criteria listed in Table 4.1 to be included.

Beam Quality Parameter	Minimum	Maximum
Spill POT	2.00×10^{12}	
Horn Current	-202kA	-198kA
Beam X and Y position on target	0.02mm	2.00mm
Beam X and Y width	0.57mm	1.58mm
Time to nearest beam spill		0.5ns

TABLE 4.1: Beam quality cuts applied to each spill to ensure proper data taking conditions. This table is originally from tech note [46].

Data Quality Cuts

Two data quality cuts are applied to simulated and real data for both detectors. These two cuts are summarized in Table 4.2. Furthermore, a run filter has been employed to filter out runs recognized as bad runs and kept them in a list. Finally, there is a cut to remove the spills which do not include any continuous segments spanning more than four detector di-blocks.

Data Quality Parameter	Detector	Metric for Spill to Pass
Number of Missing DCMs	ND	= 0
Lights On Effect Hit Fraction	ND	≤ 0.45
Missing DCMs from LiveGeometry	FD	= 0
DCM Edge Match Fraction	FD	≥ 0.2

TABLE 4.2: Data quality cuts applied to each spill to ensure proper data-taking conditions. These cuts are motivated and described in Refs. [47, 48].

4.1.3 Analysis Mask

This simple cut is used to ensure that at least four diblocks of the FD are active for an event to be considered further in the analysis, in case this criterion is not applied at the reconstruction stage. Table 4.3 presents the number of FD events which pass this cut.

Cut Type	NC	ν_μ CC	ν_e CC	ν_τ CC	Cosmic
FD Analysis Mask	612.0	950.5	119.2	9.7	4.77×10^7

TABLE 4.3: The number of FD events after the Analysis Mask cut is applied

4.1.4 Cosmic Veto

Cosmic rays rejection cuts [49] are applied during FD MC and data production to speed-up the processing by removing obvious cosmic events. The Cosmic Veto is the first slice-level cut used. Once a slice is filtered out by these cuts, this slice is not further reconstructed and does not undergo PID selection. The applied version of the Cosmic Veto for our analyses was originally developed by the NOvA ν_μ CC group, and was therefore optimized for acceptance of CC interactions. As a result, about 6% of the NC signal is removed by this cut. In this thesis, we have developed an NC-optimized version of this cut. The following is the list of the Cosmic veto cuts [50].

1. If one slice is near to any of the detector edges, an angle-related variable limit has been required $|(\cos\theta_{beam})(\cos\theta_y + 1)| > 0.3$. This cut will eliminate very vertical events which are close to the edges of the detector;
2. If the hit number of one reconstructed cosmic track is more than 80% of the entire slice hit number, the requirement is that the forward and backward

projected distances of the cosmic track should be 35 cm or more away from detector edges;

3. The entire slice hit number is required to be between 20 and 250;
4. The plane number of the slice is required to be between 5 and 125;

Table 4.4 presents the number of FD events that pass the above two cuts. The surviving FD events are the ones which are the interactions of interest for further analysis.

Cut Level	NC	ν_μ CC	ν_e CC	ν_τ CC	Cosmic
FD Analysis Mask	612.0	950.5	119.2	9.7	4.77×10^7
+ Cosmic Veto	473.3	924.9	116.4	9.5	5.08×10^6

TABLE 4.4: Number of surviving events before and after application of Cosmic Veto at the FD.

4.1.5 Event Quality

Event quality cuts are applied to individual events to ensure that there are no obvious reconstruction failures and the events have enough reconstructed information to be analyzed correctly [51]. Among the four applied cuts, the first two require the presence of a reconstructed vertex (interaction vertex) and a reconstructed FuzzyK prong (particle track) object. The reconstructed vertex and prong information for each event will be used more extensively in later stages, such as input variables for the PID training. Therefore, the above two cuts have been applied to make sure that information is available. High-energy cosmic rays in the FD are more likely to trigger events with a high number of hits per plane and therefore cause so-called **FEB Flashers** in both ND and FD. A threshold on the variable is placed to remove these events. Likewise, detected events

which have a low number of contiguous planes are most often produced by very vertical cosmic rays in the FD. As a result, it was required that the event must traverse at least two continuous planes. Otherwise, they are categorized as junk slices. The full suite of event quality cuts is summarized in Table 4.5, listing the exact cut values used.

Event Quality Metric	Metric for Event to Pass
Number of reconstructed vertex object	>0
Number of reconstructed prong objects	>0
Number of hits per plane	<8
Number of contiguous planes	>2

TABLE 4.5: Event Quality cuts applied to individual events to ensure properly reconstructed quantities. This table is taken from tech note [46].

Also, the numbers of survived events before and after application of this suite of cuts are listed in Table 4.6. Figure 4.1 shows the FD energy spectra and Fig. 4.2 shows the ND one.

Cut Level	NC	ν_μ CC	ν_e CC	ν_τ CC	Cosmic
FD Analysis Mask	612.0	950.5	119.2	9.7	4.77×10^7
+ Cosmic Veto	473.3	924.9	116.4	9.5	5.08×10^6
+ Event Quality	443.8	901.9	114.3	9.2	4.34×10^6

TABLE 4.6: The number of FD events that pass the event quality cuts. This table is originally from tech note [46].

Cut Level	NC	ν_μ CC	ν_e CC
ND: ($\times 10^3$)			
Data Quality	15466	114020	1660
+ Event Quality	11919	103828	1301

TABLE 4.7: The number of ND events that pass the event quality cuts.

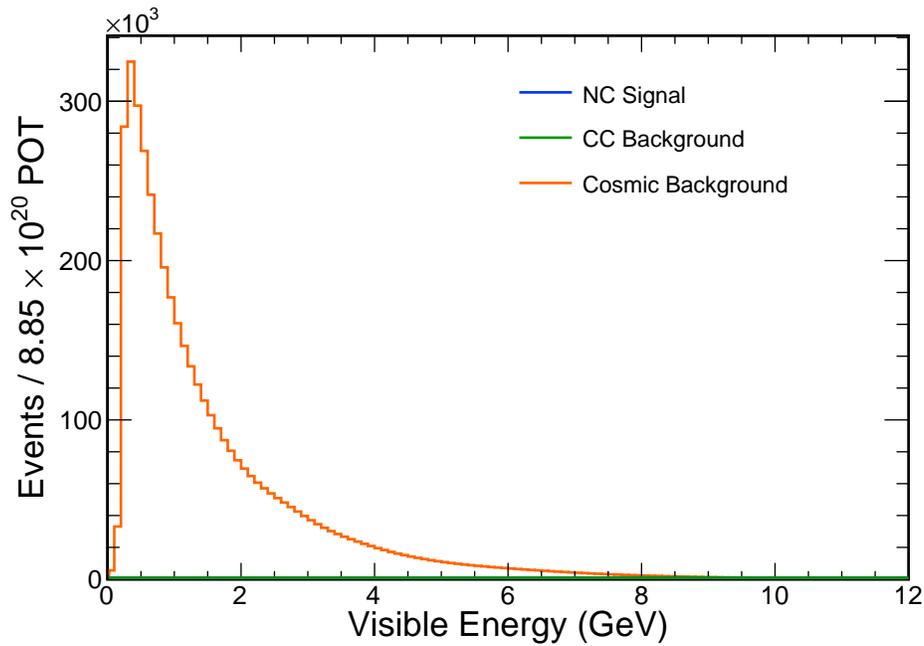


FIGURE 4.1: FD Energy spectra after applied Event Quality cuts.

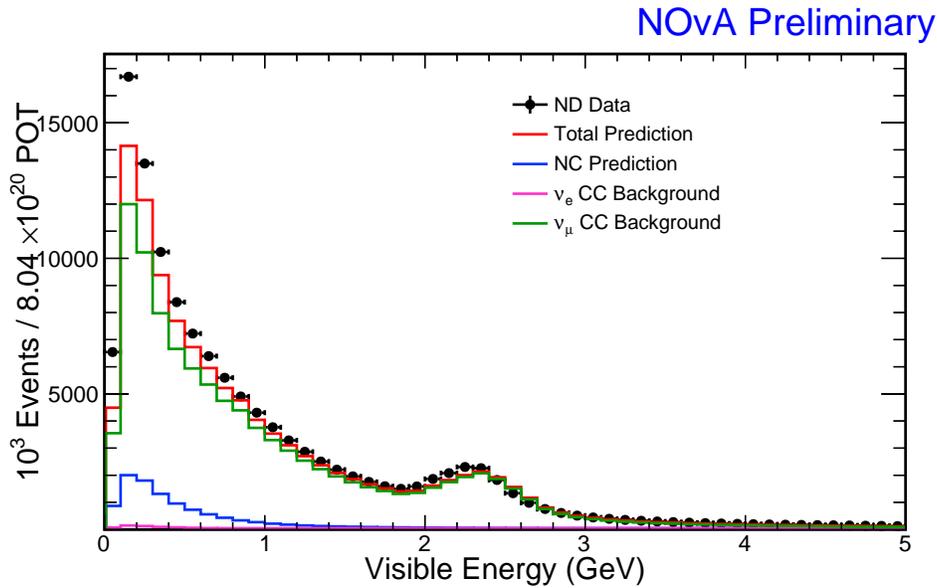


FIGURE 4.2: ND Energy spectra after applied Event Quality cuts

4.1.6 Fiducial Volume and Containment

Containment cuts (for both ND and FD) and fiducial volume cut (only applied on ND) have been employed to reject events originating outside of each detector

volume, and to guarantee the detected interactions arising inside the detector do not include activity that escapes the detector. The containment cut considers the distance of all reconstructed FuzzyK prongs of one event from the six edges (Top, Bottom, Front, Back, East, West) of the detector. For each detector edge, the limit is placed on the minimum distance of the start/stop point from the detector face in question, after considering all prongs. The fiducial volume cut is based on the position of the reconstructed interaction vertex. The specific suite of cuts is set separately for the two detectors.

Far Detector

Only the containment suite of cuts is applied to the FD. They not only improve event reconstruction, but can also help to reject cosmic events, especially the ones that enter the FD from the top face. To highlight the abundance of cosmic rays entering from the top, Fig 4.3 shows the distribution of the minimum distance of all FuzzyK prongs start/stop positions at the Top face of the FD. That also explains why a corresponding stringent limit was used for the detector top, and a looser one was used for the detector bottom. The complete suite of containment cuts are summarized in Table 4.8, listing the exact cut values.

Reconstructed Quantity	Metric for Event to Pass
Minimum Distance from all prong start/stop distances to detector Top	>100cm
Minimum Distance from all prong start/stop distances to detector Bottom	>10cm
Minimum Distance from all prong start/stop distances to detector East	>50cm
Minimum Distance from all prong start/stop distances to detector West	>50cm
Minimum Distance from all prong start/stop distances to detector Front	>50cm
Minimum Distance from all prong start/stop distances to detector Back	>50cm

TABLE 4.8: Containment Cuts Applied to Events in the FD.

Also, the numbers of events selected before and after application of this suite of cuts are listed in Table 4.9. Figure 4.4 shows the FD energy spectra.

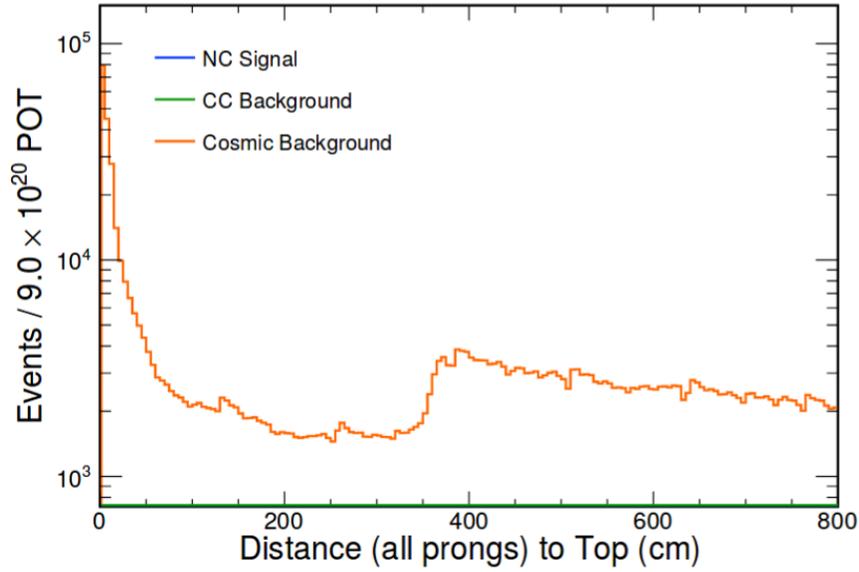


FIGURE 4.3: Distribution of the minimum distance of all prong start/stop positions to the Top of the detector (before the containment criteria are applied).

Cut Level	NC	ν_μ CC	ν_e CC	ν_τ CC	Cosmic
FD Analysis Mask	612.0	950.5	119.2	9.7	4.77×10^7
+ Cosmic Veto	473.3	924.9	116.4	9.5	5.08×10^6
+ Event Quality	443.8	901.9	114.3	9.2	4.34×10^6
+ Containment	290.6	263.8	73.6	5.1	21123.8

TABLE 4.9: Number of survived events before and after application of containment cuts at the FD.

Near Detector

The NOvA near detector is located about 100 meters underground, so there is no large rate of cosmic ray interactions to eliminate. On the other hand, plenty of events interact in the rock material outside of the ND that can leak daughter particles into the detector. Moreover, the small size of the ND means that many detected events originating inside the detector have daughter particle(s)

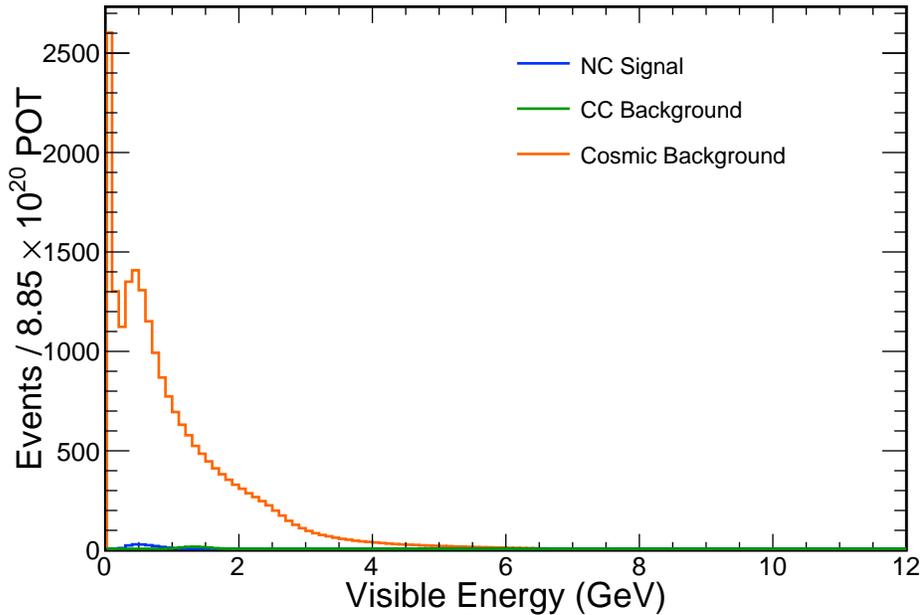


FIGURE 4.4: FD Energy spectra after applied containment cuts.

that escape the detector. The ND fiducial volume and containment cuts are intended to combat these effects. In the same fashion as for the FD, the ND containment cut considers the distance of all FuzzyK prongs in one slice from the six edges (*Top, Bottom, Front, Back, East, West*) of ND. For each detector face, the minimum distance of the start/stop point from the detector face in question is required to be higher than 25cm symmetrically for all detector faces, after considering all FuzzyK prongs. The fiducial volume cuts on the X and Y coordinates of the interaction vertex were applied symmetrically, with a modestly large cut to reject events that originate in the rock zone. The vertex cut on Z coordinate removes a large part of the detector to reject rock events which leak into the front of the ND. The full suite of cuts is summarized in the table. The full suite of cuts are summarized in Table 4.10.

Also, the numbers of ND events selected before and after application of this

Reconstructed Quantity	Metric for Event to Pass
Reconstructed Vertex X Coordinate	$ vtx(X) \leq 100\text{cm}$
Reconstructed Vertex Y Coordinate	$ vtx(Y) \leq 100\text{cm}$
Reconstructed Vertex Z Coordinate	$150\text{cm} \leq vtx(Z) \leq 1000\text{cm}$
Minimum Distance from all prong start/stop distances to detector Top	$>25\text{cm}$
Minimum Distance from all prong start/stop distances to detector Bottom	$>25\text{cm}$
Minimum Distance from all prong start/stop distances to detector East	$>25\text{cm}$
Minimum Distance from all prong start/stop distances to detector West	$>25\text{cm}$
Minimum Distance from all prong start/stop distances to detector Front	$>25\text{cm}$
Minimum Distance from all prong start/stop distances to detector Back	$>25\text{cm}$

TABLE 4.10: Containment and Fiducial Volume cuts applied to events in the ND.

suite of cuts are listed in table 4.11. Figure 4.5 shows the ND energy spectra after the applied fiducial volume cuts and Fig. 4.14 shows the spectra after the applied containment cuts.

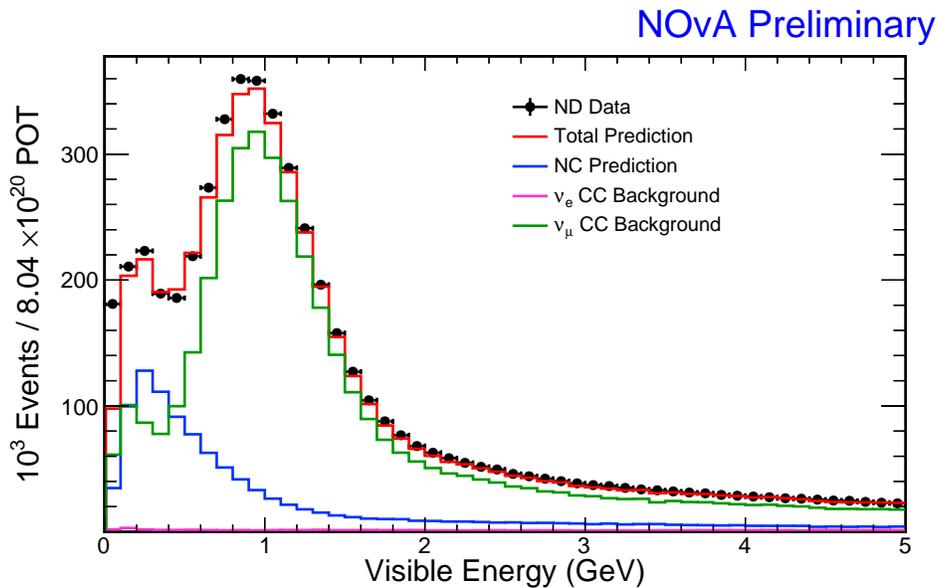


FIGURE 4.5: ND Energy spectra after applied Fiducial Volume cuts.

Cut Level	NC	ν_μ CC	ν_e CC
ND Data Quality	15466	114020	1660
+ Event Quality	11919	103828	1301
+ Fiducial Volume	1143	4402	99
+ Containment	630	853	39

TABLE 4.11: Number of ND events before and after application of containment and Fiducial Volume cuts.

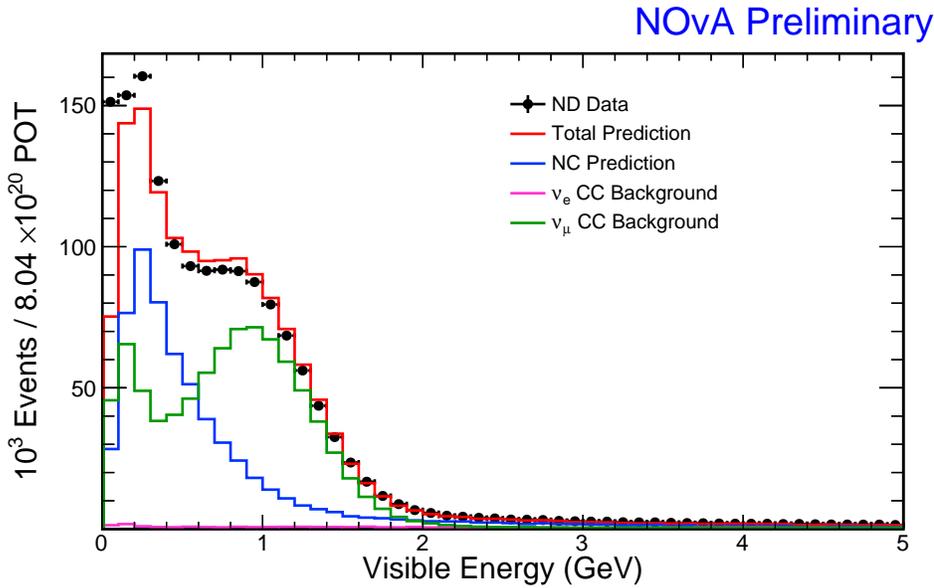


FIGURE 4.6: ND Energy spectra after applied Pre-Selection cuts.

4.1.7 Cosmic Rejection Pre-Selection Cuts

Cosmic rejection pre-selection cuts are only applied to the FD events, as the ND is situated deep underground. Therefore, the cosmogenic backgrounds are negligible at the ND selection. On the other hand, the NOvA FD is placed on the surface, about 300 m above sea level, therefore it is exposed to 148 kHz of cosmogenic particles. The most challenging components of these external events are the most horizontal secondary cosmogenic muons and neutrons generated in the FD overburden, which can mimic the NC's characteristic signature. Before applying further selection cuts, NOvA reconstructs 74,000 cosmogenic events

for one beam neutrino event in the $10\mu\text{s}$ beam spill (time) window at the FD on average. This is recorded by the NC disappearance 2016 summer analysis, the first search for sterile neutrinos within. A large portion of the FD volume, top 5 m in the YZ-view, was removed to decrease the cosmogenic events for NC first analysis. However, NC signals in the removed zone of the FD cannot be counted. Meanwhile, the 2016 analysis employed a CC-dedicated cosmic rejection classifier, which misclassifies part of the NC signal as cosmic events. Reducing cosmic backgrounds, therefore, becomes one of the main goals of these analyses.

As one of the main objectives of the work described in this thesis, a NC-dedicated cosmic rejection classifier has been produced for the 2017 summer analyses. The classifier uses a Boosted Decision Tree (BDT) algorithm trained on 13 discriminating variables, such as the CVN Cosmic identifier [52, 53], shower information, and slice hit information. The training details of the algorithm are introduced in the next section and outlined in Ref. [54]. Except for the specially trained algorithm, three other cuts, listed in table 4.12, have been employed as the FD cosmic rejection pre-selection cuts, which not only decrease the background, but also help the algorithm to focus on rejecting NC-like cosmic events in the training phase.

Cosmic Photon Rejection

We employ the *Backward Photon Cut* to reject cosmic photons entering the FD from the backside, where the overburden is thin. The cut was originally designed to separate ν_e CC events [50] in the back of the detector from the backward photon background. It has been proved that the cut works equally well for separating the cosmic photons from the NC signal. The shape feature of the electromagnetic showers from electrons is that it starts from narrow ionizing hits,

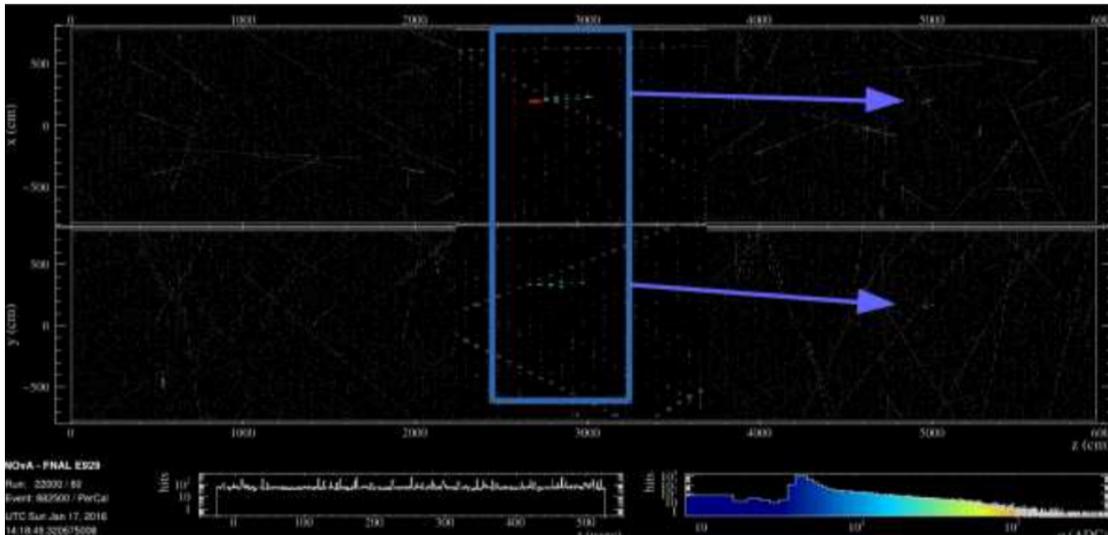


FIGURE 4.7: Example of a cosmic ray event selected as an NC interaction by the 2016 analysis.

expand in the middle and grow sparse towards the end of the electromagnetic shower, with an entirely reversed profile for backward cosmic photon showers. A sparseness asymmetry, defined as the variance divided by the sum of hitless planes in the first eight and the last eight planes of showers, is used to classify the beam events and the backward photon.

Combined Transverse Momentum Fraction and Shower Position Cut

Further effort has been made to classify the remaining cosmic events with an event transverse momentum fraction related cut. In more detail, the cut is based on carving out space in the transverse momentum versus shower maximum position in Y coordinate [43]. The idea is to clean up cosmic backgrounds by their different distribution in the 2-dimension space of the transverse momentum fraction and maximum shower position in the Y coordinate, as shown in Figs. 4.15, 4.12, and 4.13. The cut is aimed at removing some of the horizontal-like neutron-induced cosmogenic activity at the top of the detector [43].

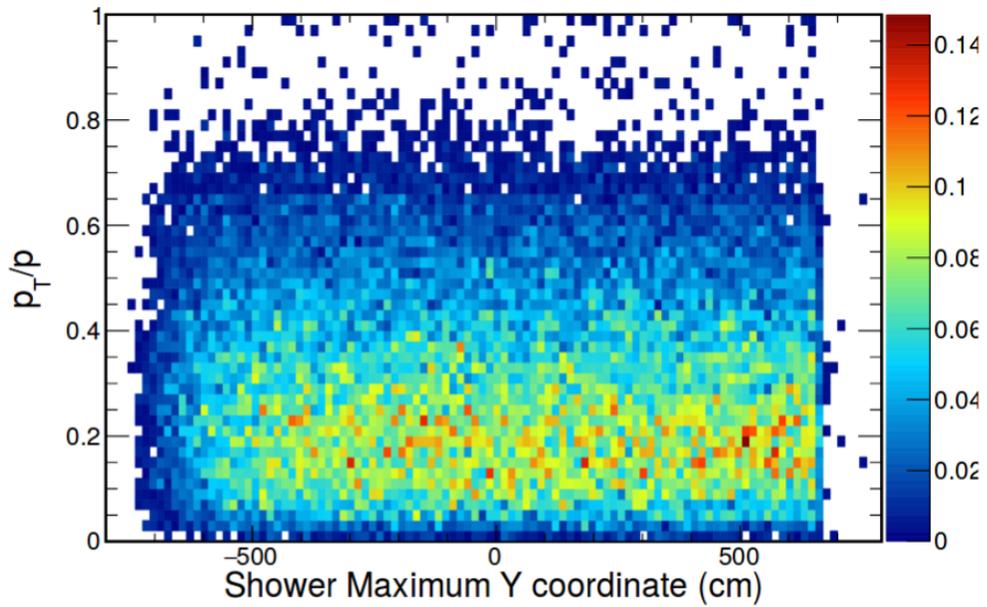


FIGURE 4.8: NC 2D distribution of the maximum Y-coordinate of showers versus the transverse momentum fraction [43]

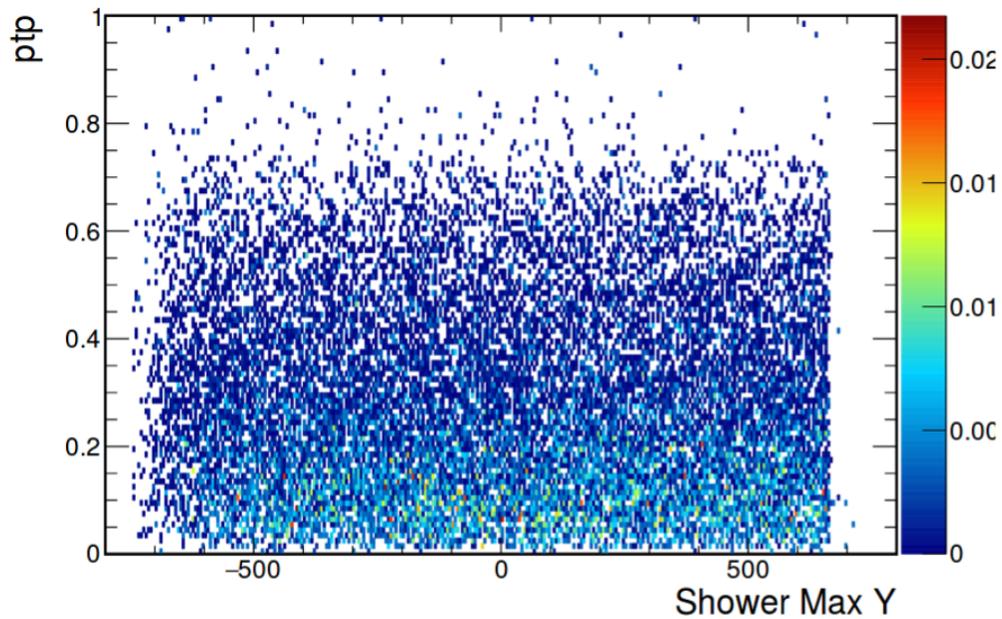


FIGURE 4.9: CC 2D distribution of the maximum Y-coordinate of showers versus the transverse momentum fraction [43]

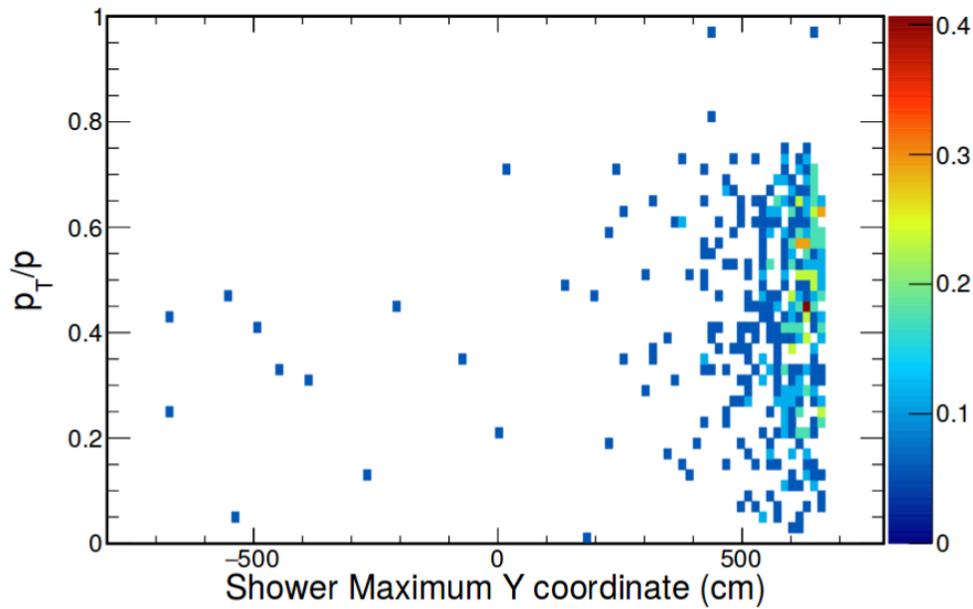


FIGURE 4.10: Cosmic 2D distribution of the maximum Y-coordinate of showers versus the transverse momentum fraction [43]

Time Gap Cut

By studying the NC First analysis selected results, we identified several of the events have a coincident slice with cosmogenic activity [55], as shown in Fig. 4.7. These slices are indicated as a delayed time window separation of the cosmogenic muons and their corresponding hit showers. Then, the shower is reconstructed as a separate slice. So, this cut looks at the time discrepancy between the slice being considered and its nearest-in-time slice, the distance of the nearest-in-time slice to the top face of FD, and the minimum distance between a slice and its nearest-in-time slice. After selecting a tight region with a prominent cosmic peak in the time difference between slices space, a bounding box is then chosen in the distance from top and distance to nearest-in-time slice 2D space [43].

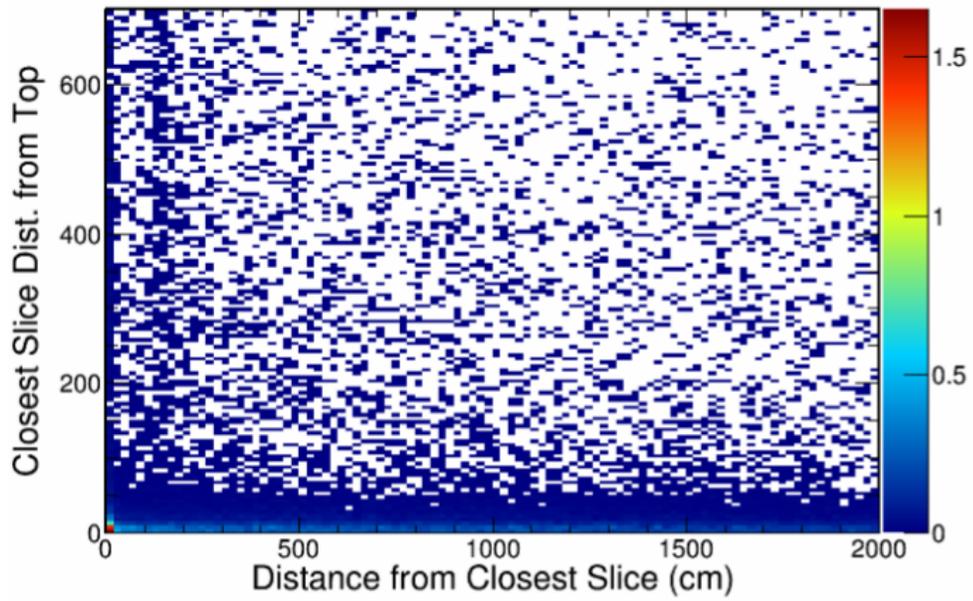


FIGURE 4.11: NC 2D distribution of the distance from closest slice versus the closest slice distance from the top [43].

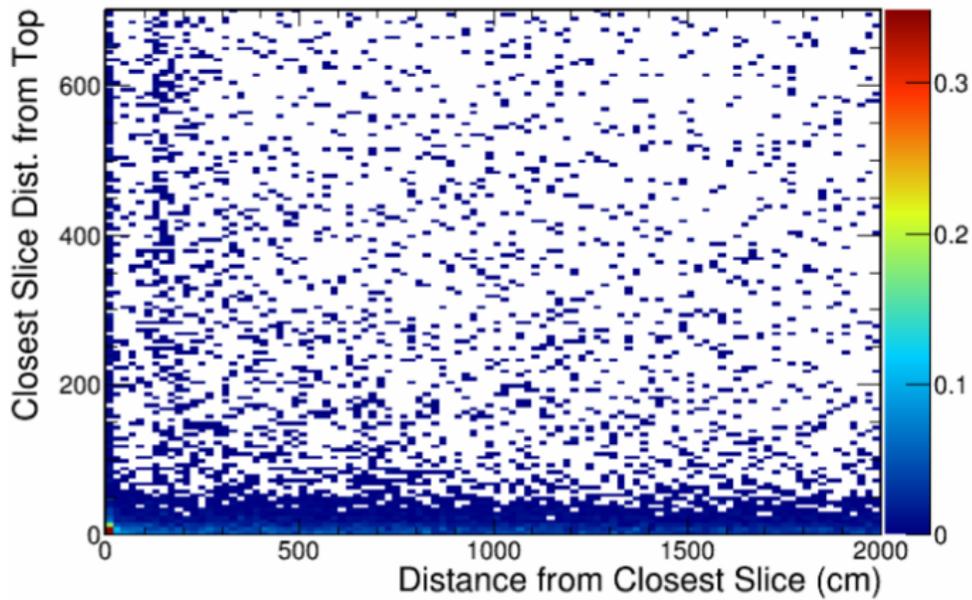


FIGURE 4.12: CC 2D distribution of the distance from closest slice versus the closest slice distance from the top [43].

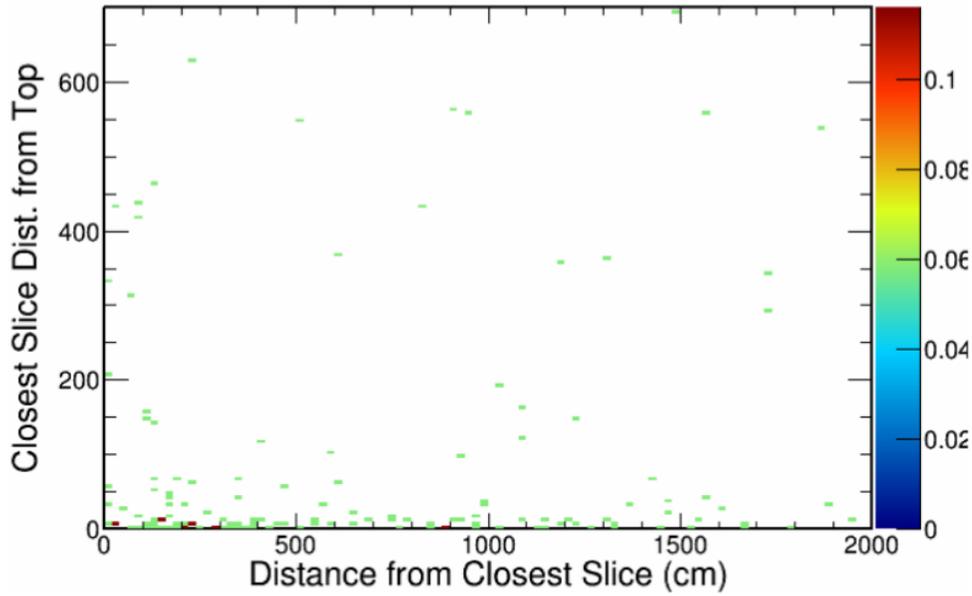


FIGURE 4.13: Cosmic 2D distribution of the Distance from Closest Slice versus the Closest Distance from Top [43].

Reconstructed Quantity	Metric for Event to Pass
Backward Photon Cut	$(kDistAllBack \geq 200 \text{ cm} \ \&\& \ kSparsenessAssym \geq 0.1)$ $\ (kDistAllBack < 200 \text{ cm})$
p_T/p and Maximum Shower Y coordinate	$((maxY > 580\text{cm} \ \&\& \ p_T/p > 0.2)$ $\&\& \ maxY > 540\text{cm} \ \&\& \ p_T/p > 0.4))$
Nearest-in-time Slice Time Gap/Distance	$!((-150 \ \mu\text{s} < \text{time gap} < 50 \ \mu\text{s}) \ \&\& \ (distance \ to \ top < 100 \ \text{cm} \ \&\& \ \text{closest slice distance} < 500 \ \text{cm}))$

TABLE 4.12: FD Cosmic Rejection Pre-Selection Cuts.

4.1.8 NC/CC Separation Pre-Selection Cuts

Another primary signal selection challenge is to categorize beam events into NCs and CCs with reasonable separation purity and efficiency. To achieve this purpose, two different but related PIDs have been separately trained based on convolutional neural networks and boosted decision tree algorithms. The two PIDs will be introduced in the next two sections. Furthermore, three independent cuts have been employed to fulfill this aim. They are listed in Table 4.14.

Cut Level	NC	ν_μ CC	ν_e CC	ν_τ CC	Cosmic
FD:					
Analysis Mask	612.0	950.5	119.2	9.7	4.77×10^7
+ Cosmic Veto	473.3	924.9	116.4	9.5	5.08×10^6
+ Event Quality	443.8	901.9	114.3	9.2	4.34×10^6
+ Containment	290.6	263.8	73.6	5.1	21123.8
+ Backward Photon Cut	286.3	256.7	73.2	5.0	20406
+ Time Gap Related Cut	285.2	255.9	73.0	5.0	20016
+ Transverse Momentum Related Cut	266.4	247.8	70.0	4.8	14832

TABLE 4.13: Number of FD events after applied (cumulative) Transverse Momentum Related Cuts.

Number of Hits

A limit has been placed on the number of hits for every single slice/event. The number is required to be higher than (or equal to) 25, which means there should be at least 25 hits in one event so that it may be considered as an NC signal.

Visible Energy

A cut on the calorimetric energy has been used to require all events must have reconstructed energy greater than 0.25 GeV for both detector selections. The above two cuts are meant to steer clear of the very low-efficiency region where threshold effects are present.[43]

ND p_T/p Cut

An additional cut is applied at the ND, on the transverse momentum fraction, to reject rock events. The cut has been decided so that it does not carve out any space in the Far Detector transverse momentum region [43].

Selection Parameter	Metric for Event to Pass
Number of Hits	≥ 25
Visible Energy	> 0.25 GeV
ND p_T/p	≤ 0.8

TABLE 4.14: NC/CC Separation Pre-Selection Cuts.

Cut Level	NC	ν_μ CC	ν_e CC	ν_τ CC	Cosmic
FD Analysis Mask	612.0	950.5	119.2	9.7	4.77×10^7
+ Cosmic Veto	473.3	924.9	116.4	9.5	5.08×10^6
+ Event Quality	443.8	901.9	114.3	9.2	4.34×10^6
+ Containment	290.6	263.8	73.6	5.1	21123
+ Cosmic Pre-Selection Cuts	266.4	247.8	70.0	4.8	14832
+ Number of Hits Cut	240.2	245.7	69.7	4.8	12780.4
+ Energy Cut	236.6	230.2	68.1	4.7	10675

TABLE 4.15: Number of FD events before and after applied (cumulative) energy cuts.

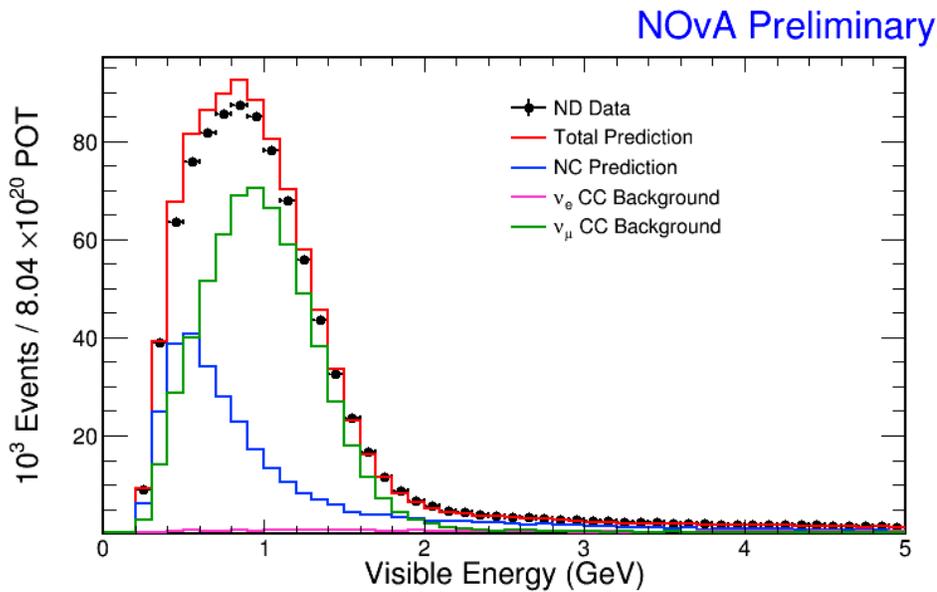


FIGURE 4.14: ND Energy spectra after applied containment cuts.

Cut Level	NC	ν_μ CC	ν_e CC
ND Data Quality	15466	114020	1660
+ Event Quality	11919	103828	1301
+ Fiducial Volume	1143	4402	99
+ Containment	630	853	39
+ ND p_T/p	474	727	35
+Number of Hits Cut	339	643	33
+ Energy Cut	329	640	32

TABLE 4.16: Number of ND Events after applying all cuts.

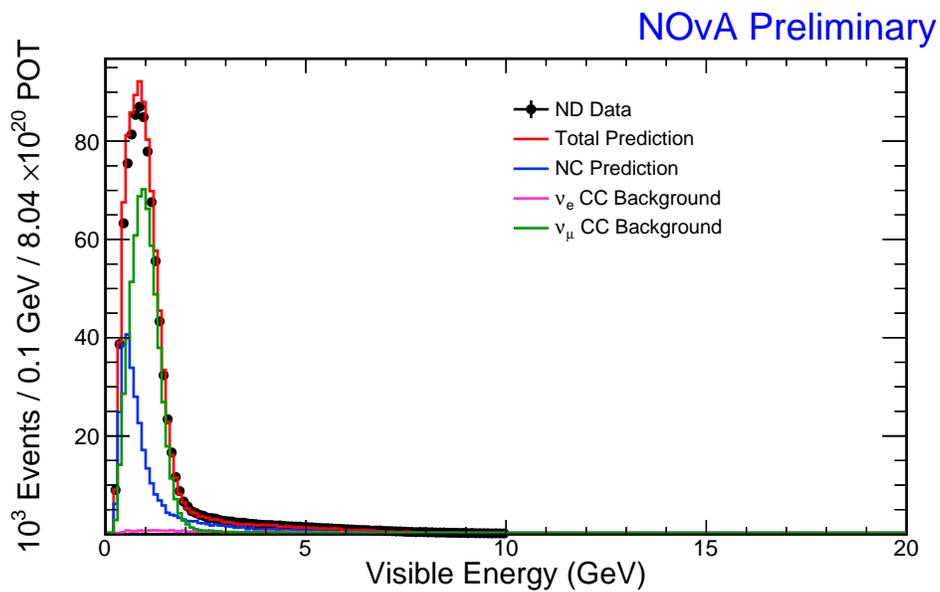


FIGURE 4.15: ND Energy Spectra after applying Distance from Closest Slice Cut [43].

4.2 NC-Dedicated Cosmic Rejection PID

4.2.1 Cosmic Rays and NOvA Far Detector

Apart from the high-energy particles (predominantly protons) associated with solar flares, cosmic rays come from the external solar system, even from distant galaxies[56]. These highly energetic particles, also called primary cosmic rays, are created in various astrophysical processes. Of all primary cosmic rays, about 99% are the nuclei of well-known atoms with lifetimes of order 10^6 years or longer, and the other 1% are principally electrons [57]. The nuclei consist of 90% protons (i.e., hydrogen nuclei), 9% alpha particles (identical to helium), and the other 1% are the nuclei of heavier elements, such as carbon, oxygen, iron, and other nuclei synthesized in stars [58, 59, 60]. Stable particles of anti-matter, such as positrons and antiprotons, account for a tiny proportion of the total abundance of cosmic rays. Magnetic monopoles are also expected as one of the elements of cosmic rays. NOvA's large and sensitive far detector gives scientists a unique ability to search for this hypothetical particle [61]. These primary cosmic rays can scatter at the top of the Earth's atmosphere, by colliding with atoms, and produce a cascade of ionized particles and electromagnetic radiation [62, 63], which are also called secondary cosmic ray air showers, as shown in Fig. 4.16.

When a primary cosmic ray, which could be a proton or a nucleus, strikes an atom's nucleus in the atmosphere, it produces various energetic daughter particles. The most likely daughter particles of the leading level interaction are pions, kaons, and baryons. However, these are not stable particles. They quickly decay in the air into the secondary daughter particles, such as muons, protons,

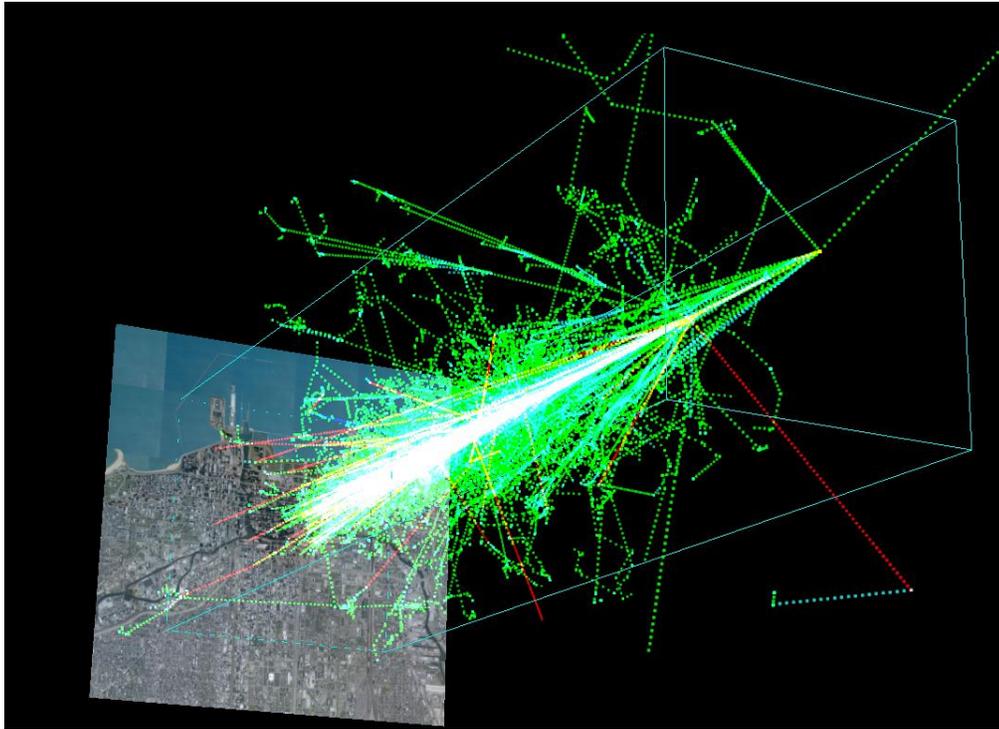


FIGURE 4.16: A simulated air shower created by a primary cosmic proton (1 TeV energy).

neutrons, and x-rays. The π^\pm primarily decay into neutrinos and muons, with a branching fraction of 0.999877, in the processes $\pi^+ \rightarrow \mu^+ + \nu$ and $\pi^- \rightarrow \mu^- + \bar{\nu}$. These are the main source of muons and neutrinos in the air shower. The π^0 preferentially decays into two photons, with a branching ratio of 0.98823. The daughter photons, therefore, form an electromagnetic cascade by producing more daughter particles, including photons, protons, antiprotons, electrons, and positrons[62].

Then, the secondary cosmic rays, such as protons, muons, alpha particles, pions, electrons, as well as neutrons, may rain down to reach the earth's surface, as shown in Fig. 4.17, and therefore create the main backgrounds in the NOvA Far Detector for the NC disappearance analysis. The secondary cosmic rays

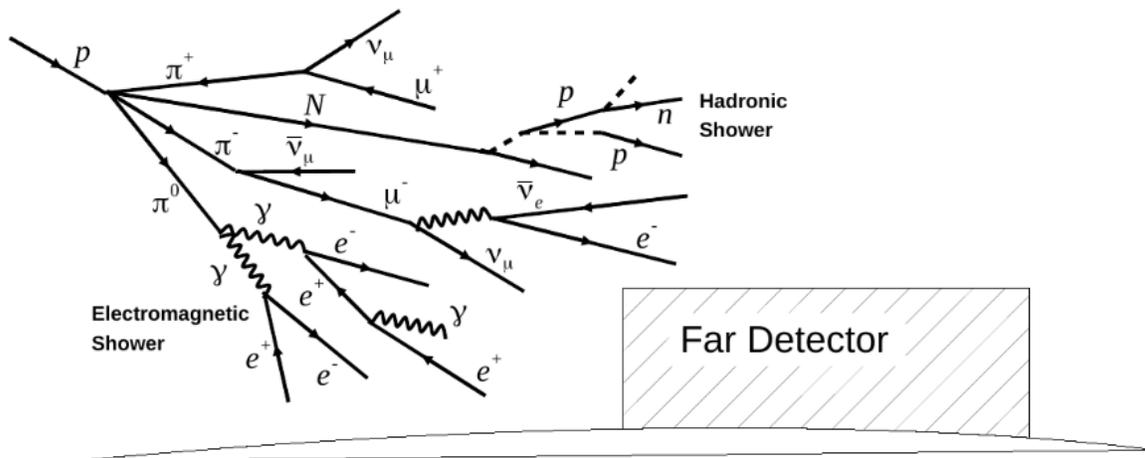


FIGURE 4.17: Secondary Cosmic Rays and the NOvA Far Detector [64].

dose is mainly from muons, electrons, and neutrons, with a dose rate that depends on the geomagnetic field, altitude, and solar cycle and therefore varies in different areas of the world[63]. The NOvA Far Detector is sited on (49°N, 92°W) and exposed to the above particles in 148 kHz. These cosmic rays, in particular, the most horizontal ones and cosmogenic neutrons produced in the FD overburden can mimic NC interactions as shown in Fig 4.18, so they are critical backgrounds to our analyses. On average, NOvA reconstructs one beam event from the NuMI beam for a total of 74,000 cosmogenic events reconstructed in the FD, which indicates that we are facing a binary classification problem with extensive background contamination of the signal.

4.2.2 Machine Learning Algorithm Selection

A core problem in our analyses is how to correctly categorize the particle interactions recorded in the NOvA detectors as signal and backgrounds. The first step

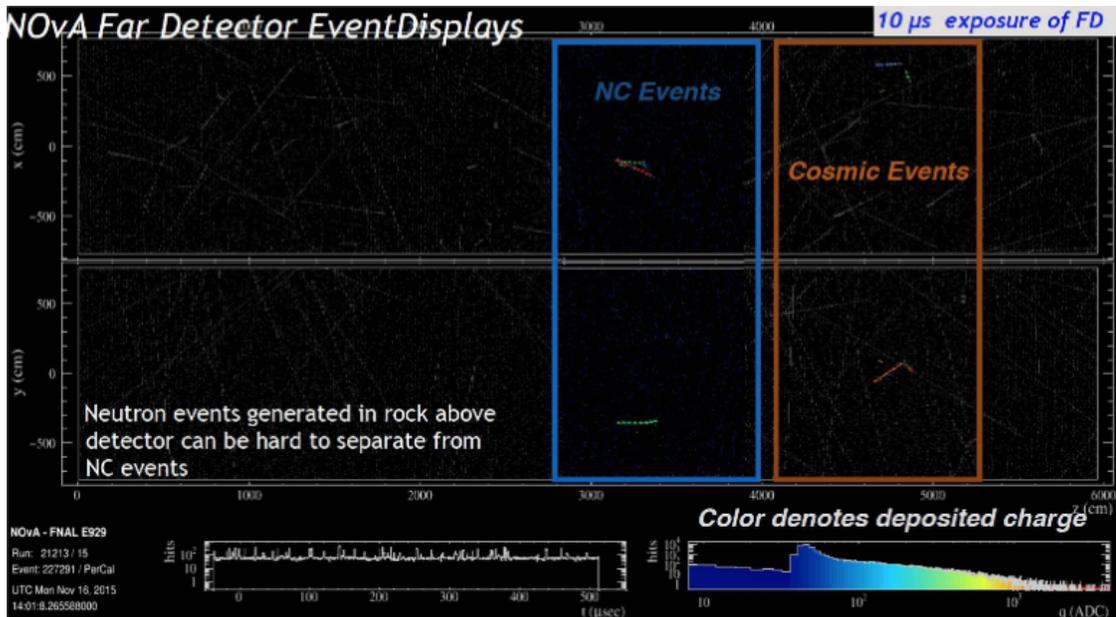


FIGURE 4.18: FD NC and cosmic ray interaction comparison.

to perform the categorization process is to reconstruct high-level physics components such as hit clusters, showers, tracks/prongs associated with particle interactions recorded by the detector and summarize the energy and timing information of these objects with a few quantities. Commonly, these quantities, also called event variables, are directly selected to separate signal from backgrounds by placing limits on them, such as the pre-selection cuts which have been employed in our analyses. While this direct selection method has been proven to work well, it is prone to two potential failings: 1) reconstruction uncertainty when reconstructing high-level physics features from the raw data (recorded hits) can lead to incorrect categorization of the interaction types; 2) given the enormous backgrounds (compared to the signal), the NC selection in our analyses is too complicated to use linear cuts on event variables as a classification method. This makes a detailed study of the classification problem necessary. Various useful (machine learning) algorithms exist to differentiate and analyze

diverse types of events by combining multiple physics features into the classification algorithms [29], for example, boosted decision trees (BDT) [65], and neural networks (NN) [66, 67].

The current study refers to two dissimilar but popular machine learning algorithms: 1) Convolutional Neural Network (CNN); the NOvA CVN group [52] produced a PID based on this algorithm; and 2) Boosted Decision Trees [68], which is the algorithm used for the NOvA cosmic rejection PID. The output of the CNN will be fed into a BDT as a root discriminating variable. This idea is supported by the successful experience using **Kaggle Competitions**, which is a professional platform built for data science. Boosted decision tree algorithms gain favor in particle physics because the selection logic they provide is more straightforward to be understood [69], as demonstrated in the simple example shown in Fig. 4.19. A decision tree is a binary structured classifier, which includes a series of decisions, so-called decision nodes, on one input data, where each decision is made on a single discriminating event-variable. The sorting process will terminate at a leaf node, which labels the input data as signal or background. The input data phase space is therefore nonlinearly cut into binary (signal and background) regions. This algorithm is widely known to have high performance in nonlinear classification problems. Also, it requires little tuning of the configuration to achieve satisfactory results. Therefore, the NOvA ν_μ CC disappearance group employed this algorithm to separate the ν_μ CC (signal) from the cosmic event background in the FD for their analysis based on 2016 Summer data set.

On the other hand, a decision tree regularly suffers from overtraining issues,

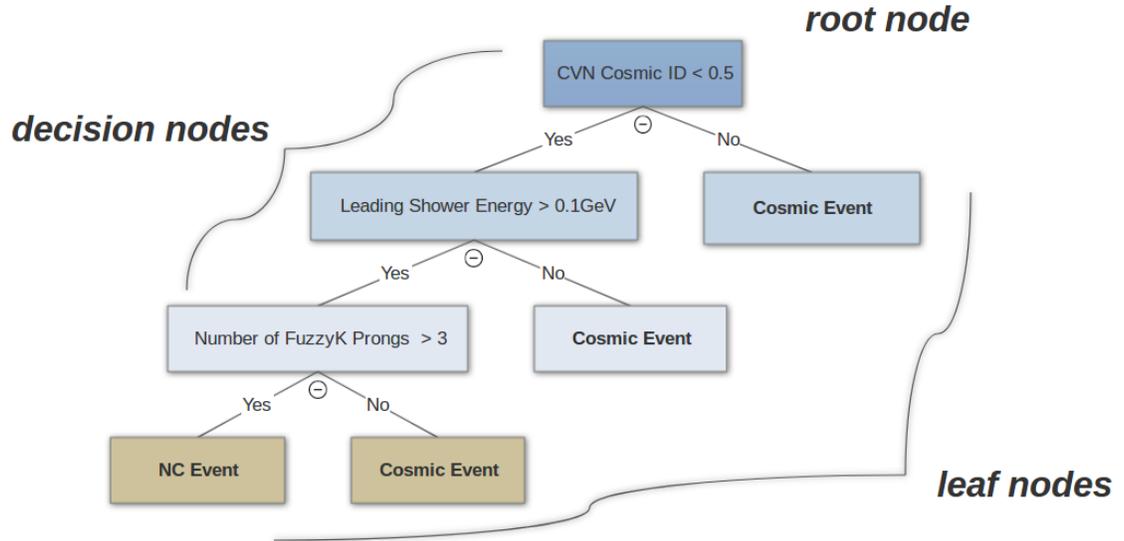


FIGURE 4.19: A simple Decision Tree example for NC-dedicated Cosmic rejection. The employed decision nodes' variables will be detailed in the following sections.

which means the trained algorithm is overly optimized to learn statistical fluctuations from the training input data, which are not real features of the distributions. Another critical challenge is to select suitable variables with high discrimination power to feed into the algorithm. Given the pros and cons of decision tree algorithms, and following the ν_μ CC-dedicated cosmic rejection PID, we trained and tested a series of NC-dedicated cosmic rejection PIDs to isolate neutral current interactions from cosmic interactions with the tree-based algorithm. PID training and application phases details will be introduced in the following parts of this section.

4.2.3 Training Phase

Combining algorithms

In challenging binary classification problems with a high requirement for majorization, it is often useful to combine a variety of machine learning models [65] [69]. We combine two types of algorithms, decision tree and boosting, for producing a NC-dedicated cosmic rejection PID. A decision tree learns from input data to create an easy for humans to understand result based on a series of if-then-else decision rules. However, one decision tree is susceptible to statistical fluctuations of the training input. Boosting, on the other hand, is an algorithm which can increase the stability concerning statistical variations in the training input by sequentially applying one (or more) chosen algorithm(s) to reweighted (boosted) versions of the training input and then taking a weighted majority vote of the sequence of the classifiers thus produced. Also, this method can improve the final classification performance by combining the individual classifiers. Consequently, we trained the PID by boosting decision trees to create the decision tree forest.

Boosting Algorithm The boosting algorithm, a method that can be applied to any machine learning algorithm, has been employed to improve the classification performance of the tree-based algorithm. The core concept of the boosting algorithm is to repeat the same or different classifying algorithm(s) several times to obtain a succession of classifiers, each of them focusing on the data which have been misclassified by the previous one(s). Among the three types of boosting methods provided by TMVA, we employed and tested two of them:

- Adaptive boosting :

Also called AdaBoost [70], it is the first realization of boosting algorithm that achieved great success in application. It produces a stronger classifier by training one or more algorithm(s) (decision tree in our case) sequentially in a way that they learn from the errors of previous ones [71]. The first tree was grown based on the initial configuration, and all the training input events are weighted with the same value. The sum of event weights remains constant in the training phase for all the trees. The fraction of misclassified events, f_{error} , is determined and all misclassified events will be reweighed with a higher event weight, α , which is defined as $(1 - f_{error}) / f_{error}$. Then, the event weights of the entire training input are renormalised to keep the sum of weights constant. A second tree is then trained with the boosted (reweighted) training input [72, 73]. This process will continue until we stop it by setting a fixed number of trees or a value for the signal purity in leaf nodes. The summed classifiers is then referred to as a forest [74]. The prediction result of one classifier is defined as $h(\mathbf{x})$, where \mathbf{x} is the tuple of event variables, which is encoded as 1 for signal and -1 for background. The final prediction for one event, $y_{Boost}(\mathbf{x})$, is therefore given by

$$y_{Boost}(\mathbf{x}) = \frac{1}{N} \times \sum_{i=1}^N \ln(\alpha_i) \cdot h_i(\mathbf{x}) \quad (4.1)$$

where N is the number of trees. AdaBoost performance is often further improved by setting a “slow learning” and allowing a more significant number of boost steps instead. In TMVA framework, the AdaBoost learning rate is setting by the hyperparameter β by redefining the boost weight

from α to α^β . It is one of the main hyperparameters which can be controlled using the configuration option string of the selected algorithm to be boosted, which will be described below.

- Gradient boosting :

AdaBoost was further developed by Jerome H. Friedman in what is termed Gradient Boosting. The basic idea is to consider boosting as an additive expansion approach. The final prediction function, $F(x)$, is supposed to be a weighted sum of parameterized base functions $f(x; a_m)$, each one corresponding to one trained classifier (decision tree for our analysis).

$$F(x; P) = \sum_{m=0}^M \beta_m f(x; a_m); P \in \{\beta_m; a_m\}_0^M \quad (4.2)$$

From a technical point of view, any classification algorithms could benefit from this approach, but the decision tree algorithm can profit the most. Therefore, the boosting procedure is now focused on how to adjust the parameters \mathbf{P} to minimize the divergence between the additive classifiers prediction $F(x)$ and the labeled value y obtained from the training sample. The loss function $L(F, y)$, which measures the divergence, can adequately determine the boosting approach. The AdaBoost used the exponential function as the loss function, $L(F, y) = e^{-F(x)y}$. The exponential loss function has a weakness in that it lacks robustness in the presence of outlier (unlabeled) or mislabelled samples. To rectify this deficiency, a variety of loss functions have been exercised to increase the robustness while keeping the good out-of-the-box performance of AdaBoost. In the TMVA framework, a binomial log-likelihood loss function, $L(F, y) = \ln(1 + e^{-2F(x)y})$, has been employed for the classification problem. By calculating the current gradient

of the loss function, the boosting algorithm performs the steepest-descent approach to doing the minimization. Iterating the steepest-descent procedure produces the wanted forest which minimizes the loss function.

Boosted Decision Tree Algorithm The decision tree is a well-known algorithm that enables a straightforward interpretation as a simple two-dimensional tree structure can visualize it. The algorithm sequentially cuts on event variables by repeating the yes/no judgment which is made based on one event variable at a time until a stop rule is met. By this way, the input event is processed through the binary structure yielding a final classifier response of signal-like or background-like. The tree-based method then splits the space of event variables, reconstructed physics variables in our analyses, into many rectangular regions, which are labeled by signal or background separately. The path down the decision tree to the leaf node describes a cut sequence that chooses signal or background based on the type of the leaf node. The trees are trained progressively, i.e., starting with the root node. The cut value and event-variable at each node are chosen, such that it provides the best separation ability between signal and background. The training stops at a node as soon as the critical lower bound of events is reached at that node.

The main shortcoming of the decision tree algorithm is the instability concerning statistical fluctuations in the training sample from which the classifier is created. For example, if two event variables show comparable separation ability, a fluctuation in the training input events may lead the tree to grow by splitting on one event variable, while the other one could be used without that fluctuation. Therefore, the entire tree structure is changed downstream of this decision node, possibly resulting also in a considerably altered classifier response. The boosting algorithm can solve this concern by building a forest of decision trees.

The trees are created from the same training input events, with the misclassified events being subsequently increased in their weights before deriving the next tree. The final classification for every single event is based on the majority vote of the prediction results from each tree in the forest. The forest is also called Boosted Decision Trees (BDT). Compared to one decision tree, BDT has stronger statistical stability and better separation performance.

Input Event-Variable Selection

The event variables input to the selection algorithm play a vital role in the training phase to obtain a classifier with a good performance. When selecting the event variables, we focus on two characteristics of them: 1) Data and MC agreement; 2) Discriminating capability. FD MC is used as a signal in the training phase. Hence, we require the data distribution of the selected event variables to be consistent with the MC ones, which was examined in the ND data and MC files. The FD cosmic sample is used as the background in the training phase. We consequently require the selected event variables to help separate backgrounds and signal. After a detailed discussion within the NC Disappearance group, 13 reconstructed variables have been selected as the input event variables. These variables are based on slice hits, reconstructed shower, and CVN information.

Data Vs MC Agreement Study

Employed Files: 1,000 NOvA ND Data and Monte Carlo files have been used to study the Data vs MC agreement.

Employed Cuts:

- Event Quality Cuts;

- ND Fiducial Volume Cuts;
- ND Containment Cuts;
- Energy Cut;

Discriminating Capability Study

Employed Files: 100,000 cosmic data files and 1,000 NOvA Monte Carlo files have been used for the training input.

Employed Cuts: The samples used in the PID training study include standard spill cuts, which ensure good quality of spills and cosmic veto, which remove the events that look like a single muon entering the detector from the top, and are also used by other analyses in NOvA. Furthermore, some NC 2016 Analysis Official cuts have been applied to produce the training input files. These are:

- Event Quality Cuts;
- FD Fiducial Volume Cuts;
- FD Containment Cuts;
- Energy Cut;
- Muon Removal Cut (We require the event's RemID be larger than or equal to 0.9.)

To have a better picture of these event variables, we describe them one by one below.

- **CVN CosmicID:** This is based on CVN classification results. For every event, the CVN gives it a probability to be a specific event type. This is the probability of one event to be a cosmic.

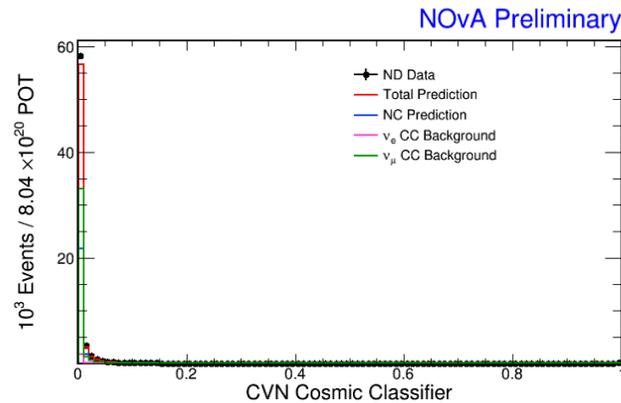


FIGURE 4.20: ND Data and MC match with each other. Both NC and CC events have a very low probability to be recognized as a cosmic event based on CVN output.

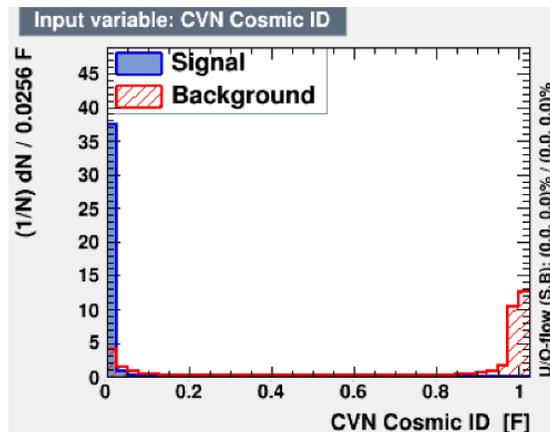


FIGURE 4.21: The blue histogram contains the FD MC NC events which have very low probability to be classified as a cosmic event, as expected. On the other hand, the red histogram contains the FD Cosmic Data. Most of them have a much higher probability to be recognized as a cosmic event. Based on the distribution plot of the signal and backgrounds, the CVN Cosmic ID gives us the most robust discrimination ability among the 13 input event-variables, and therefore it has been employed as the root splitting variable in the BDT.

- **NumberShowers:** Number of LID showers reconstructed for this event.

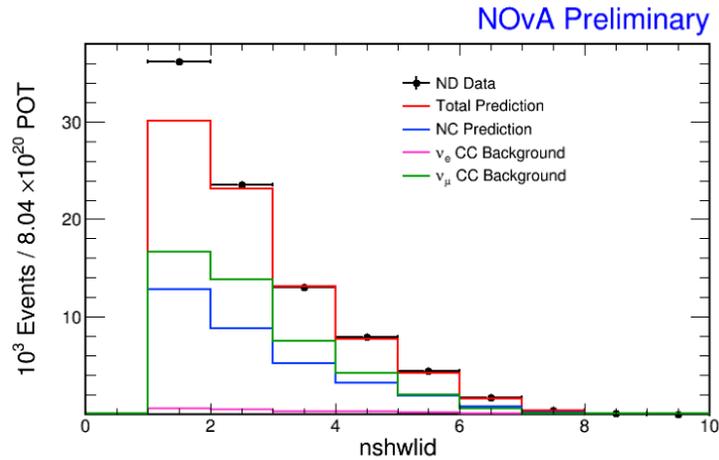


FIGURE 4.22: ND Data Vs MC Simulated FD Number of Showers. ND Data and MC match better when the event has more than two reconstructed showers. We also expect a reconstruction improvement for the event shower based variables.

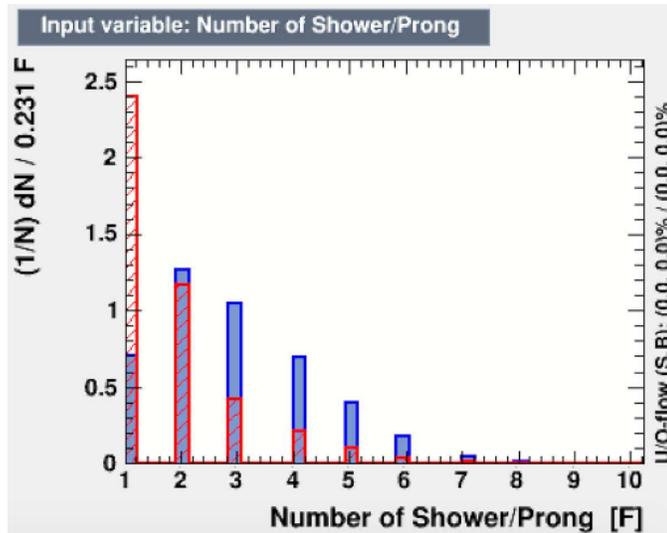


FIGURE 4.23: FD Signal vs Background Number of Showers. The Blue plot indicates the FD MC NC events, and the red one represents the cosmic backgrounds. We can see that cosmic events tend to have fewer reconstructed showers compared to NC events.

- **ShowerYDirection:** A shower is a collection of hits with a reconstructed path through them. Y direction of the leading shower (daughter particle) can indicate the incoming direction of the mother particles.

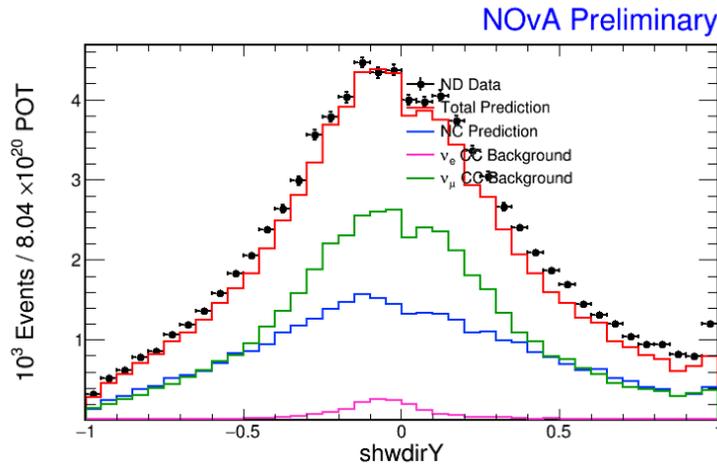


FIGURE 4.24: ND Data Vs MC Simulated FD Leading Shower Y Direction. ND Data and MC strongly agree with each other.

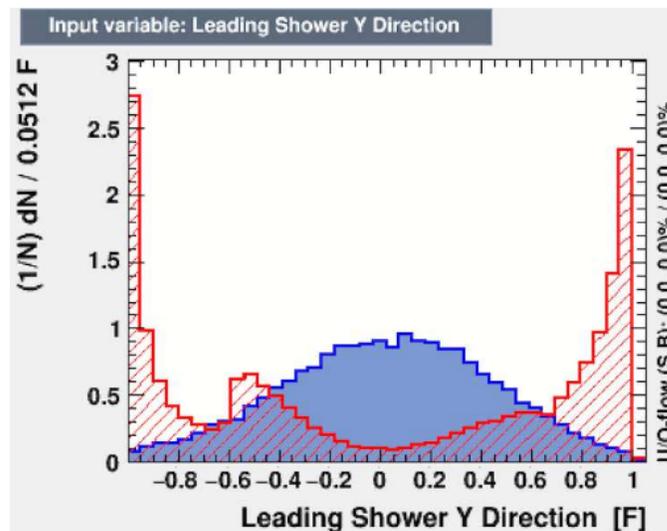


FIGURE 4.25: FD Signal Vs Background Leading Shower Y Direction. Blue ones are the FD MC NC events, and red ones are cosmic backgrounds. The distribution plots show that this variable has a very strong discrimination power as well as good agreement between data and MC.

- **LengthShower:** The reconstructed length of the leading shower.

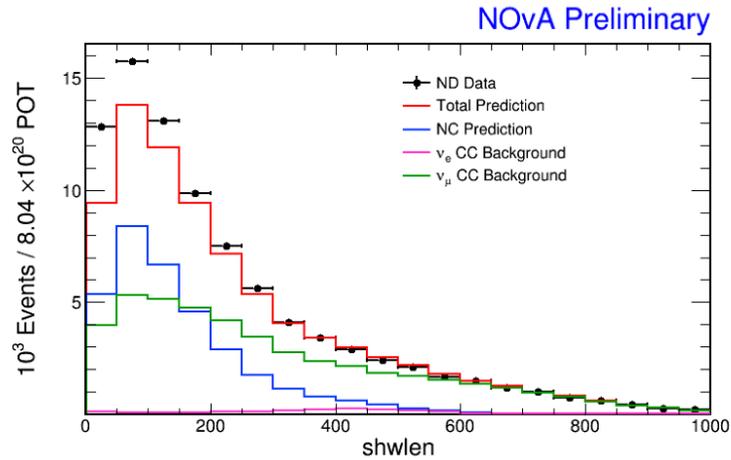


FIGURE 4.26: ND Data Vs MC Length of Shower. The longer the leading shower, the better the Data and MC agreement is. It implies that there is still room for improvement on low-energy daughter particle reconstruction.

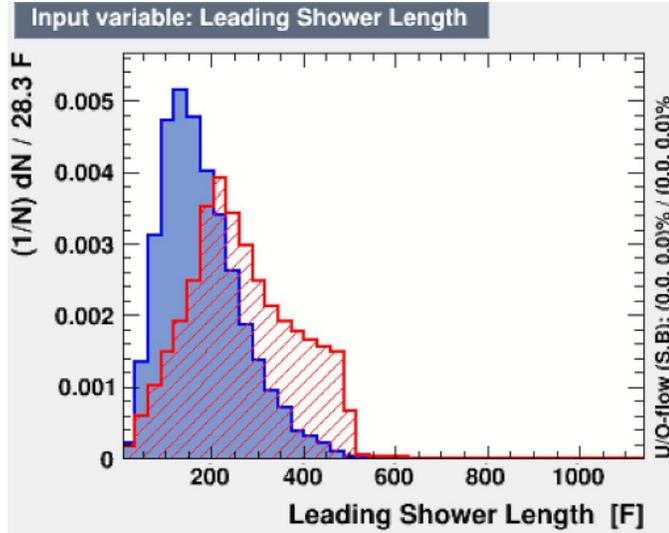


FIGURE 4.27: FD Signal Vs Background Length of Shower. The blue plot indicates the FD MC NC events, and the red one represents the cosmic backgrounds. The distribution plots show the discrimination capability of the variable.

- **Particle Pt/P:** Fraction of Transverse Momentum is also a indicator of mother particle moving direction, but from the point of view of momentum.

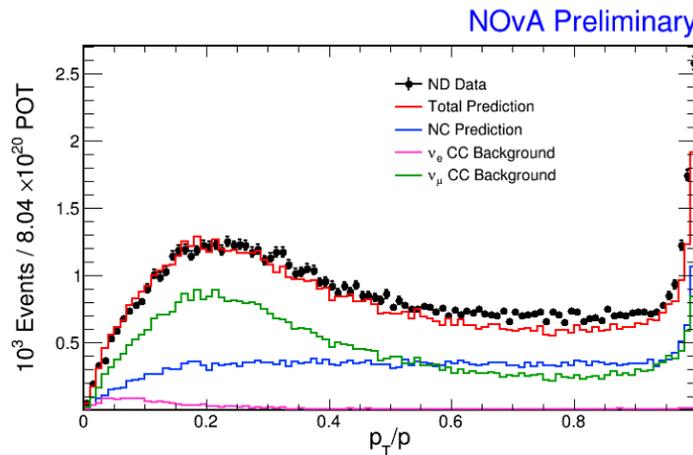


FIGURE 4.28: ND Data Vs MC Fraction of Transverse Momentum. ND Data and MC strongly agree with each other.

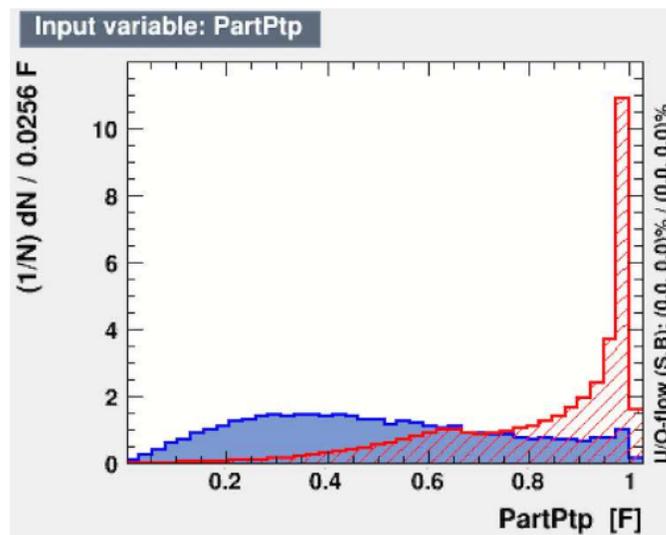


FIGURE 4.29: FD Signal Vs Background Fraction of Transverse Momentum. The blue plot indicates the FD MC NC events, and the red one represents the cosmic backgrounds. The high peak of the cosmic sample is around 1, which indicates the strong separation ability of the variable.

- **ShwWidth:** This is another reconstructed shower feature, which present the width of the leading shower.

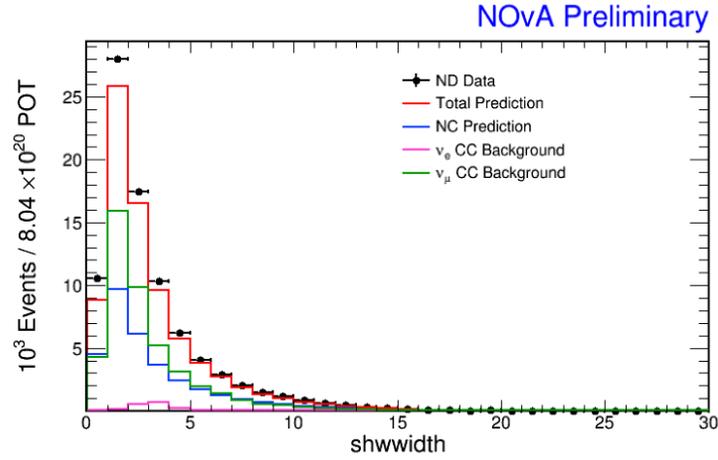


FIGURE 4.30: Wider the leading shower, better Data and MC agreement.

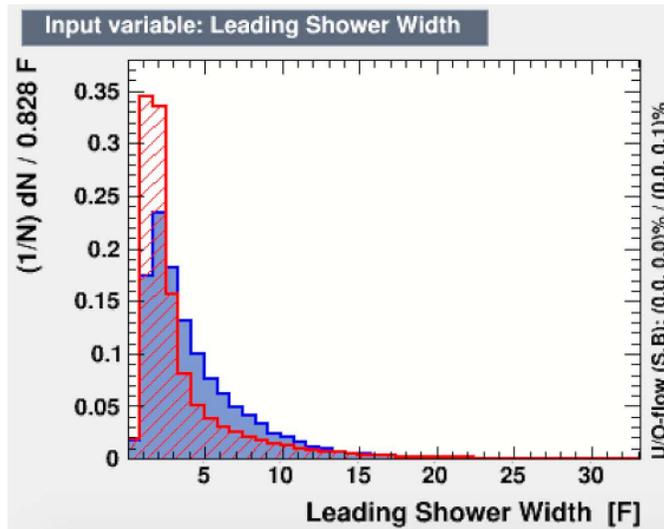


FIGURE 4.31: FD Signal Vs Background Reconstructed Shower Width. The blue plot indicates the FD MC NC events, and the red one represents the cosmic backgrounds. The high peak of the cosmic is around 0, which has a small gap with the NC peak.

- **ShwNHit**: The number of hits of the leading shower.

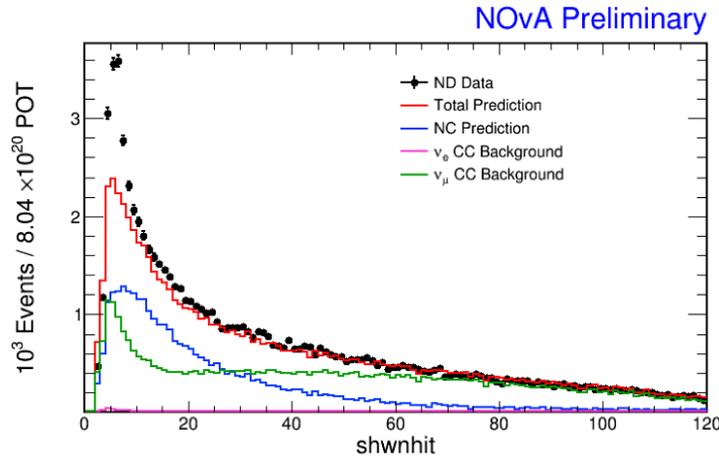


FIGURE 4.32: More hits, better agreement. It also indicates that there is still room for improvement on low energy daughter particle reconstruction.

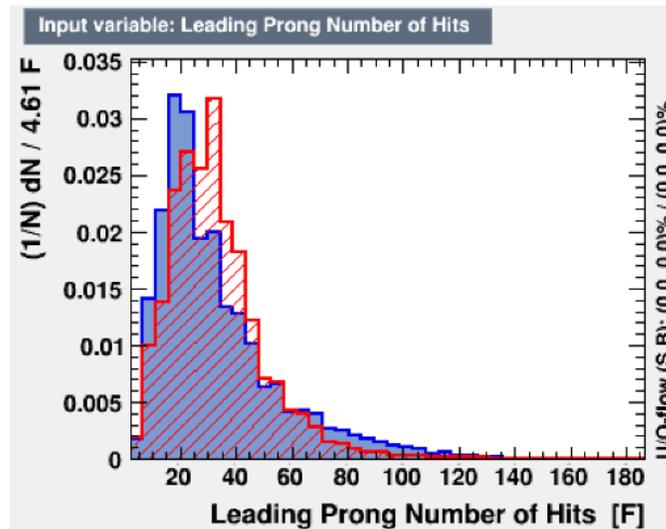


FIGURE 4.33: FD Signal Vs Background Number of Hits. The blue plot indicates the FD MC NC events, and the red one represents the cosmic backgrounds. The high peak of the NC is around 20, which has a small gap with the cosmic peak.

- **XplusY**: Top view slice hit number plus side view slice hit number.

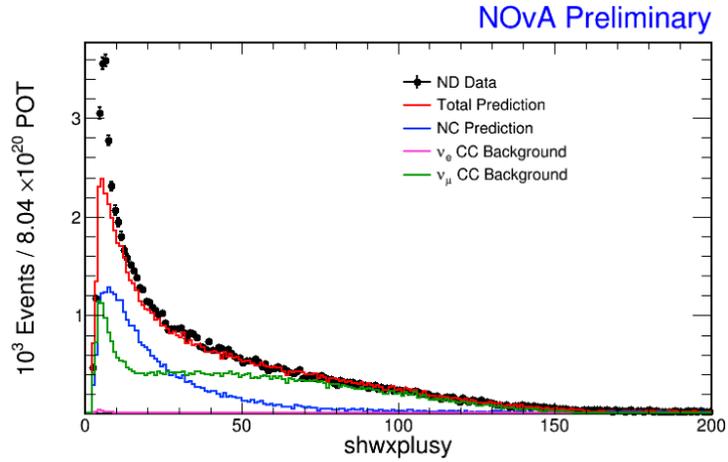


FIGURE 4.34: The ND Data and MC agreement is becoming better while the amount of hits is growing.

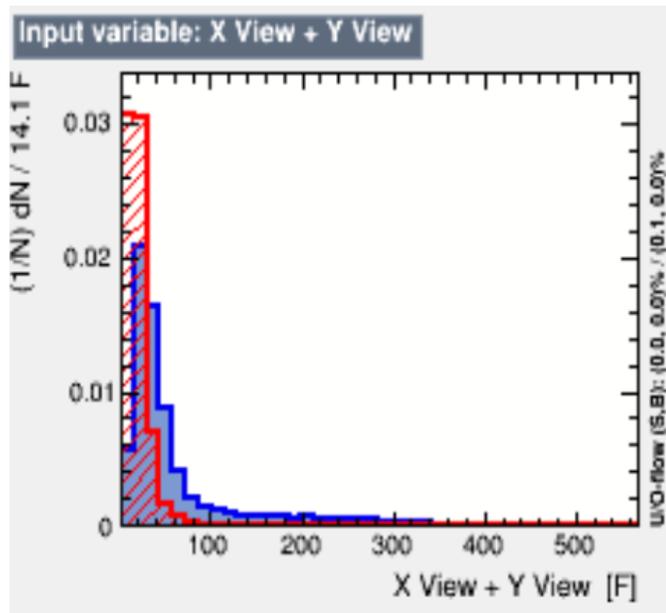


FIGURE 4.35: FD Signal Vs Background XplusY. The blue plot indicates the FD MC NC events, and the red one represents the cosmic backgrounds. The high peak of the cosmic is around 0, which has a small gap with the NC peak.

- **XminusY**: Top view slice hit number minus side view slice hit number.

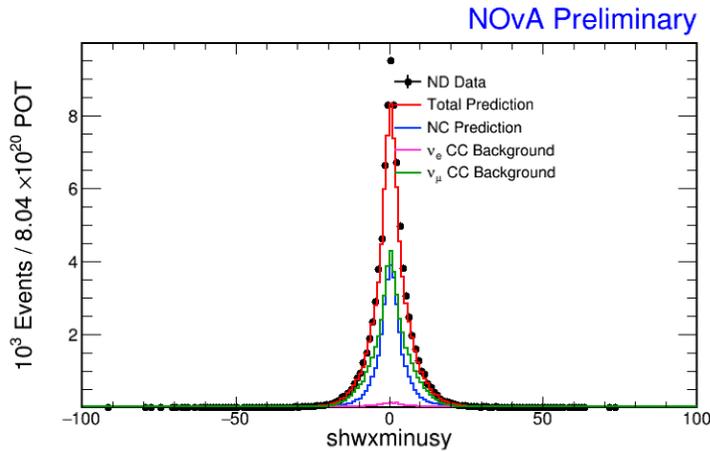


FIGURE 4.36: The ND Data and MC agreement is good when the hit number of the top view and side view are different.

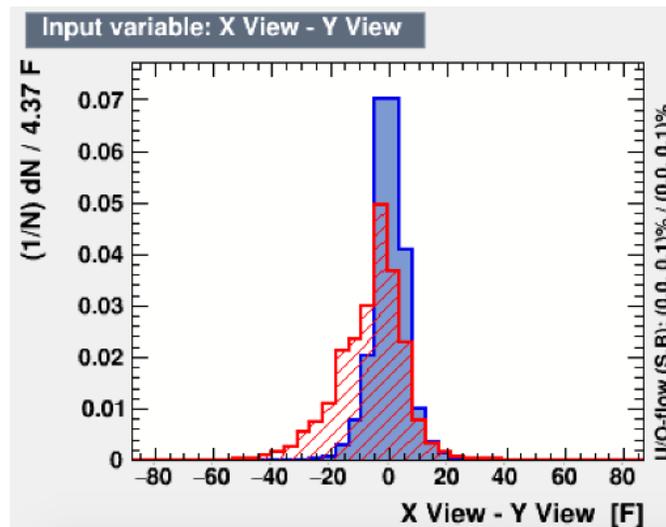


FIGURE 4.37: FD Signal Vs Background XminusY. The blue plot indicates the FD MC NC events, and the red one represents the cosmic backgrounds. The different distribution areas indicate the general separation ability of the variable.

- **XoverY:**

(Top view slice hit number plus side view slice hit number)/(Top view slice hit number plus side view slice hit number)

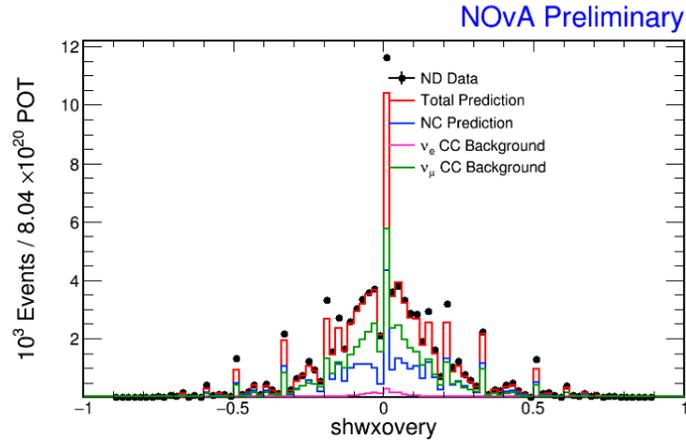


FIGURE 4.38: General ND Data and MC Agreement

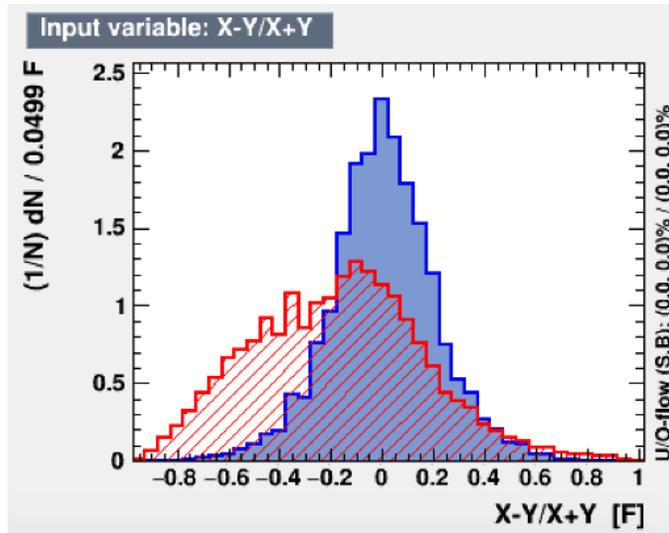


FIGURE 4.39: FD Signal Vs Background XoverY. The blue plot indicates the FD MC NC events, and the red one represents the cosmic backgrounds.

- MIP:** Number of minimum ionizing particle (mip) hits in one slice. A mip loses a constant fraction of its energy when traveling through material. For NOvA, with a liquid scintillator density of 0.88 g/cm^3 , this amounts to an average energy deposit of 1.7 MeV/cm traveled at normal incidence. Therefore, the number of mip hits with energy characteristic of a mip in one event can be used to classify electromagnetic particles from muons.

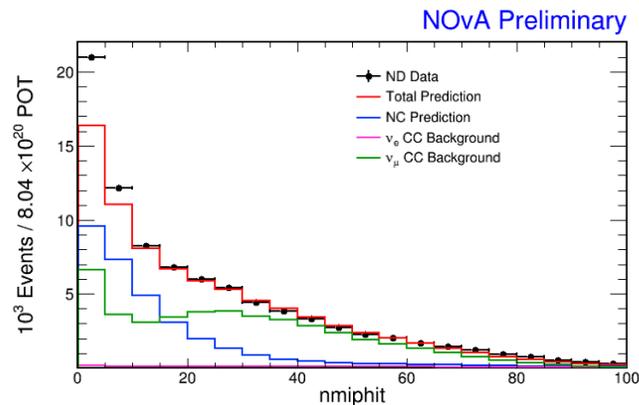


FIGURE 4.40: ND Data Vs MC MIP. ND Data and MC agree well when the mip number of hits is more than 10.

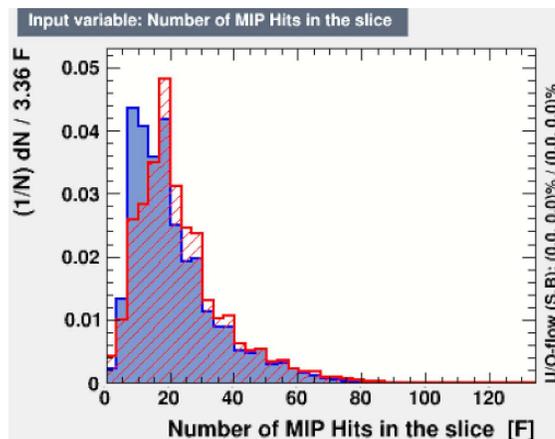


FIGURE 4.41: FD Signal Vs Background MIP. The blue plot indicates the FD MC NC events, and the red one represents the cosmic backgrounds. The high peak of the NC is around 10. On the other hand, cosmic peak is around 25.

- **ShwGap:** The reconstructed gap between the vertex and the starting point of the leading shower.

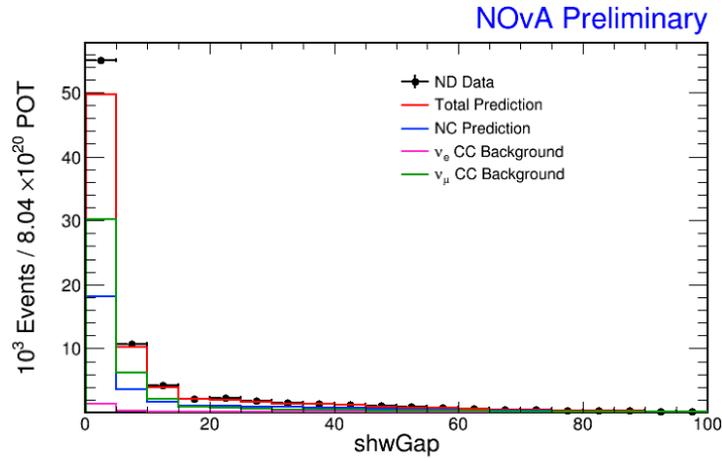


FIGURE 4.42: ND Data Vs MC Shower of Gap. ND Data and MC agrees with each other, especially when the gap is larger than 10 cm.

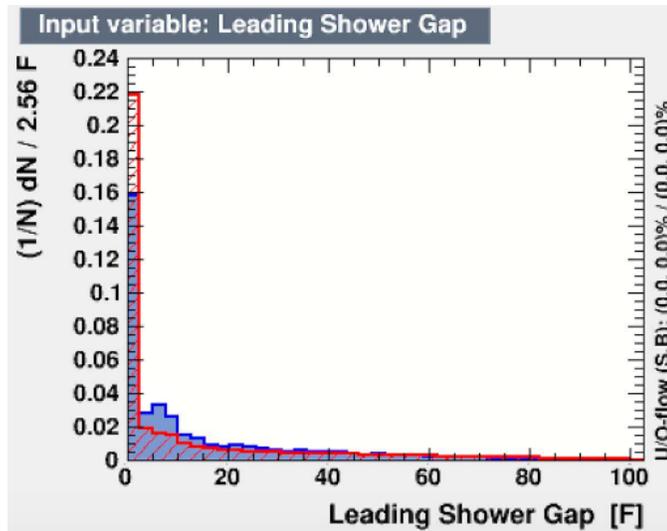


FIGURE 4.43: FD Signal Vs Background Shower Gap. The blue plot indicates the FD MC NC events, and the red one represents the cosmic backgrounds.

- **ShwCalE:**

The reconstructed energy of the leading shower.

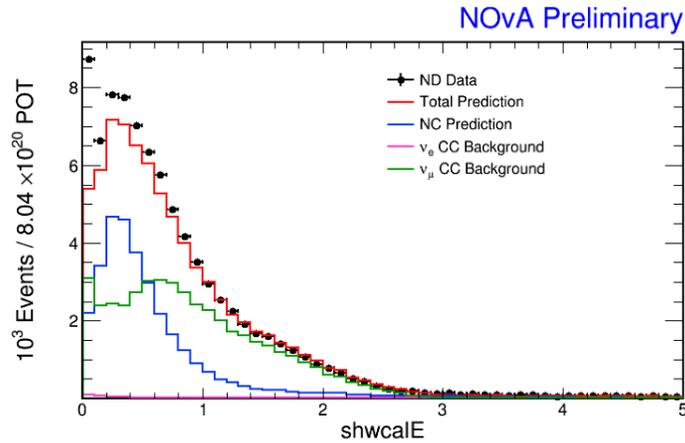


FIGURE 4.44: ND Data Vs MC Shower of Energy. ND Data and MC agree well, especially for the energy range between 1 to 5 GeV.

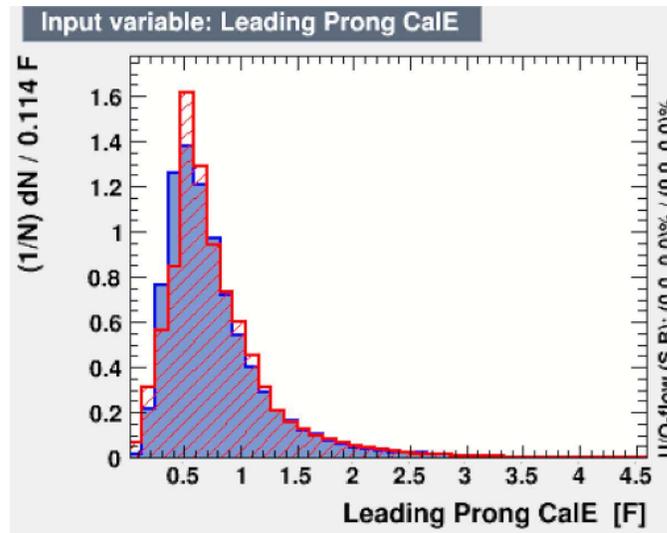


FIGURE 4.45: FD Signal Vs Background Shower Energy. The blue plot indicates the FD MC NC events, and the red one represents the cosmic backgrounds.

Feeding the above thirteen selected event variables into TMVA, the framework then return us the unspecific method separation ability of them. The ranking is the same as we learned from the discriminating ability study. We exhibit them in table 4.17.

Input Variable	Separation (method unspecific)
CVN CosmicID	8.331e-01
NumberShowers	3.461e-01
ShowerYDirection	2.982e-01
LengthShower	2.658e-01
Particle P_T/P	1.966e-01
ShwWidth	1.937e-01
ShwNHit	1.119e-01
XplusY	1.119e-01
XminusY	9.115e-02
XoverY	8.803e-02
MIP	8.224e-02
ShwGap	5.013e-02
ShwCalE	6.441e-02

TABLE 4.17: TMVA unspecific method separation ability

Growing the Decision Tree

The growing of a decision tree is the procedure that determines the splitting criteria for each decision node. The process begins with the root node, where the first splitting criterion for the entire training dataset is defined. Then the separated two subsets of training events go through the same algorithm of defining the next splitting iteration separately. This step is repeated until the decision

tree is growing. At each decision node, the event-variable and its cut value were chosen based on the best discrimination between signal and background. The splitting process stops once it meets one of the stop criteria which is specified in the BDT configuration. Five separation criteria have been offered by TMVA to evaluate the separation power of an event-variable and a specific cut value. We employ the Gini Index as the separation criterion for our training.

- Gini Index: $p \times (1 - p)$, where p is the selection purity.

Gini Index is the metric to gauge how often a chosen element would be incorrectly identified. It means an event-variable with the lowest Gini Index should be chosen.

Overtraining and Pruning

When a build tree is depends too closely on a specific group of training events, similar events may have a different response from the decision tree. This is the *overtraining* problem for decision tree algorithms. Overtraining is problematic as the events can be classified differently in the application phase from the similar ones in the training phase. Therefore the BDT response distribution cannot be employed to define the optimal cut value. Tree-based algorithms are in general sensitive to this issue. One possible solution, "pruning," has been developed to remove statistically insignificant decision nodes and hence decrease the overtraining of the decision tree. The pruning is the procedure of cutting back a decision tree from the bottom up after it has been growing to the maximum size. For the BDT which we trained, however, pruning is not necessary since we drastically limit the tree depth. Instead of pruning the trained trees, we check the overtraining issue by equally dividing the training events into two subsets, where one is used for the overtraining checking.

Final Hyper-Parameter Setting

In this section, we list all the hyper-parameters which we employed for the BDT training.

- Boost Type: Real Adaptive Boosting;
- Number of Trees in the forest: 500;
- Learning Rate : 0.5;
- Number of Cuts: 40; (Number of splitting points for one variable used in searching the best cut.)
- MaxDepth : 5; (The possible highest depth of decision trees in the forest.)
- MinNodeSize : 2%; (The minimum rate of input training events asked in one leaf node.)
- NodePurityLimit : 0.9; (The required final leaf node signal purity percentage.)
- SeparationType : Gini Index; (Separation criterion for decision node splitting)
- PruneMethod : No Pruning;

BDT Output

After the hyper-parameter was defined, the BDT training was started and return the following outputs: algorithm-specific event-variable separation ranking, and BDT response.

Input Variable	Separation (method specific)
CVN Cosmic ID	2.954e-01
Leading Shower Y Direction	1.155e-01
Number of MIP Hits in the slice	8.868e-02
Leading Shower Width	7.357e-02
Leading Shower Length	6.989e-02
Leading Shower Energy	6.989e-02
Particle P_T/P	5.752e-02
Leading Shower Number of Hits	5.637e-02
Showerxovery	5.224e-02
Showrminus	5.185e-01
Leading Shower Gap	5.013e-02
Number of Leading Showers	1.889e-02
Shxpluxy	0.000e-02

TABLE 4.18: TMVA method specific separation ability.

Event-Variable Ranking The ranking is derived by calculating how often the event-variables are employed to split decision nodes, by weighting each split occurrence by the separation gain squared, and by the number of events in the leaf node. This measure of the event-variable importance can be used for the forest. As present in the Table 4.18, the last event variable can not contribute for the trained BDT.

BDT Response

4.2.4 Application Phase

Let's go back to CAF framework. The NCID package relies on LID Shower, CVN, and Slice, and thus should be run after these in the reconstruction chain. The corresponding producer and analyzer (validation) modules can be found

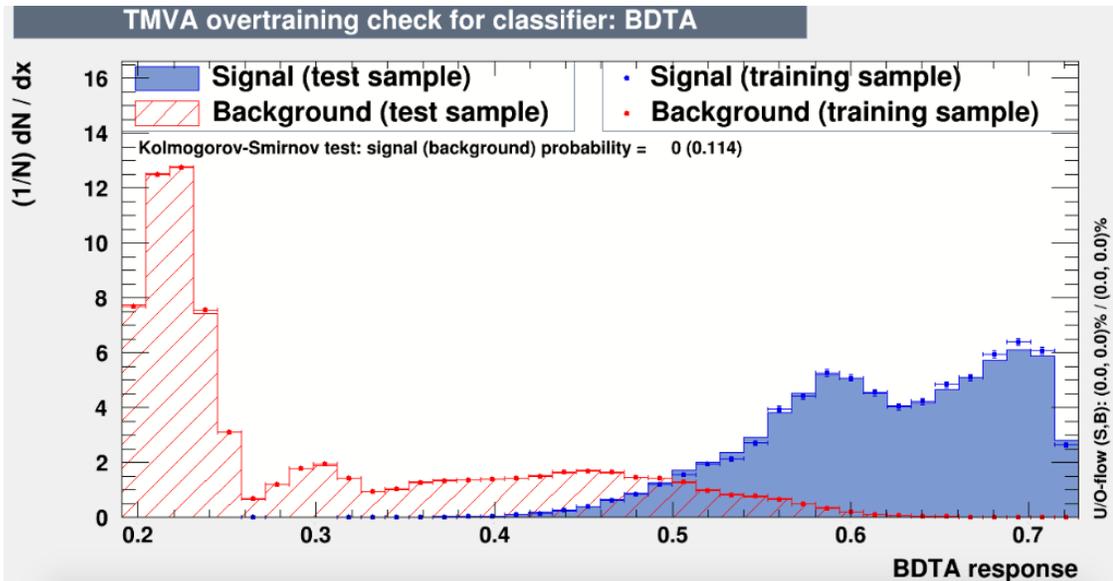


FIGURE 4.46: TMVA-Based NC-dedicated Cosmic Rejection PID Distribution.

in the NCID package. After adding the new PID into CAF framework, we also produced the application plot, normalized to 9.0×10^{20} POT.

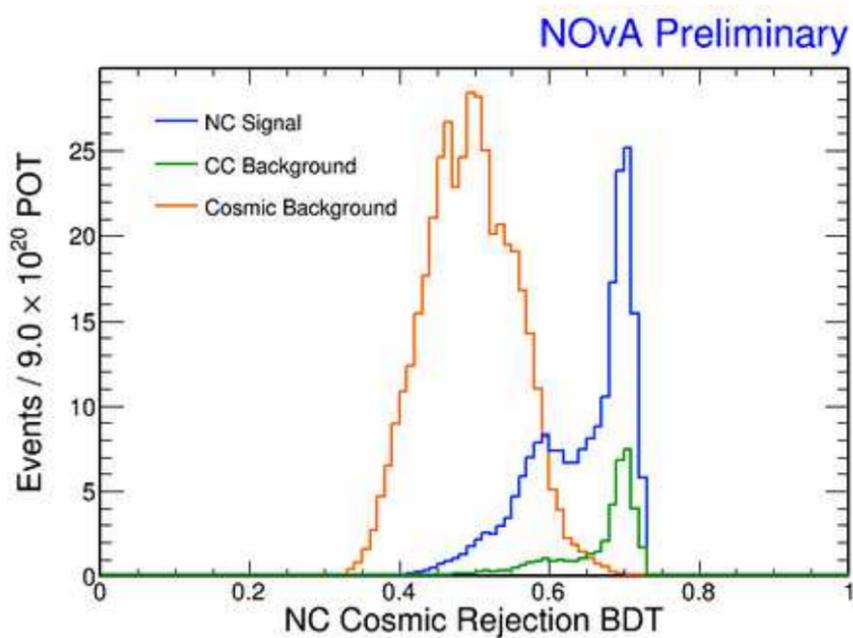


FIGURE 4.47: CAF-Based NC-dedicated Cosmic Rejection PID Distribution.

4.3 NC/CC Binary Classifier

To train an advanced binary classification algorithm, the NC/CC Classifier, for separating Neutral Current (NC) interactions from backgrounds (mainly ν_μ CC events), we produce a series of selection results based on different PID cut setting and compare the corresponding sensitivities based on the various selections to decide the final selection cuts. the NC-dedicated Cosmic Rejection BDT is the first PID which we trained for NC selection which mostly follows the ν_μ CC-Dedicated Cosmic Rejection BDT training philosophy on both input event-variable selection and on algorithm hyper-parameter setting.

4.3.1 Input Event-Variables Selection

- **CVN NCID:** This is based on CVN classification results. For every event, the CVN gives it a probability to be a specific event type. This is the probability of one event to be a cosmic.

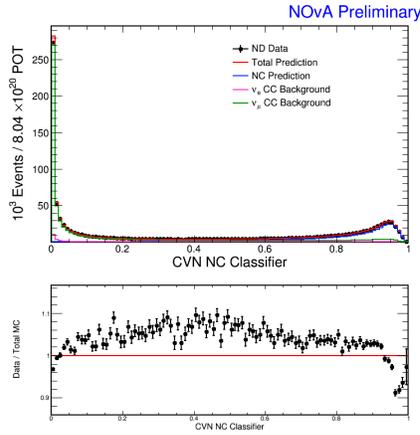


FIGURE 4.48: ND Ratio CVN CosmicID. ND Data and MC match with each other. Both NC and CC events have a very low probability to be recognized as a cosmic event based on CVN output.

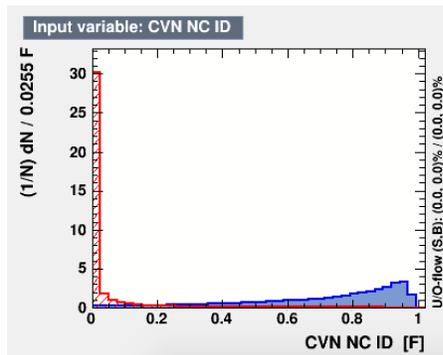


FIGURE 4.49: FD Signal vs Background CVN CosmicID. Based on the distribution plot of the signal and backgrounds, the CVN Cosmic ID gives us the most robust discrimination ability among the 13 input event-variables, and therefore it has been employed as the root splitting variable in the BDT.

- **Leading Prong Length:** The reconstructed length of the leading shower.

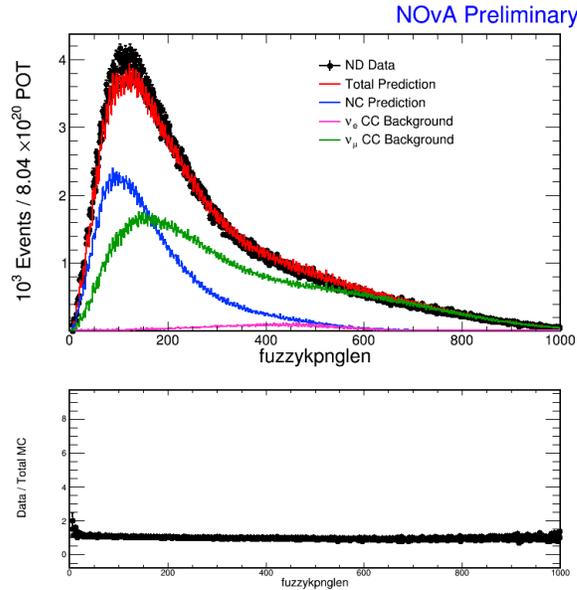


FIGURE 4.50: Longer the leading shower, better Data and MC agreement. It implies that there is still room for improvement on low energy daughter particle reconstruction.

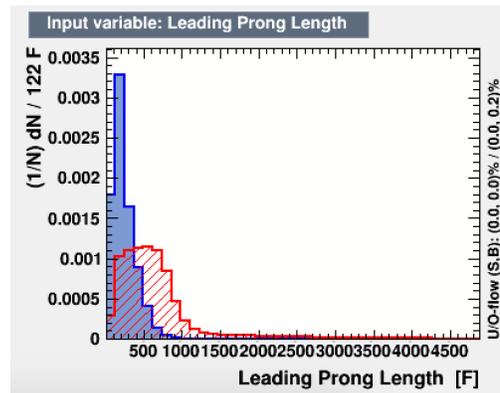


FIGURE 4.51: The blue plot indicates the FD MC NC events, and the red one represents the cosmic backgrounds. The distribution plots show the discrimination capability of the variable.

- **Number of MIP Hits:** How many minimum ionizing particle (mip) hits in one slice.

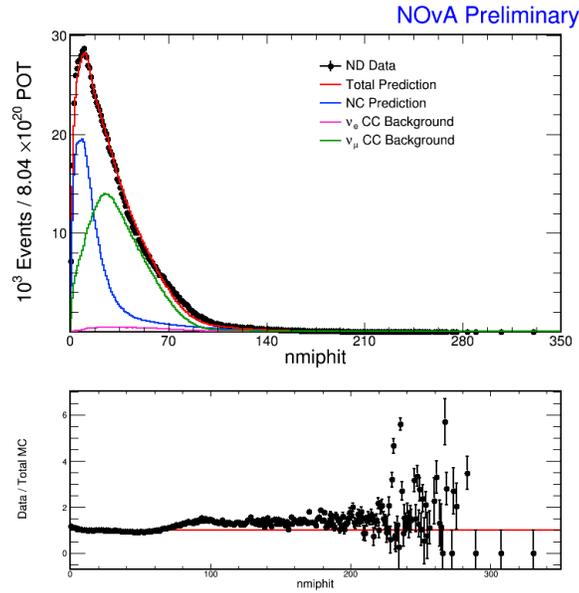


FIGURE 4.52: ND Data and MC agree well when the MIP hits is more than 10.

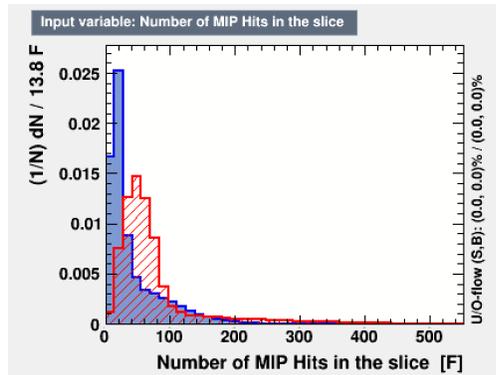


FIGURE 4.53: The blue plot indicates the FD MC NC events, and the red one represents the cosmic backgrounds. The high peak of the NC is around 10. On the other hand, cosmic peak is around 25. This indicates the general separation ability of the variable

- **Number of Calibrated Hits:**

How many minimum ionizing particle (mip) hits in one slice.

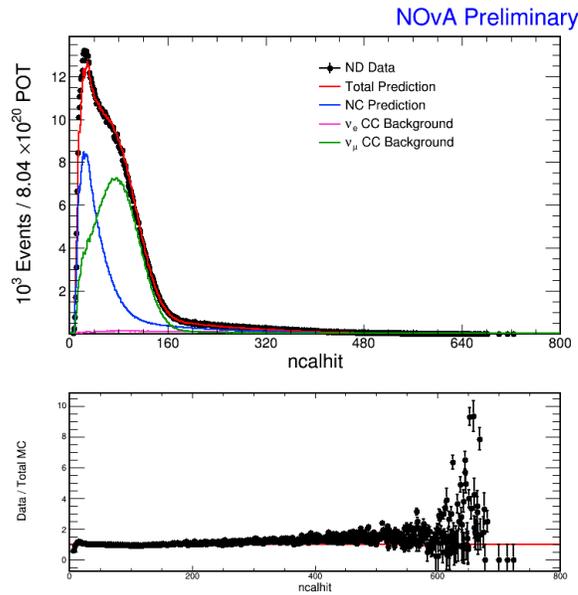


FIGURE 4.54: ND Data and MC agree well when the MIP hits is more than 10.

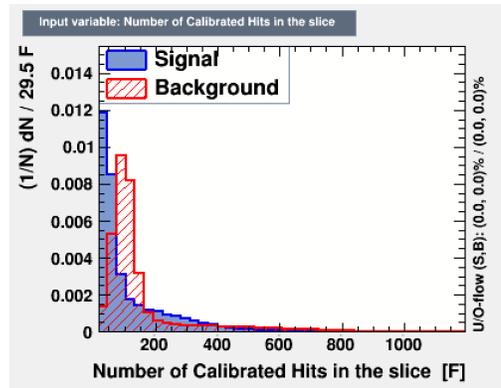


FIGURE 4.55: The high peak of the NC is around 10. On the other hand, cosmic peak is around 25. This indicates the general separation ability of the variable

- **Number of Hits in the Leading Prongs:**

The number of hits of the leading shower.

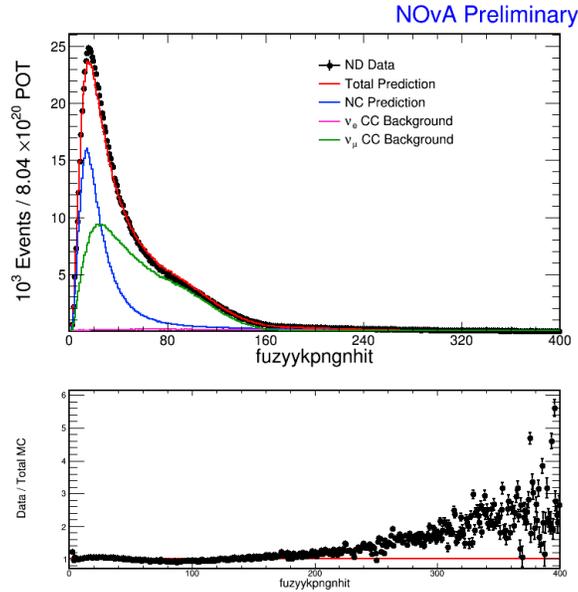


FIGURE 4.56: More hits, better agreement. It also indicates that there is still room for improvement on low energy daughter particle reconstruction.

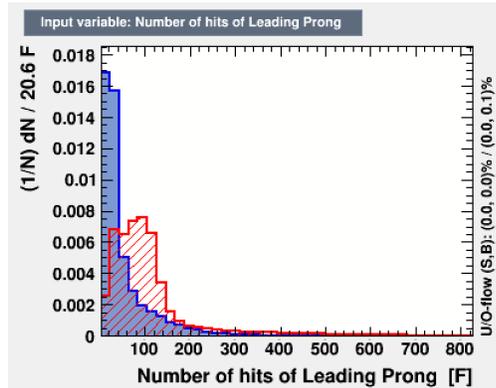


FIGURE 4.57: The high peak of the NC is around 20, which has a small gap with the cosmic peak. This indicates the general separation ability of the variable

- **X View Number of Hits in the Leading Prongs:** The number of hits of the leading shower.

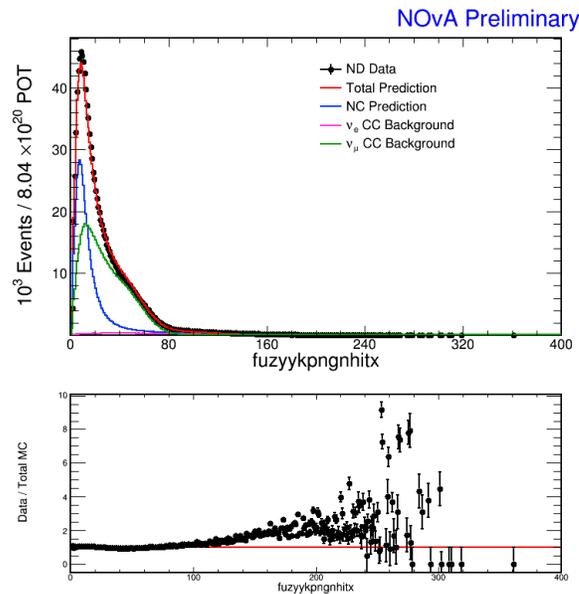


FIGURE 4.58: More hits, better agreement. It also indicates that there is still room for improvement on low energy daughter particle reconstruction.

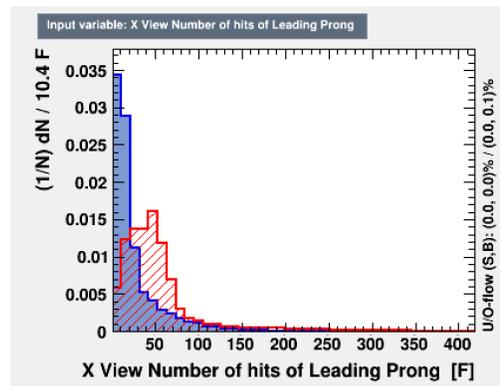


FIGURE 4.59: The high peak of the NC is around 20, which has a small gap with the cosmic peak. This indicates the general separation ability of the variable

- **Y View Number of Hits in the Leading Prongs:** The number of hits of the leading shower.

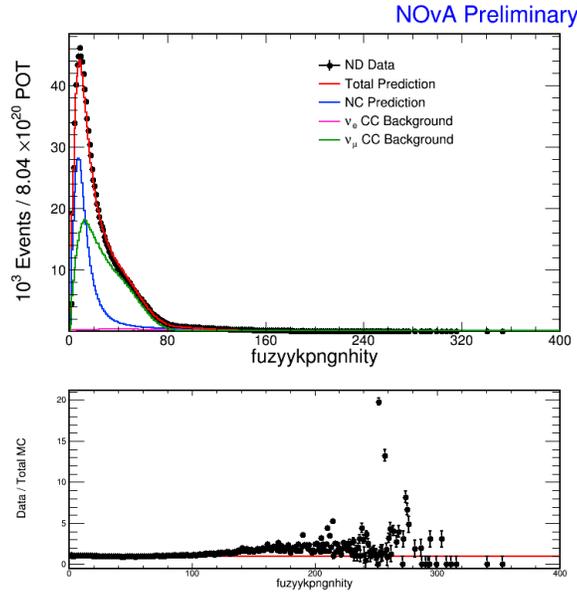


FIGURE 4.60: More hits, better agreement. It also indicates that there is still room for improvement on low energy daughter particle reconstruction.

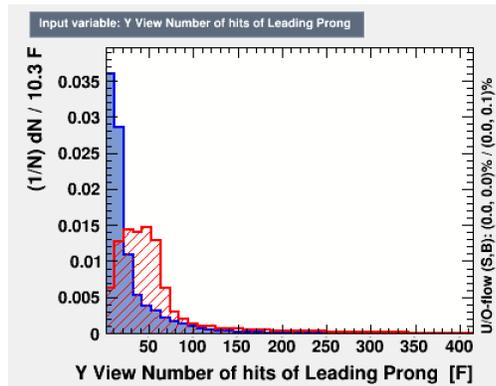


FIGURE 4.61: The high peak of the NC is around 20, which has a small gap with the cosmic peak. This indicates the general separation ability of the variable

- **Leading Prong Calibrated Energy:** The reconstructed energy of the leading shower.

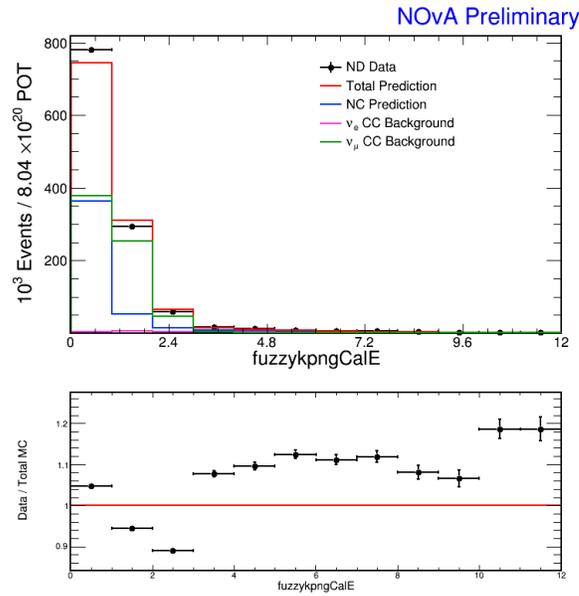


FIGURE 4.62: ND Data and MC agree well, especially for the energy range between 1 to 5 GeV

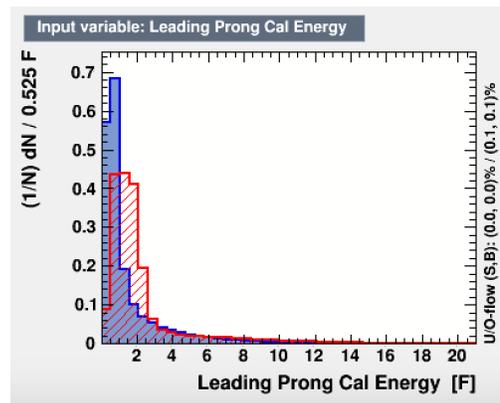


FIGURE 4.63: The different distribution areas indicate the weak separation ability of the variable

- **Energy Weighted Summarized X View Hits:** The reconstructed energy of the leading shower.

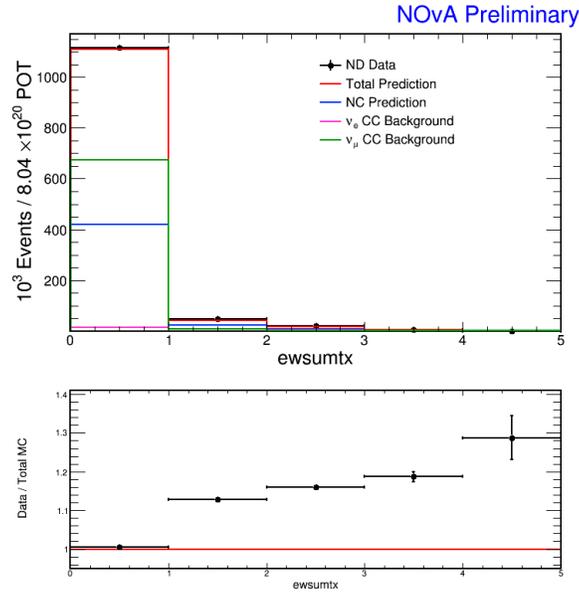


FIGURE 4.64: ND Data and MC agree well, especially for the energy range between 1 to 5 GeV

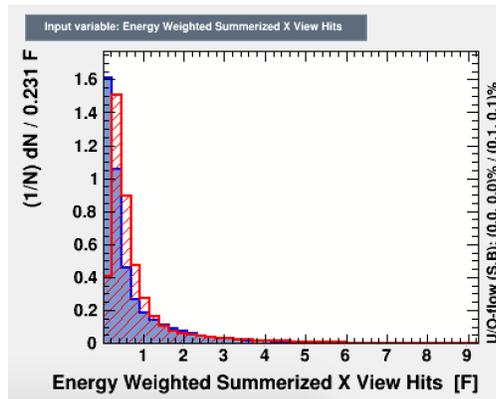


FIGURE 4.65: The different distribution areas indicate the weak separation ability of the variable

- **Energy Weighted Summarized Y View Hits:** The reconstructed energy of the leading shower.

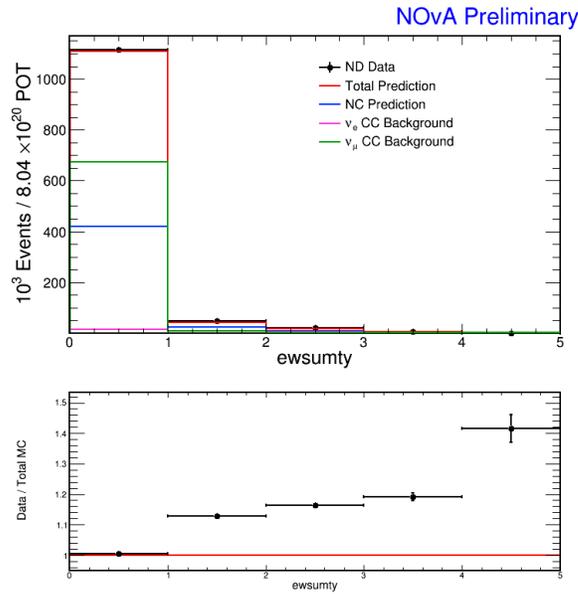


FIGURE 4.66: ND Data and MC agree well, especially for the energy range between 1 to 5 GeV

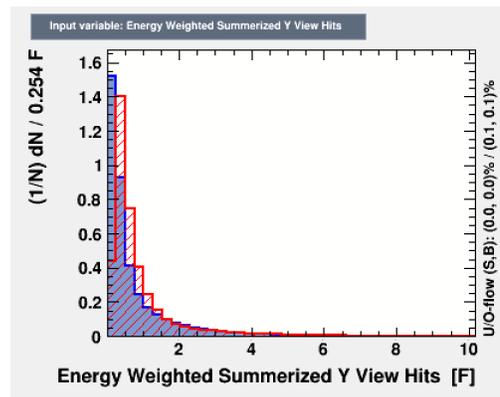


FIGURE 4.67: The different distributing areas indicate the weak separation ability of the variable

- **Summarized X View Hits:** The reconstructed energy of the leading shower.

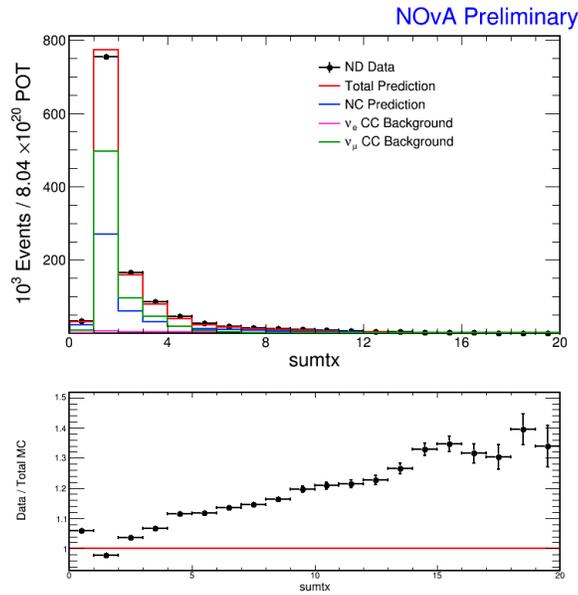


FIGURE 4.68: ND Data and MC agree well, especially for the energy range between 1 to 5 GeV

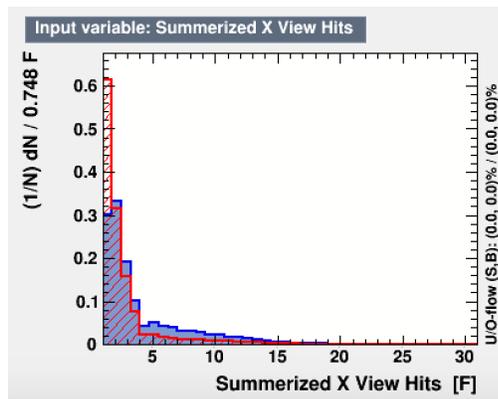


FIGURE 4.69: The different distribution areas indicate the weak separation ability of the variable

- **Summarized Y View Hits:** The reconstructed energy of the leading shower.

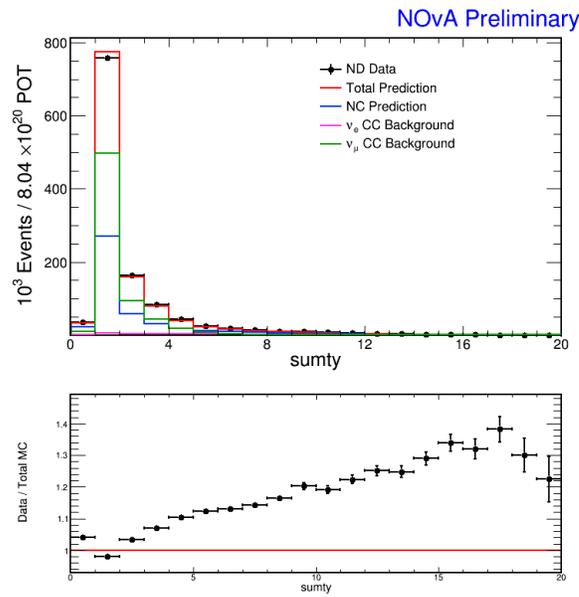


FIGURE 4.70: ND Data and MC agree well, especially for the energy range between 1 to 5 GeV

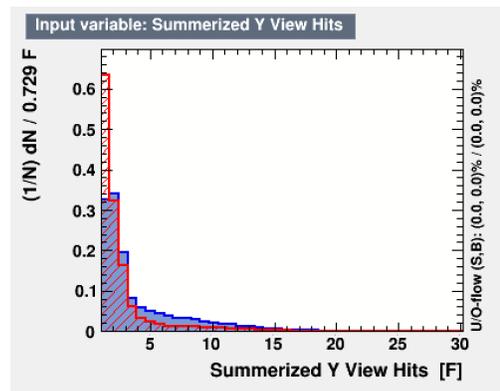


FIGURE 4.71: The different distribution areas indicate the weak separation ability of the variable

- **Number of FuzzyK Prongs:** Number of LID showers be reconstructed for this event.

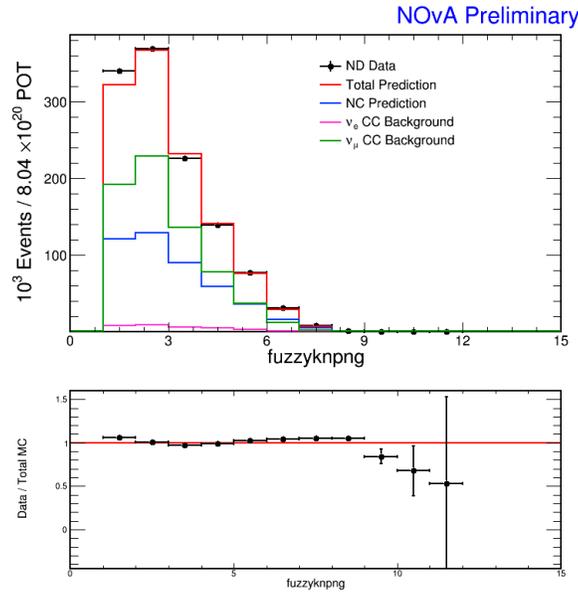


FIGURE 4.72: ND Data and MC match better when the event has more than two reconstructed showers. We also expect a reconstruction improvement for the event shower based variables.

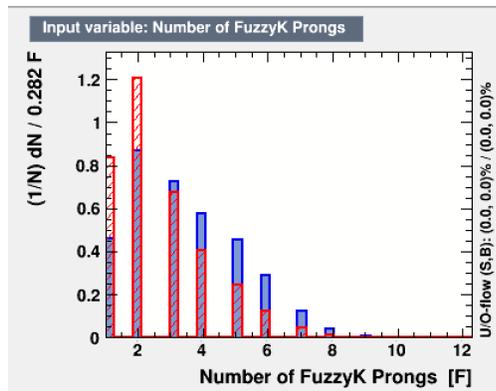


FIGURE 4.73: The blue plot indicates the FD MC NC events, and the red one represents the cosmic backgrounds. We can see that cosmic event tend to have less reconstructed showers compared to NC events.

- **Leading Prong Z Direction:** A shower is a collection of hits with a reconstructed path through them. Y direction of the leading shower (daughter particle) can indicate the coming direction of the mother particles.

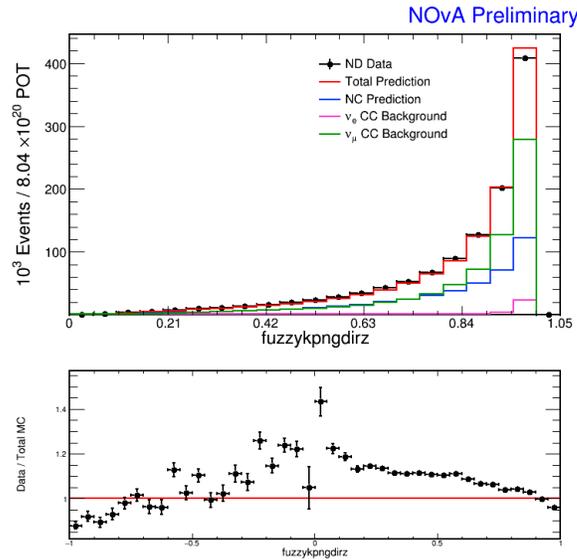


FIGURE 4.74: ND Data and MC strongly agree with each other.

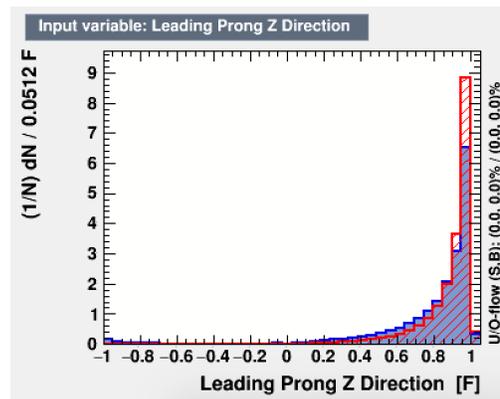


FIGURE 4.75: The blue ones are the FD MC NC events, and red ones are cosmic backgrounds. The distribution plots show that this variable has a very strong discrimination power as well as the agreement between data and MC, as shown in the above plot.

- **Maximum Y in the box:** The reconstructed gap between the vertex and the starting point of the leading shower.

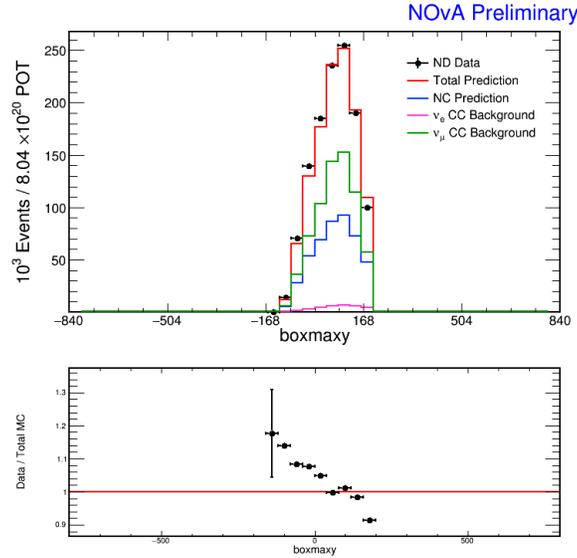


FIGURE 4.76: ND Data and MC agrees with each other, especially when the gap is larger than 10 cm.

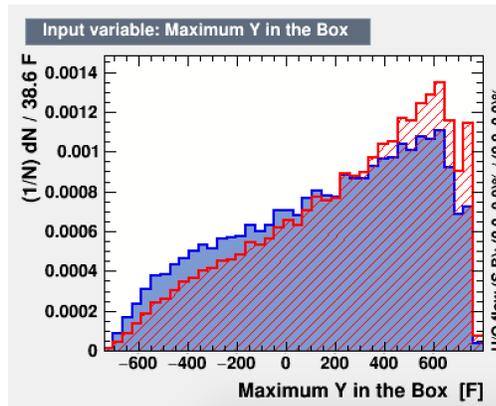


FIGURE 4.77: The blue plot indicates the FD MC NC events, and the red one represents the cosmic backgrounds. The different distribution areas indicate the weak separation ability of the variable

- **Minimum Y in the box:** The reconstructed gap between the vertex and the starting point of the leading shower.

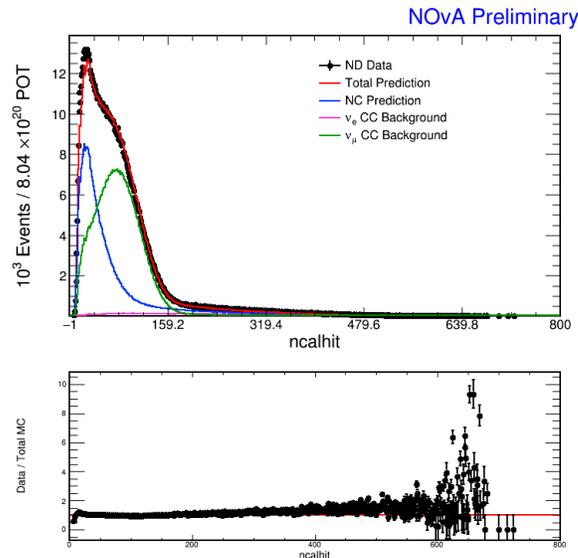


FIGURE 4.78: ND Data and MC agrees with each other, especially when the gap is larger than 10 cm.

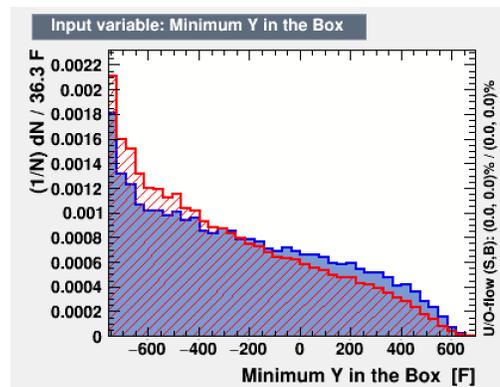


FIGURE 4.79: The blue plot indicates the FD MC NC events, and the red one represents the cosmic backgrounds. The different distribution areas indicate the weak separation ability of the variable

4.3.2 Training and Application Results

Recently, particle physicists have begun to explore the potential of convolutional neural networks (CNN). Several tools that implement these methods are available for use in classifying high-energy particle interactions, such as Caffe, Keras. In this section, we show what we learned from the last BDT training and how we can improve the training phase with a feature which is an output of a specific CNN architecture.

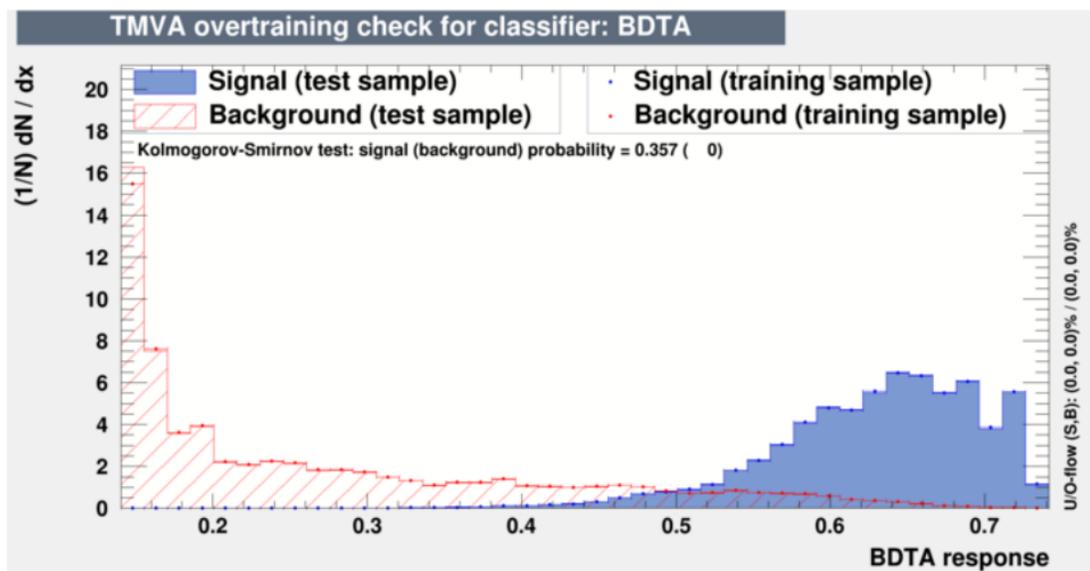


FIGURE 4.80: TMVA-Based NC/CC Classifier Response

Input Variable	Separation
CVN NC ID	7.589e-01
Leading Prong Length	3.511e-01
Number of MIP Hits in the slice	2.845e-01
Number of Calibrated Hits in the slice	2.768e-01
Number of hits of Leading Prong	2.702e-01
X View Number of hits of Leading Prong	2.634e-01
Y View Number of hits of Leading Prong	2.628e-01
Leading Prong Cal Energy	2.023e-01
Energy Weighted Summarized X View Hits	1.182e-01
Energy Weighted Summarized Y View Hits	1.086e-01
Summarized X View Hits	7.857e-02
Summarized Y View Hits	7.436e-02
Number of FuzzyK Prongs	5.431e-02
Leading Prong Z Direction	3.758e-02
Maximum Y in the Box	1.044e-02
Minimum Y in the Box	8.231e-03

TABLE 4.19: TMVA method specific separation ability (NC/CC).

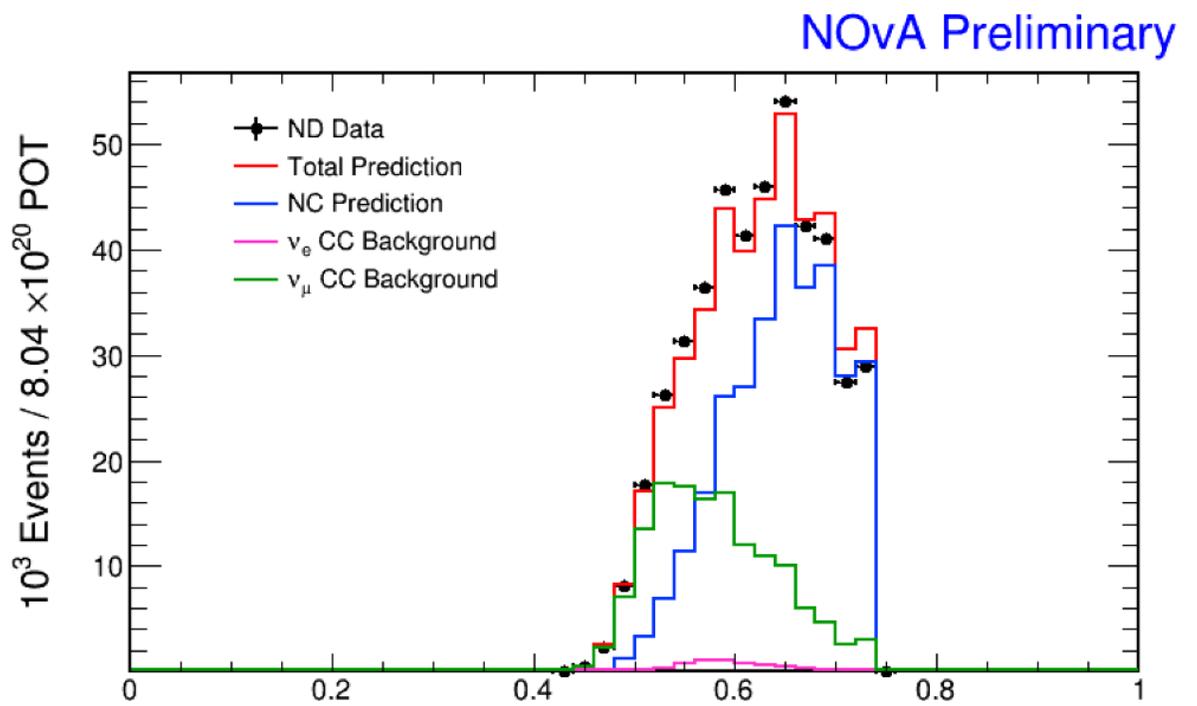


FIGURE 4.81: CAF-Based NC/CC Classifier Distribution.

4.4 PID Tuning for Nus17 Official and Re-Analysis Selections

After understanding the classification ability of the PIDs, including the CVN which is employed as one independent PID for the 2017 Official analysis, we finalize the final selection cuts for both the 2017 Official and the 2017 Re-Analysis. Here are the details:

4.4.1 2017 Official

The NOvA NC/Sterile analysis group decides the final selection cuts for the 2017 Official Analysis which are listed below for FD and ND separately:

FD

Applied three group of cuts:

- Standard Cuts: All of the above FD Pre-Selection Cuts;
- Cosmic BDT Cuts: Selected Events Cosmic BDT Values are required to be larger than 0.62;
- CVN Cuts: Selected Events CVN Values are required to be larger than 0.2;

ND

Applied two groups of cuts:

- Standard Cuts: All of the above ND Pre-Selection Cuts;
- CVN Cuts: All Selected Events CVN Values are required to be larger than 0.2;

4.4.2 2017 Re-Analysis

For the 2017 Re-Analysis, we employed a more complex method to select the cuts. We employed a list of values of the two PIDs (Cosmic BDT and NC/CC Classifier) as selection cuts along with the standard cuts as introduced above. Then, we compared the series of results with 2017 Official selection results to decide the best one for the 2017 Re-Analysis. The results are listed in Table 4.20.

After a detailed comparison, we decided to adopt the following cuts for ND and FD:

FD

Applied three group of cuts:

- Standard Cuts: All of the above FD Pre-Selection Cuts;
- Cosmic BDT Cuts: All Selected Events Cosmic BDT Values are required to be larger than 0.60;
- NC/CC Classifier: All Selected Events NC/CC Classifier are required to be larger than 0.54;

ND

Applied two groups of cuts:

- Standard Cuts: All of the above ND Pre-Selection Cuts;
- NC/CC Classifier: All Selected Events NC/CC Classifier are required to be larger than 0.54;

	Cosmic BDT	NC/CC Classifier	NC	CC	Cosmic	FoM
1	0.60	0.50	160.1	40.8	16.9	10.8448
2	0.60	0.52	156.6	35.2	16.7	10.8449
3	0.60	0.54	151.4	29.5	16.3	10.7804
4	0.60	0.56	143.6	23.5	15.7	10.6223
5	0.60	0.58	132.6	17.9	14.8	10.3149
6	0.62	0.50	144.1	38.9	8.2	10.4206
7	0.62	0.52	140.8	33.5	8.0	10.4289
8	0.62	0.54	136.1	28.2	7.8	10.3759
9	0.62	0.56	129.0	22.5	7.4	10.2342
10	0.62	0.58	118.9	17.1	7.0	9.94755
11	0.64	0.50	129.8	37.1	4.1	9.92694
12	0.64	0.52	126.8	31.9	3.9	9.94031
13	0.64	0.54	122.5	26.9	3.8	9.89578
14	0.64	0.56	116.1	21.4	3.7	9.76911
15	0.64	0.58	106.9	16.3	3.4	9.50426
16	0.66	0.50	113.3	34.6	1.7	9.26006
17	0.66	0.52	110.7	29.8	1.6	9.28007
18	0.66	0.54	106.9	25.2	1.6	9.2476
19	0.66	0.56	101.4	20.2	1.5	9.13833
20	0.66	0.58	93.5	15.4	1.4	8.90049
21	0.68	0.50	92.8	30.8	0.5	8.33205
22	0.68	0.52	90.8	26.6	0.5	8.35993
23	0.68	0.54	87.9	22.7	0.5	8.34163
24	0.68	0.56	83.5	18.2	0.4	8.25903
25	0.68	0.58	77.0	13.9	0.4	8.05
Nus17 Official			141.8	32.8	8.2	10.5

TABLE 4.20: NC Selection PID Tuning for the 2017 Re-Analysis.

4.5 Predicted Near and Far Detector Spectra

We summarize the direct selection of FD NC-like events obtained for the 2017 Official and 2017 Re-Analysis cases in Table ???. In the table, we also show the results for the **Standard** selection, which consists of the summarized FD Pre-Selection cuts, and for the **SideBand** selection, which employs an inverse NC/CC Classifier cut.

	NC Signal	CC Background			Cosmic	FoM
		ν_μ CC	ν_e CC	ν_τ CC		
Standard	236.6	230.3	68.1	4.7	10675.4	2.2
Nus17 Official	141.8	20.5	9.5	2.8	8.2	10.5
Nus17 Re-Analysis	151.4	18.2	8.4	2.8	16.2	10.8
SideBand	18.5	187.7	57.1	1.6	1.8	1.1

TABLE 4.21: FD Direct Selection Results.

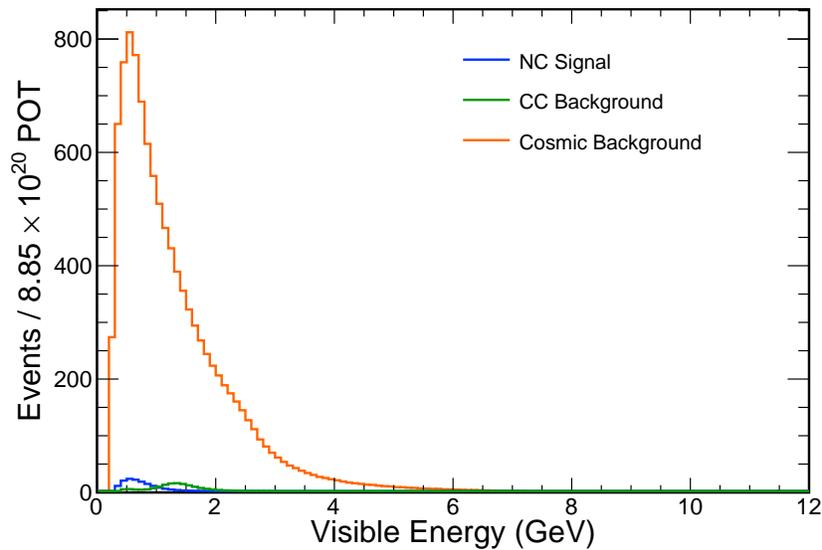


FIGURE 4.82: FD Standard Energy Spectra (Direct Selection Method).

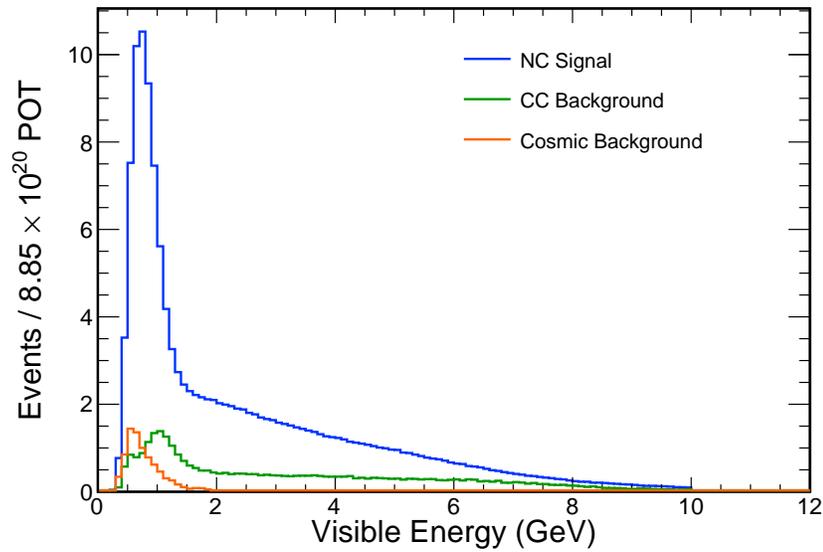


FIGURE 4.83: FD 2017 Official Energy Spectra (Direct Selection Method).

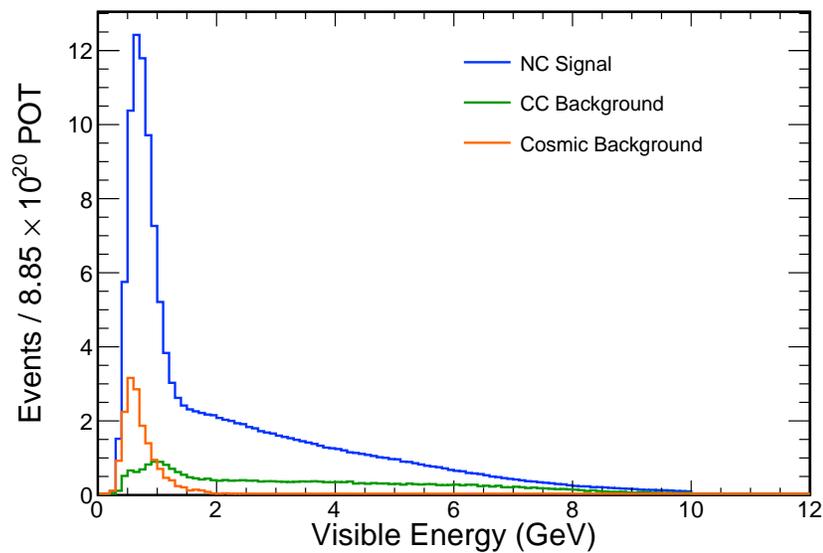


FIGURE 4.84: FD 2017 Re-Analysis Energy Spectra (Direct Selection Method).

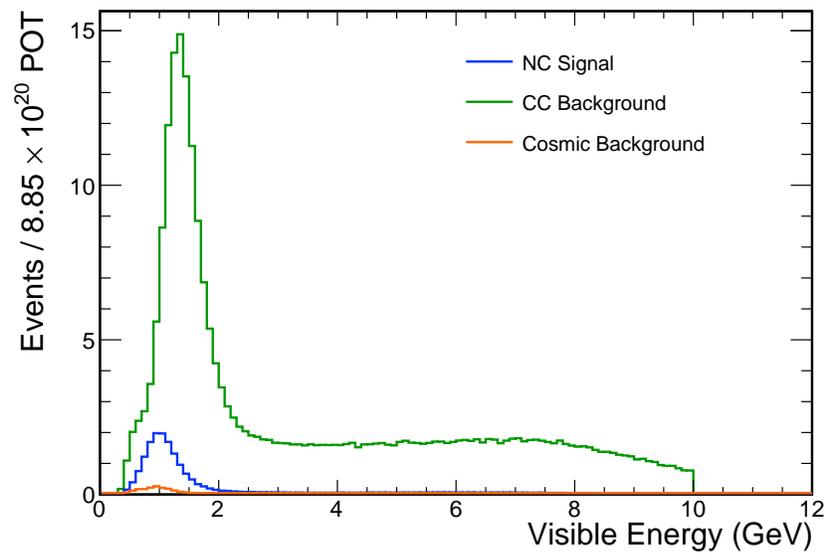


FIGURE 4.85: FD 2017 Re-Analysis Side Band Study Energy Spectra (Direct Selection Method).

Chapter 5

Extraction of Sterile Mixing

Parameters

“There’s two possible outcomes: if the result confirms the hypothesis, then you’ve made a discovery. If the result is contrary to the hypothesis, then you’ve made a discovery.”

— Enrico Fermi

The analysis presented in this dissertation was designed to make several measurements. As we have introduced in Chapter 1, several experimental results cannot be explained by the three-neutrino flavor model. The NOvA NC disappearance analysis is contributing to the global picture of the light sterile mixing parameters in the framework of 3+1 neutrino mixing since 2016. Within NOvA’s 810 km baseline, potential sterile neutrino mixing effects from $\nu_\mu \rightarrow \nu_s$ on the NOvA NC data are primarily driven by θ_{34} and θ_{24} , over a large range of values of Δm_{41}^2 . The FD data used in the NC disappearance analyses began to be collected on 6th February 2014. Until May 2016, 6.69×10^{20} POT or 6.05×10^{20} full

detector equivalent POT were collected and the corresponding results were published in PRD [24]. Until 20th February 2017, 9.0×10^{20} POT FD data have been collected, corresponding to 8.85×10^{20} full detector equivalent POT. The analysis introduced in this dissertation is based on the data collected until 20th February 2017, which is why we call them 2017 analyses. The details of the method to extract sterile mixing parameters is introduced in the following sections of this chapter.

Figure 5.1 displays the analysis flow chart. The comparison between the measured (unblinded) FD data spectra with the prediction allows us to make inferences about the underlying active-sterile neutrino oscillation physics. As it is shown, it includes two steps: FD energy spectrum prediction; and fitting the predicted and measured FD spectra. The FD spectrum prediction chain consists of three adjoining steps. They are the ND Decomposition, Near-to-Far Extrapolation, and FD Prediction. Therefore, the FD spectrum is constructed by merging information from the ND measured data with the ND and FD simulated data.

Estimation of Statistical and systematic uncertainties plays an essential role in extracting physical quantities in a high-energy physics experiment. The systematic uncertainties chiefly arise from the measurement inaccuracy, limitation of theoretical assumptions, or shortcomings in simulating the particle interactions. The NC disappearance analysis is impacted by a series of systematic uncertainties. The functionally-identical detector design between ND and FD is essential in constraining the FD prediction, as cross-section and neutrino beam flux uncertainties change the energy spectra of both NOvA detectors in a similar

way and therefore largely cancel off. On the other hand, there are also some effects which are different in ND and FD. Furthermore, different systematic uncertainties are expected to have a disparate impact on signal and backgrounds. Accordingly, uncertainties from signal and backgrounds need to be estimated separately. The main references describing systematic uncertainties in the NO ν A oscillation analyses are [75, 76, 77, 78].

The final analysis step is fitting the data, which compares the measured FD spectra with the prediction extrapolated from the ND measurement to extract the 3+1 mixing matrix parameters. The techniques for fitting the binned data, including the treatment of oscillation parameters, and systematic uncertainties are described below. Simulated FD data is also employed to assess the sensitivity of the analysis.

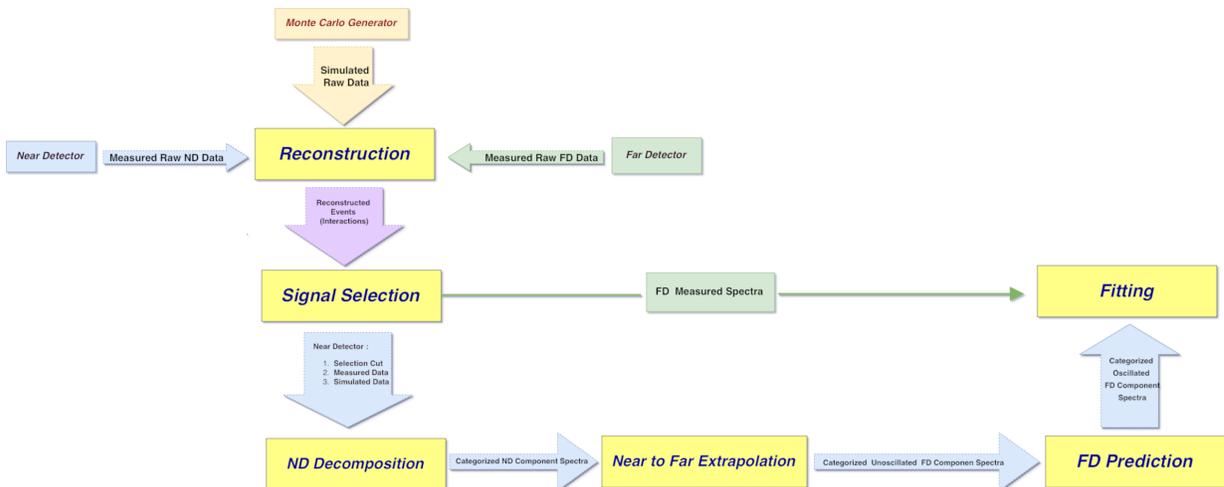


FIGURE 5.1: NO ν A NC Disappearance Analysis flow chart. The blue flow shows how the ND data is employed for the FD spectrum prediction, and the green flow shows how the FD data is used for extracting the mixing parameters by fitting the predicted and measured FD spectra.

Finally, in the first part of the analysis result section, we present the existing

NOvA 2017 Official analysis results and compare them with the 2016 published results. The other primary goal of the analyses was to demonstrate the strong ability of the state-of-the-art machine learning algorithms to identify NC events. Therefore, in the second part of the analysis result section, we present the improved results from the 2017 Re-Analysis compared to the 2017 Official one.

Chapter Organization

- Section 1: 3 + 1 Model
- Section 2: Far Detector Spectra Prediction
- Section 3: Systematic Uncertainty Analysis
- Section 4: Fitting Procedure
- Section 5: Analysis Results

Employed Data Sets and Normalization Factors

Near Detector ND data and Monte Carlo files were produced by the NOvA production group as detailed in the Data Production chapter, Chapter 3.

Far Detector As introduced in the Data Production Chapter, there are three types of FD files: Cosmic Background file, Beam Simulation (FD Monte Carlo) file, and FD data file.

5.1 3 + 1 Model

The NOvA NC analysis considers sterile neutrinos in a 3 + 1 model [79], which adds one more flavor state, ν_s , and one more mass state, ν_4 , therefore expanding the PMNS matrix to a 4×4 matrix, as shown in Fig 5.2.

$$U = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} & U_{e4} \\ U_{\mu1} & U_{\mu2} & U_{\mu3} & U_{\mu4} \\ U_{\tau1} & U_{\tau2} & U_{\tau3} & U_{\tau4} \\ U_{s1} & U_{s2} & U_{s3} & U_{s4} \end{pmatrix}$$

FIGURE 5.2: The PMNS matrix (in the red box) and the expanded 4×4 mixing matrix.

Expanding the PMNS matrix from 3×3 to 4×4 results in six new parameters: one mass splitting (Δm_{41}^2), two CPV phases (δ_{14} , δ_{24}) three mixing angles (θ_{14} , θ_{24} , and θ_{34}).

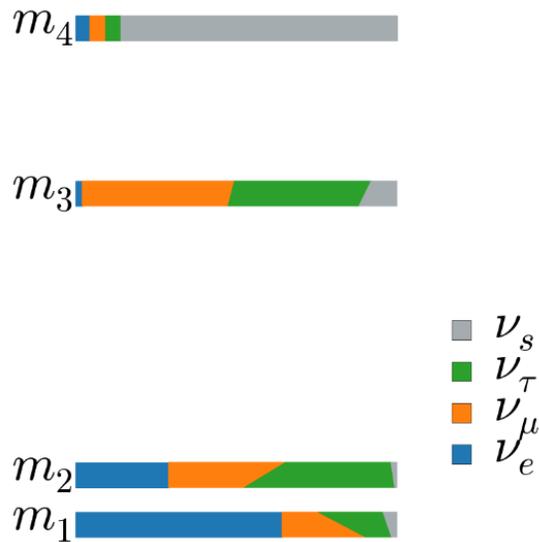


FIGURE 5.3: Example of a 3+1 neutrino flavor model mass diagram.

The rate of NC interactions is constant for mixing among three neutrino flavors, however if oscillations into a sterile neutrino occur within the NOvA baseline, a depletion of the NC interactions would be seen in the FD with respect to the ND measurement. This depletion corresponds to the $1 - P(\nu_\mu \rightarrow \nu_s)$ oscillation probability, which is approximately described by:

$$\begin{aligned}
1 - P(\nu_\mu \rightarrow \nu_s) \approx & 1 - 4 |U_{\mu 3}|^2 |U_{s 3}|^2 \sin^2 \Delta_{31} \\
& - 4 |U_{\mu 4}|^2 |U_{s 4}|^2 \sin^2 \Delta_{41} \\
& - 8 \Re(Z) \sin \Delta_{31} \cos \Delta_{43} \sin \Delta_{41} \\
& - 8 \Im(Z) \sin \Delta_{31} \sin \Delta_{43} \sin \Delta_{41}
\end{aligned} \tag{5.1}$$

where

$$Z = U_{\mu 4}^* U_{s 4} U_{\mu 3} U_{\mu 3}^*$$

A detailed derivation of the exact four-flavor sterile neutrino appearance probability used for fitting the sterile mixing parameters by the NOvA NC Disappearance analysis is presented in Ref. [79]. In the following part of this section, we discuss the effect of different mixing matrix elements by showing how the $1 - P(\nu_\mu \rightarrow \nu_s)$ oscillation probability behave within the L/E and neutrino energies sampled by the NOvA Near and Far Detectors. In the following figures the effects of the various mixing parameters on the oscillation probability are shown. In Figure 5.4, θ_{23} controls the amplitude of the oscillation maximum; In Figure 5.5, θ_{24} modulates the amplitude of oscillations away from the maximum if the sterile mixing parameters have nonzero values; In Figure 5.6, the amplitude of oscillations at the oscillation maximum is primarily determined by θ_{34} ; and in Figure 5.7, the CP-violating phase δ_{24} can reduce the NC disappearance and shift the position of the oscillation maximum.

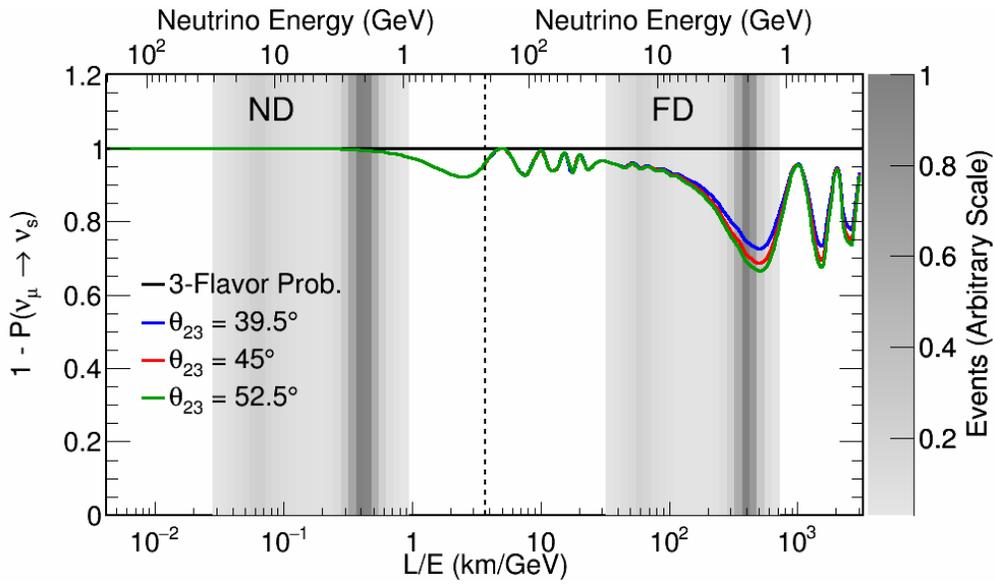


FIGURE 5.4: θ_{23} influence on NC Disappearance

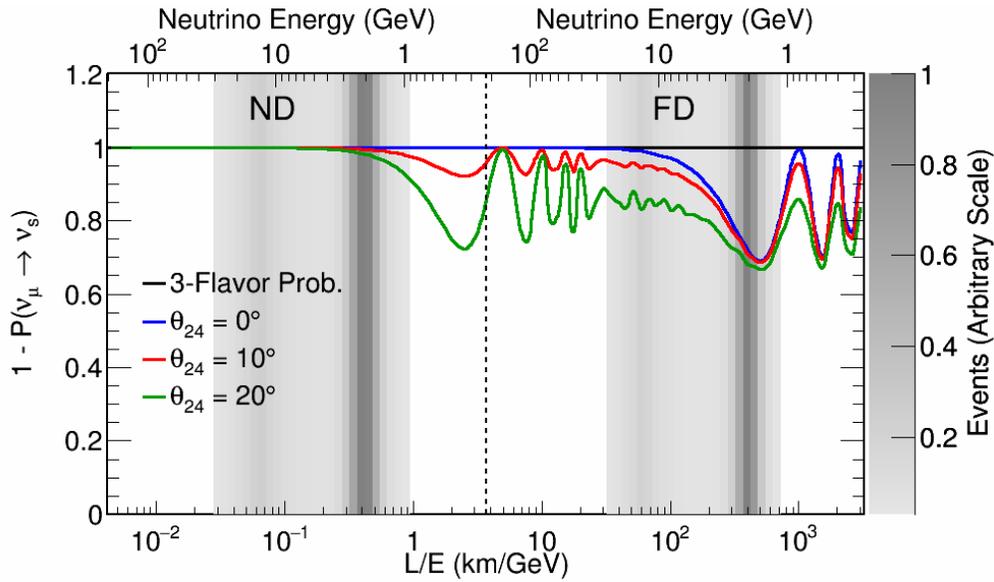
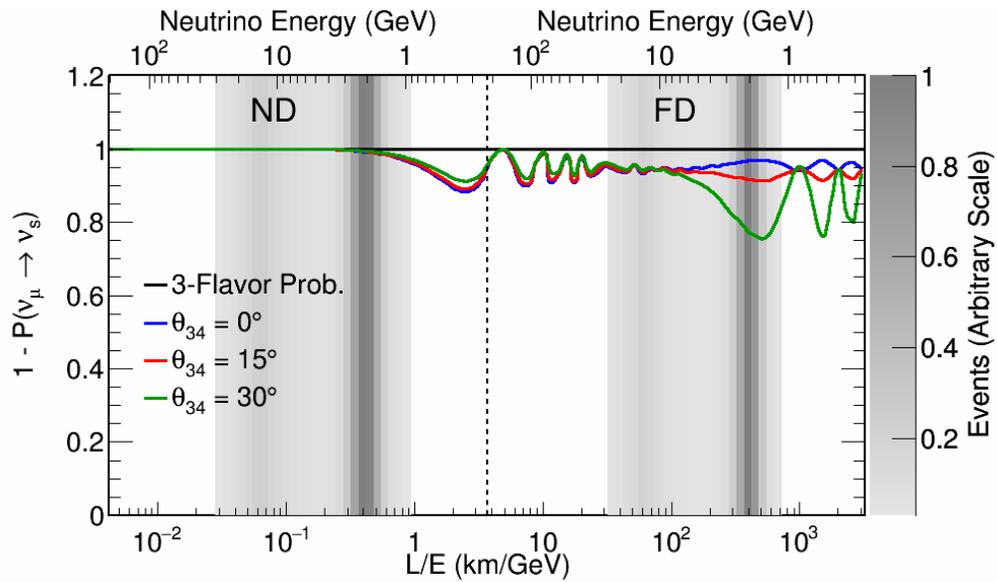
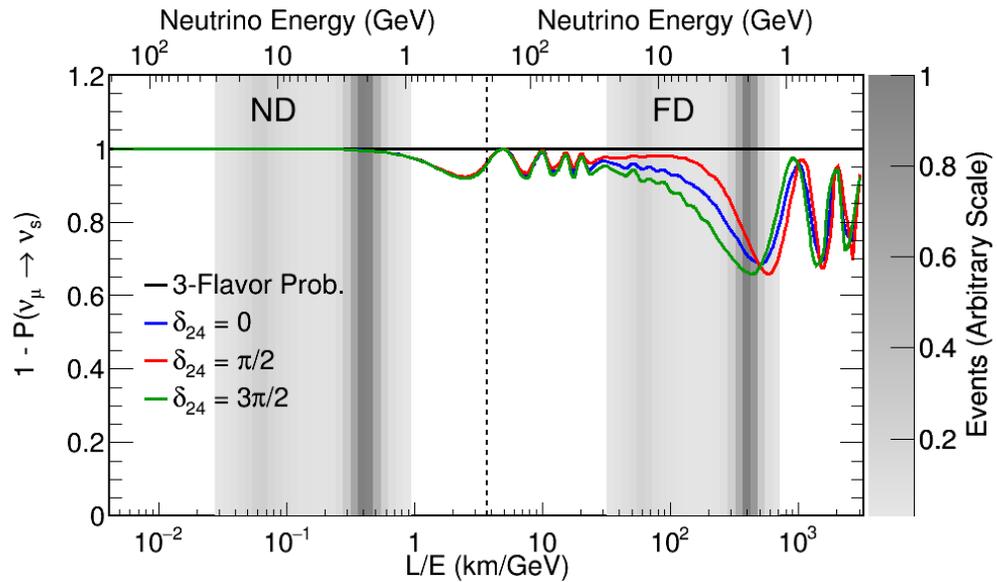


FIGURE 5.5: θ_{24} influence on NC Disappearance

FIGURE 5.6: θ_{34} influence on NC DisappearanceFIGURE 5.7: δ_{24} influence on NC Disappearance

5.2 Far Detector Spectra Prediction

The prediction chain, as shown in Fig 5.8, is performed by three steps in the NOvA Common Analysis Framework (CAF). The first link of the prediction chain is the ND decomposition, which categorizes survived ND data spectrum into a set of component (neutrino flavor) spectra based on the corresponding survived simulated ND data spectra. The next step is called Near-to-Far Extrapolation, which employed the measured and simulated ND data and the simulated FD data to produce a set of unoscillated FD component spectra. Finally, in the FD prediction step, the oscillation weights have been applied on the unoscillated spectra and return a merged FD prediction spectrum in bins of reconstructed energy.

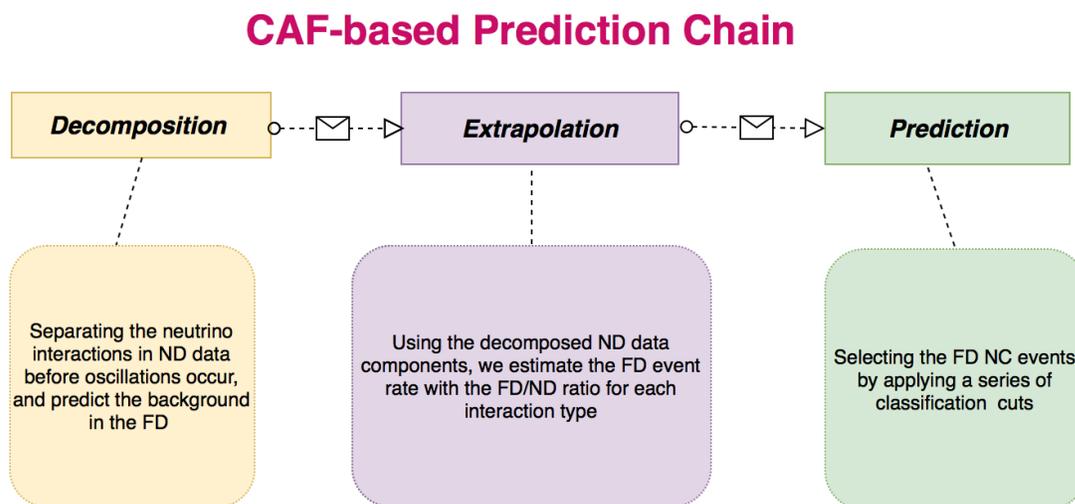


FIGURE 5.8: The CAF-based Prediction Chain includes three steps: Decomposition, Extrapolation and Prediction. The plot shows the basic function of each step and the corresponding info which to be sent to the next step.

5.2.1 Near Detector Decomposition

Different neutrino flavors and CP-conjugate flavor states have different oscillation weights, due to the different values of the mixing angles, and matter effects, respectively. This requires categorizing the ND measured data into different components before oscillations are applied [80]. In the ND Decomposition, we classify the surviving data that passed our selection cuts into five component spectra, NC, ν_μ CC, $\bar{\nu}_\mu$ CC, ν_e CC, and $\bar{\nu}_e$ CC (given there is no ν_τ or $\bar{\nu}_\tau$ in the ND) [81]. In the CAF framework, it is performed by decomposing the measured data proportionally to the simulated data that passed the same selection cuts. The complete normalization of the data is respected for each energy bin, but the percentage of events attributed to each category was set to that found in the simulated ones.

5.2.2 Near-to-Far Extrapolation

The following step in the analysis chain is to produce each oscillation component in bins of true energy and calorimetric (reconstructed) energy by extrapolating the measured ND data to the FD using the simulated data of both ND and FD [82]. This process is called the Near-to-Far Extrapolation. Until our 2017 analyses, we employ the Far-Over-Near (F/N) direct extrapolation method as shown in the following function.

$$FD^{Pred} = ND^{Data} \frac{FD^{MC}}{ND^{MC}} = FD^{MC} \frac{ND^{Data}}{ND^{MC}} \quad (5.2)$$

The basic idea is to apply an F/N ratio to the simulated FD data. The main advantage to employ this method is that it can cancel the systematic effects to first order due to NOvA using two functionally-identical detectors. To actually perform the extrapolation, we categorize the oscillation components into three

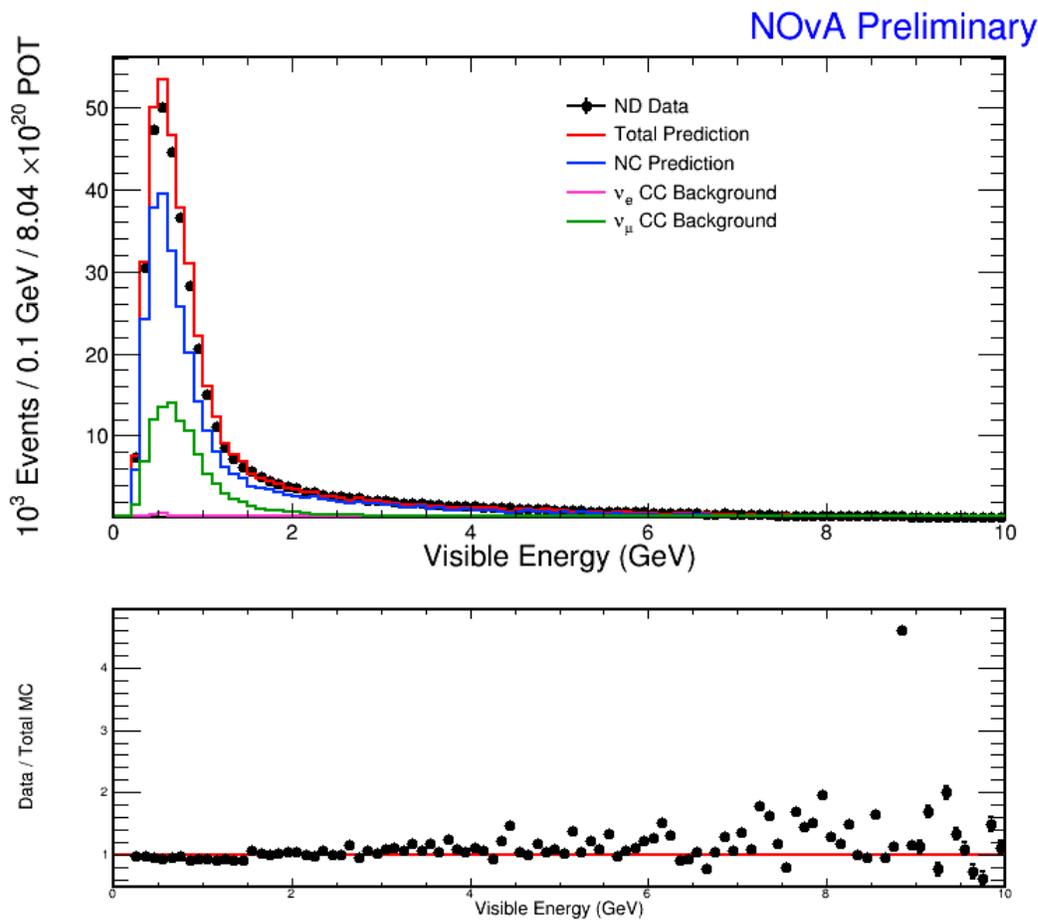


FIGURE 5.9: ND Data vs MC reconstructed energy spectra. In the simulated spectra, we only present the ν_μ CC and ν_e CC spectra, since the number of antineutrino interactions is small in the NOvA ND. Also, the above plot shows the NC ND measured and simulated data agreement. In the important energy range (0.25 - 5 GeV), it shows very good agreement between the simulation and real data.

groups based on their statistics and energy resolution.

Truth Energy Extrapolation

The first group is called true extrapolation, which is designed for $\nu_\mu \rightarrow \nu_e$ and $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$ oscillation. This appearance component is the primary goal of NOvA's

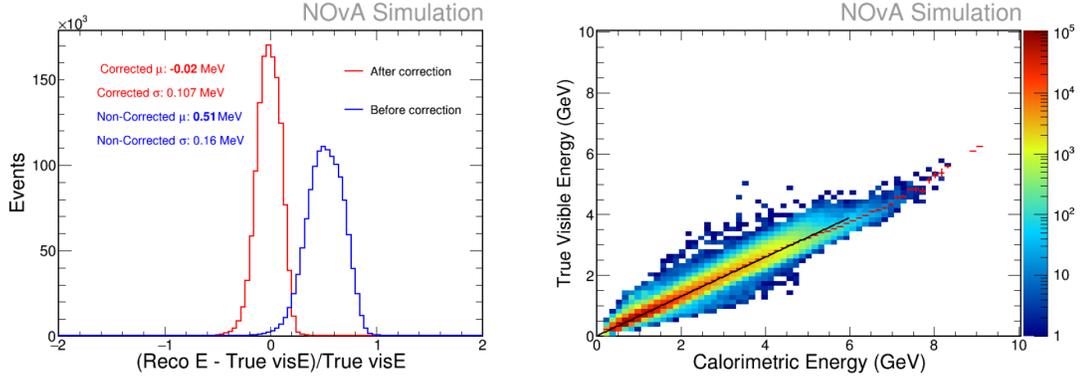


FIGURE 5.10: ND 2D plot of comparison of reco energy to the true energy (right), and the energy resolution (ND) before and after correction (left). For the 2016 analysis, the sliced reco energy has been employed, which is defined as the sum of the energy of the reconstructed cells times a simple scale factor, to correct for dead material and thresholds. A new correction factor was determined using a linear fit to adjust the sliced reconstructed energy for the 2017 analyses. Therefore, the reconstructed energy matches the true energy of NC events better. Separate correction factors were determined for both the Near and Far Detectors. [83].

three-flavor oscillation measurements. The ν_μ selection cuts, produced by the NOvA ν_μ disappearance analysis group, are employed on the ND to select the events. These components have enough statistics so that the F/N ratio can be applied in bins of true energy as it shown in the following equations:

$$FD_{\alpha \rightarrow \beta}^{Pred}(S_{NC}; E_i^R, E_j^T) = \frac{FD_{\alpha \rightarrow \beta}^{MC}(S_{NC}; E_i^R, E_j^T) \cdot ND_{\alpha}^{Data}(S_{NC}; E_i^R)}{ND_{\alpha}^{MC}(S_{NC}; E_i^R)} \quad (5.3)$$

$$ND_{\alpha}^{Pred}(S_{\nu_\mu}; E_j^T) = \sum_k \frac{ND_{\alpha}^{MC}(S_{\nu_\mu}; E_k^R, E_j^T) \cdot ND_{\alpha}^{Data}(S_{\nu_\mu}; E_i^R)}{ND_{\alpha}^{MC}(S_{\nu_\mu}; E_i^R)} \quad (5.4)$$

In the above equation, the combination of terms inside the sum normalizes the calorimetric energy columns to the data values, then sums across the true energy

row to find the ND true energy predicted value.

Reco Energy Extrapolation

When the components' statistics or energy resolution degrades, another F/N extrapolation method has been applied. It is called Reco Energy Extrapolation, due to the F/N ratio being applied only in bins of reconstructed energy. This method is used for the NC signal, $\nu_\mu \rightarrow \nu_\mu$ and $\nu_e \rightarrow \nu_e$ CC survival components. The small contribution from ν_e are the intrinsic beam ν_e , as introduced in the NOvA chapter.

$$FD_{\alpha \rightarrow \beta}^{Pred}(S_{NC}; E_i^R, E_j^T) = \frac{FD_{\alpha \rightarrow \beta}^{MC}(S_{NC}; E_i^R, E_j^T) \cdot ND_{\alpha}^{Data}(S_{NC}; E_i^R)}{ND_{\alpha}^{MC}(S_{NC}; E_i^R)} \quad (5.5)$$

The goal of the extrapolation is to predict the number of events from each component that are selected as NC-like. Therefore, for survival components, the NC ND and FD selection cuts are employed.

Base MC Extrapolation

This is the simplest case, the MC is taken as the prediction.

$$FD_{\alpha \rightarrow \beta}^{Pred}(S_{NC}; E_i^R, E_j^T) = FD_{\alpha \rightarrow \beta}^{MC}(S_{NC}; E_i^R, E_j^T) \quad (5.6)$$

where FD^{MC} denotes a quantity calculated from the MC. It is applied for the smallest CC background components with the lowest statistics, $\nu_\mu \rightarrow \nu_\tau$, $\bar{\nu}_\mu \rightarrow \bar{\nu}_\tau$, $\nu_e \rightarrow \nu_\mu$, $\bar{\nu}_e \rightarrow \bar{\nu}_\mu$, $\bar{\nu}_\mu \rightarrow \bar{\nu}_\mu$, and $\bar{\nu}_e \rightarrow \bar{\nu}_e$.

Final Extrapolation Output

Though we list the NC spectrum as independent, only the flavor eigenstates affect the oscillations. Based on the three-flavor model, NC interactions are not affected by flavor oscillations. However, when sterile states are considered, oscillations do affect the NC rate in ND and FD, and the neutrino flavor breakdown of the NC interactions does matter. The flavor state does not matter if the final state is one of the three active ones. Therefore, it suffices to split the selected NC events based on their four initial flavor states. We perform this process by what is described above and splitting the FD result with flavor proportions constructed from the simulated FD data.

$$FD_{NC,\alpha}^{Pred}(S_{NC}; E_i^R, E_j^T) = FD_{NC}^{Pred}(S_{NC}; E_i^R, E_j^T) \frac{FD_{NC,\alpha}^{MC}(S_{NC}; E_i^R)}{FD_{NC}^{MC}(S_{NC}; E_i^R)} \quad (5.7)$$

5.2.3 Far Detector Spectra Prediction

The final step is called FD Spectra Prediction, in which the corresponding oscillation weights are applied to each component spectra. The extrapolation output contains true energy information due to the oscillation weights being applied in bins of true energy. For a given component, the predicted spectrum is then calculated by summing over the bins of true energy, returning a one-dimensional spectrum in reco energy bins.

$$FD_{\alpha \rightarrow \beta}^{Pred}(S_{NC}; E_i^R) = \sum_j FD_{\alpha \rightarrow \beta}^{Pred}(S_{NC}; E_i^R, E_j^T) \cdot P(\nu_\alpha \rightarrow \nu_\beta, E_j^T) \quad (5.8)$$

The selected NC-like events includes four components given the initial neutrino eigenstates. The oscillation probability of the four type neutrinos are considered and applied separately.

$$FD_{NC,\alpha}^{Pred}(S_{NC}; E_i^R) = \sum_j FD_{NC,\alpha}^{Pred}(S_{NC}; E_i^R, E_j^T) \cdot P(\nu_\alpha \rightarrow \nu_{active}, E_j^T) \quad (5.9)$$

The final step is to sum over all the selected NC-like spectra which in fact include signal and backgrounds as we introduced in the signal selection chapter.

$$FD^{Pred}(S_{NC}; E_i^R) = \sum_{Component} FD_{Component}^{Pred}(S_{NC}; E_i^R) \quad (5.10)$$

Based on the description above, the NOvA CAF-based prediction chain produces the FD-predicted spectra, which will be used to compare with the measured FD spectrum to extract the mixing matrix parameters. This fitting procedure will be described in the following section.

5.3 Systematic Uncertainty Analysis

The method we employed to analyze NC disappearance systematic uncertainties is to run the full extrapolation process and produce a predicted spectrum with and without a systematic effect applied. Each systematic effect was used to shift the simulated data at one or both detectors as appropriate. The separation between the predicted and shifted spectra was quantified as a systematic uncertainty.

5.3.1 Mixing Parameter Values Used in Systematic Studies

5.3.2 Beam Systematic Uncertainty

The beam-related uncertainties come from two sources: Beam transport and Package to Predict the FluX (PPFX) simulation [37, 38], as introduced in the

Oscillation Parameter	Value
ρ	$2.84\text{g}/\text{cm}^3$
Δm_{21}^2	$7.53 \times 10^{-5} eV^2$
$\sin^2 2\theta_{12}$	0.846
Δm_{32}^2	$2.67 \times 10^{-3} eV^2$
θ_{23}	0.404 (0.623)
$\sin^2 2\theta_{13}$	0.085
δ	1.48π

following part.

Beam Transport Systematic

The NOvA simulation chain aims to produce the most realistic MC possible. Any mismatch between reality and simulation may produce systematic errors. The beam related uncertainties will be propagated from the ND to the FD. Consequently, they are of much importance in our analyses. The Beam Group studies the potential effects caused by the uncertainties in the neutrino flux simulation. The technique employed to quantify the variety of uncertainties is by changing one of the geometry parameters at a time while simulating NuMI beam and then finding the effect of the change. The considered geometric variables are :

- Horn current: by $100 \pm 2\text{kA}$ (For FHC) and $-100 \pm 2\text{kA}$ (For RHC);
- The magnetic horn-1 position: shifted by $\pm 3\text{mm}$ in X and Y separately;
- The magnetic horn-2 position: shifted by $\pm 3\text{mm}$ in X and Y separately;
- The beam position in the target: shifted $\pm 1\text{mm}$ in X and Y separately;
- The beam spot size: changed $\pm 0.2\text{mm}$ in X and Y (based on the nominal value of beam spot size which is 1.3 mm in X and Y);

- Horn water layer: changed by $\pm 1\text{mm}$ (the nominal water layer is 1mm);
- Target z position: shifted along z by $\pm 7\text{mm}$ (the nominal value is -143.3 cm);
- Full stat simulation files with magnetic field in the decay pipe.

Systematic	NC Difference (%)	Background Difference (%)
Horn Current	0.15	0.25
Magnetic horn-1 Position X	0.57	0.69
Magnetic horn-1 Position Y	0.31	0.53
Magnetic horn-2 Position X	0.29	0.33
Magnetic horn-2 Position Y	0.16	0.18
Beam Position in Target X	0.08	0.08
Beam Position in Target Y	0.02	0.03
Spot Size	0.07	0.19
Horn Water Layer	0.65	0.84
Target Z Position	0.04	0.10
Combined Beam Transport Systematics	0.85	0.93

TABLE 5.1: The percentage difference between the shifted and nominal predictions for the number of extrapolated FD events due to beam transport systematic effects.

The beam transport systematic effects on ND spectra

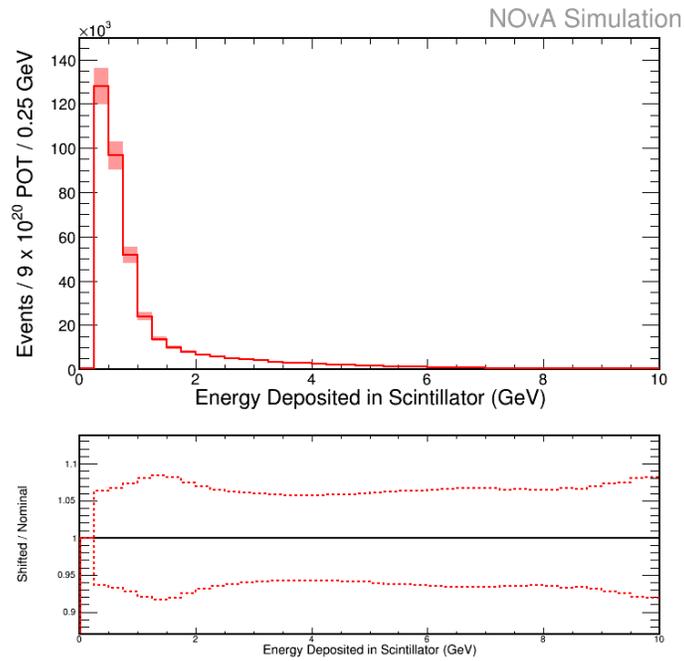


FIGURE 5.11: Beam Transport Uncertainty on ND Signal

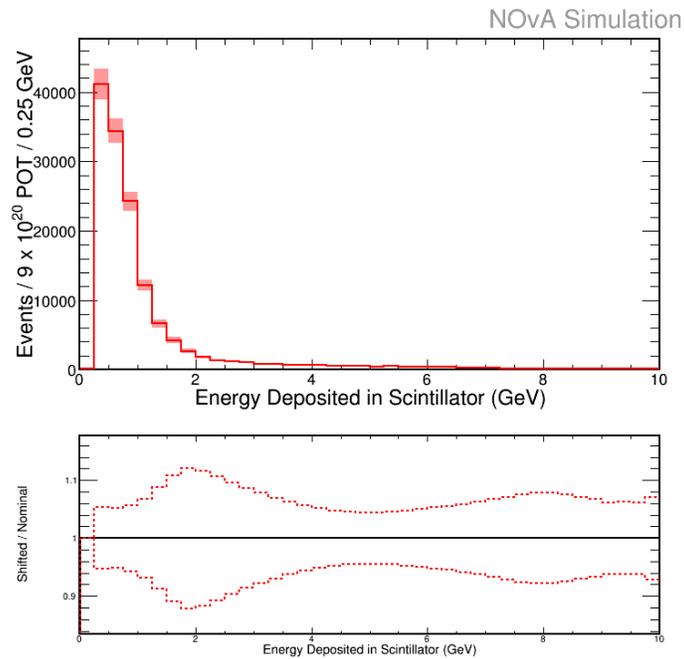


FIGURE 5.12: Beam Transport Uncertainty on ND Background

The beam transport systematic effects on FD spectra (the direct selection results)

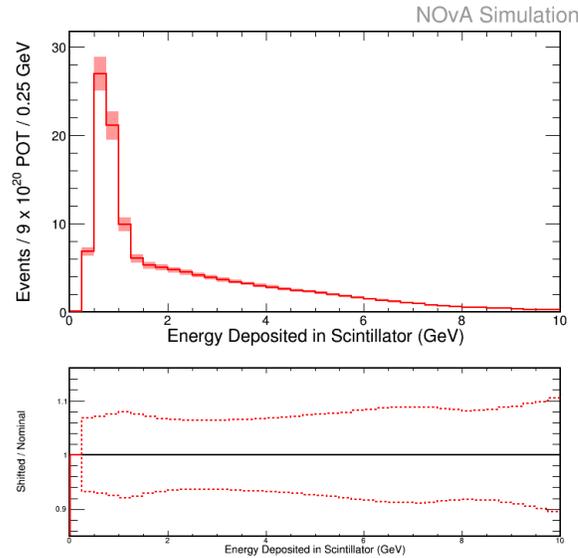


FIGURE 5.13: Beam Transport Uncertainty on FD Signal (direct selection)

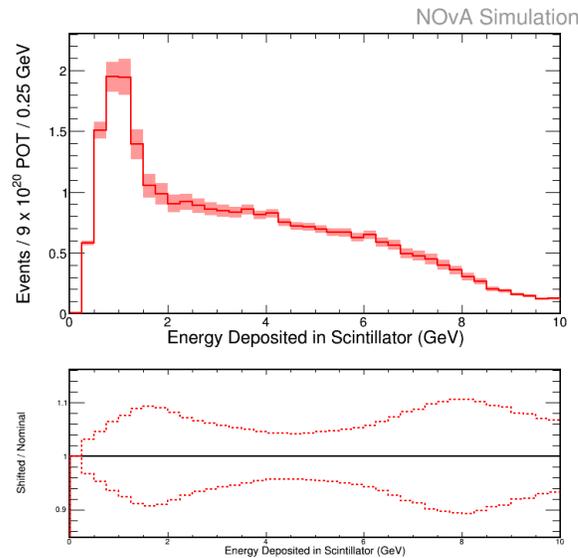


FIGURE 5.14: Beam Transport Uncertainty on FD Background (direct selection)

The beam transport systematic effects on FD spectra (the extrapolated selection result)

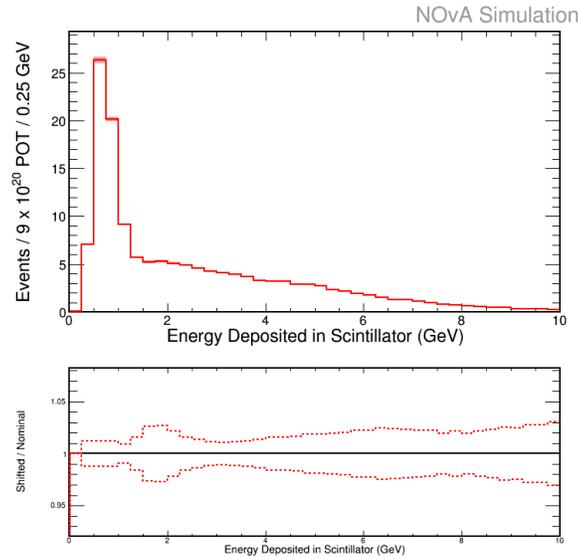


FIGURE 5.15: Beam Transport Uncertainty on FD Signal (extrapolated)

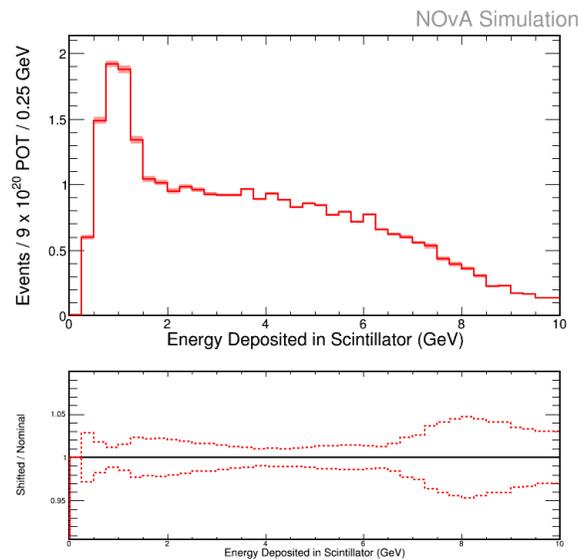


FIGURE 5.16: Beam Transport Uncertainty on FD Background (extrapolated)

Package to Predict the FluX Systematic

The PPFX uncertainties [38] handle the errors which may be generated during the hadron production simulation process in the target zone. The updated data-driven technique has been employed to find the correction value between the simulated data and detected data. The correction from PPFX takes two issues into account: the probability of a specific interaction to happen and the probability of a particular final state.

We applied the Multi-universe technique to propagate uncertainties for NC disappearance analysis. Each universe has been created by using a random value for each free parameter:

- An ensemble of 100 Universes was created for the ND, then the 1σ variation from the central value of the universe is found. Then the shift of the 1σ boundary from a nominal ND spectra was quantified;
- Repeated the same procedure for the FD spectra and found the FD PPFX systematic error;
- The ensemble of 100 ND Universes are passed through the extrapolation framework and then the FD 1σ variation is found. A nominal ND spectra is extrapolated to the FD to get the FD extrapolated nominal spectra. We then quantify the shift of the 1σ boundary from the extrapolated prediction of the nominal spectra.

The PPFX systematic effects on ND spectra.

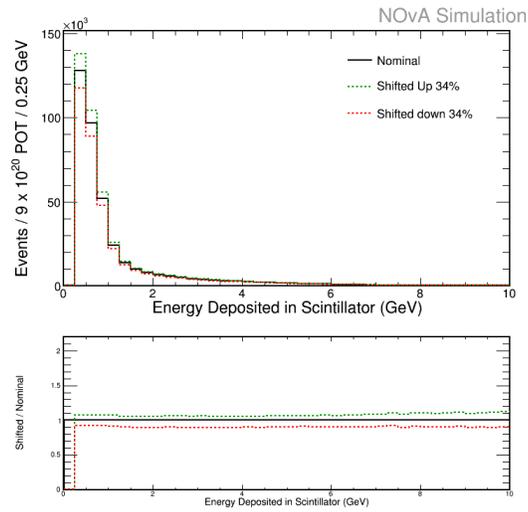


FIGURE 5.17: PPFX Uncertainty on ND Signal.

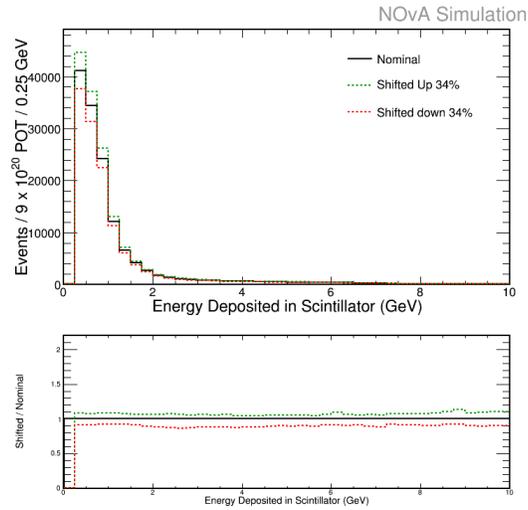


FIGURE 5.18: PPFX Uncertainty on ND Background.

Systematic	NC Difference (%)	Background Difference (%)
ND Up	7.25	7.80
ND Low	8.41	8.67

TABLE 5.2: The percentage difference between the shifted and nominal predictions for the number of ND events due to PPFX systematic effects.

The PPFX systematic effects on FD spectra (direct selection results)

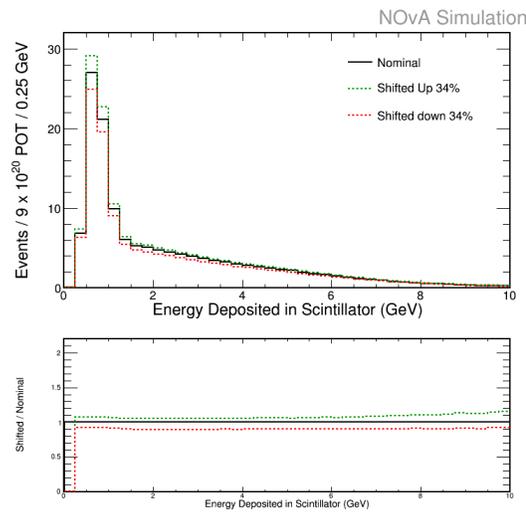


FIGURE 5.19: PPFX Uncertainty on FD Signal (direct selection)

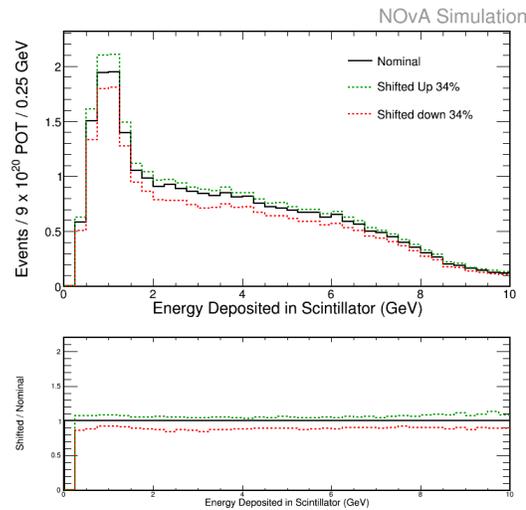


FIGURE 5.20: PPFX Uncertainty on FD Background (direct selection)

Systematic	NC Difference (%)	Background Difference (%)
FD Up	6.70	6.09
FD Low	9.12	10.89

TABLE 5.3: The percentage difference between the shifted and nominal predictions for the number of FD events due to PPFX systematic effects.

The PPFX systematic effects on FD spectra (the extrapolated selection results)

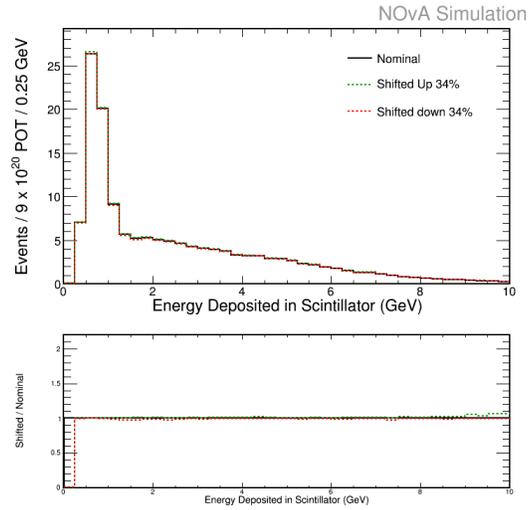


FIGURE 5.21: PPFX Uncertainty on FD Signal (extrapolated)

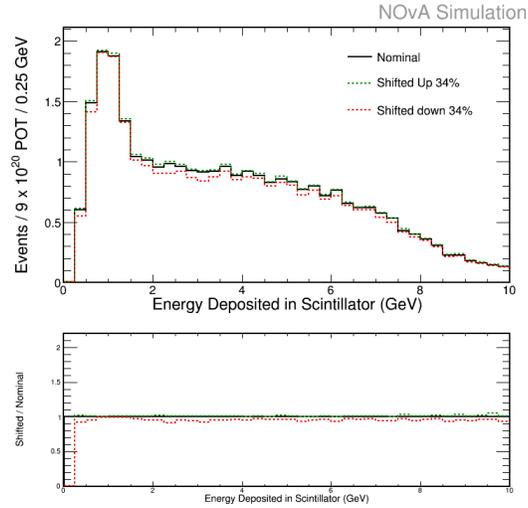


FIGURE 5.22: PPFX Uncertainty on FD Background (extrapolated)

Systematic	NC Difference (%)	Background Difference (%)
FD Up (Extrapolated)	0.86	1.36
FD Low (Extrapolated)	0.99	4.11

TABLE 5.4: The percentage difference between the shifted and nominal predictions for the number of extrapolated FD events due to PPFT systematic effects.

5.3.3 GENIE Simulation

The NOvA simulation employs GENIE to generate neutrino-nuclear interactions which involve a detailed physics modeling of the hadronization, interaction cross-section, and final state interactions (FSI). We studied the NOvA specific list of GENIE systematic uncertainties. Several parameters play pivotal roles in neutrino interaction production. These parameters can decide the daughter particles of that particular interaction. The GENIE developers group has provided the standard deviation for each of these parameters. An event re-weighting scheme introduces the corresponding shifts to a specific interaction. The figure ?? shows the combined GENIE systematic uncertainty in FD.

Systematic	NC Difference (%)	Background Difference (%)
ND	24.13	17.38
FD (direct Selection)	24.29	20.32
FD (Extrapolated)	2.75	6.07

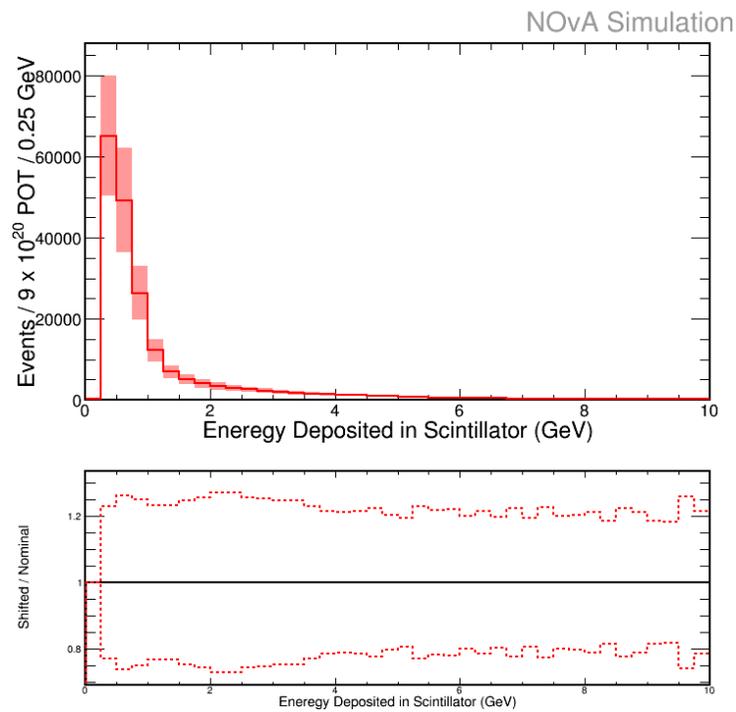
The GENIE systematic effects on ND spectra.

FIGURE 5.23: GENIE Uncertainty on ND Signal

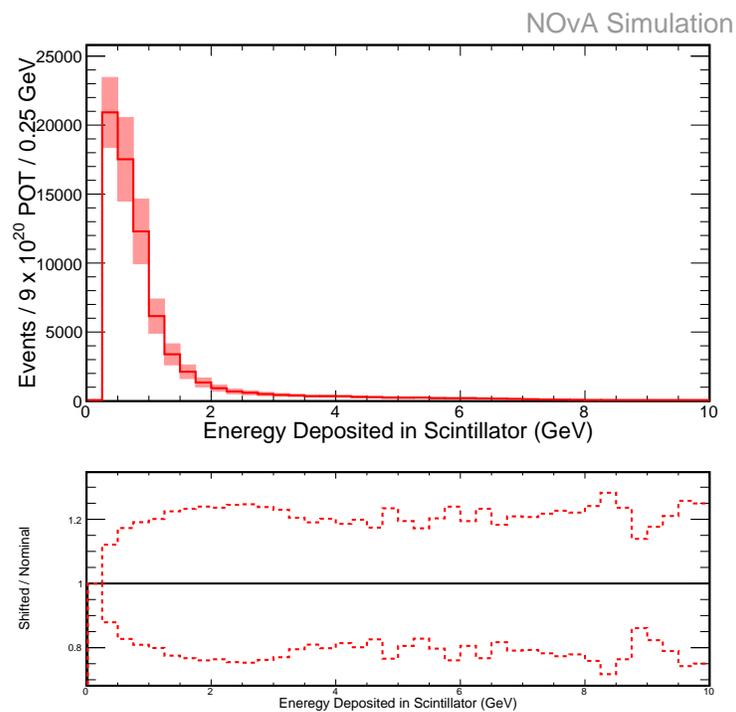


FIGURE 5.24: GENIE Uncertainty on ND Background

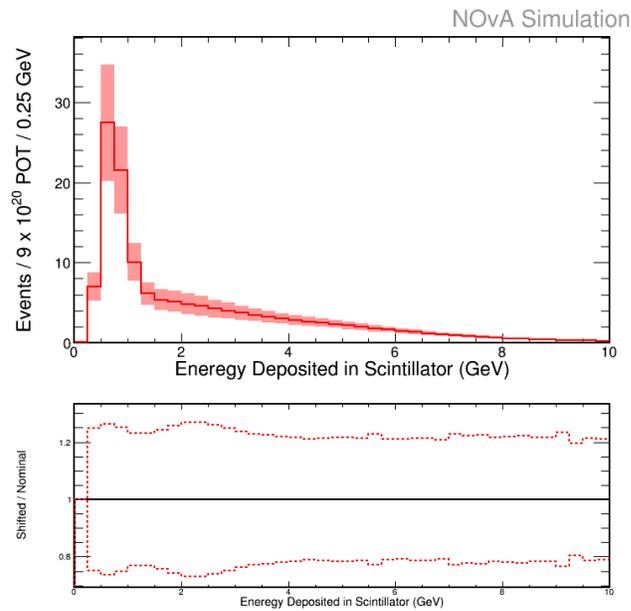
The GENIE systematic effects on FD spectra (the direct selection results)

FIGURE 5.25: GENIE Uncertainty on FD Signal (direct selection)

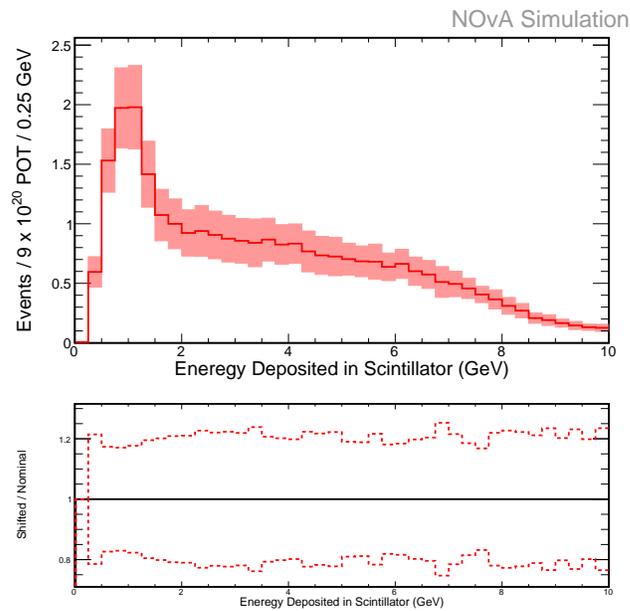


FIGURE 5.26: GENIE Uncertainty on FD Background (direct selection)

The GENIE systematic effects on FD spectra (the extrapolated selection results)

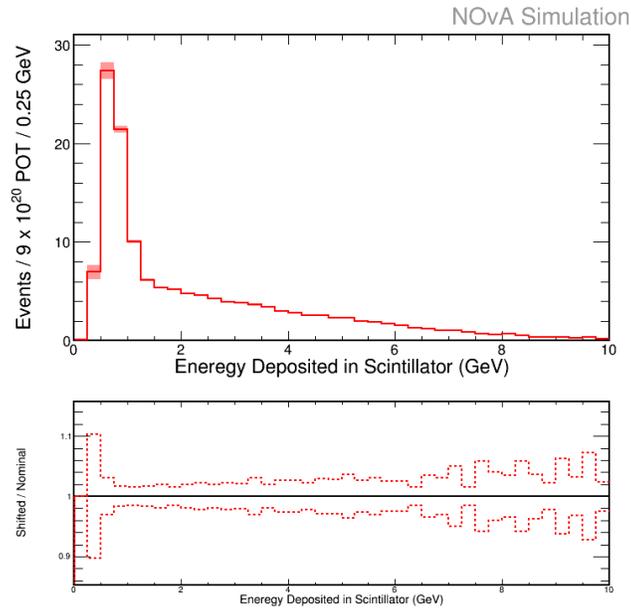


FIGURE 5.27: GENIE Uncertainty on FD Signal (extrapolated)

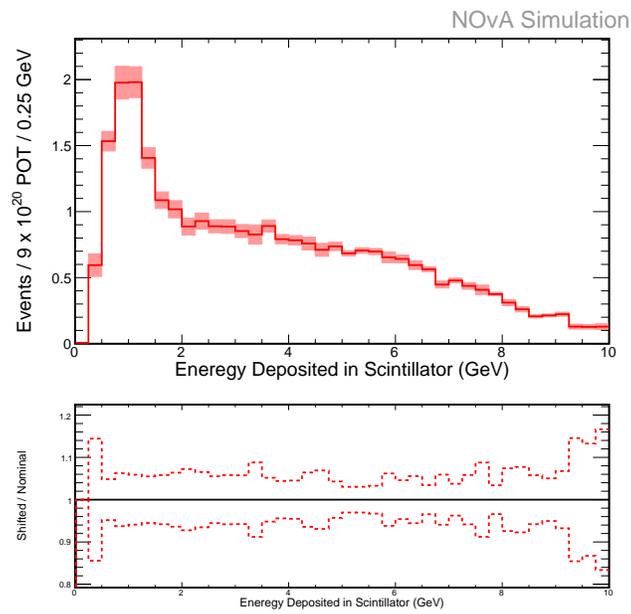


FIGURE 5.28: GENIE Uncertainty on FD Background (extrapolated)

5.3.4 Calibration Systematic

The energy response of the detectors are not uniform through the two detectors.

We categorize the calibration related systematic as:

1. Absolute Calibration
2. Relative Calibration
3. Calibration Shape

Absolute Calibration

Miscalibration may happen in the two detectors separately. The Absolute calibration uncertainty study aims to find the shifts caused by this type of miscalibration. Uncertainties in each detector are calculated by seeing the energy shift from the nominal spectra to the one produced out of fabricated miscalibration files.

Systematic	NC Difference (%)	Background Difference (%)
Flat Scale Up	4.53	2.54
Flat Scale Down	5.91	3.57
Calibration shape	2.25	2.26

The Absolute Calibration systematic effects on ND spectra

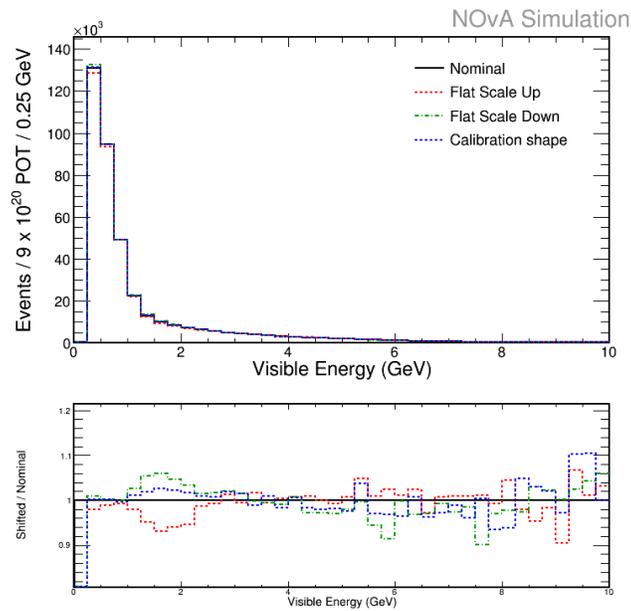


FIGURE 5.29: Absolute Calibration Uncertainty on ND Signal

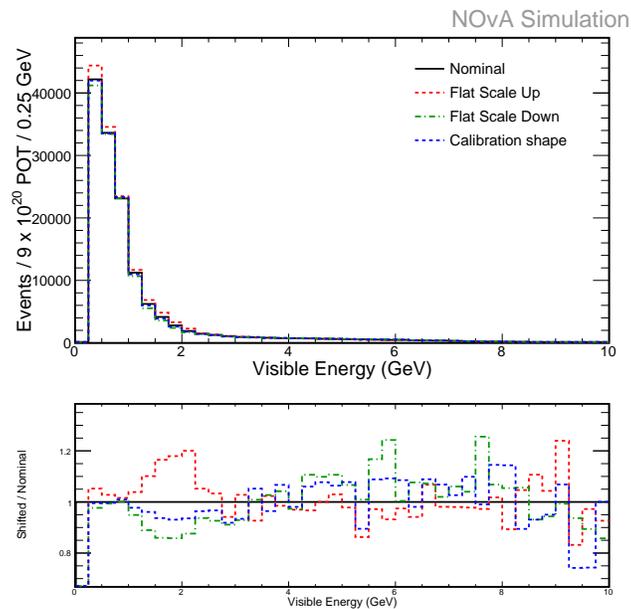


FIGURE 5.30: Absolute Calibration Uncertainty on ND Background

The Absolute Calibration systematic effects on ND spectra (the direct selection results)

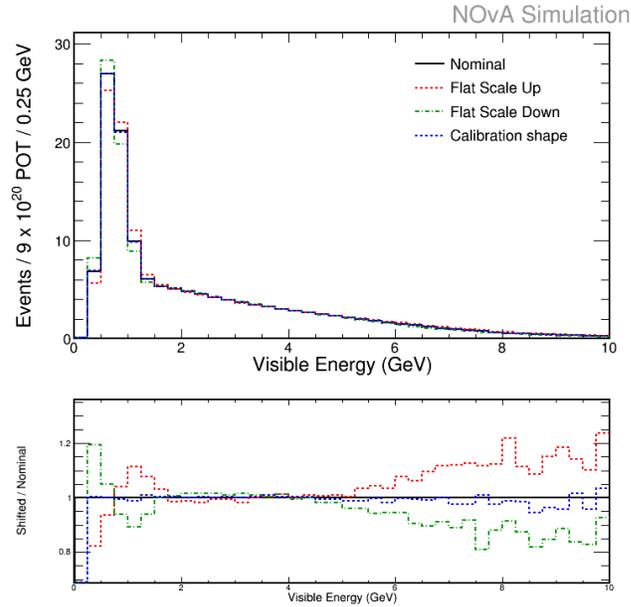


FIGURE 5.31: Absolute Calibration Uncertainty on FD Signal (direct selection)

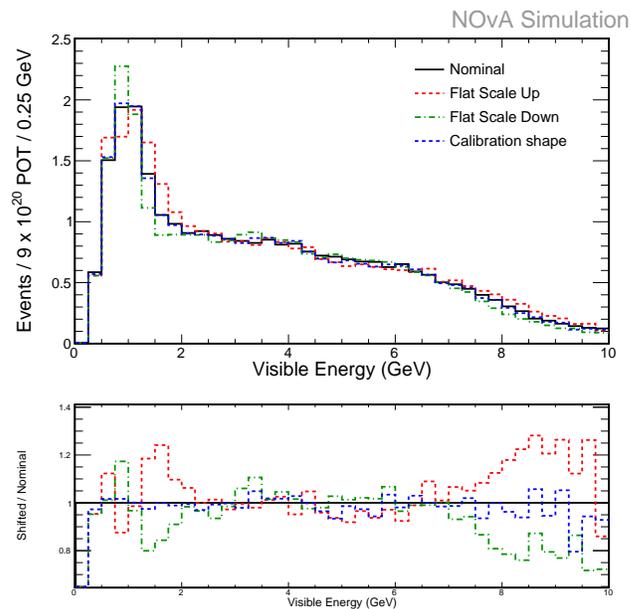


FIGURE 5.32: Absolute Calibration Uncertainty on FD Background (direct selection)

The Absolute Calibration systematic effects on ND spectra (the extrapolated selection results)

Relative Calibration

Another type of miscalibration may arise due to effects in only one of the detectors. Relative calibration studies shift the predicted spectra from the nominal one then quantify the uncertainties that arise as a result of the difference in energy difference between the detectors. Similarly to the absolute calibration study, we have corresponding files for the relative calibration. These fabricated files are having a constant 5% calibration scale up or down. The resulting files are ND 5% calibration up, ND 5% calibration down, FD 5% calibration up and FD 5% calibration down. We apply these shifts one at the time and use the maximum systematic shift from the nominal spectra at ND, FD and FD extrapolate prediction. Repeating the procedure produces corresponding shifts.

Systematic	NC Difference (%)	Background Difference (%)
ND Scale Up	5.83	5.43
ND Scale Down	6.31	5.70
FD Scale Up	0.95	2.33
FD Scale Down	0.70	2.51

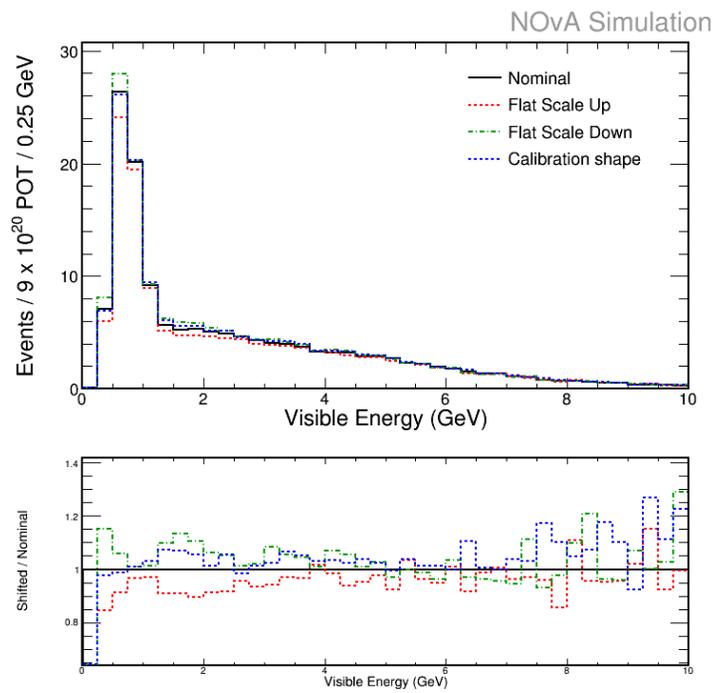


FIGURE 5.33: Absolute Calibration Uncertainty on FD Signal (extrapolated)

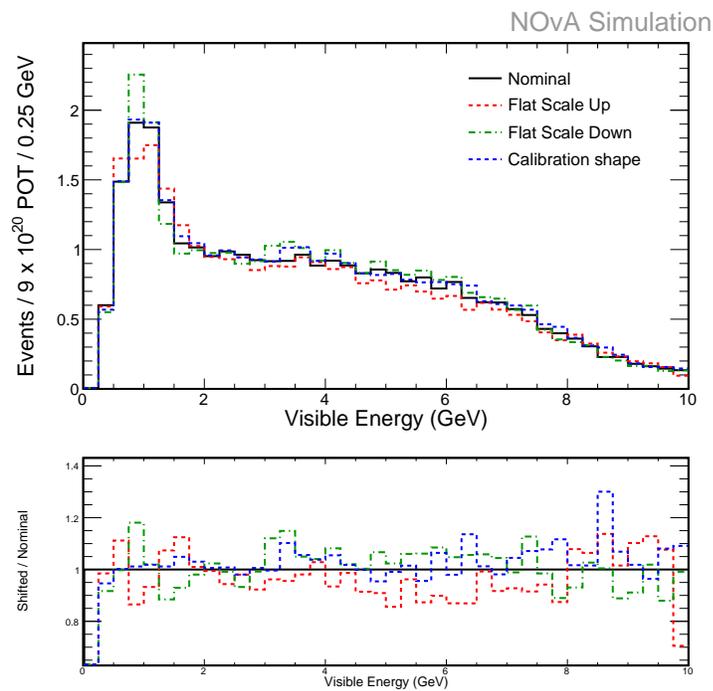


FIGURE 5.34: Absolute Calibration Uncertainty on FD Background (extrapolated)

The Relative Calibration systematic effects on ND spectra

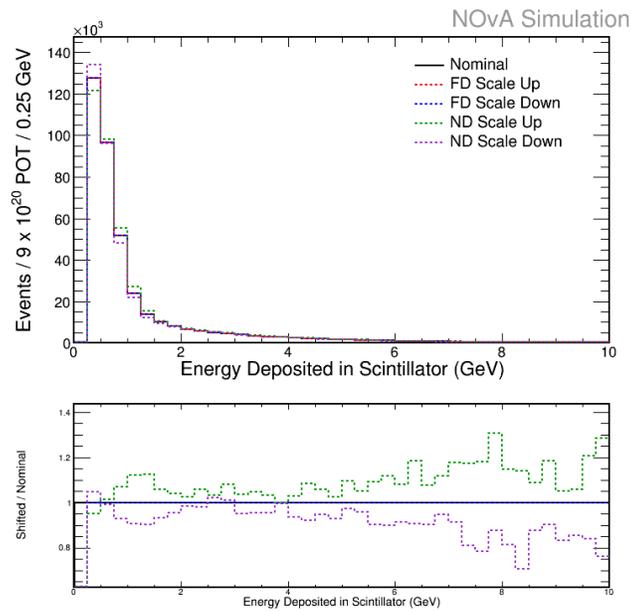


FIGURE 5.35: Relative Calibration Uncertainty on ND Signal

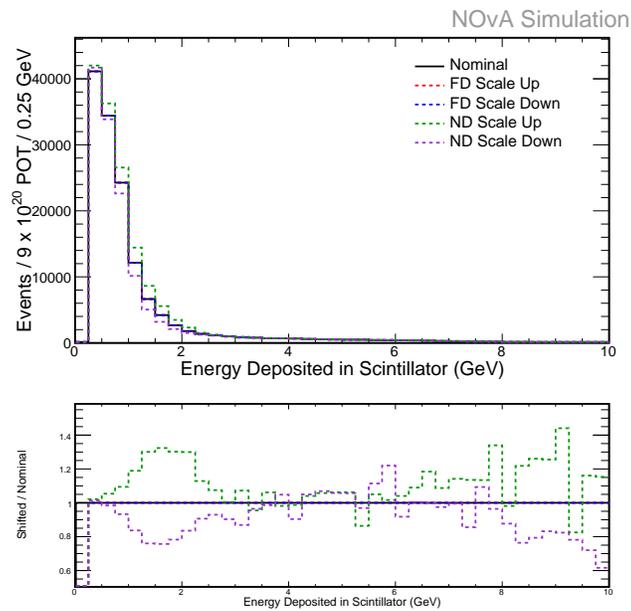


FIGURE 5.36: Relative Calibration Uncertainty on ND Background

The Relative Calibration systematic effects on FD spectra (the direct selection results)

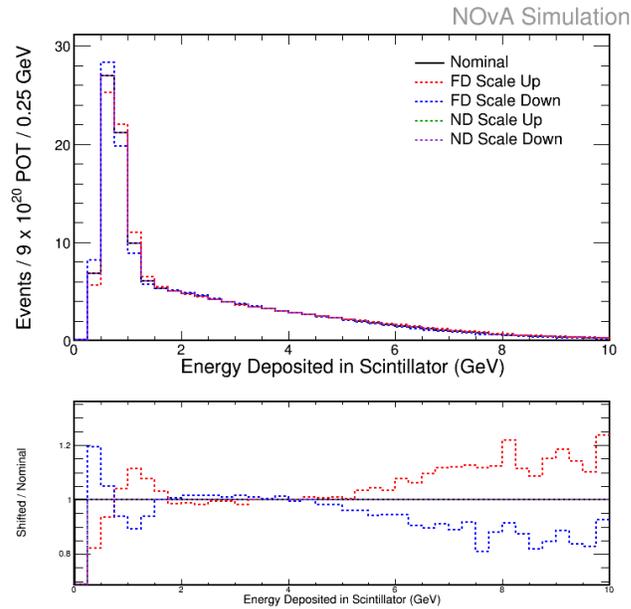


FIGURE 5.37: Relative Calibration Uncertainty on FD Signal (direct selection)

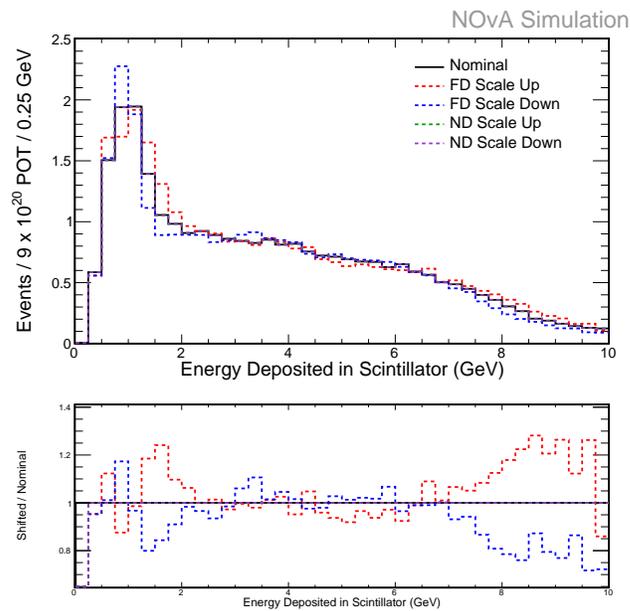


FIGURE 5.38: Relative Calibration Uncertainty on FD Background (direct selection)

The Relative Calibration systematic effects on FD spectra (the extrapolated selection results)

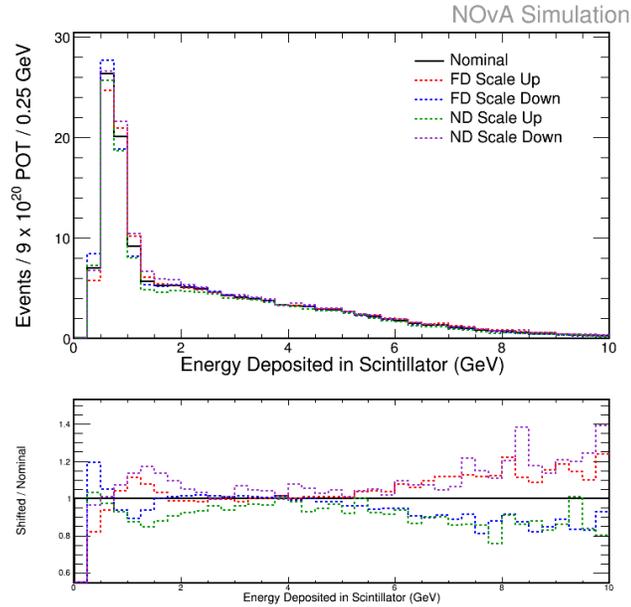


FIGURE 5.39: Relative Calibration Uncertainty on FD Signal (extrapolated)

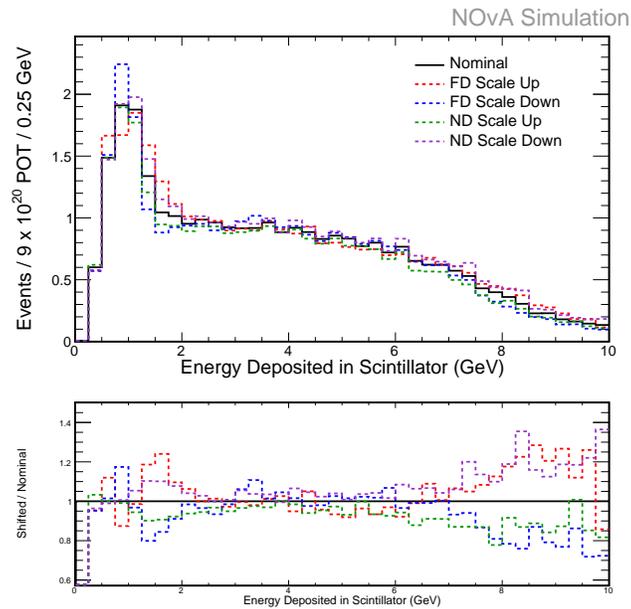


FIGURE 5.40: Relative Calibration Uncertainty on FD Background (extrapolated)

Cherenkov Systematic

The NOvA simulation chain employs the Birks-Chou Law to model the relationship between scintillator light yield and deposited energy. The Cherenkov light response of the detector for the energy deposition by electrons are not similar to that of the protons. We then consider the potential systematic effects caused by the Cherenkov light modeling.

Systematic	NC Difference (%)	Background Difference (%)
Cherenkov	0.76	0.06
Light Level Up	1.74	0.66
Light Level Down	0.0002	0.71

The Cherenkov and Light Level systematic effects on ND spectra

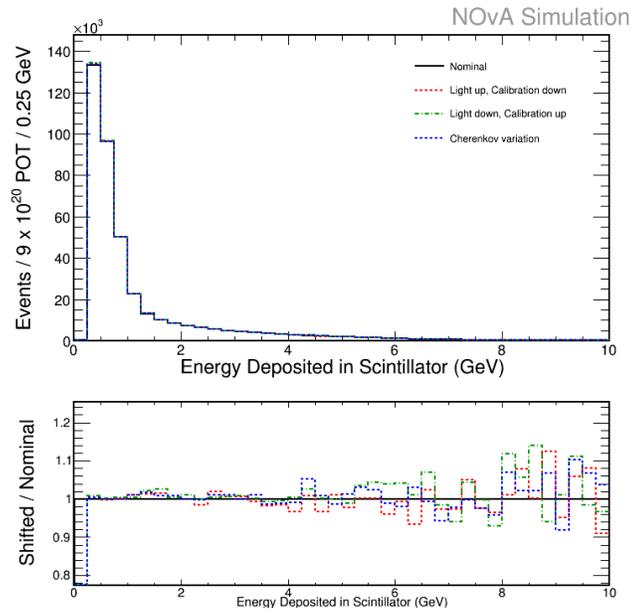


FIGURE 5.41: Cherenkov Systematic Uncertainty on ND Signal

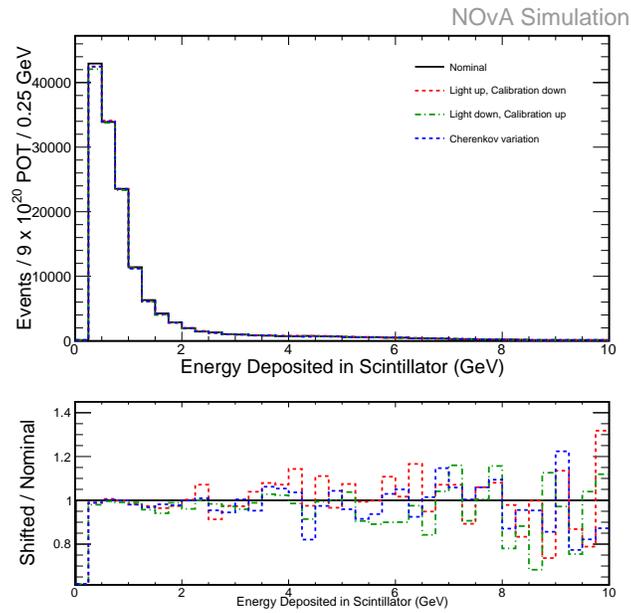


FIGURE 5.42: Cherenkov Systematic Uncertainty on ND Background

The Cherenkov and Light Level systematic effects on FD spectra (the direct selection results)

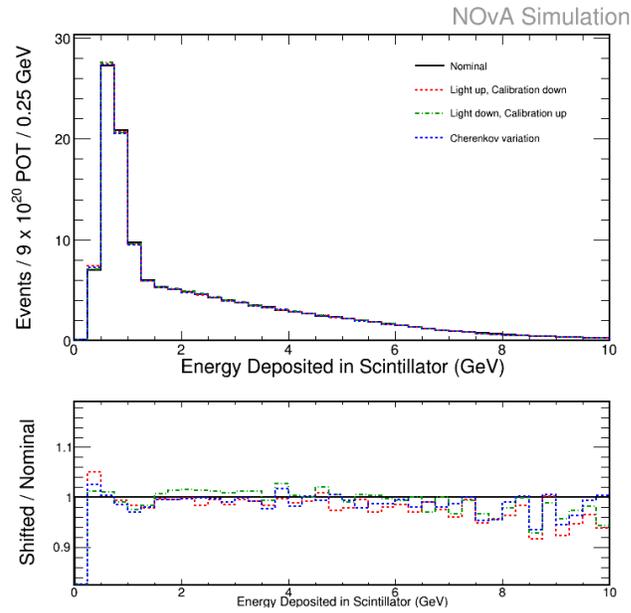


FIGURE 5.43: Cherenkov Systematic Uncertainty on ND Signal (direct selection)

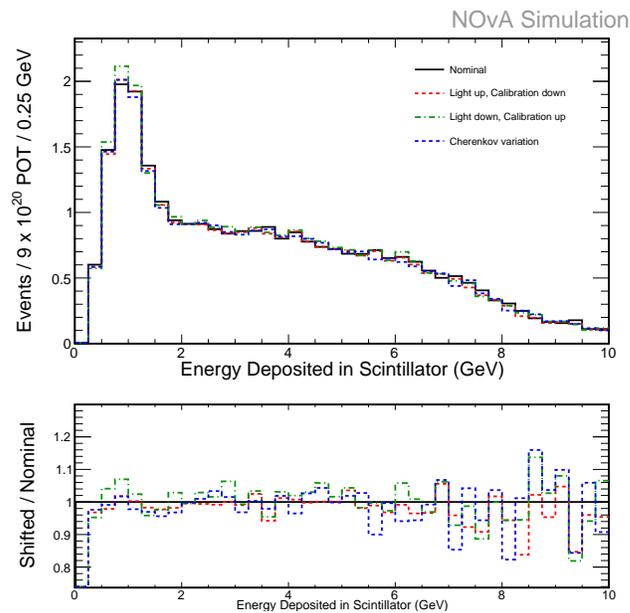


FIGURE 5.44: Cherenkov Systematic Uncertainty on FD Background (direct selection)

The Cherenkov and Light Level systematic effects on FD spectra (the extrapolated selection results)

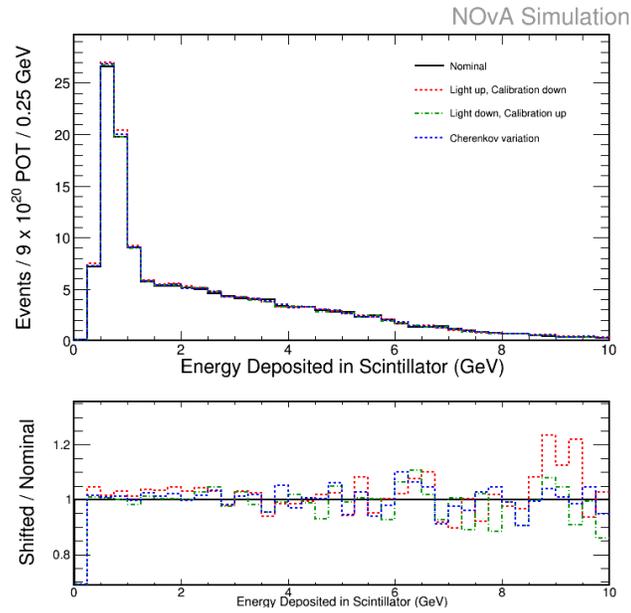


FIGURE 5.45: Cherenkov Systematic Uncertainty on ND Signal (extrapolated)

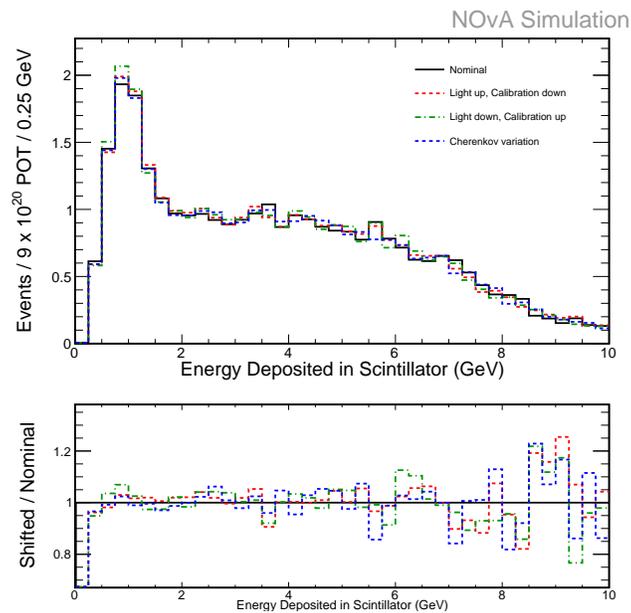


FIGURE 5.46: Cherenkov Systematic Uncertainty on FD Background (extrapolated)

5.3.5 ND MC Decomposition Systematic

There is an obvious mismatch of detected and simulated data in the ND. We proportionally decompose the ND data to get the FD simulated prediction, which will be compared with the FD data. In proportional decomposition, the ND data is decomposed into NC signal, charged-current muon neutrino background, and charged-current electron neutrino background. These data are employed to get the FD MC prediction. The lack of understanding of which interaction type contributes to the observed data MC discrepancy introduces uncertainties in the analysis.

The ND MC Decomposition systematic effects on ND spectra

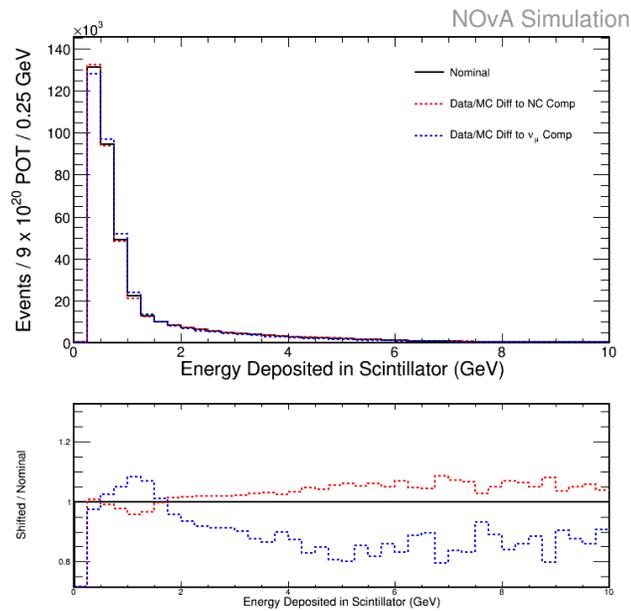


FIGURE 5.47: ND MC Decomposition Uncertainty on ND Signal

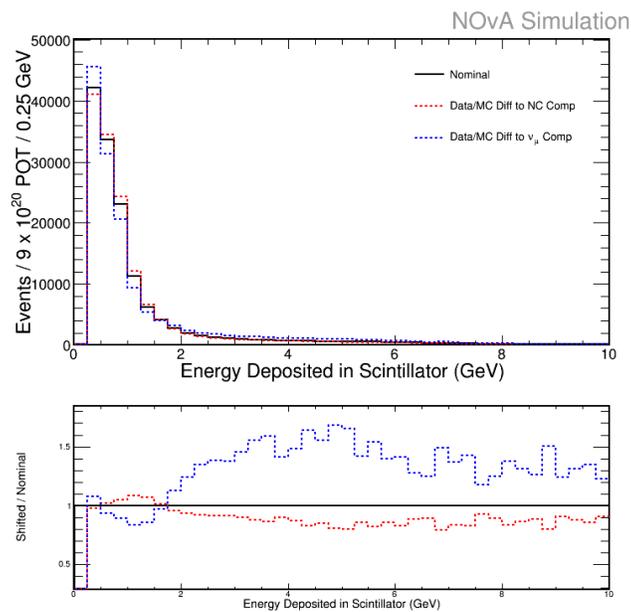


FIGURE 5.48: ND MC Decomposition Uncertainty on ND Background

The ND MC Decomposition systematic effects on FD spectra (the direct selection results)

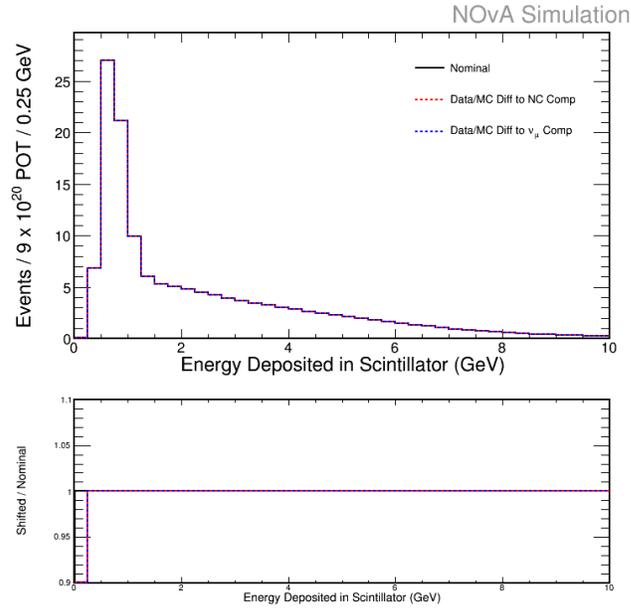


FIGURE 5.49: ND MC Decomposition Uncertainty on FD Signal (direct selection)

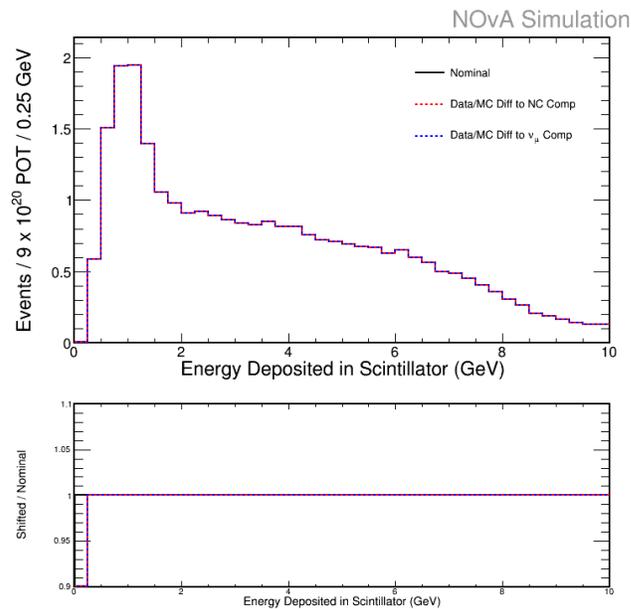


FIGURE 5.50: ND MC Decomposition Uncertainty on FD Background (direct selection)

The ND MC Decomposition systematic effects on FD spectra (the extrapolated selection results)

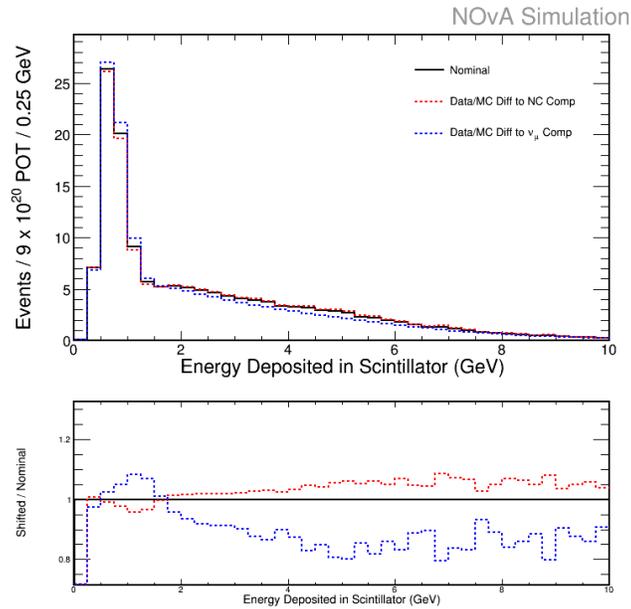


FIGURE 5.51: ND MC Decomposition Uncertainty on FD Signal (extrapolated)

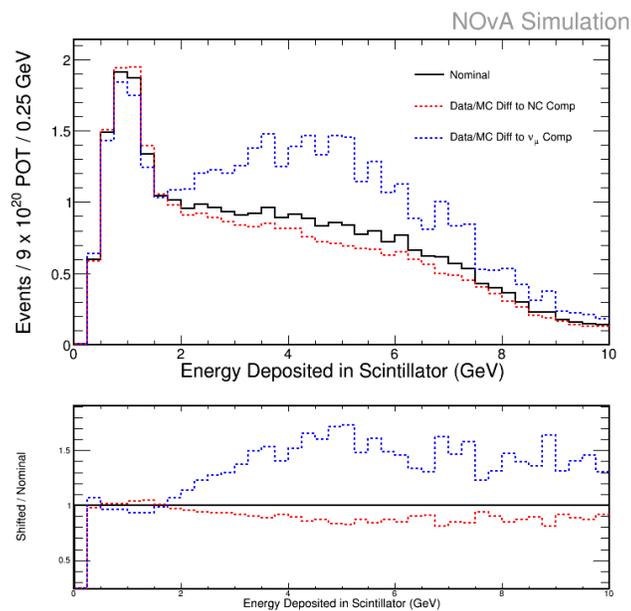


FIGURE 5.52: ND MC Decomposition Uncertainty on FD Background (extrapolated)

5.3.6 Acceptance Study

The NOvA detectors are sited 14.6 mrad off from the NuMI beam and are apart from each other about 810 km. The neutrino beam cone dramatically grows up during its travel. The acceptance of the detectors therefore plays an important role in our analysis. For this uncertainty, we divide the ND into four quadrants; cut through X and Z plane then extrapolate each of the four ND spectra to FD separately to find the predicted spectra. The systematic shift is calculated by quantifying the difference between the nominal extrapolated prediction and extrapolated prediction from each of the four ND sections.

Systematic	NC Difference (%)	Background Difference (%)
ND One	1.64	1.25
ND Two	0.85	1.22
ND Three	2.50	1.90
ND Four	0.87	0.86

The Acceptance systematic effects on ND One spectra

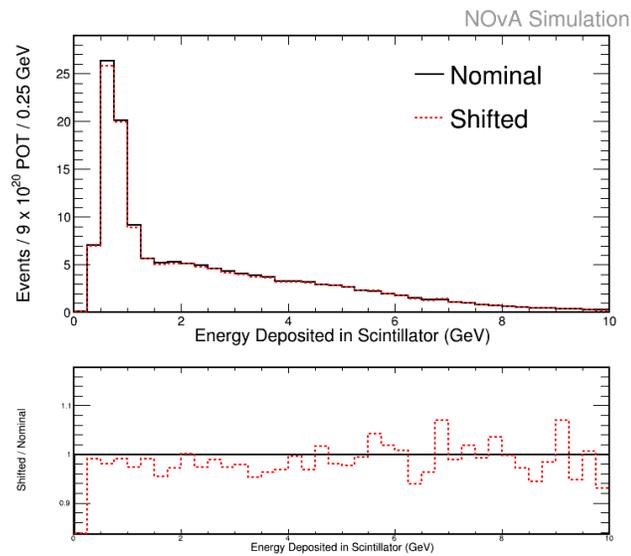


FIGURE 5.53: Acceptance Uncertainty on ND Signal (One spectra)

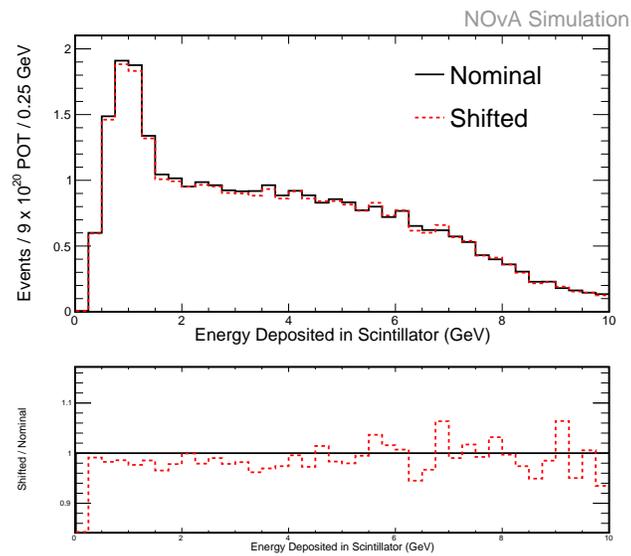


FIGURE 5.54: Acceptance Uncertainty on Background (One spectra)

The Acceptance systematic effects on ND Two spectra

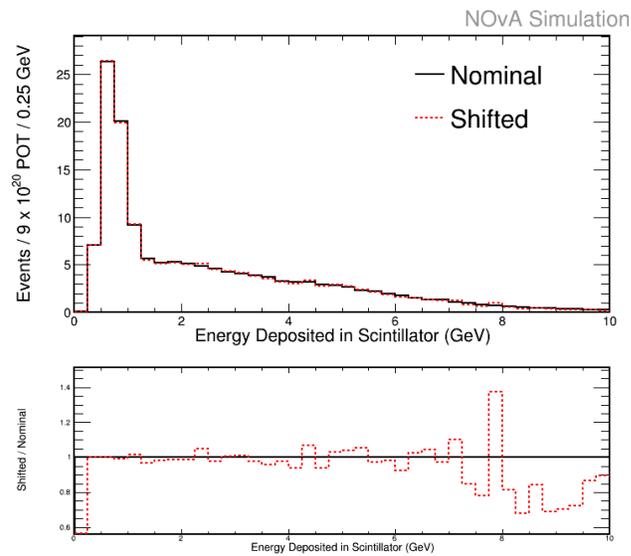


FIGURE 5.55: Acceptance Uncertainty on ND Signal (Two spectra)

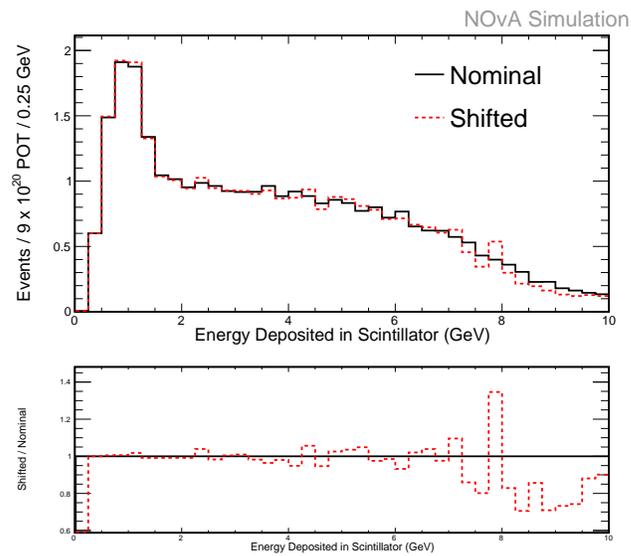


FIGURE 5.56: Acceptance Uncertainty on Background (Two spectra)

The Acceptance systematic effects on ND Three spectra

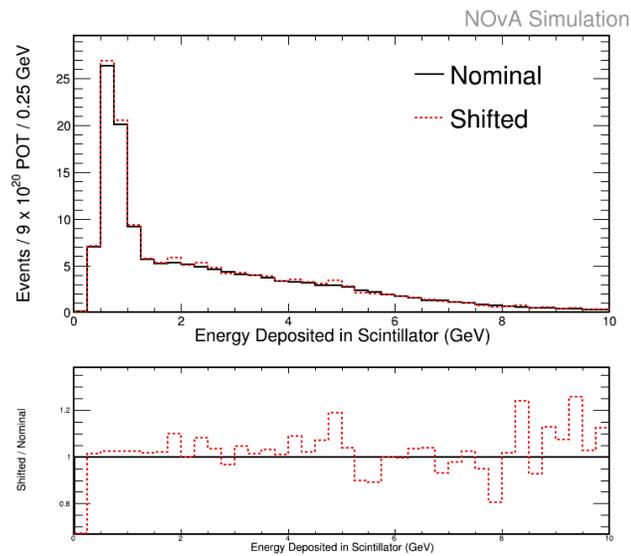


FIGURE 5.57: Acceptance Uncertainty on ND Signal (Three spectra)

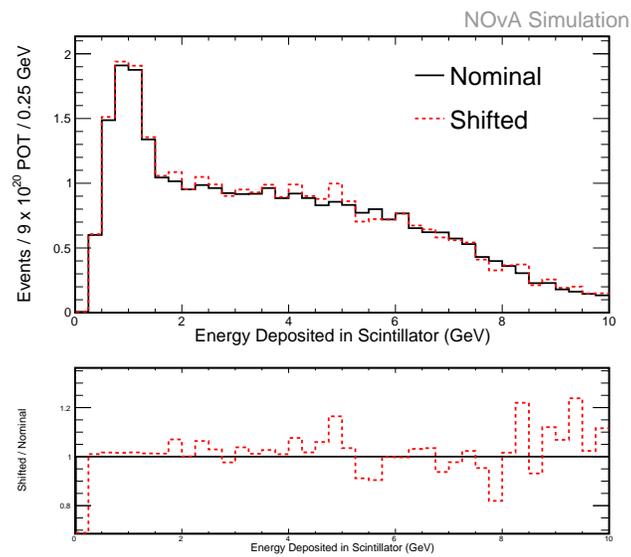


FIGURE 5.58: Acceptance Uncertainty on Background (Three spectra)

The Acceptance systematic effects on ND Four spectra

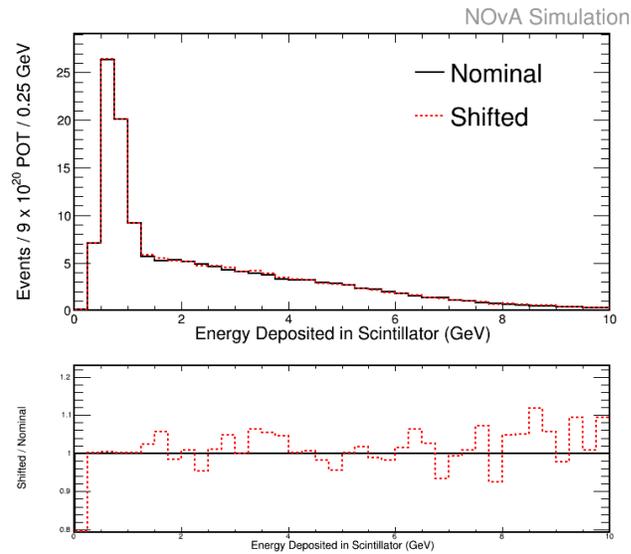


FIGURE 5.59: Acceptance Uncertainty on ND Signal (Four spectra)

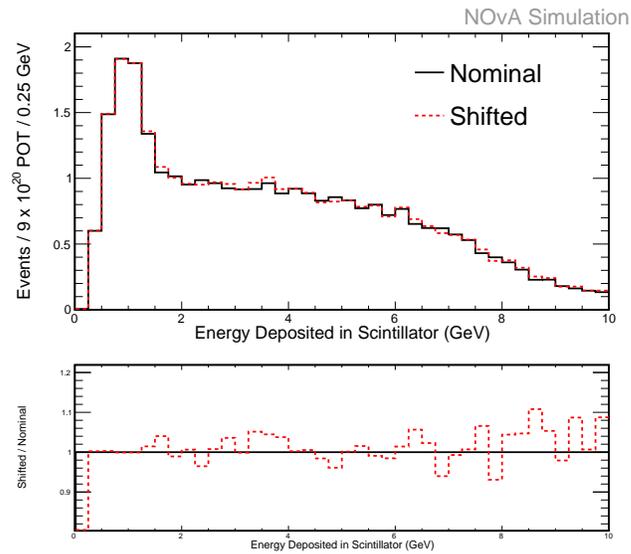


FIGURE 5.60: Acceptance Uncertainty on Background (Four spectra)

5.3.7 3-flavor oscillation parameters

The NC disappearance analysis includes an uncertainty in the 3-flavor oscillation parameters. This is included as a normalization uncertainty based on the allowed range of signal and background by varying the 3-flavor oscillation parameters within the PDG allowed uncertainties. For the background extrapolated prediction Oscillation parameter uncertainty is the dominant one and it contributes 10.7 percent. But for the NC signal, it is only 0.7 percent.

5.3.8 The normalization uncertainty

The normalization uncertainty includes contributions from the POT and mass uncertainties. There is a 0.5 percent error on the POT counting which arises from a small difference in the two toroids that determine the POT in a spill.

1. FD: ± 0.28 percent, dominated by PVC weighing.
2. ND: ± 0.28 percent, dominated by PVC lot variations.
3. Muon Catcher: ± 0.34 percent, dominated by PVC lot variations.
4. FD/ND ratio: ± 0.26 percent, strongly dominated by PVC lot variations.

5.3.9 Uncertainty Summary

The Total statistical uncertainty for NC signal is 8.2% and 15.2% for backgrounds, while the systematic uncertainty values are 9.3% for NC signal and 16.5% for backgrounds.

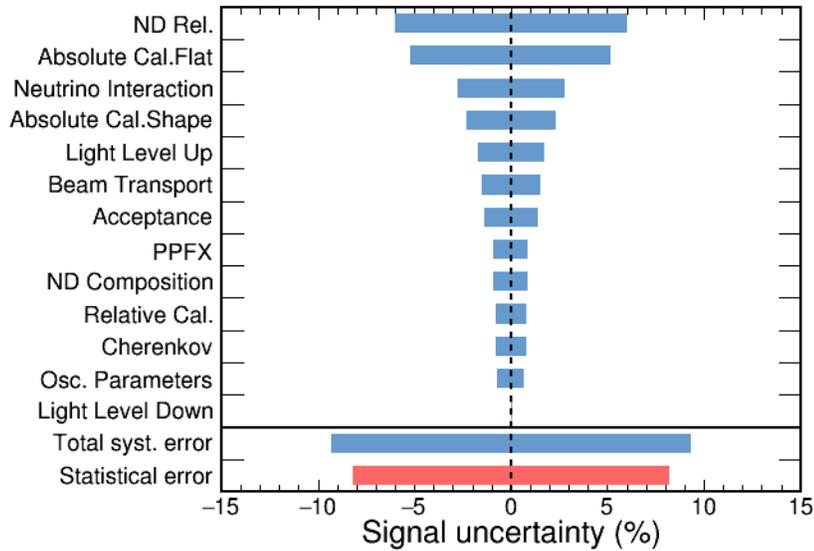


FIGURE 5.61: Signal Uncertainty Summary

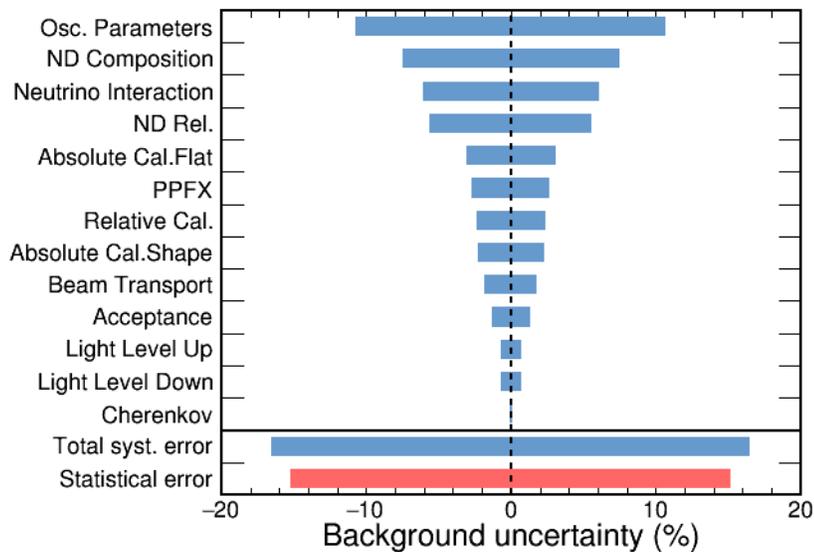


FIGURE 5.62: Background Uncertainty Summary

5.4 Fitting Procedure

The 2016 analysis was a cut-and-count analysis. Therefore, though the extrapolation and prediction were performed in 250 MeV bins of reco energy, the 2016 analysis compares the FD NC event rate with predicted rate and fitted for the integrated events. Improved modeling of the detector and the neutrino interaction cross-section enable us to employ a better fitting method, which takes into account the spectral shape information of the selected events, and will be introduced in detail in the first subsection. These improvements include: better detector modeling; more accurate threshold modeling from data; Cherenkov light modeling; and the improved cross-section modeling of the NC events. The fitting procedure was performed by allowing some mixing parameters to float and some to be fixed, while others were set to 0 as NOvA has no sensitivity to them. The treatment of mixing parameters will be presented in the second subsection. Furthermore, in the third subsection will present how the systematic uncertainties affect the fit results.

5.4.1 Shape Fitting

In order to better extract the mixing parameters from the measured data, the 2017 analyses employ a frequentist statistical approach [84]. Considered a binned prediction for a vector of oscillation parameters $\vec{\theta}$, we calculate the Poisson log-likelihood function as:

$$-2\ln\lambda(\vec{\theta}) = 2 \sum_{i=1}^{bins} [v_i(\vec{\theta}) - n_i + n_i \ln \frac{n_i}{v_i(\vec{\theta})}] \quad (5.11)$$

where v_i and n_i is the predicted and measured number of events in the each bin. The sum runs over all the bins (from 0.25 GeV to 10 GeV for our 2017 analyses).

By defining χ^2 as

$$\chi^2 \equiv -2\ln\lambda(\vec{\theta}), \quad (5.12)$$

the vector of parameters that minimizes the above equation is the best-fit result $\hat{\theta} \equiv \vec{\theta}_{best}$,

$$\chi_{best}^2 \equiv \chi^2(\vec{\theta}) = \min(\chi^2(\vec{\theta})) \quad (5.13)$$

and we then define the test statistic $\Delta\chi^2$ relative to the minimum χ^2 as

$$\Delta\chi^2(\vec{\theta}) = \chi^2(\vec{\theta}) - \chi^2(\hat{\theta}) \quad (5.14)$$

With the above definition, $\Delta\chi^2$ is non-negative and ordered so that if $\Delta\chi^2(\hat{\theta}_a) < \Delta\chi^2(\hat{\theta}_b)$, the combination of parameters $\hat{\theta}_a$ is a better fit to the data than $\hat{\theta}_b$. Strictly, it is the asymptotic distribution of $-2\ln\lambda(\vec{\theta})$ that follows a χ^2 distribution, given some regularity conditions.

5.4.2 Treatment of PMNS Mixing Parameters

As we stated in the first section, the mixing parameter setting plays an important role in our analyses. Three different sets of parameters are used by the NOvA sterile neutrino analysis in extracting results: the 2014 PDG best fits, the NOvA 2017 best fits and NOvA 2018 best fits. The 2016 analysis results are based on 2014 PDG best fits, and the 2017 analyses results are presented based on both the NOvA 2017 and NOvA 2018 best fits, as will be the results from the 2017 Re-Analysis that represent the primary results of this thesis.

Official Analysis Setting The number of degrees of freedom is reduced by holding the following three-flavor mixing parameters fixed for the 2016 and 2017 sterile analyses separately:

Oscillation Parameter	2016 Employed Value	2017 Employed Value	2018 Employed Value
ρ	$2.84g/cm^3$	$2.84g/cm^3$	$2.84g/cm^3$
Δm_{21}^2	$7.53 \times 10^{-5}eV^2$	$7.53 \times 10^{-5}eV^2$	$7.53 \times 10^{-5}eV^2$
$\sin^2 2\theta_{12}$	0.846	0.846	0.846
Δm_{32}^2	$2.37 \times 10^{-3}eV^2$	$2.67 \times 10^{-3}eV^2$	$2.44 \times 10^{-3}eV^2$
$\sin^2 \theta_{23}$	1	0.404 (0.623)	0.558
$\sin^2 2\theta_{13}$	0.085	0.085	0.085
δ	0	1.48π (0.74π)	1.21π

TABLE 5.5: The sets of three-flavor oscillation parameters used in the NOvA sterile analyses presented in this thesis.

ρ is the density of the earth estimated with the CRUST 2.0 model [85], using the average depth underground between the two detectors(9.38 km). δ is the CP violation phase, which is the δ_{13} in the 3+1 model. Small variations in any of these parameters have negligible effects in our results.

5.4.3 Treatment of Extra Mixing Parameters in 3+1 Model

As introduced before, one of the major analysis goal was to measure two 3+1 model mixing angles: θ_{24} and θ_{34} . They are allowed to float between 0° to 45° . Values outside of this range are either equivalent to this range through redefinition of the angles, or already disfavored by previous experiments and in a region difficult for fitting due to degenerate local minima. We also set $\Delta m_{41}^2 = 0.5eV^2$, and $\rho_{14} = 0$. ρ_{24} has been profiled for the 2017 analyses. We also study the ρ_{24} dependence for the 2017 Re-Analysis. Finally, in order to present the results, we construct one and two dimensional graphs based on the $\Delta\chi^2$ distributions,

where the remaining parameters are treated as nuisance parameters, being profiled over.

5.4.4 Treatment of Systematic Uncertainties

Systematic uncertainties are incorporated in the likelihood function as nuisance parameters, each one of them is associated with an additional free term in the fit, which requires an additional measurement or external constraint. As mentioned earlier, the near-to-far extrapolation procedure aids in this respect: predictions are created from the systematically-shifted MC samples and the ND data, effectively correcting the central values. Any leftover discrepancies between the nominal and the shifted predictions determine the allowed ranges for the systematic pulls. Penalty terms are added to the χ^2 to keep their pulls in the 1σ ranges.

$$\chi^2(\vec{\theta}) = \min(\chi^2(\vec{\theta}, \vec{s})) + \sum_{i=1}^{\text{uncertainties}} \frac{s_i^2}{\sigma_i^2} \quad (5.15)$$

s_i are the values of each systematic uncertainty. σ_i are the 1-sigma ranges employed to assign the penalty term for each systematic shift. This method assumes that all sources of uncertainty are independent of each other. Alternative approaches could use a covariance matrix method, or a simultaneous two-detector fit.

5.5 Analysis Results

The main impact of NC Disappearance analyses is on the limits on the largely unconstrained θ_{34} and θ_{24} . In this section, the Re-Analysis results obtained

through the work described in this thesis are shown along with the 2017 official results from the NOvA NC/Sterile group. In the first subsection, we summarize the 2017 Official results that used the cosmic rejection method described in this thesis, and compare them with some of the 2016 analysis results which have been published in PRD [24]. In the second subsection, we will present the 2017 Re-Analysis results and compare them with the 2017 Official ones.

5.5.1 Official 2017 Results

The 2016 analysis was a counting experiment and so fitted only rate information from selected NC-like data. In the 2017 analysis, additional shape information of the FD spectrum is included into a rate + shape fit.

2017 Official Selected NC-like Data Distributions The 2017 Official analysis identified 214 NC neutrino interaction candidates in a 8.85×10^{20} POT-equivalent FD data exposure, for an expected three-flavor prediction of $191.2 \pm 13.8(\text{stat.}) \pm 22.0(\text{syst.})$. Therefore, no evidence for a deficit of NC interactions is seen in the FD data. Figures 5.63 and 5.64 show the vertex position of the neutrino interactions selected as NC-like in the NOvA FD, along with the fiducial cuts applied.

The reconstructed energy spectrum of selected NC events in the FD is shown in Fig. 5.65. The three-flavor NC prediction is computed assuming the NOvA three-flavor best fits published in [86].

2017 Official R-Ratio The agreement of the rate-only results with the null-sterile mixing hypothesis can be quantified by using the R-Ratio, a model-independent measure of NC disappearance, which is defined by equation 5.16:

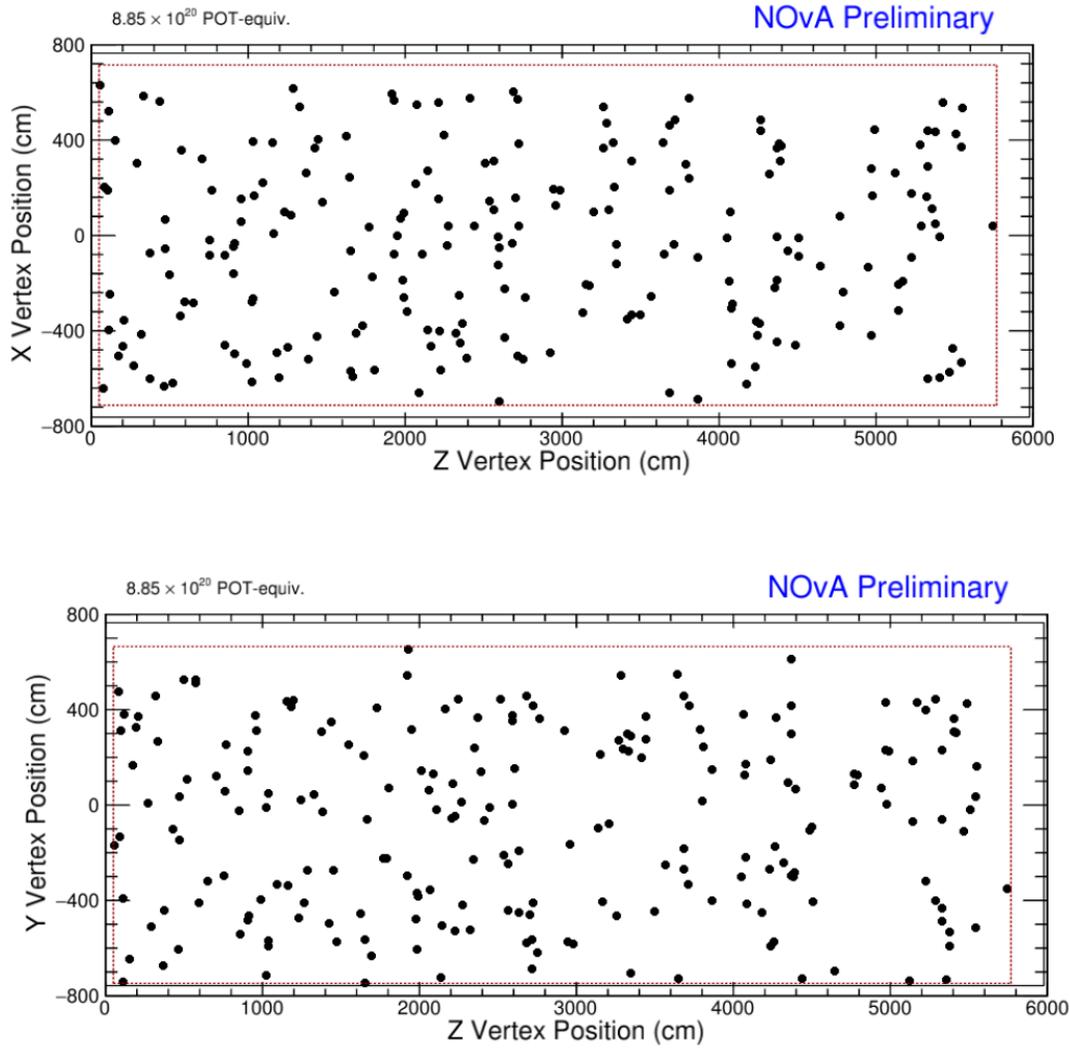


FIGURE 5.63: The upper one is the X-Z view of the FD and the lower one is the Y-Z view of the FD. The red dashed line shows the corresponding Fiducial Cuts.

$$R_{NC} \equiv \frac{F^{data} - \sum F^{pred}(background)}{F^{pred}(NC)} \quad (5.16)$$

Table 5.6 shows the R-ratio computed for three regions in reconstructed neutrino energy. The 0-2.5 GeV region is where an expected NC deficit would be likely observed if there was mixing with sterile neutrinos, the 2.5-10 GeV region is expected to be consistent with the three-flavor hypothesis, and the 0-10 GeV

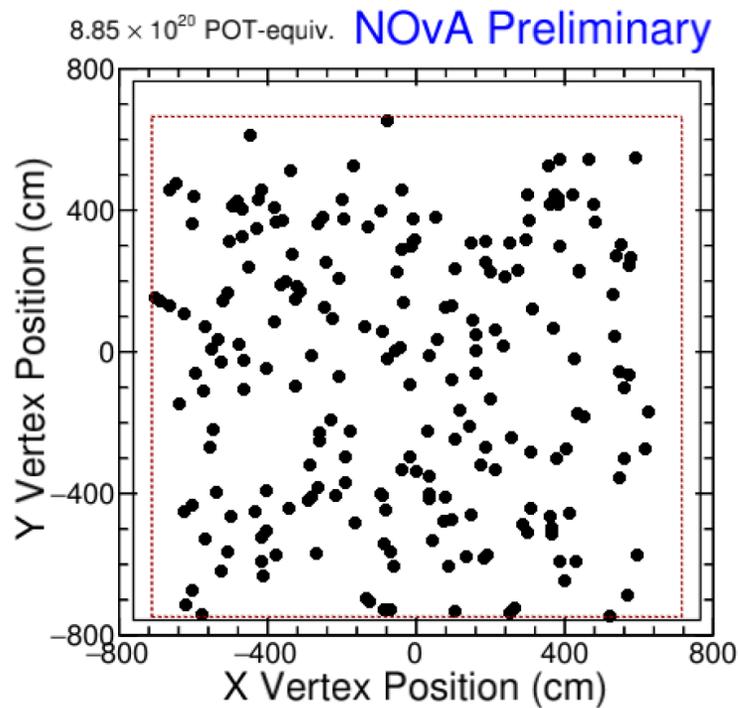


FIGURE 5.64: The X-Y view of the FD, the red dashed line shows the corresponding Fiducial Cuts.

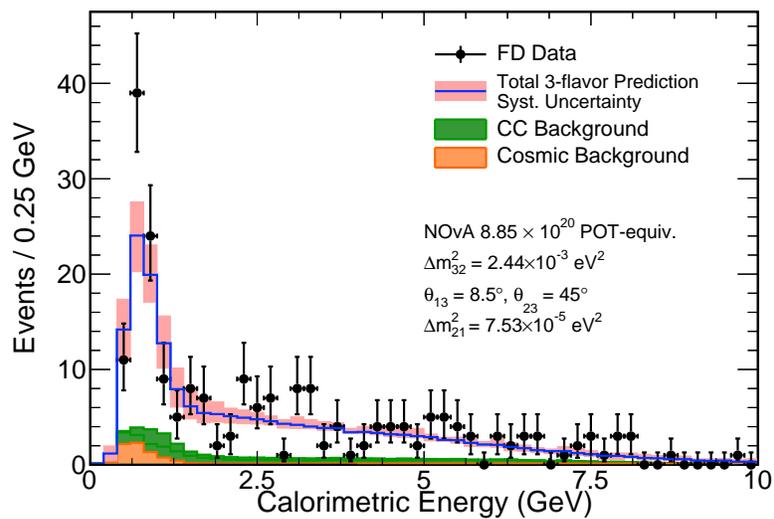


FIGURE 5.65: Reconstructed neutrino energy in the FD for NC-like data events, shown along the predicted spectra for NC signal and backgrounds.

region includes the full range of reconstructed neutrino energies analyzed.

2017 Official	0 - 2.5 GeV	2.5 - 10 GeV	0 - 10 GeV
R-Ratio	$1.19 \pm 0.14 \pm 0.12$	$1.08 \pm 0.14 \pm 0.12$	$1.15 \pm 0.14 \pm 0.12$
Selected Events	119	95	214

TABLE 5.6: The R-Ratio for the 2017 Official Analysis in the three reconstructed neutrino energy ranges considered. For reference, the value found for the 0-19 GeV range for the 2016 analysis was $1.19 \pm 0.16 \pm 0.10$.

In all energy ranges, the R-Ratio is larger than 1, which means that the 2017 Official results show no evidence of NC disappearance, and therefore no evidence for active to sterile neutrino oscillations happening in NOvA's long baseline.

2017 Official 2D Contour on θ_{34} and θ_{24}

By doing a shape+rate fit to the data using the 3+1 model, one can extract values and confidence intervals for the sterile mixing angles θ_{24} and θ_{34} . The 68% and 90% CL. 2D allowed regions obtained for those mixing angles for a fixed value of $\Delta m_{41}^2 = 0.5 \text{ eV}^2$ are shown in Fig. 5.66. The regions are computed fixing the atmospheric mixing parameters to the NOvA three-flavor best fit values published in [86].

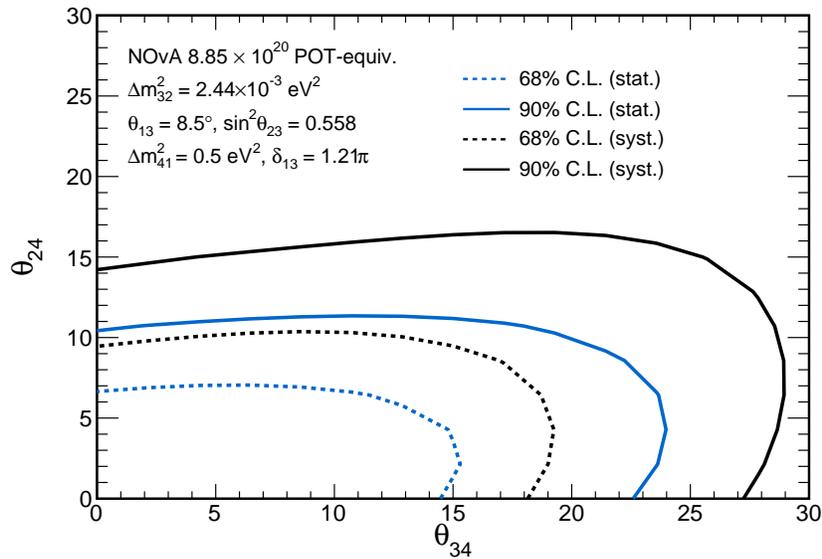


FIGURE 5.66: The NOvA 2017 Official 68% and 90% CL. 2D allowed regions for θ_{24} and θ_{34} . All values given in degrees.

Global Picture on θ_{34} and θ_{24}

We can compare the NOvA results with current results from other experiments by extracting 1D limits on θ_{24} and θ_{34} . The same 1D limits can be expressed in

terms of the relevant matrix elements

$$|U_{\mu 4}|^2 = \cos^2 \theta_{14} \sin^2 \theta_{24} \quad (5.17)$$

$$|U_{\tau 4}|^2 = \cos^2 \theta_{14} \cos^2 \theta_{24} \sin^2 \theta_{34}, \quad (5.18)$$

by conservatively assuming $\cos^2 \theta_{14} = 1$ in both cases, as we know θ_{14} is small from reactor measurements, and $\cos^2 \theta_{24} = 1$ in the $|U_{\tau 4}|^2$ case.

The rate+shape 2017 Official analysis improves over the rate-only 2016 limits for θ_{24} by 4.6° , and for θ_{34} by 1.4° . This result is competitive with results from other experiments with just 1/4 of the total NOvA planned exposure.

	θ_{24}	θ_{34}	$ U_{\mu 4} ^2$	$ U_{\tau 4} ^2$
NOvA 2017	16.2°	29.8°	0.078	0.228
NOvA 2016	20.8°	31.2°	0.126	0.268
MINOS	7.3°	26.6°	0.016	0.20
SuperK	11.7°	25.1°	0.041	0.18
IceCube	4.1°	-	0.005	-
IceCube-DeepCore	19.4°	22.8°	0.11	0.15

TABLE 5.7: 2017 Official NOvA 90% CL. 1D limits on θ_{24} and θ_{34} , compared to the NOvA 2016 results, as well as with results from searches for sterile neutrinos from other experiments.

5.5.2 2017 Re-Analysis vs 2017 Official Analysis

As presented in the motivation chapter, Chapter 1, we tried to improve on the 2017 Official analysis results based on improved machine learning algorithms. The improvements made in the 2017 Re-Analysis are introduced in Chapter 4 - Neutrino Interaction Classification and Signal Selection. The results based on the Re-Analysis selection cuts are presented below, along with comparisons to the 2017 Official analysis.

2017 NC Selection Comparison

We have developed a new selection result based on the Re-Analysis cuts. The Re-Analysis selection provides a better result based on the Figure of Merit (FoM), $Signal / \sqrt{Signal + Background}$, as we will show below. From looking at Fig. 5.69, one can see the Re-Analysis selection gets more NC signal, less CC background, but more cosmic background. In the following parts of this section, we compare the results for 2017 Official and Re-Analysis using the NOvA three-flavor best fit values published in [86] to fix the atmospheric oscillation parameters.

Figure 5.67 shows a simulation-based comparison of the selection efficiency and purity for the 2017 Official and Re-Analysis case.

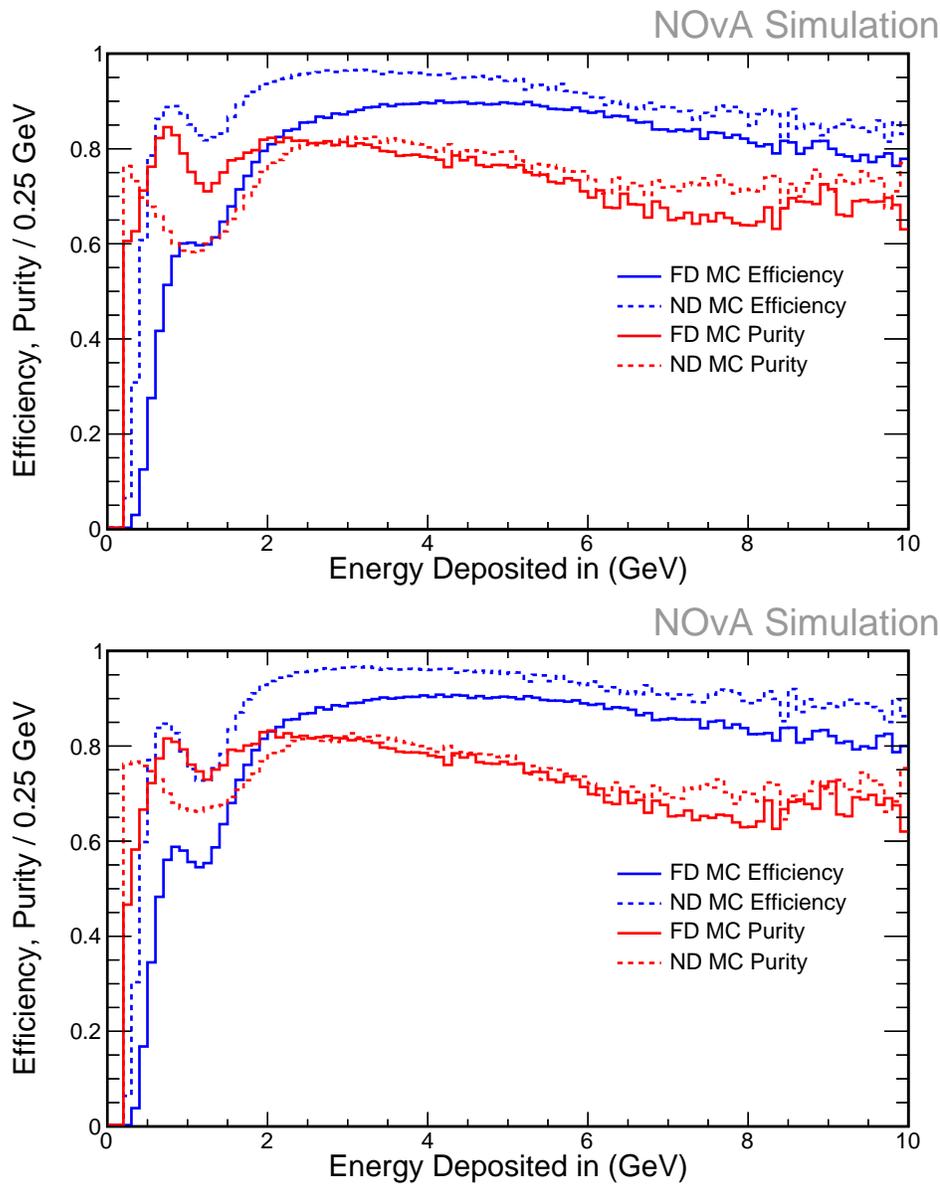


FIGURE 5.67: Comparison of 2017 Official (top) and Re-Analysis (Bottom) Selection Efficiency and Purity.

Figure 5.68 shows a simulation-based comparison of the ND reconstructed energy spectrum for NC signal and backgrounds for the 2017 Official and Re-Analysis case.

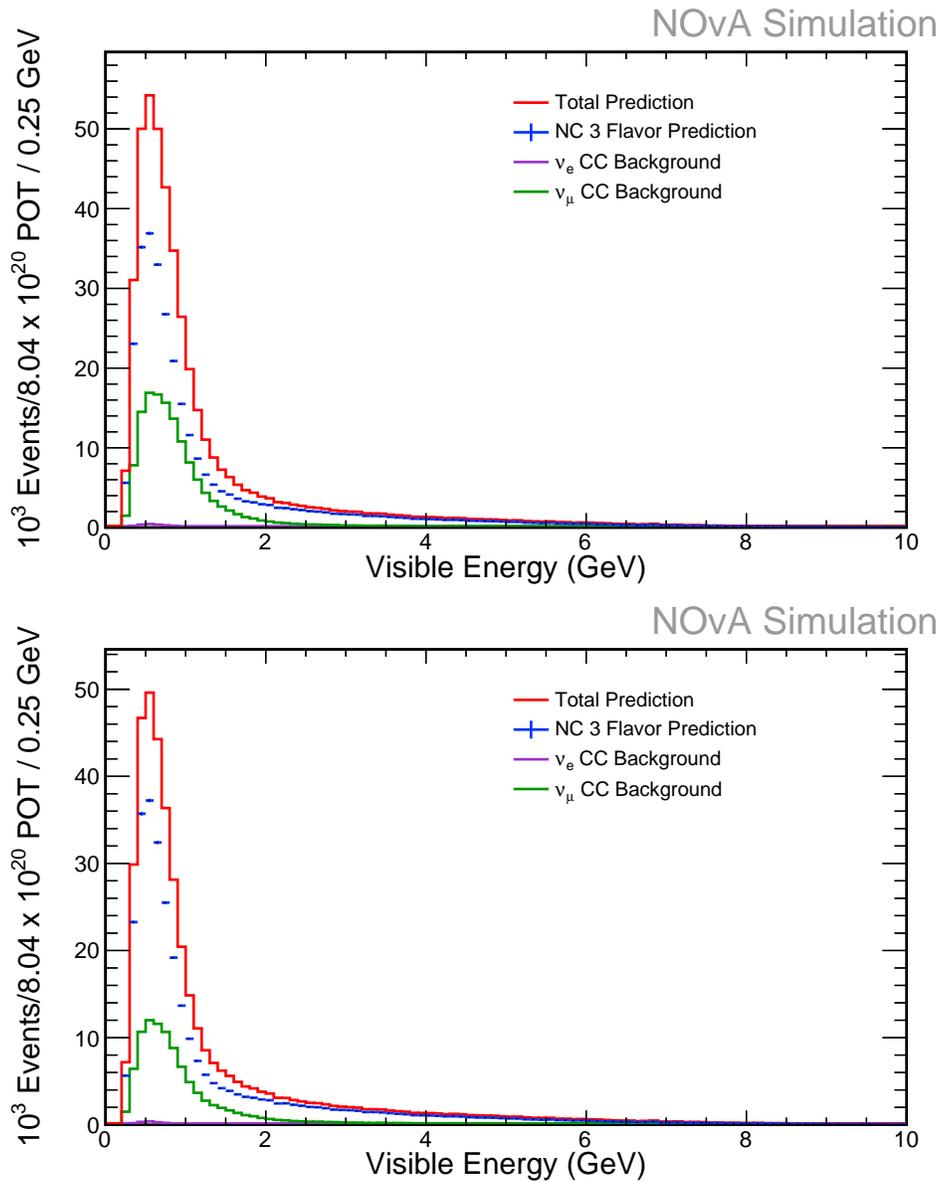


FIGURE 5.68: Comparison of 2017 Official (top) and Re-Analysis (Bottom) ND reconstructed energy spectra.

Figure 5.69 shows a simulation-based comparison of the FD reconstructed energy spectrum for NC signal and backgrounds for the 2017 Official and Re-Analysis case.

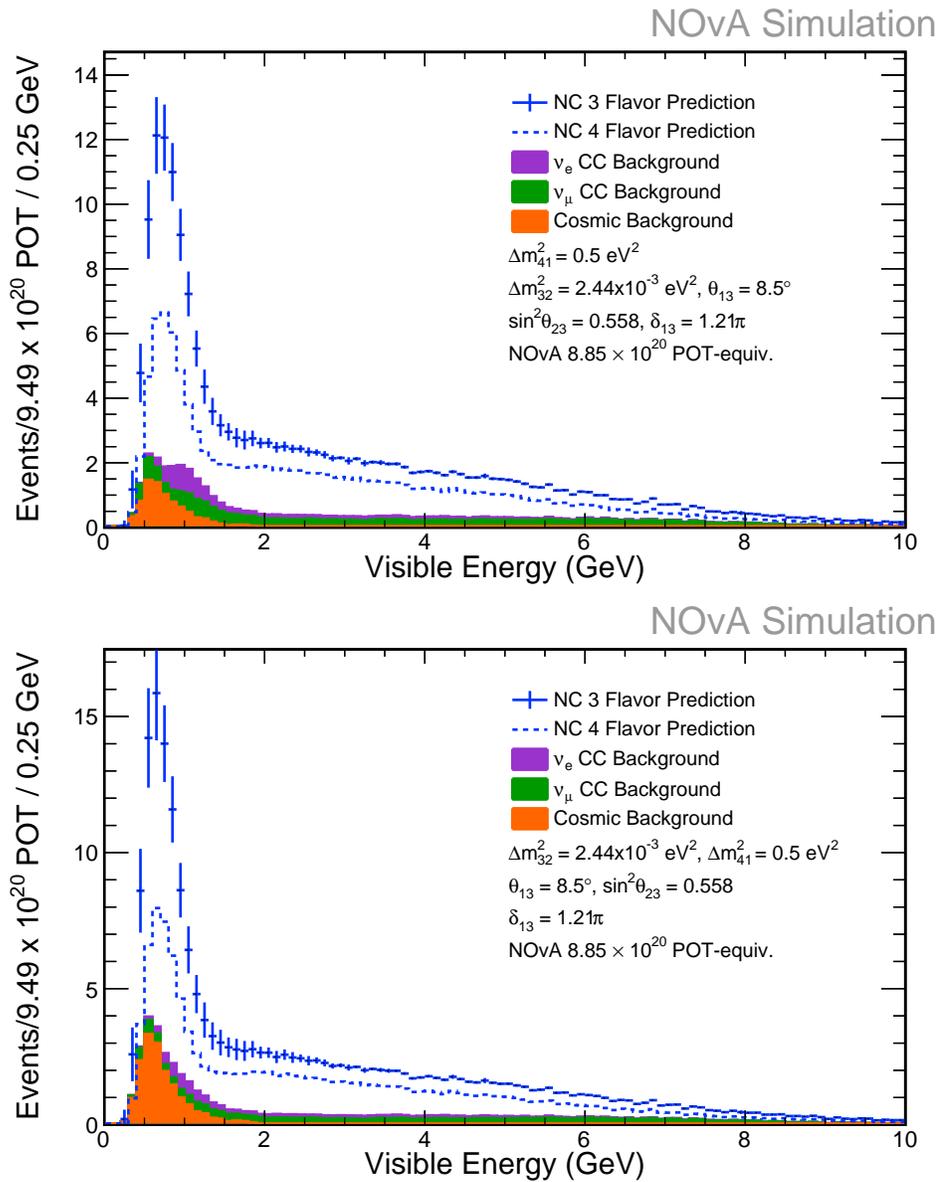


FIGURE 5.69: Comparison of 2017 Official (top) and Re-Analysis (Bottom) FD reconstructed energy spectra.

Finally, Figure 5.70 shows a stats-only sensitivity comparison of 2D allowed regions for the 2017 Official and Re-Analysis case.

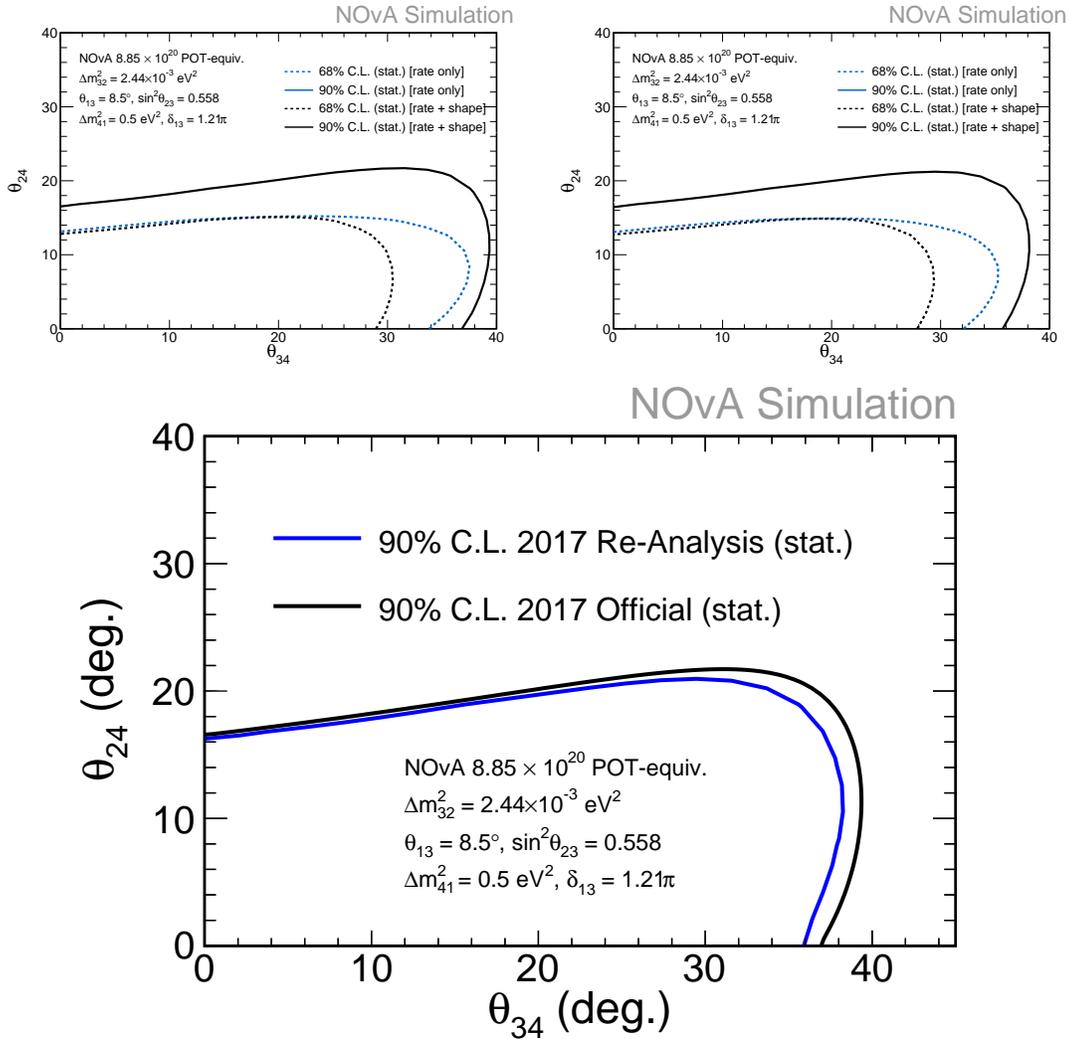


FIGURE 5.70: Comparison of 2017 Official (Top Left) and Re-Analysis (Top Right) stats-only sensitivities, showing the 2D 68% CL. and 90% CL. allowed regions. Comparison for contours using rate-only or rate+shape are also shown to illustrate improvements from including shape information. The Bottom plot displays a direct comparison of the 90% CL. sensitivities. It shows We can expect a better limit on θ_{34} from the 2017 Re-Analysis and a similar limit on θ_{24} .

Re-Analysis Selection Sideband Study

To avoid biases, we performed a blind analysis, whereby the FD data was only analyzed after all analysis procedures were studied using ND data, and frozen. As a final check, a sideband study was performed before we open the box for the 2017 Re-Analysis. We replace the NC/CC classifier cut with its inverse setting. The same study has been done in last chapter based on the FD direct selection. Here, we present the result based on the extrapolated selection method. The selected results are shown in Fig. 5.71.

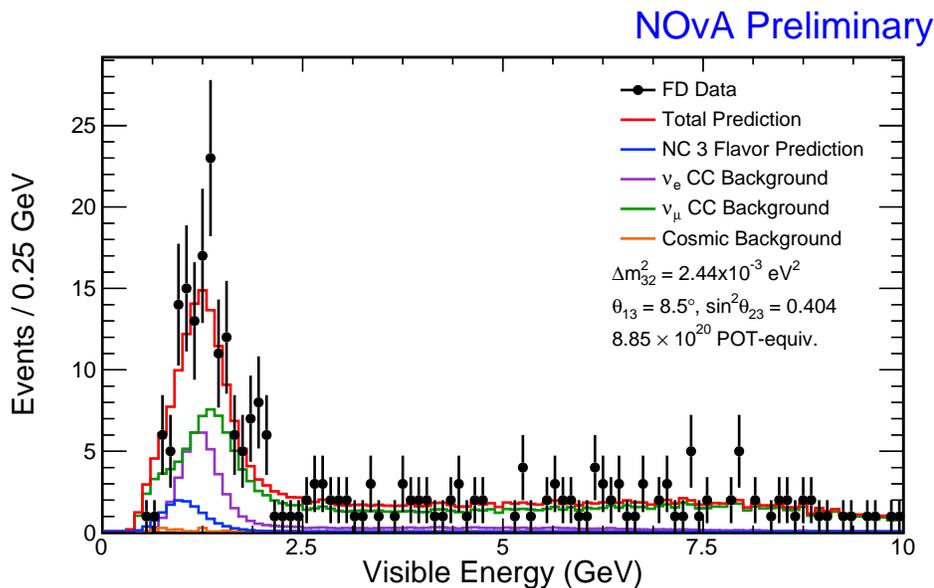


FIGURE 5.71: 2017 Re-Analysis sideband energy spectra. As expected, after applying the inverse NC/CC classifier cut, we observe a dramatically increased number of selected CC events.

2017 Re-Analysis Selection Results

Using the 2017 Re-Analysis Selection Cuts, we select 237 NC-like events in the same 8.85×10^{20} POT-equivalent FD data exposure as the 2017 Official analysis. The predicted number of events in the absence of light sterile neutrino mixing is $206.1 \pm 14.3(\text{stat.}) \pm 25.8(\text{syst.})$. Again, if active neutrinos mixed with sterile neutrinos, the number of measured NC-like events in the data would have been depleted with respect to the prediction. A comparison between NC-like selected events in the 2017 Official and 2017 Re-Analysis is shown in Table 5.8.

	Total Prediction	NC Signal	CC Background			Cosmics	FD Data
			ν_μ	ν_e	ν_τ		
2017 Re-Analysis	206.1	157.9	20.3	9.6	2.8	15.5	237
2017 Official	192.3	148.3	22.5	10.8	2.8	7.89	214

TABLE 5.8: Comparison of NC-selected events for the 2017 Re-Analysis and 2017 Official cases.

2017 Re-Analysis Selected NC-like Data Distributions

Figures 5.72 and 5.73 show the vertex position of the neutrino interactions selected by the Re-Analysis as NC-like in the NOvA FD, along with the fiducial cuts applied.

The reconstructed energy spectrum of Re-Analysis selected NC events in the FD is shown in Fig. 5.74 along with a comparison with the 2017 Official spectra. The three-flavor NC prediction is computed assuming the NOvA three-flavor best fits published in [86].

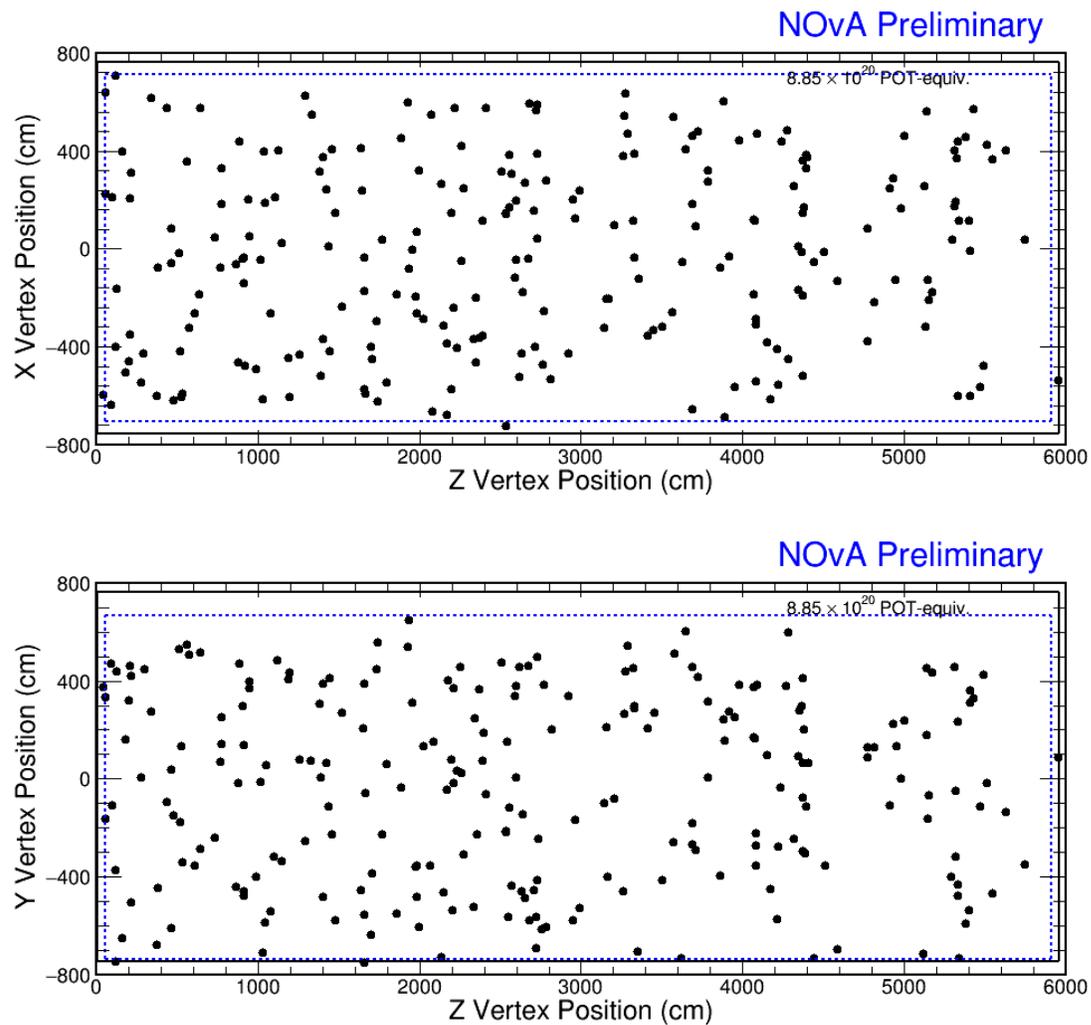


FIGURE 5.72: 2017 Re-Analysis FD data interaction vertex distributions. The upper one is the X-Z view of the FD and the lower one is the Y-Z view. The blue dashed line shows the corresponding fiducial cuts.

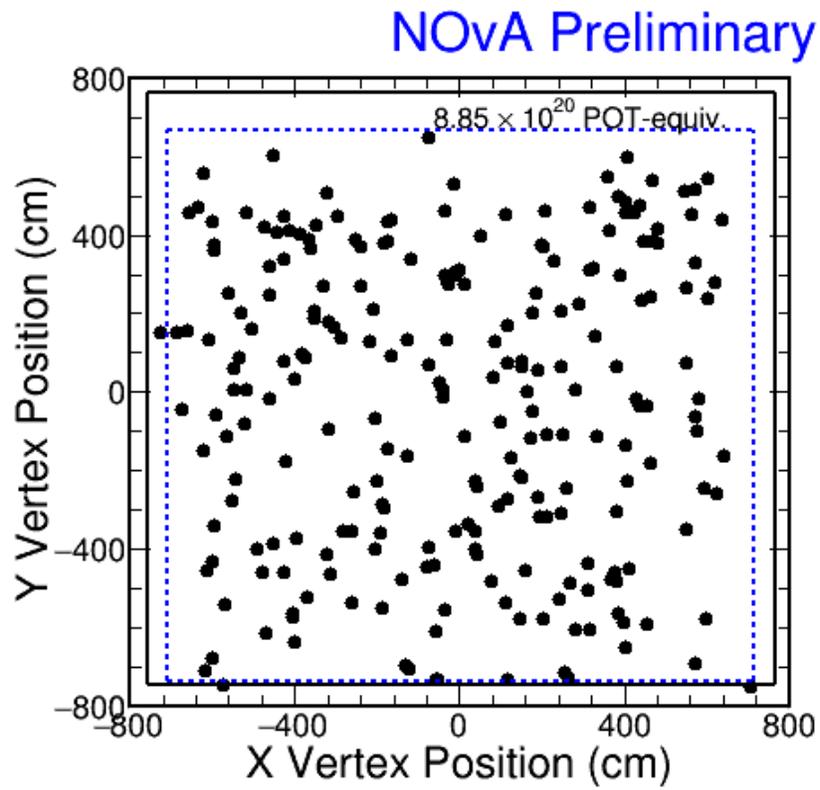


FIGURE 5.73: 2017 Re-Analysis FD data interaction vertex XY distribution. The blue dashed line shows the corresponding fiducial cuts.

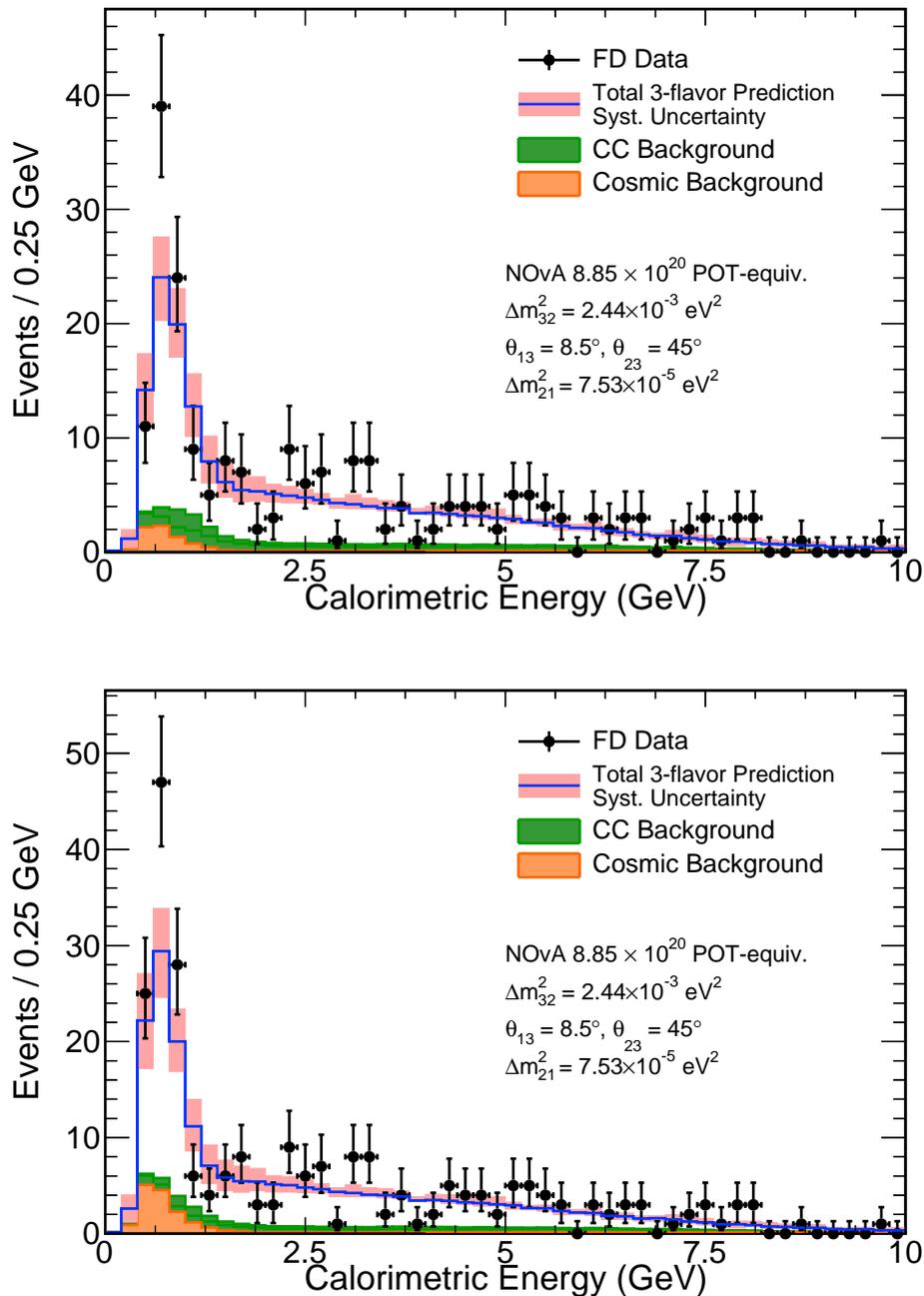


FIGURE 5.74: 2017 Official (top) and 2017 Re-Analysis (bottom) reconstructed neutrino energy spectra in the FD for NC-like data events, shown along the predicted spectra for NC signal and backgrounds.

2017 Re-Analysis R-Ratio

As shown above, we can quantify the agreement of the results with the three-flavor hypothesis through the R-Ratio, as summarized in Table 5.9. The 2017 Official R-Ratio values are repeated in the table, for ease of comparison.

Analysis	0 - 2.5 GeV	2.5 - 10 GeV	0 - 10 GeV
2017 Re-Analysis R-Ratio	$1.182 \pm 0.135 \pm 0.116$	$1.213 \pm 0.137 \pm 0.147$	$1.195 \pm 0.13 \pm 0.12$
2017 Re-Analysis Selected Events	141	96	237
2017 Official R-Ratio	$1.19 \pm 0.14 \pm 0.12$	$1.08 \pm 0.14 \pm 0.12$	$1.15 \pm 0.14 \pm 0.12$
2017 Official Selected Events	119	95	214

TABLE 5.9: 2017 Re-Analyses R-Ratio, compared with the 2017 Official equivalent result.

The R-ratio values obtained are all consistent with no mixing with light sterile neutrinos. Therefore, in the following, we will extract limits on the sterile mixing parameters.

2017 Re-Analysis 2D Contour on θ_{34} and θ_{24}

Here, we show the limits on sterile mixing parameters extracted using the 2017 Re-Analysis, and present comparisons between the 2D contours for this analysis and the 2017 Official Analysis. The 68% CL. and 90% CL. 2D allowed regions obtained by using the Re-Analysis and fitting a 3+1 model to the FD Data, using the same fitting methodology used in the 2017 Official analysis, are shown in Fig. 5.75. In all cases, we fixed the value of $\Delta m_{41}^2 = 0.5 \text{ eV}^2$.

The comparison between the 2017 Re-Analysis and the 2017 Official analysis is shown in Fig. 5.76.

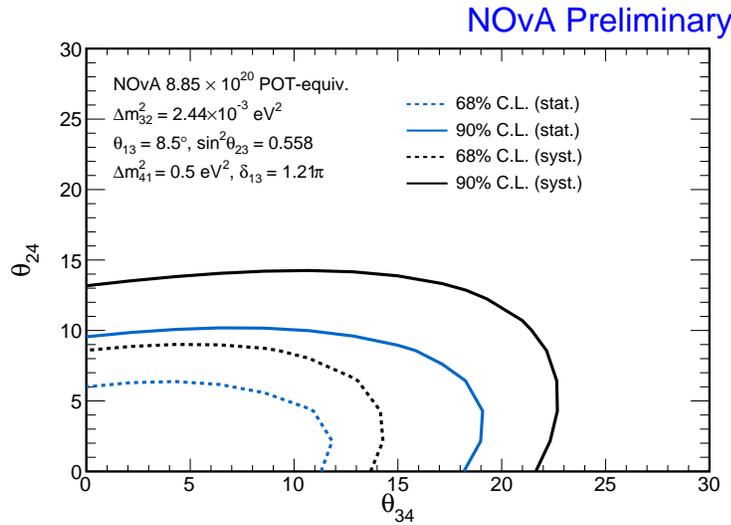


FIGURE 5.75: The NOvA 2017 Re-Analysis 68% and 90% CL. 2D allowed regions for θ_{24} and θ_{34} , showing results from a rate+shape fit of the 3+1 model to data not including the effect of systematic uncertainties (dashed lines), and including systematics (solid lines). All values given in degrees.

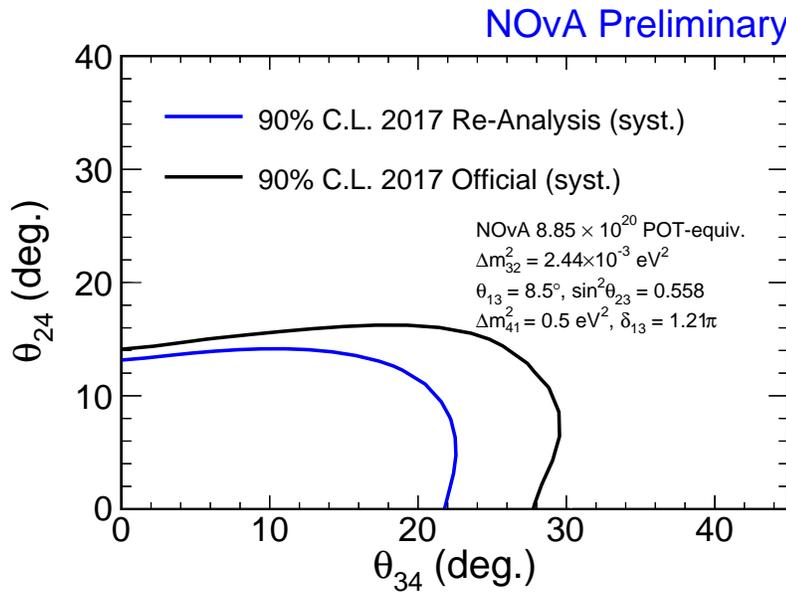


FIGURE 5.76: Comparison of 2D allowed regions in the sterile mixing angles for the 2017 Re-Analysis and the 2017 Official analysis. The atmospheric mixing parameters are fixed to the values shown in the plot for both cases.

Finally, we show a comparison of the 2017 Re-Analysis results with the 2017 Official and the published 2016 analysis in Fig. 5.77. Note that while the 2017 Re-Analysis and Official regions are produced at the same values of the atmospheric mixing parameters as the 2018 NOvA best-fit values [86], the 2016 results used the 2014 PDG values, which leads to very small differences with using the NOvA best-fit values.

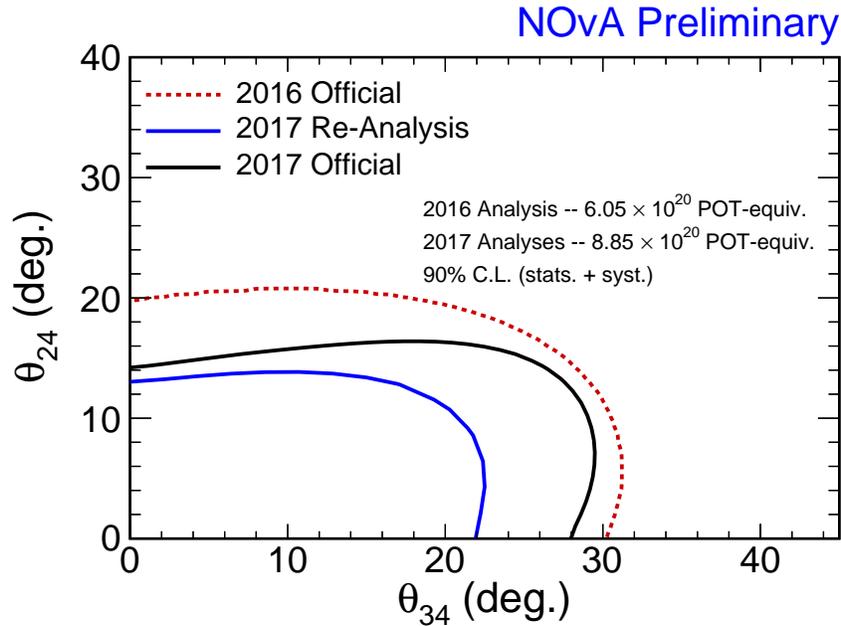


FIGURE 5.77: Comparison of 2D allowed regions in the sterile mixing angles for the 2017 Re-Analysis, the 2017 Official analysis, and the 2016 published analysis.

Global Picture on θ_{34} and θ_{24}

To place the limits obtained with the 2017 Re-Analysis within the global perspective of searches for sterile neutrinos, we show 1D projections of the 2D contours for θ_{24} and θ_{34} , while profiling over the sterile phase δ_{24} . From the plots shown in Fig. 5.78, we can extract the 90% CL. limits on the sterile mixing angles and associated mixing matrix elements, which are shown in Table 5.10. It should

be noted that for consistency with the values reported for the 2017 Official results, the 1D plots were computed from fits assuming the 2017 best-fits for the atmospheric mixing parameters, with the values shown in Fig. 5.78.

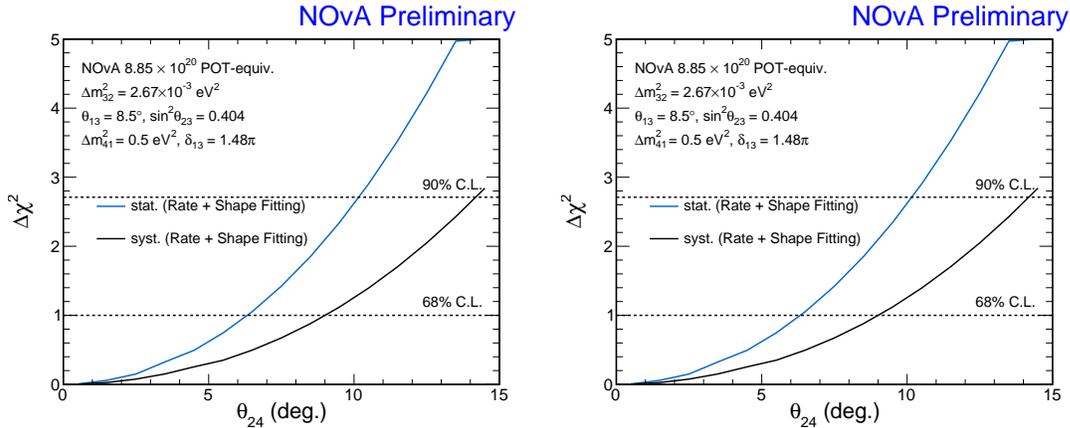


FIGURE 5.78: The 1D χ^2 profiles for the two sterile mixing angles.

	θ_{24}	θ_{34}	$ U_{\mu 4} ^2$	$ U_{\tau 4} ^2$
NOvA 2017 Re-Analysis	14.1°	23.1°	0.059	0.144
NOvA 2017 Official	16.2°	29.8°	0.078	0.228
NOvA 2016	20.8°	31.2°	0.126	0.268
MINOS/MINOS+	4.4°	23.6°	0.006	0.16
SuperK	11.7°	25.1°	0.041	0.18
IceCube	4.1°	-	0.005	-
IceCube-DeepCore	19.4°	22.8°	0.11	0.15

TABLE 5.10: Comparison of 1D limits on sterile mixing parameters from the 2017 Re-Analysis with the other NOvA results and results from sterile neutrino searches conducted with other experiments.

As can be seen in Table 5.10, the 2017 Re-Analysis improves over the 2017 Official limits for θ_{24} by 2.1° , and for θ_{34} by 6.7° . These results are competitive with present results from other experiments using just 1/4 of NOvA's final planned exposure.

2017 Re-Analysis Results Conclusion The results from our search for light sterile neutrino mixing through NC disappearance show no evidence of active to sterile oscillations. The results are consistent with the predicted NC rate and spectral shape if only mixing between three active neutrinos is considered. Furthermore, the results are consistent with the official results from NOvA searches for sterile neutrinos, and with null results from other long-baseline neutrino experiments, like MINOS and T2K, as well as atmospheric experiments like SuperK and IceCube. Therefore, these results add to the existing tension between null results from disappearance searches and the electron appearance signals reported by LSND and MiniBooNE.

Chapter 6

Conclusion and Future

Improvements

Starting in 2015, The NOvA NC Disappearance analysis has been searching for active to light sterile neutrino oscillations. Our series of analysis results is consistent with other disappearance experiments, such as MINOS/MINOS+, in finding no extra neutrino flavor(s). The sterile mixing hypothesis arose from appearance oscillation experiments, such as LSND, and appearance is quadratically suppressed with respect to disappearance effects, so if LSND is explained by light sterile neutrino mixing, we should see large effects in long-baseline experiments. Our results contribute to the present Appearance-Disappearance tension, and place upper limits on the allowed values of two sterile mixing angles in the context of a 3+1 model. These limits are competitive with other experiments using only about one quarter of the NOvA total planned exposure.

Furthermore, by applying state-of-the-art machine learning algorithms, we obtained improvements based on the same data sets, as shown by comparing the 2017 Re-Analysis with the 2017 Official analysis. There are also several known avenues for producing further improvements being explored, which will extend

these results to a large range of values of Δm_{41}^2 , and potentially reduce the effects of systematic uncertainties on the NOvA sterile searches.

One inspiring idea, proposed by Prof. Alex Sousa in 2016, in employing advanced ML algorithms, is to apply a Generative Adversarial Network (GAN), which is a type of Deep Neural Network (DNN), to produce neutrino-nuclear simulation data for the NOvA detectors by using NOvA Test Beam Detector data as the input information. The NOvA Test Beam Detector has just started collecting first beam line data, so developing this idea will be possible in the next year or so.

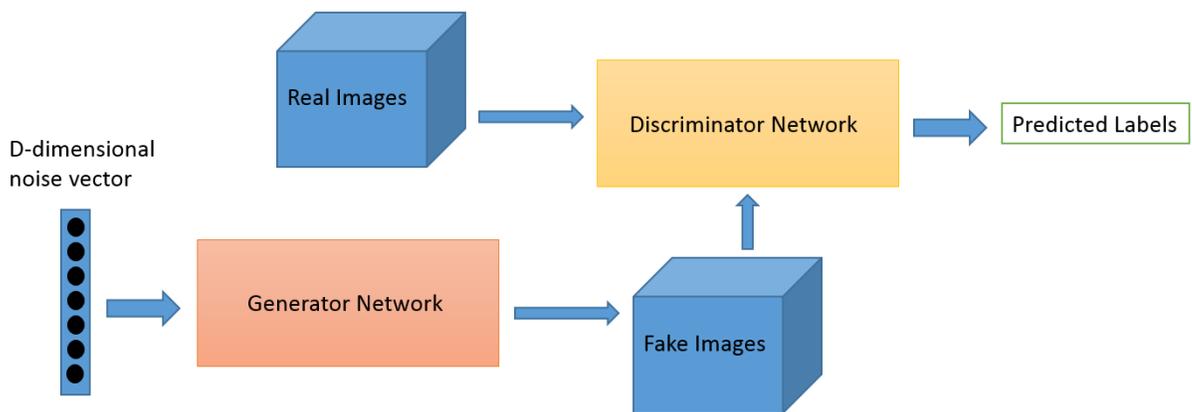


FIGURE 6.1: NOvA measured data can be transformed into images, essentially event displays, in NOvA. Therefore, the connection between NOvA data and GAN input images can be built. the NOvA Test Beam detector [87, 88] will collect millions of tagged charged particles with known type and momentum. These are the same particles created by neutrino interactions in the NOvA detectors, so one can employ the GAN network to produce the simulated data for ND and FD by using the tagged particles from the Test Beam. The above GAN architecture shows a detailed pipeline how the Test Beam data can be used to produce ND and FD simulation data.

Producing simulation analysis files based on Deep Neural Networks, instead of purely using a Monte Carlo generator, may bring a revolution in high-energy physics experimental analysis, and therefore improve prospects for future light sterile neutrino searches.

Appendix A

2017 Re-Analysis Based NC Disappearance δ_{24} Dependence Study

To get a better understanding of how the analysis results depend on the parameter (δ_{24}), which is an extra CP violation phase in the 3+1 Model, we produced a series of neutrino oscillation probabilities as a function of neutrino travel distance (L) and neutrino energy (E), also called **LowerE** plots, and Energy Spectra based on simulated data (also called Fake data).

Fixed Oscillation Parameters For this study, we fixed the following 3+1 model parameters to understand the influence from δ_{24} on the active to sterile oscillations over the NOvA long baseline.

Oscillation Probabilities We use four different values for δ_{24} and obtain the corresponding four oscillation probabilities, as shown in Fig. A.1.

Further, we set $\delta_{13} = 0, \pi/2, \pi$ and $3\pi/2$ separately to understand possible impacts in the results. This is shown in Fig. ??.

Oscillation Parameter	Employed Value
Δm_{21}^2	$7.53 \times 10^{-5} eV^2$
$\sin^2 2\theta_{12}$	0.846
Δm_{32}^2	$2.44 \times 10^{-3} eV^2$
θ_{23}	45°
θ_{13}	8.5°
Δm_{41}^2	$0.5 eV^2$
θ_{24}	30°
θ_{34}	31.2°
δ_{13}	0

TABLE A.1: The oscillation parameter settings used in the δ_{24} dependence study.

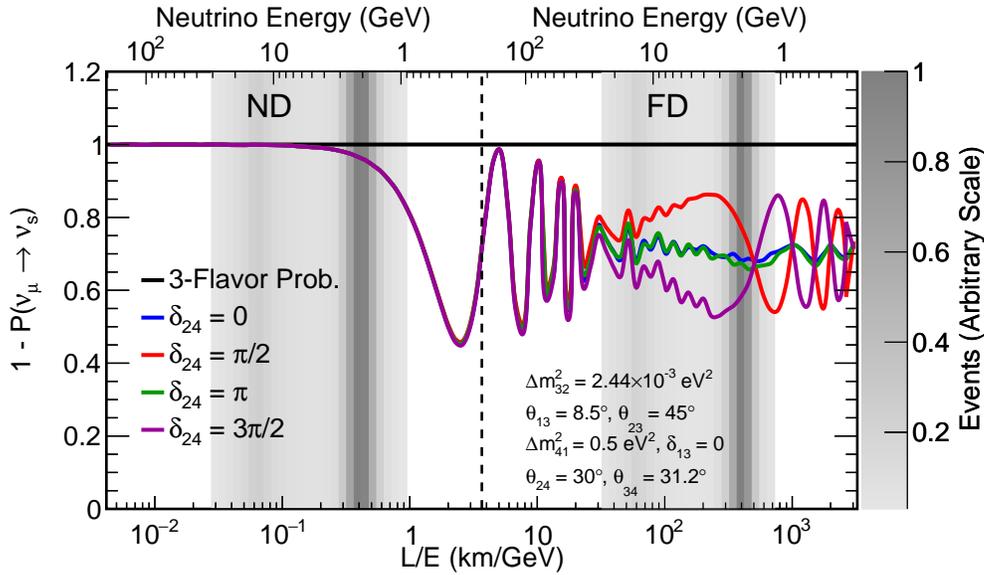


FIGURE A.1: NC Disappearance δ_{24} dependence. We use four different values for δ_{24} and obtain the corresponding four oscillation probabilities.

Energy Spectrum To understand the impact on the neutrino energy spectrum, we present a series of predicted FD spectra assuming different values of δ_{24} . Figure A.3 shows the energy spectrum with $\theta_{24} = 30^\circ$ and $\theta_{34} = 31.2^\circ$. Figure ?? shows the energy spectrum for various values of δ_{24} . These are combined in Fig. A.5, which shows the energy spectra with fixed δ_{24} ($0, \pi/2, \pi, 3\pi/2$).

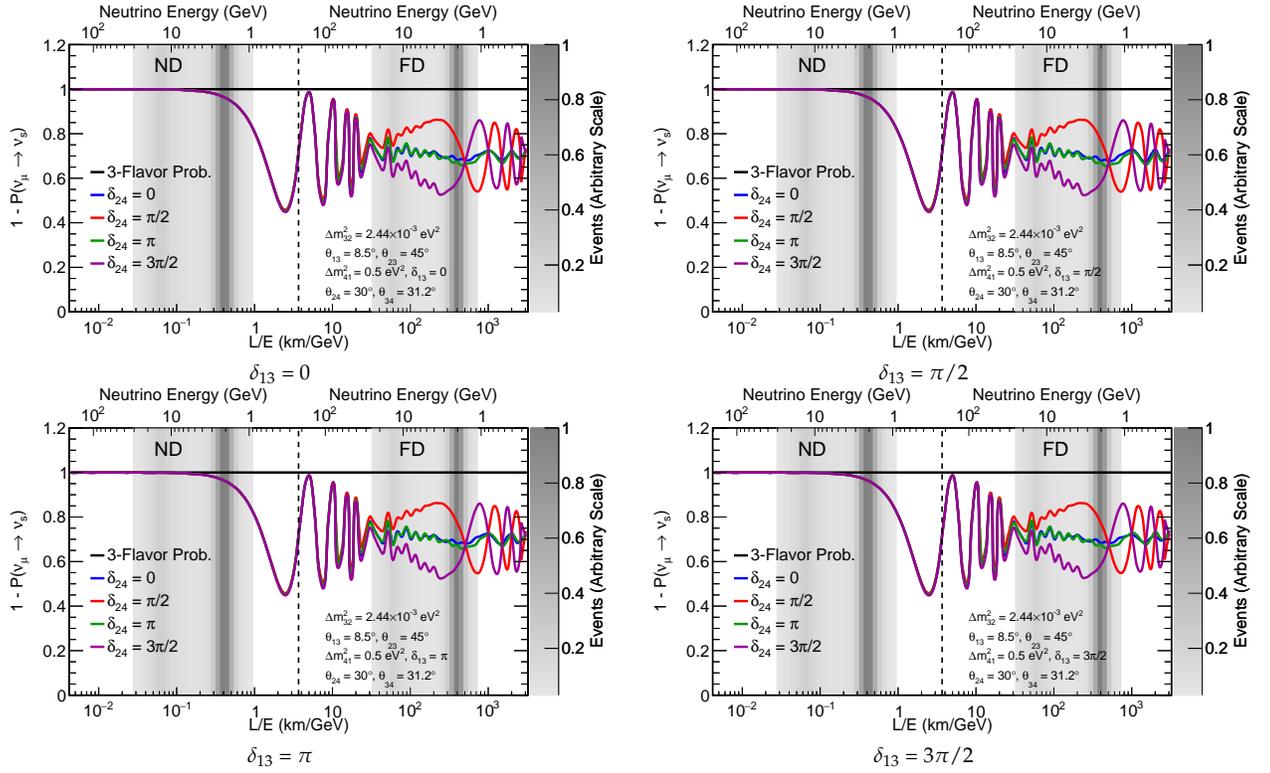


FIGURE A.2: NC Disappearance δ_{24} dependence. We use four different values for δ_{24} and obtain the corresponding four oscillation probabilities.

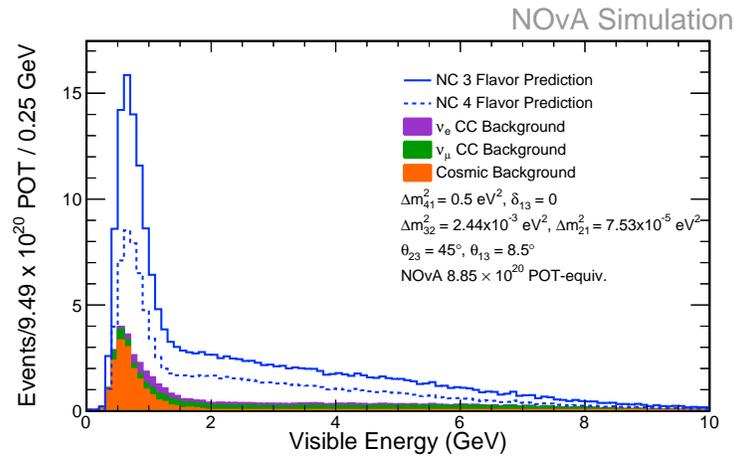
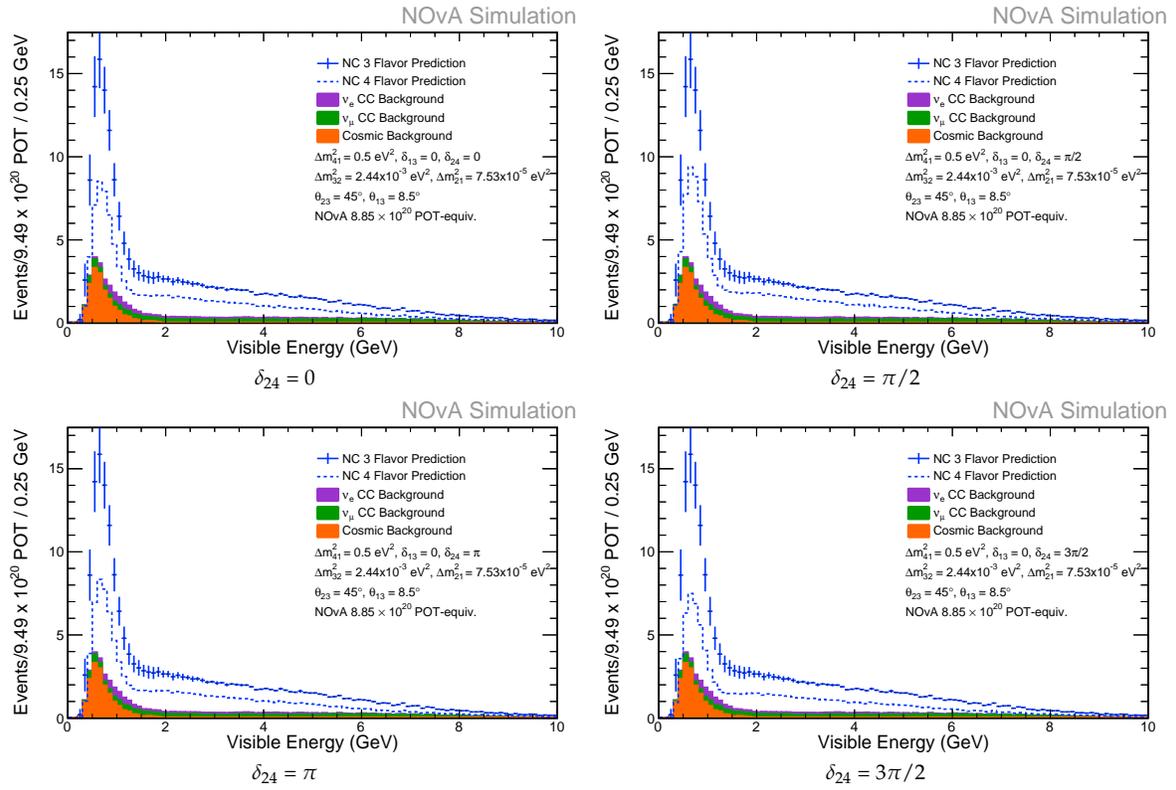
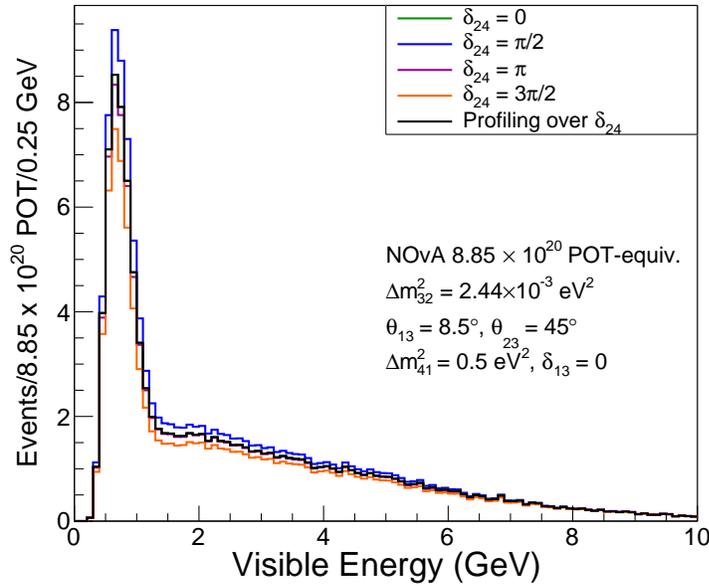


FIGURE A.3: FD Energy Spectrum after we set $\theta_{24} = 30^\circ$ and $\theta_{34} = 31.2^\circ$.

Study


 FIGURE A.4: Energy Spectra with fixed δ_{24} ($0, \pi/2, \pi, 3\pi/2$).

 FIGURE A.5: Comparison of Energy Spectra with fixed δ_{24} ($0, \pi/2, \pi, 3\pi/2$).

Impact on Sensitivity We also present a series of 2D contours obtained for different δ_{24} values. Figure A.6 shows the nominal 2D Contour for θ_{24} and θ_{34} . Figure A.7 shows a series of 2D contours with fixed δ_{24} ($0, \pi/2, \pi, 3\pi/2$), which are then combined into Fig. A.8

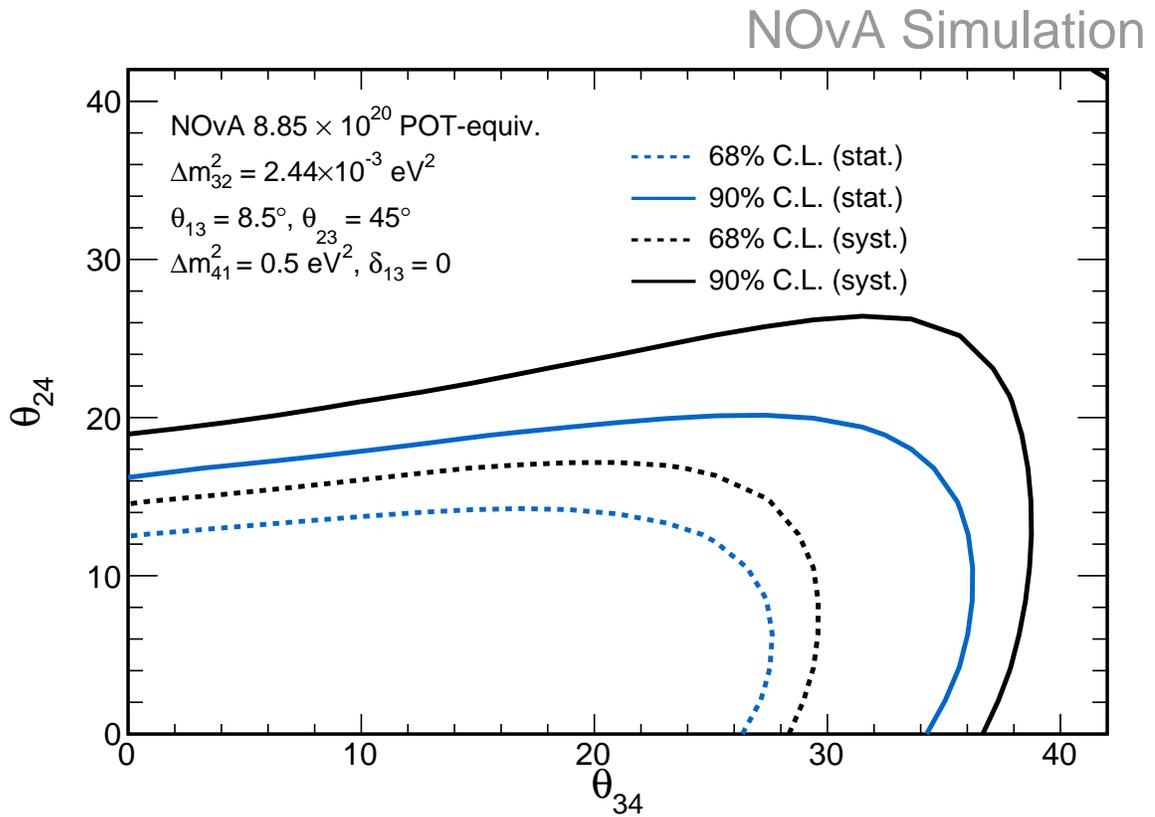
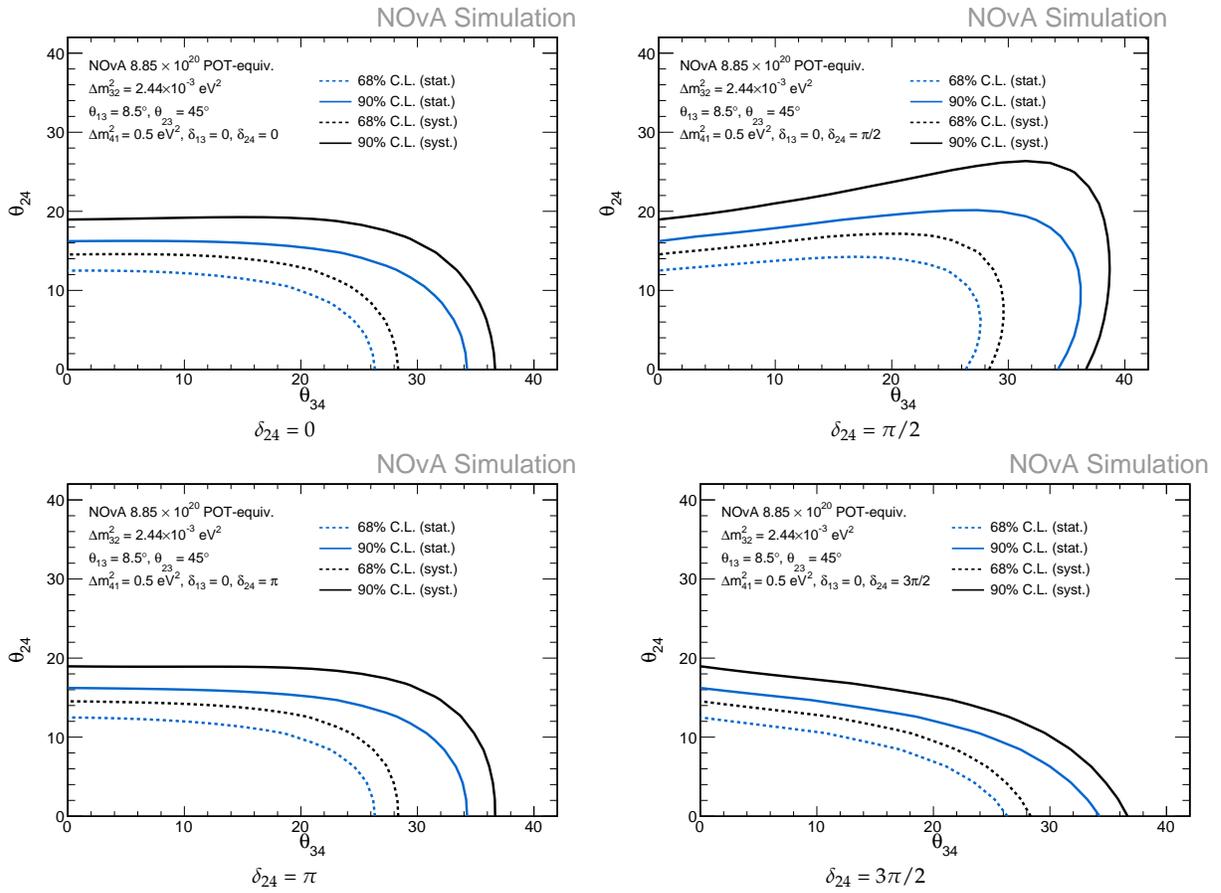
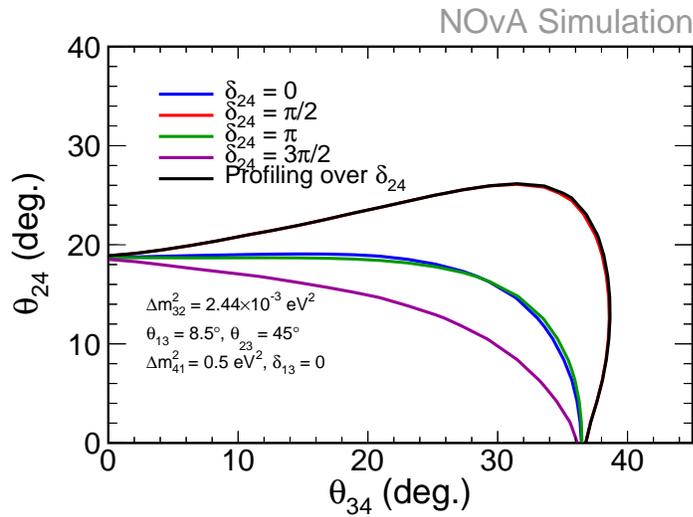


FIGURE A.6: 2D Contour while profiling over δ_{24} .

Study


 FIGURE A.7: 2D contours for fixed values of δ_{24} ($0, \pi/2, \pi, 3\pi/2$).

 FIGURE A.8: Comparison of 2D contours with fixed δ_{24} ($0, \pi/2, \pi, 3\pi/2$).

Appendix B

2017 Official and 2017 Re-Analysis

Per-Period Study

NOvA started to collect data when the FD was still under construction, which inspired the Per-Period Study shown in this Appendix.

B.1 NOvA Per-Period Data Set Details

	Period One	Period Two	Period Three	Period Five	Full DataSet
Far Detector (POT)	0.63×10^{20}	2.62×10^{20}	3.84×10^{20}	1.76×10^{20}	8.85 e20
Near Detector (POT)	1.25×10^{20}	2.52×10^{20}	3.65×10^{20}	1.73×10^{20}	8.04×10^{20}
Cosmic (Live Time)	128.1s	119.1s	139s	51.9s	438.1 s

TABLE B.1: Period DataSet Details.

B.2 2017 Official Vs 2017 Re-Analysis Per-Period Selection Results

Period	Epoch	FD Run #	Selected Events #	Selected # by Period	Equivalent POT	# scaled to 8.85
1		12942 – 17139	27	27	0.63	379.28
2	a	17891 – 19096	32	63	2.62	212.80
	b	19097 – 19586	24			
	c	19587 – 19746	7			
3	b	20923 – 21230	2	76	3.84	175.15
	c	21231 – 22019	17			
	d	22020 – 22900	37			
	e	22901 – 23419	21			
5	a	24614 – 25035	27	48	1.76	241.36
	b	25036 – 25412	21			

FIGURE B.1: 2017 Official Per-Period Selection Results.

Period	Epoch	FD Run #	Selected Events #	Selected # by Period	Equivalent POT	# scaled to 8.85
1		12942 – 17139	27	27	0.63	379.28
2	a	17891 – 19096	40	74	2.62	249.96
	b	19097 – 19586	26			
	c	19587 – 19746	8			
3	b	20923 – 21230	2	86	3.84	198.20
	c	21231 – 22019	19			
	d	22020 – 22900	41			
	e	22901 – 23419	24			
5	a	24614 – 25035	29	50	1.76	251.42
	b	25036 – 25412	21			

FIGURE B.2: 2017 Re-Analysis Per-Period Selection Results.

B.3 2017 Re-Analysis Per-Period Selected Data Distribution

B.3.1 Period One

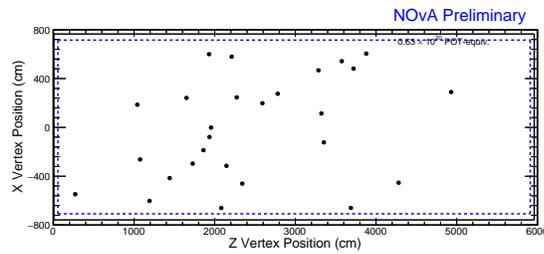


FIGURE B.3: Period One Selected Data Distribution XZ View.

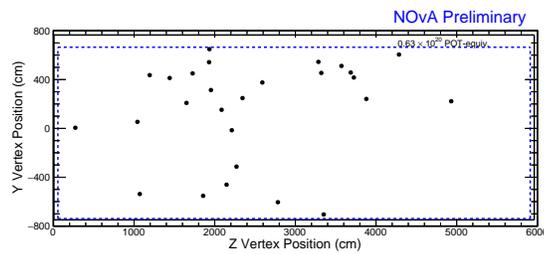


FIGURE B.4: Period One Selected Data Distribution YZ View.

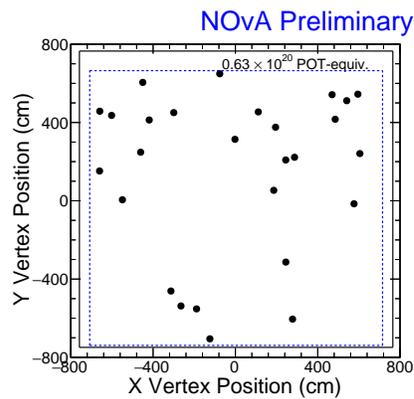


FIGURE B.5: Period One Selected Data Distribution XY View.

B.3.2 Period Two

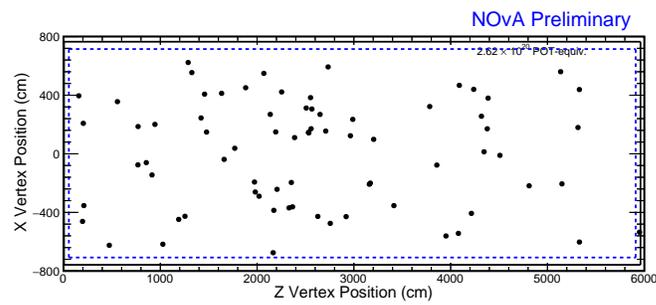


FIGURE B.6: Period Two Selected Data Distribution XZ View.

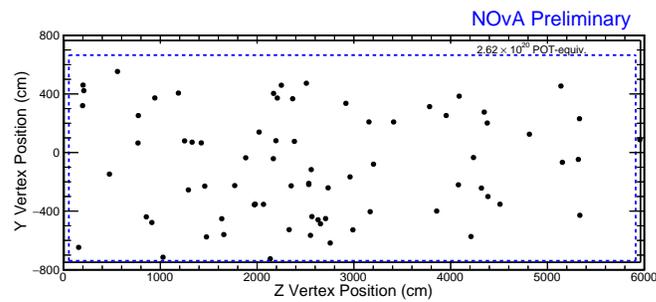


FIGURE B.7: Period Two Selected Data Distribution YZ View.

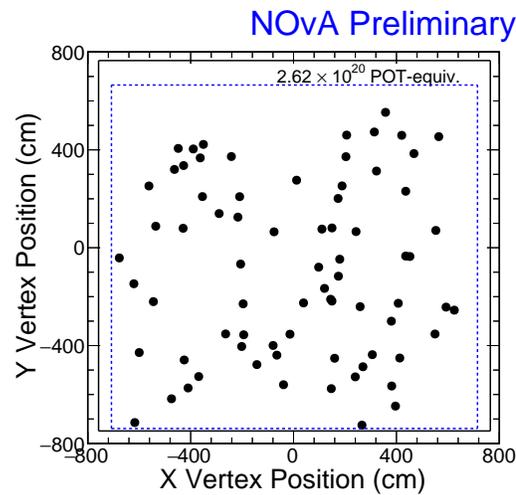


FIGURE B.8: Period Two Selected Data Distribution XY View.

B.3.3 Period Three

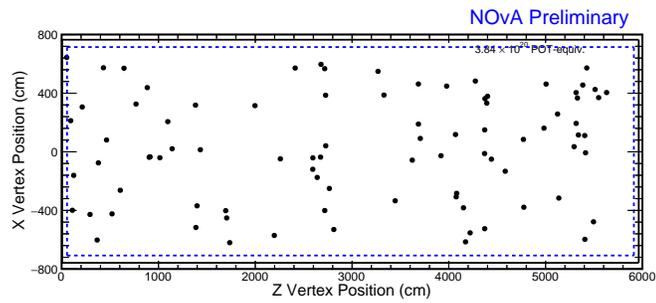


FIGURE B.9: Period Three Selected Data Distribution XZ View.

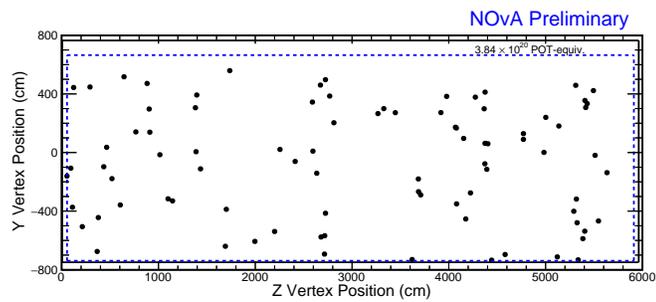


FIGURE B.10: Period Three Selected Data Distribution YZ View.

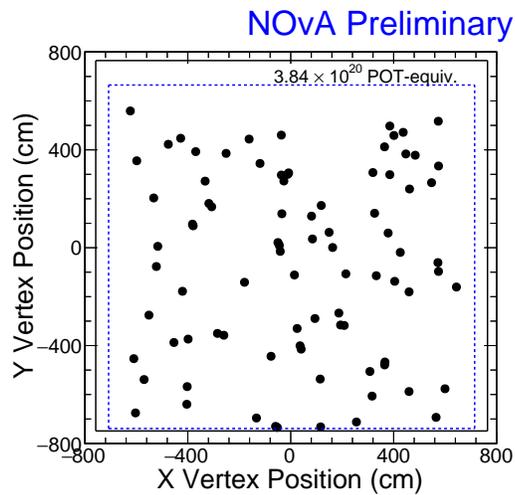


FIGURE B.11: Period Three Selected Data Distribution XY View.

B.3.4 Period Five

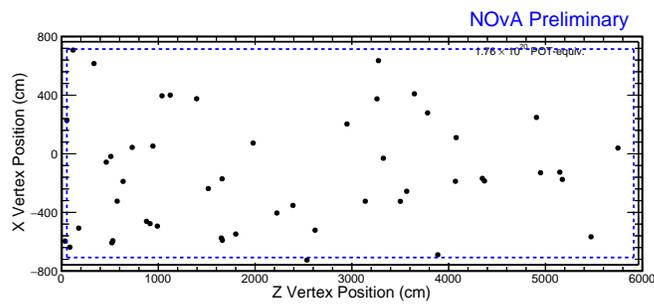


FIGURE B.12: Period Five Selected Data Distribution XZ View.

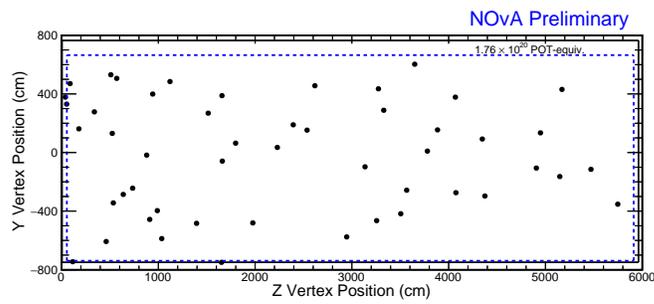


FIGURE B.13: Period Five Selected Data Distribution YZ View.

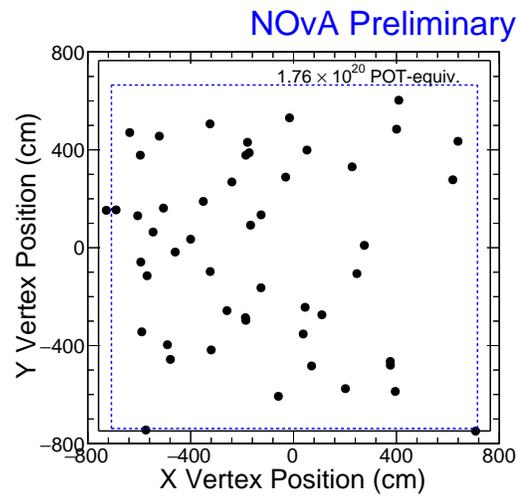


FIGURE B.14: Period Five Selected Data Distribution XY View.

Bibliography

- [1] Collaboration T2K. *About Neutrinos*. 2013.
- [2] Wolfgang Pauli. "Pauli letter collection: letter to Lise Meitner". Typed copy.
URL: <https://cds.cern.ch/record/83282>.
- [3] E. Fermi. "An attempt of a theory of beta radiation. 1." In: *Z. Phys.* 88 (1934), pp. 161–177. DOI: 10.1007/BF01351864.
- [4] S. M. Bilenky. "Neutrino. History of a unique particle". In: *Eur. Phys. J.* H38 (2013), pp. 345–404. DOI: 10.1140/epjh/e2012-20068-9. arXiv: 1210.3065 [hep-ph].
- [5] H. Bethe and R. Peierls. "The 'neutrino'". In: *Nature* 133 (1934), p. 532. DOI: 10.1038/133532a0.
- [6] B. Pontecorvo. "Inverse beta process". In: *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.* 1 (1991). [,97(1946)], pp. 25–31.
- [7] B. T. Cleveland et al. "Measurement of the solar electron neutrino flux with the Homestake chlorine detector". In: *Astrophys. J.* 496 (1998), pp. 505–526. DOI: 10.1086/305343.
- [8] F. Reines and C. L. Cowan. "Detection of the free neutrino". In: *Phys. Rev.* 92 (1953), pp. 830–831. DOI: 10.1103/PhysRev.92.830.
- [9] G. Danby et al. "Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos". In: *Phys. Rev. Lett.* 9 (1962), pp. 36–44. DOI: 10.1103/PhysRevLett.9.36.

- [10] F. J. Hasert et al. "Observation of Neutrino Like Interactions Without Muon Or Electron in the Gargamelle Neutrino Experiment". In: *Phys. Lett.* B46 (1973). [5.15(1973)], pp. 138–140. DOI: 10.1016/0370-2693(73)90499-1.
- [11] K. Kodama et al. "Observation of tau neutrino interactions". In: *Phys. Lett.* B504 (2001), pp. 218–224. DOI: 10.1016/S0370-2693(01)00307-0. arXiv: hep-ex/0012035 [hep-ex].
- [12] Y. Fukuda et al. "Evidence for oscillation of atmospheric neutrinos". In: *Phys. Rev. Lett.* 81 (1998), pp. 1562–1567. DOI: 10.1103/PhysRevLett.81.1562. arXiv: hep-ex/9807003 [hep-ex].
- [13] Q. R. Ahmad et al. "Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory". In: *Phys. Rev. Lett.* 89 (2002), p. 011301. DOI: 10.1103/PhysRevLett.89.011301. arXiv: nucl-ex/0204008 [nucl-ex].
- [14] M. Altmann et al. "Complete results for five years of GNO solar neutrino observations". In: *Phys. Lett.* B616 (2005). [Zh. Eksp. Teor. Fiz.122,211(2002)], pp. 174–190. DOI: 10.1016/j.physletb.2005.04.068. arXiv: hep-ex/0504037 (astro-ph).
- [15] T. Araki et al. "Measurement of neutrino oscillation with KamLAND: Evidence of spectral distortion". In: *Phys. Rev. Lett.* 94 (2005), p. 081801. DOI: 10.1103/PhysRevLett.94.081801. arXiv: hep-ex/0406035 [hep-ex].
- [16] G. Bellini et al. "Neutrinos from the primary proton–proton fusion process in the Sun". In: *Nature* 512.7515 (2014), pp. 383–386. DOI: 10.1038/nature13702.

- [17] P. Adamson et al. "Measurement of Neutrino and Antineutrino Oscillations Using Beam and Atmospheric Data in MINOS". In: *Phys. Rev. Lett.* 110.25 (2013), p. 251801. DOI: 10.1103/PhysRevLett.110.251801. arXiv: 1304.6335 [hep-ex].
- [18] M. H. Ahn et al. "Measurement of Neutrino Oscillation by the K2K Experiment". In: *Phys. Rev. D* 74 (2006), p. 072003. DOI: 10.1103/PhysRevD.74.072003. arXiv: hep-ex/0606032 [hep-ex].
- [19] K. Abe et al. "Precise Measurement of the Neutrino Mixing Parameter θ_{23} from Muon Neutrino Disappearance in an Off-Axis Beam". In: *Phys. Rev. Lett.* 112.18 (2014), p. 181801. DOI: 10.1103/PhysRevLett.112.181801. arXiv: 1403.1532 [hep-ex].
- [20] F. P. An et al. "Spectral measurement of electron antineutrino oscillation amplitude and frequency at Daya Bay". In: *Phys. Rev. Lett.* 112 (2014), p. 061801. DOI: 10.1103/PhysRevLett.112.061801. arXiv: 1310.6732 [hep-ex].
- [21] J. K. Ahn et al. "Observation of Reactor Electron Antineutrino Disappearance in the RENO Experiment". In: *Phys. Rev. Lett.* 108 (2012), p. 191802. DOI: 10.1103/PhysRevLett.108.191802. arXiv: 1204.0626 [hep-ex].
- [22] Y. Abe et al. "First Measurement of θ_{13} from Delayed Neutron Capture on Hydrogen in the Double Chooz Experiment". In: *Phys. Lett. B* 723 (2013), pp. 66–70. DOI: 10.1016/j.physletb.2013.04.050. arXiv: 1301.2948 [hep-ex].
- [23] Wikipedia. *Standard Model of Particles*. 2011.

- [24] P. Adamson et al. "Search for active-sterile neutrino mixing using neutral-current interactions in NOvA". In: *Phys. Rev. D* 96.7 (2017), p. 072006. DOI: 10.1103/PhysRevD.96.072006. arXiv: 1706.04592 [hep-ex].
- [25] A. Aguilar-Arevalo et al. "Evidence for neutrino oscillations from the observation of anti-neutrino(electron) appearance in a anti-neutrino(muon) beam". In: *Phys. Rev. D* 64 (2001), p. 112007. DOI: 10.1103/PhysRevD.64.112007. arXiv: hep-ex/0104049 [hep-ex].
- [26] C Athanassopoulos et al. "The Liquid scintillator neutrino detector and LAMPF neutrino source". In: *Nucl. Instrum. Meth.* A388 (1997), pp. 149–172. DOI: 10.1016/S0168-9002(96)01155-2. arXiv: nucl-ex/9605002 [nucl-ex].
- [27] K. N. Abazajian et al. "Light Sterile Neutrinos: A White Paper". In: (2012). arXiv: 1204.5379 [hep-ph].
- [28] A. A. Aguilar-Arevalo et al. "Significant Excess of ElectronLike Events in the MiniBooNE Short-Baseline Neutrino Experiment". In: *Phys. Rev. Lett.* 121.22 (2018), p. 221801. DOI: 10.1103/PhysRevLett.121.221801. arXiv: 1805.12028 [hep-ex].
- [29] Kim Albertsson et al. "Machine Learning in High Energy Physics Community White Paper". In: *J. Phys. Conf. Ser.* 1085.2 (2018), p. 022008. DOI: 10.1088/1742-6596/1085/2/022008. arXiv: 1807.02876 [physics.comp-ph].
- [30] D. S. Ayres et al. "The NOvA Technical Design Report". In: (2007). DOI: 10.2172/935497.
- [31] J. Vasel. *NP 2016 BLESSING PACKAGE: TIMING SYSTEM PLOTS*. NOvA Internal Document DocDB 15631. 2016.

- [32] M. Tamsett. *DDT blessed plots Neutrino*. NOvA Internal Document DocDB 11471. 2015.
- [33] M. Frank. *DDT Trigger Table*. NOvA Internal Document DocDB 10064. 2015.
- [34] A. Sheshukov. *Supernova DDT schemes blessing package*. NOvA Internal Document DocDB 23643-v3. 2017.
- [35] A. Back. *All-time POT exposure (to May 2018)*. NOvA Internal Document DocDB 30195-v5. 2017.
- [36] S. Agostinelli et al. "GEANT4: A Simulation toolkit". In: *Nucl. Instrum. Meth.* A506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.
- [37] L. Aliaga-Soplin. *Neutrino Flux Predictions for the NuMI Beamline*. PhD thesis, William-Mary Coll. 2016.
- [38] L. Aliaga et al. "Neutrino Flux Predictions for the NuMI Beam". In: *Phys. Rev.* D94.9 (2016). [Addendum: *Phys. Rev.* D95,no.3,039903(2017)], p. 092005. DOI: 10.1103/PhysRevD.94.092005, 10.1103/PhysRevD.95.039903. arXiv: 1607.00704 [hep-ex].
- [39] C. Andreopoulos et al. "The GENIE Neutrino Monte Carlo Generator". In: *Nucl. Instrum. Meth.* A614 (2010), pp. 87–104. DOI: 10.1016/j.nima.2009.12.009. arXiv: 0905.2517 [hep-ph].
- [40] K. Bays. *Cosmic rejection technical note*. NOvA Internal Document DocDB 11205. 2015.
- [41] H. Voss et al. "TMVA, the Toolkit for Multivariate Data Analysis with ROOT". In: *PoS ACAT* (2007), p. 040. DOI: 10.22323/1.050.0040.
- [42] C. Backhouse et al. *The CAFAna Framework*. NOvA Internal Document DocDB 9222. 2015.

- [43] G. S. Davies et al. *Event Selection for the Nus17 NC Disappearance Analysis*. NOvA Internal Note DocDB-22439. 2017.
- [44] A. Sousa and L. Suter. *2017 NC Disappearance Executive Summary : neutral-current-executive*. NOvA Internal Note DocDB-22514. 2017.
- [45] A. Sousa and L. Suter. *2017 NC Disappearance Executive Summary : Additional Information for box opening*. NOvA Internal Note DocDB-22514. 2017.
- [46] L. Goodenough and S. Phan-Budd. *NOvA Beam Monitoring for 2015 Summer Analysis*. NOvA Internal Note DocDB-13572. 2015.
- [47] K. Bays. *LiveGeometry Tech Note*. NOvA Internal Note DocDB-11470. 2015.
- [48] S. Lein. *Tech Note: DCM Edge Metric*. NOvA Internal Note DocDB-13527. 2015.
- [49] K. Sachdev. *Tech Note: Cosmic Veto*. NOvA Internal Note DocDB-14654. 2015.
- [50] K. Sachdev. *Tech Note: FD Selection Cuts for Second Analysis*. NOvA Internal Note DocDB-15387. 2016.
- [51] G. Kafka. *Tech Note: A Technote Describing the Event Selection for the NC Disappearance Analysis*. NOvA Internal Note DocDB-15340. 2016.
- [52] A. Aurisano et al. "A Convolutional Neural Network Neutrino Event Classifier". In: *JINST* 11.09 (2016), P09001. DOI: 10.1088/1748-0221/11/09/P09001. arXiv: 1604.01444 [hep-ex].
- [53] A. Tsaris et al. *Tech Note: CVN*. NOvA Internal Note DocDB-27467. 2016.
- [54] G. S. Davies S. Yang E. Song. *Tech Note: NC Cosmic Rejection BDT*. NOvA Internal Note DocDB-2511. 2017.
- [55] M. Goodman. *Tech Note: Scanning high cvn neutral current background*. NOvA Internal Note DocDB-15319. 2017.

-
- [56] K. Christensen. *Detecting cosmic rays from a galaxy far, far away*. Science Daily. 2017.
- [57] Goddard Space Flight Center NASA. *What are cosmic rays?* Science Daily. 2012.
- [58] C. Patrignani et al. "Review of Particle Physics". In: *Chin. Phys. C* 40.10 (2016), p. 100001. DOI: 10.1088/1674-1137/40/10/100001.
- [59] M. Aguilar et al. "First Result from the Alpha Magnetic Spectrometer on the International Space Station: Precision Measurement of the Positron Fraction in Primary Cosmic Rays of 0.5–350 GeV". In: *Phys. Rev. Lett.* 110 (2013), p. 141102. DOI: 10.1103/PhysRevLett.110.141102.
- [60] O. Adriani et al. "Measurements of cosmic-ray proton and helium spectra with the PAMELA calorimeter". In: *Adv. Space Res.* 51 (2013), pp. 219–226. DOI: 10.1016/j.asr.2012.09.029.
- [61] A. Antoshkin et al. *NOvA Slow Monopole Analysis*. NOvA Internal Note DocDB-22610-v8. 2015.
- [62] M. Rao. *Extensive Air Showers*. World Scientific, p. 5, ISBN 9789810228880. 1998.
- [63] M. Rao. *Extensive Air Showers*. World Scientific, p. 10, ISBN 9789810228880. 1998.
- [64] T. Xin and J. Bian. *Selection Cuts for ν_e Appearance Analysis technical note*. NOvA Internal Note DocDB-13592-v8. 2015.
- [65] A. Géron. *Hands-On Machine Learning with Scikit-Learn and TensorFlow*. Google. 2015.

- [66] Jürgen Schmidhuber. “Deep Learning in Neural Networks: An Overview”. In: *CoRR* abs/1404.7828 (2014). arXiv: 1404.7828. URL: <http://arxiv.org/abs/1404.7828>.
- [67] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. “ImageNet Classification with Deep Convolutional Neural Networks”. In: *Commun. ACM* 60.6 (May 2017), pp. 84–90. ISSN: 0001-0782. DOI: 10.1145/3065386. URL: <http://doi.acm.org/10.1145/3065386>.
- [68] Tianqi Chen and Carlos Guestrin. “XGBoost: A Scalable Tree Boosting System”. In: (2016). DOI: 10.1145/2939672.2939785. arXiv: 1603.02754 [cs.LG].
- [69] Alexander Radovic et al. “Machine learning at the energy and intensity frontiers of particle physics”. In: *Nature* 560.7716 (2018), pp. 41–48. DOI: 10.1038/s41586-018-0361-2.
- [70] S. Joglekar. *Sachin Joglekar’s blog*. codesachin.wordpress.com. Retrieved. 2016.
- [71] R. Rojas. *AdaBoost and the super bowl of classifiers a tutorial introduction to adaptive boosting*. 2009.
- [72] Robert E. Schapire. “A Brief Introduction to Boosting”. In: *Proceedings of the 16th International Joint Conference on Artificial Intelligence - Volume 2*. IJCAI’99. Stockholm, Sweden: Morgan Kaufmann Publishers Inc., 1999, pp. 1401–1406. URL: <http://dl.acm.org/citation.cfm?id=1624312.1624417>.
- [73] Tong Zhang. “Statistical Behavior and Consistency of Classification Methods based on Convex Risk Minimization”. In: *Annals of Statistics* 32 (Dec. 2002). DOI: 10.1214/aos/1079120130.

-
- [74] Christino Tamon and Jie Xiang. "On the Boosting Pruning Problem". In: *ECML*. 2000.
- [75] G. Kafka. *A Technote for the NC Disappearance Analysis Systematic Errors*. NOvA Internal Document Docdb: 15330. 2016.
- [76] E. Catano-Mur. *Technote Nue 2017 Systematics and Cross-checks*. NOvA Internal Document Docdb-23213. 2017.
- [77] T. Warburton. *2017 Analysis - NuMu Systematics Technote*. NOvA Internal Document Docdb-23138. 2017.
- [78] E. Catano-Mur. *Technote Nue and Numu Systematics 2017*. NOvA Internal Document Docdb-24029. 2017.
- [79] A. Aurisano. *Validating the Approximation of Sterile Appearance*. NOvA Internal Document Docdb-16140. 2016.
- [80] E. Catano-Mur. *BEN decomposition technote*. NOvA Internal Document Docdb-15392. 2017.
- [81] D. Pershey. *MichelDecomp for the SA*. NOvA Internal Document Docdb-15395. 2016.
- [82] J. Lozier. *ModularExtrap technical note*. NOvA Internal Document Docdb-12563. 2016.
- [83] S. Edayath. *NC Visible Energy Correction*. NOvA Internal Document Docdb-22508-v3. 2016.
- [84] M. Tanabashi et al. "Review of Particle Physics". In: *Phys. Rev. D* 98.3 (2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001.
- [85] Chantal Bassin. "The Current Limits of resolution for surface wave tomography in North America". In: 2000.

- [86] M. A. Acero et al. “New constraints on oscillation parameters from ν_e appearance and ν_μ disappearance in the NOvA experiment”. In: *Phys. Rev. D* 98 (2018), p. 032012. DOI: 10.1103/PhysRevD.98.032012. arXiv: 1806.00096 [hep-ex].
- [87] Alex Sousa. *NOvA Test Beam Introduction*. NOvA Internal Document Docdb-7344-v2. 2012.
- [88] Alex Sousa. *NOvA Test Beam Task Force Intro*. NOvA Internal Document Docdb 14909. 2016.