

Classification of particles with a convolutional neural network for neutrino and antineutrino events in the NOvA experiment

To be submitted in the partial fulfilment for the degree of
DOCTOR OF PHILOSOPHY IN PHYSICS

BY

CHATLA AKSHAY

16PHPH08



Under the supervision of

Prof. BINDU A BAMBAH

&

Prof. RUKMANI MOHANTA (Co-Supervisor)

School of Physics

University of Hyderabad

Hyderabad 500 046, INDIA

January 2024

DECLARATION

I hereby declare that, this thesis titled **Classification of particles with a convolutional neural network for neutrino and antineutrino events in the NOvA experiment** submitted by me, under the guidance and supervision of **Prof. Bindu A Bambah** and **Prof. Rukmani Mohanta** (Co-Supervisor), is a bonafide research work and is free from plagiarism. I also declare that it has not been submitted previously, in part or in full to this University or any other University or Institution, for the award of any degree or diploma. I hereby agree that my thesis can be deposited in Shodhganga/INFLIBNET.

A report on plagiarism statistics from the University Librarian is enclosed.

Hyderabad

Date: 23-01-2024



(Chatla Akshay)

Reg. No.: 16PHPH08



CERTIFICATE

This is to certify that the thesis entitled **Classification of particles with a convolutional neural network for neutrino and antineutrino events in the NOvA experiment** submitted by Chatla Akshay bearing registration number 16PHPH08 in partial fulfillment of the requirements for award of Doctor of Philosophy in the School of Physics is a bonafide work carried out by his, under my supervision and guidance.

This thesis is free from plagiarism and has not been submitted previously in part or in full to this or any other University or Institution for award of any degree or diploma.

Further, the student has the following publications before the submission of the thesis for adjudication.

Thesis publications:

1. M. A. Acero *et al.* [NOvA], “*The Profiled Feldman-Cousins technique for confidence interval construction in the presence of nuisance parameters,*” [arXiv:2207.14353 [hep-ex]]
2. M. A. Acero *et al.* [NOvA], “*Measurement of the ν_e –Nucleus Charged-Current Double-Differential Cross Section at $\langle E_\nu \rangle = 2.4$ GeV using NOvA,*” Phys. Rev. Lett. **130** (2023) no.5, 051802, doi:10.1103/PhysRevLett.130.051802, [arXiv:2206.10585 [hep-ex]].
3. **Akshay Chatla**, Sahithi Rudrabhatla and Bindu A Bambah, “*Degeneracy Resolution Capabilities of NOvA and DUNE in the Presence of Light Sterile Neutrino*”, **Adv.High Energy Phys.** **2018 (2018) 2547358**, [arXiv:1804.02818 [hep-ph]].

Proceedings:

1. Akshay Chatla, Sahithi Rudrabhatla and Bindu A Bambah, "*Effect of Sterile Neutrinos on Degeneracy Resolution Capacities of NOvA and DUNE*", **Springer Proc.Phys.** **261** (2021) 453-458, [DOI:10.1007/978-981-33-4408-2_63].
2. Akshay Chatla and Bindu A Bambah, "*Effect of sterile phases on degeneracy resolution capabilities of LBL experiments*" *Phys.At.Nucl.*84 (2021) 3, 377-380 , **arXiv:2010.06321** [hep-ph]

Other publications:

1. Abhishek Kumar Jha, Akshay Chatla , "*Quantum studies of neutrinos on IBMQ processors*," *Eur. Phys. J. ST* **231** (2022) no.2, 141-149, doi:10.1140/epjs/s11734-021-00358-9
2. Abhishek Kumar Jha, Akshay Chatla and Bindu A Bambah, "*Neutrinos as Qubits and Qutrits*," *Eur. Phys. J. Plus* **139**, 68 (2024) [arXiv:2203.13485 [hep-ph]].

Further, the student has passed the following courses towards fulfillment of coursework requirement for Ph.D:

Course Code	Name	Credits	Pass/Fail
PY801	Research Methodology	4	Pass
PY802	Advanced Quantum Mechanics	4	Pass
PY803	Advanced Experimental Techniques	4	Pass
PY804	Advanced Condensed Matter Physics	4	Pass

Bindu A. Bambah

Prof. Bindu A Bambah
 Supervisor
 School of Physics
 University of Hyderabad

Date: 23-01-2024

Rukmani Mohanta

Prof. Rukmani Mohanta
 Co-Supervisor
 School of Physics
 University of Hyderabad

Dr. Rukmani Mohanta
 Professor
 School of Physics
 UNIVERSITY OF HYDERABAD
 Hyderabad-500 046.

Jay
 23/1/24
Prof K C James Raju
 Dean
 School of Physics
 University of Hyderabad
DEAN
School of Physics
University of Hyderabad
HYDERABAD - 500 046.

Acknowledgements

I extend my deepest gratitude to my supervisor, Prof. Bindu A Bambah for her guidance and patience through my PhD. I would like to express my gratitude to her for providing me with the invaluable opportunity to work on the NOvA experiment through the esteemed research institute, Fermi National Accelerator Laboratory, USA. I want to extend my heartfelt appreciation towards Prof. Rukmani Mohanta for her support as my co-supervisor.

I am very thankful to my collaborators at NOvA, Prof. Louise Suter, and the convenors of the Reconstruction and Deep Learning group Dr. Erica Smith, Dr. Ashley Back, Dr. Alexander Booth, and Dr. Dr. Wenjie Wu for their regular inputs and insightful feedback which have been instrumental while developing Prong CNN for the NOvA experiment. I am very grateful to Dr. Micah for showing the ropes of PandAna module which helped a lot.

I am extremely grateful to my collaborators Dr. Abhishek Jha, and Sahithi Rudrabhatla for their engaging discussions. I am extremely thankful to Prof. Soma Sanyal for her valuable suggestions during my Doctoral Committee meetings.

I would like to convey my special thanks to our Dean, School of Physics, Prof. K. C. James Raju. With the same respect, I wish to thank our former Deans Prof. Bindu A. Bambah, Prof. V. Sheshubai, Prof. Ashok Chatterjee for fostering a thriving academic environment in the department. In addition, I would like to convey my thanks to all the faculty members of our school.

I express my sincere gratitude to the Department of Science and Technology, India, and Fermilab, USA, for fostering an exceptional collaboration and providing invaluable opportunities for research.

I acknowledge all the non-teaching staff of our School. My special thank goes to Mrs. Shailaja, Mrs. Deepika, Mr. Mahesh, and Mr. Sudharshan for their patience and help in the academic activities.

I would like to appreciate my seniors Dr. Shivaramkrishna, Dr. Sivaprasad, Dr. Rijeesh, Dr. Robertson, Dr. Naveen, and Dr. Soumya for their invaluable support. A very special thanks to my senior and sister Dr. Suchismita Sahoo for being there for me during both my success and struggles. I am extremely thankful to my colleagues Aishwarya , Dr. Dinesh, Dr. Abhishek, Dr. Atasi, Dr. Rudra, and Dr. Mitesh for many memorable experiences. I wish all the very best to my juniors Dinesh, Priya, Papia, Sambit, and Dhiren for their future.

A huge thanks to my Telugu Thambis: Vishnu, Dr. Mangababu, and Sivaram for making my campus life so much fun and eventful. I am profoundly grateful to my dear friend Sriram for taking time to check up on me regularly and the fun philosophical debates we had. I have nothing but love for Bunty and Bubby, our department dogs who kept me company during the long nights of work. My deepest appreciation to Eiichiro Oda for creating “One Piece” which never fails to put smile on my face.

I would like to convey my love and gratitude to my parents Shobha and Bhumeswar, my sister Akshitha and Brother-in-law Saikrishna for being my rock during my turbulent times. I am sincerely grateful to my extended family for their unwavering support, encouragement, and love. A special thank you to my cousins gang for all the good times and bearing with my stupid jokes. Last but not least, I would like to convey my love to Chocolate, our family dog for his unconditional love.

Abstract

We are in the precision era of neutrino oscillation experiments. The NOvA experiment aims to precisely determine neutrino oscillation parameters (δ_{CP} , θ_{23} , and sign of Δm_{32}^2). High energy physics experiments like NOvA produce vast amounts of data. This makes traditional analysis methods challenging and time-consuming. Machine learning algorithms excel at extracting patterns and insights from large datasets, aiding in the identification of relevant particle signatures, event classification, and background noise reduction. This makes using machine learning in high energy physics very beneficial.

NOvA uses machine learning to develop various tools to assist the analysis. For example, Event CNN is used to classify candidate neutrino interactions. Identification of final state particles of an event is done through Prong CNN. Prong CNN is trained to identify all the final-state particles ($e^\pm, p^+, \mu, \pi^\pm, \gamma$) of a given neutrino event.

In this thesis, we gave detailed description of NOvA experiment and various algorithms used by NOvA for simulation and reconstruction. Then, we described the development of Prong CNN, and modifications made compared to previous version. We trained single Prong CNN using both neutrino and anti-neutrino events. In our updated Prong CNN, we were able to reduce runtime on CPUs without compromising performance. Our Prong CNN is now more computationally efficient, enabling higher classification efficiency within fewer training epochs. The newer prong CNN has a global efficiency greater than 86% with improved the classification efficiency by 3% compared to previous Prong CNN.

Besides developing the prong CNN for the NOvA experiment, a phenomenological study was conducted within the framework of the 3+1 sterile neutrino model using GLoBES. The study was aimed to assess the impact of sterile neutrinos on the degeneracy resolution capabilities of both the NOvA and DUNE experiments.

Contents

Declaration	iv
Certificate	v
Acknowledgements	ix
Abstract	xi
List of Figures	xiv
List of Tables	xviii
1 Introduction to Neutrinos	1
1.1 Discovery of the neutrino	2
1.2 Neutrino in the Standard Model	5
1.3 Neutrino mass: going beyond standard model	7
1.4 Neutrino Sources	9
1.5 Neutrino interactions	11
1.6 Neutrino Oscillations	14
1.6.1 Neutrino oscillations in matter	18
1.7 Neutrino experiments	20
1.8 Oscillation Parameters - Current status	22
1.9 Outline of Thesis	24
2 The NOvA Experiment	25
2.1 The NuMI Beam	26
2.2 NOvA Detectors	31
2.3 Data Acquisition System	36
2.3.1 Avalanche Photodiodes	36
2.3.2 Front End Boards	37
2.3.3 Data Concentrator Modules	38
2.4 Timing System	39
2.4.1 Triggers	42
3 NOvA Simulation and Event Reconstruction	44
3.1 Monte-Carlo simulations	44
3.2 Simulation chain of NOvA	46
3.2.1 NuMI Beamline simulation	46
3.2.2 Simulation of Neutrino Interactions	49
3.2.3 Detector simulation	50
3.2.4 Photon Transport	51

3.3	Detector Calibration	51
3.3.1	Relative calibration	53
3.3.2	Absolute calibration	55
3.4	Event Reconstruction	56
3.4.1	Clustering	57
3.4.2	Vertexing	61
3.4.3	Prong reconstruction	65
4	Prong Convolutional Neural Network	68
4.1	Introduction to Machine learning	69
4.1.1	Artificial neural networks	69
4.1.2	Training Procedure	72
4.1.3	Convolutional neural networks	75
4.2	Machine learning in NOvA	80
4.2.1	Event CNN	80
4.3	Prong CNN	81
4.4	Training Datasets:	82
4.5	Architecture	84
4.6	Performance Metrics	87
4.7	Application of Prong CNN	89
4.7.1	Calibration of π^0 mass peak	89
5	Degeneracy resolution capabilities of NOvA and DUNE in the presence of light sterile neutrino	93
5.1	Introduction to sterile neutrinos	93
5.2	3+1 Sterile Neutrino Framework	95
5.3	Experiment simulation specifications	97
5.4	Results	103
5.4.1	Effect of sterile neutrino on $\sin^2\theta_{23}$ and δ_{13}	103
5.4.2	Effect of sterile phases on $\sin^2\theta_{23}$ and δ_{13}	109
5.5	Conclusions	110
6	Conclusion and Future work	112
	Bibliography	115

List of Figures

1.1	β decay spectrum of Bi^{210} [6] which confirmed that the β decay produces a continuous energy spectrum. Bi^{210} was called Radium E in 1927.	3
1.2	Family of Standard Model Theory	7
1.3	Predicted neutrino flux from different natural sources [22].	10
1.4	The proton - proton chain reactions which produce neutrinos in stars, taken from [23]	11
1.5	CC and NC neutrino interactions are depicted through Feynman diagrams. . . .	13
1.6	Neutrino Cross sections. The neutrino(antineutrino)-nucleon cross sections plotted as a function of neutrino(antineutrino) energy are shown in the left (right) panel [24].	14
1.7	The two hierarchies of the neutrino mass states. On the left, the normal ordering ($m_1 < m_2 < m_3$) is shown, and the inverted ordering ($m_3 < m_2 < m_1$) is on the right.	23
2.1	Map depicting the off-axis beamline of NOvA originating at Fermilab. Courtesy of [46].	26
2.2	Overview of the Fermilab accelerator complex in schematic form[48].	27
2.3	A Schematic illustrating the production of a neutrino beam for NOvA. It depicts 120 GeV protons colliding the target on the left side, while produced neutrinos exit to the right [49].	29
2.4	The neutrino energy(left) and neutrino flux(right) as a function of pion energy. These are shown for various angles of θ in relation to the pion trajectory. [48]. .	30
2.5	Expected Neutrino events at FD in FHC mode are shown in the left plot and right plot is for RHC mode. The event rates shown here are corrected for neutrino cross sections in the detector and the flux of the beam, without applying oscillations [50].	31
2.6	Model of Far and Near Detectors of NOvA to scale[48].	33
2.7	A single NOvA cell with wavelength-shifting fiber looped inside it[52].	34
2.8	Left: The end points of WLS fibers from each cell, from total 32 cells. Right: The front face of an APD with 32 inputs. This will be pushed against the fiber ends[52].	34
2.9	A 3D representation of the NOvA Detector in schematic form displaying the alternating vertical and horizontal placement of planes to provide us with three-dimensional events[52].	35
2.10	Data Acquisition (DAQ) system of NOvA FD.[48]	39
2.11	Timing system schematic of NOvA.[48]	40
2.12	DCS algorithm. Two pulses of different shapes are shown. In the singlepoint timing method, both pulses will have the same ADC value ($s_i - s_{i-3}$) and the TDC value is time at s_{i-3} . We must use a multipoint timing method to differentiate between these two pulses.[53]	42
3.1	The simulation chain of NOvA [56].	47
3.2	The fractional uncertainties of the flux due to hadron production uncertainties as a function of neutrino energy at NOvA ND detector for the ν_e (left) and ν_μ (right) flux. [62].	47

3.3	The PPFX corrected simulated NuMI beam flux at NOvA ND detector for neutrino mode in the left plot and antineutrino mode in the right plot [64].	48
3.4	Mass stopping power (dE/dx) for positive muons in copper [71].	53
3.5	The tri-cell cut used for detector calibration.	53
3.6	The dE/dx of stopping muons as function of distance from the end point of track.	55
3.7	(a) Attenuation corrected dE/dx. (b) Absolute calibrated dE/dx.	56
3.8	The event display showing readout of 550 μ s of NuMI trigger at FD. The top part of the event display shows the x-view and the bottom pane shows the y-view. Hit colors correspond to ADC values.	57
3.9	The event display showing readout of FD Cosmic trigger at FD. Hit colors correspond to readout timing.	58
3.10	The event display showing readout of 10 μ s of NuMI beam spill at ND. Hit colors correspond to readout timing.	59
3.11	Comparison between 4DSlicer and TDSlicer by plotting The fraction of good slices as a function of intensity.	62
3.12	A schematic diagram illustrating the coordinate transformation of the Hough transform. The left plot depicts the line in Cartesian coordinate space; The right plot transforms the line into the Hough space.	63
3.13	The multi point Hough transform on a event in FD. The red cross is position of interaction vertex produced by elastic arms algorithm	64
3.14	The prong reconstruction of event in fig.3.13 using FuzzyK algorithm. Each of the prong identified by FuzzyK algorithm are shown in different colors.	67
4.1	Schematic overview of an Artificial Neural Network.	70
4.2	Activation Functions.	71
4.3	Schematic overview of a basic CNN architecture.	76
4.4	Operation process of a Convolutional layer.	77
4.5	Operation process of Max Pooling and Average Pooling layers.	78
4.6	The plot shows the comparison between the energy spectrum of individual hits employing 8 bit precision binning with only 256 bins and an arbitrary fine binning.	81
4.7	No: of prongs belonging to different classes after initial data preprocessing cuts are applied.	83
4.8	Depth-wise separable Convolution.	85
4.9	Comparison between Swish and Hard Swish.	85
4.10	(a) Residual connection, (b) Residual connection with SE Block, (c) Residual connection with ECA Block	86
4.11	Block diagram of New Prong CNN Architecture based on modified MobileNetV2 architecture, adding Efficient channel attention block for improved channel attention.	87
4.12	Classification matrices showing the comparison between performance of new prong CNN and old model on FD MC. They are normalized to show the efficiency of each category along the diagonal. The off-diagonal shows how each category is misclassified.	88
4.13	Classification matrices showing the comparison between performance of new prong CNN and old model on FD MC. They are normalized to show the purity of each category along the diagonal. The off-diagonals show the common backgrounds to each category.	89
4.14	The ROC curves of each class showing the comparison between performance of new prong CNN and old model on FD MC.	90
4.15	The reconstructed π^0 invariant mass using selection based on new(old) Prong CNN is on the left(right). The invariant mass is reconstructed using ND FHC. The MC is scaled to the data POT for both plots.	91
4.16	The reconstructed π^0 invariant mass using selection based on new(old) Prong CNN is on the left(right). The invariant mass is reconstructed using ND RHC. The MC is scaled to the data POT for both plots.	92

5.1	The oscillation probability $P_{\mu e}$ as a function of energy. The Top(bottom) panel is NOvA(DUNE). The bands correspond to different values of δ_{14} , ranging from -180° to 180° when $\delta_{24} = 0^\circ$. Inside each band, the probability for $\delta_{14} = 90^\circ$ ($\delta_{14} = -90^\circ$) case is shown as the solid (dashed) line. The left(right) panel corresponds to neutrinos(anti-neutrinos).	100
5.2	The oscillation probability $P_{\mu e}$ as a function of energy. The Top(bottom) panel is NOvA(DUNE). The bands correspond to different values of δ_{24} , ranging from -180° to 180° when $\delta_{14} = 0^\circ$. Inside each band, the probability for $\delta_{24} = 90^\circ$ ($\delta_{24} = -90^\circ$) case is shown as solid (dashed) line. The left(right) panel is for neutrinos(anti-neutrinos).	101
5.3	(a)The electron neutrino((b)The electron anti-neutrino) appearance probability $P_{\mu e}(P_{\bar{\mu e}})$ for NOvA as function of $L/E(\text{km}/\text{GeV})$ for various bands of $(\delta_{14} - \delta_{24})$ while varying individual values of δ_{14}, δ_{24} inside the band.	102
5.4	(a)The electron neutrino((b)The electron anti-neutrino) appearance probability $P_{\mu e}(P_{\bar{\mu e}})$ for DUNE as function of $L/E(\text{km}/\text{GeV})$ for various bands of $(\delta_{14} - \delta_{24})$ while varying individual values of δ_{14}, δ_{24} inside the band.	103
5.5	Contour plots of allowed regions in the test plane, θ_{23} vs δ_{13} , at 90% C.I with top, middle and bottom rows for NOvA runs of $3 + \bar{0}, 3 + \bar{1}$ and $3 + \bar{3}$ years respectively.	104
5.6	Contour plots of allowed regions in the test plane θ_{23} vs δ_{13} at 99% C.L with top, middle and bottom rows for DUNE runs of $1 + \bar{0}, 1 + \bar{1}$ years and DUNE[$1 + \bar{1}$]+NOvA[$3 + \bar{3}$] respectively.	107
5.7	Contour plots of allowed regions in the test plane θ_{23} vs δ_{13} at 99% C.L with top and bottom rows for DUNE[$5 + \bar{5}$]and NOvA[$3 + \bar{3}$] + DUNE[$5 + \bar{5}$] respectively.	108
5.8	Contour plots of allowed regions in the test plane, θ_{23} vs δ_{13} for different $(\delta_{14} - \delta_{24})$ values at 90% C.I regions for NOvA[$3 + \bar{3}$]. We take NH-HO(NH-LO) as true values for a(b) case.	109
5.9	Contour plots of allowed regions in the test plane, θ_{23} vs δ_{13} for different $(\delta_{14} - \delta_{24})$ values at 90% C.I regions for DUNE[$3 + \bar{3}$]. We take NH-HO(NH-LO) as true values for a(b) case.	110

List of Tables

1.1	Current global fit of three-flavor neutrino oscillations [42].	22
2.1	Prominent decay modes of charged mesons to neutrinos and their branching fractions.	30
4.1	Various Cuts used for preprocessing the data for training.	83
4.2	Comparison of efficiency and purity metrics between new and old prong CNN. . .	90
5.1	Details of experiments used for simulation using GLoBES.	99
5.2	Systematic errors associated with NOvA and DUNE	99
5.3	Oscillation parameters considered in numerical analysis. The $\sin^2 \theta_{23}$ and δ_{13} are taken from latest NOvA results[133].	102

Dedicated to

$\{\{\mathbf{Kind\ people}\} \cap \{\mathbf{Critical\ thinkers}\}\} \cup \{\mathbf{Dogs}\}$

Chapter 1

Introduction to Neutrinos

Humans have always been inquisitive about the world we live in. Our continuous and scientific study of the universe and the laws governing it has advanced human civilization to a great extent. Continuing the legacy, particle physicists are working towards a better understanding of fundamental particles and their interactions.

Neutrinos are fundamental particles with unique properties that make them interesting to study. Neutrinos stand as the most prevalent massive particles in the whole universe. Despite that, neutrinos are among the least understood because of their tiny interaction cross section. For example, we receive a neutrino flux of approximately $\approx 6.5 \times 10^{10}/cm^2/s$ [1] from the Sun, but these neutrinos pass through the Earth, barely interacting with the Earth's mass. Neutrinos emitted from astrophysical processes can travel long distances without interacting with matter, providing us with another channel for studying star formation. As a result, the study of neutrinos offers us a chance to probe the origin, expansion, and evolution of the universe.

The low interaction cross section of neutrinos poses challenges for their detection and study. Not only is it rare for neutrinos to interact, but the exact rates and processes of their interactions in current models are also subject to significant

uncertainties. Nevertheless, researchers in experimental particle physics are dedicated to gaining comprehensive knowledge about these enigmatic particles and their part in the evolution of the universe. Though, neutrino physics saw great progress since the first observations of neutrinos, many unanswered questions remain.

1.1 Discovery of the neutrino

Radioactivity was discovered by Henry Becquerel when he found an unidentified energy source forming tracks on photographic plates originating from Uranium salts [2–4]. The study of this new energy source led to a revolution in physics and was the origin of particle physics, which we love so much. The Nobel Prize for Physics of 1903 was presented to Henry Becquerel, Marie Curie and Pierre Curie for their phenomenal contribution to the study of radioactivity. Continuing the studies of radioactivity in 1914, a continuous energy spectrum for β decay was observed by James Chadwick (See Fig 1.1) instead of the discrete energy spectrum one expects from decay containing two final particles [5]. Ellis and Wooster confirmed this result in 1927 [6]. This finding posed a severe problem as it challenged the fundamental conservation laws of energy and angular momentum. Multiple theories were put forth to explain this issue, but none of them were successful.

In 1930, Wolfgang Pauli proposed a weakly interacting neutral particle carrying away the missing energy and named it neutron [7]. Enrico Fermi proposed a four point interaction theory for explaining the β decay [8] in 1933. It included the mystery particle proposed by Pauli. He also renamed the third particle involved in this process as “neutrino” which translates to “little neutral one” in Italian, with the intention of differentiating it from neutron discovered by James Chadwick. With this theory, Fermi successfully replicated the observed continuous spectrum.

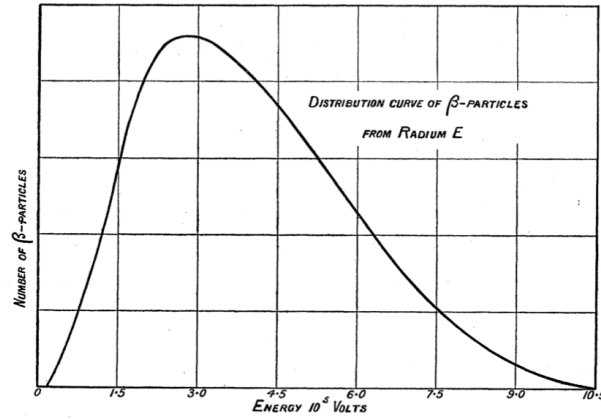


FIGURE 1.1: β decay spectrum of Bi^{210} [6] which confirmed that the β decay produces a continuous energy spectrum. Bi^{210} was called Radium E in 1927.

It took over 25 years from Pauli's initial prediction for a neutrino to be definitively observed. The neutrino interaction with a nucleus through inverse β decay was finally observed in 1956 by F.Reines and C.Cowan [9].

$$\bar{\nu} + p^+ \rightarrow n + e^+ \quad (1.1)$$

Their experiment involved detecting the scintillation light from the energy deposited by two photons released by positron annihilation and the subsequent delayed photon released by neutron capture resulting from the neutrino interactions. This experiment employed a close by nuclear reactor as the source of neutrinos. The neutrino they observed is later found to be electron antineutrino, one of the three flavours of neutrinos. For this work, Reines was awarded the Nobel Prize for Physics in 1995. Pontecorvo proposed the theory of neutrino oscillation [10], drawing an analogy to K^0 - $\bar{K}^0$ mixing in 1957. However, during that period, a single neutrino flavour, namely the electron neutrino, was known to exist. So, he suggested the prospect of a transition between a neutrino and an antineutrino.

Today, from the width of Z boson decay measured by the OPAL [11] and ALEPH detectors at the Large Electron-Positron collider (LEP), it is known that there can be only three flavours of lightweight active neutrinos [12]. Each of these

neutrinos (electron, muon, and tau neutrino) form a group with corresponding charged leptons (electron, muon, and tau) respectively. A given flavour of neutrino always generates its respective charged lepton in charged current (CC) interactions. CC interactions occur via mediation of charged W boson. In neutral current (NC) interaction case, a neutrino scatters off a nucleon. The NC interactions occur via mediation of a neutral Z boson. In 1962, muon neutrino, the second neutrino flavour, was observed at Brookhaven National Lab [13]. In 1988, the Nobel Prize for Physics was awarded to Leon M. Lederman, Melvin Schwartz, and Jack Steinberger for “the neutrino beam method and the demonstration of the doublet structure of the leptons through the discovery of the muon neutrino”. The third and final flavour, the tau neutrino, was observed by the DONUT experiment at FermiLab in 2000 [14].

In the late 1960s, Raymond Davis, Jr. and John N. Bahcall led the Homestake experiment [15, 16]. The aim of the experiment was to detect and count neutrinos produced in the Sun by nuclear fusion. Bahcall carried out the theoretical calculations for the expected neutrino detection rate while Davis designed the experiment. However, when the experiment was conducted, only one-third of the expected neutrinos were detected. This discrepancy was called the solar neutrino problem. The Homestake experiment was the first to achieve the milestone of successfully detecting and counting solar neutrinos. In 2002, the Nobel Prize for Physics was awarded to Ray Davis for “pioneering contributions to astrophysics, in particular for the detection of cosmic neutrinos.”

In 1998, the Super-Kamiokande experiment observed the proof of neutrino oscillations using atmospheric ν_μ s produced by cosmic rays interacting in a water Cherenkov detector [17]. This resolved the solar neutrino problem by revealing that approximately two-thirds of the solar ν_e s oscillated into ν_μ s or ν_τ s.

As the Homestake experiment was not sensitive to ν_μ s or ν_τ s, it could not resolve the solar neutrino problem because it The SNO (Sudbury Neutrino Observatory)

[18] and Super-Kamiokande experiments confirmed the phenomenon of neutrino oscillations. Takaaki Kajita and Arthur B. McDonald were jointly awarded the Nobel Prize in Physics 2015 “for the discovery of neutrino oscillations, which shows that neutrinos have mass.”

The three flavour neutrino oscillation is governed by the “Pontecorvo-Maki-Nakagawa-Sakata (PMNS)” matrix. Numerous neutrino oscillation experiments were constructed with the aim of accurately measuring the oscillation parameters. Today we have accurate measurements of most oscillation parameters pushing neutrino physics into its precision era.

1.2 Neutrino in the Standard Model

The Standard Model (SM) of particle physics is a well-established theoretical framework that describes the fundamental particles and their interactions. It is the culmination of decades of experimental observations and theoretical developments in the field of particle physics. The SM is based on a symmetry principle of local gauge invariance with a symmetry group $SU(3)_C \times SU(2)_L \times U(1)_Y$, where C, L, and Y represent the colour, left-handed chirality, and hypercharge respectively.

The Standard Model contains a total of seventeen elementary particles divided into two types: fermions and bosons.

- Fermions: Fermions are the building blocks of matter and come in two categories: quarks and leptons.

1. Quarks are bound together by the strong force to form protons and neutrons, which are the constituents of atomic nuclei. The six quarks can be classified into three generations:

- I generation: up, down.
- II generation: charm, strange.

- III generation: top, bottom.
- 2. Leptons consist of three types: electron, muon, and tau, each with its corresponding neutrino. Electrons are familiar as they are found in atoms, while muons and taus are heavier counterparts of electrons.
- Bosons: Bosons are force-carrying particles responsible for mediating interactions between fermions. The Standard Model includes the following bosons:
 1. Photon: Mediates electromagnetic interactions.
 2. Z^0 and $W^\pm$ bosons: Responsible for weak nuclear interactions.
 3. Gluons: Carry the strong force, binding quarks together within protons, neutrons, and other hadrons.
 4. Higgs boson: The last elementary particle to be discovered, its associated field imparts other particles with mass through the Higgs mechanism.

In 2012, Large Hadron Collider (LHC) [19] discovered the Higgs boson which completed the SM. The Higgs mechanism is responsible for generating masses for fundamental particles. The Yukawa coupling between Higgs field and fermion field generates fermion mass (except for neutrino). Refer to Fig 1.2 to see the complete family of SM theory.

Electromagnetic, strong, and weak nuclear forces (three of the four known forces) are successfully explained by SM, describing most of the known experimental results in particle physics with remarkable precision. However, it has its limitations and does not incorporate gravity, dark matter, or dark energy, which are major components of the universe. Scientists continue to work on extending the Standard Model to address these unanswered questions, and it continues to be an active research domain within the field of particle physics.

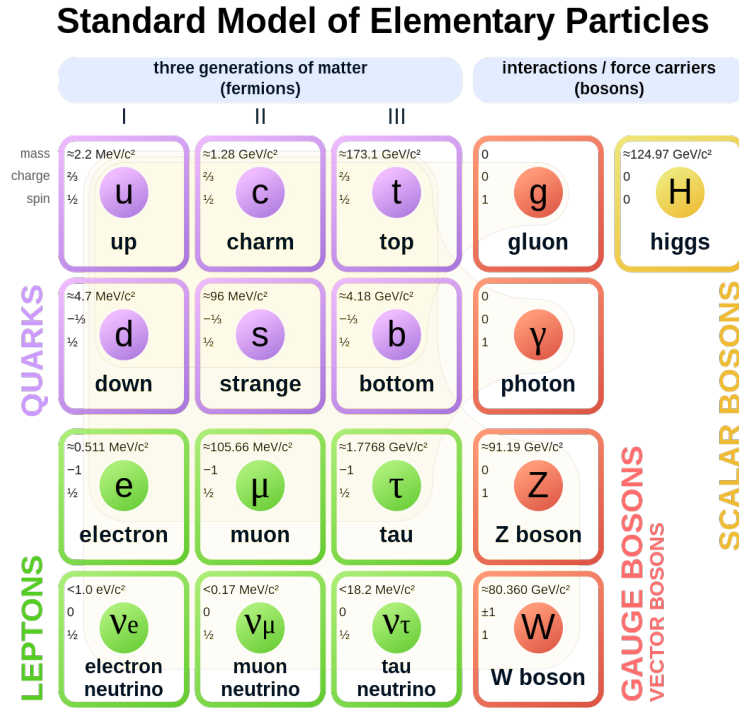


FIGURE 1.2: Family of Standard Model Theory

1.3 Neutrino mass: going beyond standard model

Each active neutrino is linked with its corresponding charged lepton in an isospin doublet. All the fermions are massless before spontaneous symmetry breaking. Nevertheless, only leptons get mass from Yukawa coupling with the Higgs field after the symmetry breaking. absence of right handed neutrinos causes neutrinos to remain massless in SM. Neutrino is a chargeless, massless particle according to SM theory. Nevertheless, for neutrino oscillations to occur, neutrinos need mass. This makes neutrinos special as they go against SM predictions.

Derived from cosmological data, the total sum of neutrino (Σm_ν) masses is calculated to be less than 120 meV [20]. Compared to the electron mass ($m_e = 0.511 \text{ MeV}$), this is six orders of magnitude smaller. Direct measurements of the neutrino mass are obtained through beta decay detection by the KATRIN experiment, put the present upper limit being less than 800 meV at 90% C.L[21]. To explain

this tiny mass of neutrinos, we require special mass modelling like the Seesaw mechanism.

Neutrinos of SM are Dirac particles which means particles and antiparticles of neutrinos are distinguishable. The seesaw mechanism requires a Majorana neutrino, which is a hypothetical type of neutrino that is its own antiparticle. So, in the seesaw mechanism, neutrino can have both Dirac and Majorana mass.

Here we show standard model Lagrangian :

$$L_{SM} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\bar{\psi}\not{D}\psi + h.c. + \bar{\psi}_i y_{ij} \psi_j \phi + h.c. + |D_\mu \phi|^2 - V(\phi) \quad (1.2)$$

We need to add following neutrino mass terms to L_{SM} :

A Dirac mass term:

$$m_D \bar{\psi} \psi = m_D \bar{\psi}_L \psi_R + m_D \bar{\psi}_R \psi_L \quad (1.3)$$

Now, Majorana mass term is derived by substituting ψ_R with $\psi_L^C = C\bar{\psi}_L^T$:

$$m_M \bar{\psi} \psi = m_L \bar{\psi}_L^C \psi_L + m_R \bar{\psi}_R^C \psi_R \quad (1.4)$$

Total Lagrangian :

$$L_m = \frac{1}{2}(L_L^M + L_L^D + L_R^M + L_R^D) + h.c. = \quad (1.5)$$

$$= m_D \bar{\psi}_L \psi_R + m_L \bar{\psi}_L^C \psi_L + m_D \bar{\psi}_R \psi_L + m_R \bar{\psi}_R^C \psi_R \quad (1.6)$$

This can be rewritten in matrix form as

$$L_m \approx \begin{pmatrix} \bar{\psi}_L^C & \bar{\psi}_R \end{pmatrix} \begin{pmatrix} m_L & m_D \\ m_D & m_R \end{pmatrix} \begin{pmatrix} \psi_L \\ \psi_R^C \end{pmatrix} \quad (1.7)$$

When we diagonalize this matrix, the mass terms can be written using m_R, m_L and m_D as

$$m_{1,2} = \frac{1}{2} [(m_L + m_R) \pm \sqrt{(m_L - m_R)^2 + 4m_D^2}] \quad (1.8)$$

Supposing that $m_L = 0$ and $m_D \ll m_R$, the eqn. 1.8 gives

$$m_1 = \frac{m_D^2}{m_R} (\text{mass of } \nu_1 \text{ field}) \quad (1.9)$$

$$m_2 = m_R (1 + \frac{m_D^2}{m_R^2}) \approx m_R (\text{mass of } \nu_2 \text{ field}) \quad (1.10)$$

The extreme light mass of neutrino can be explained by the presence of a suppression factor $\frac{1}{m_R}$. The suppression factor lets neutrino to have a tiny mass if the other neutrino has a large mass. The low mass of active neutrinos is well-explained by this mechanism.

This introduction of neutrino masses in the SM leads to a question about the type of neutrinos. While Dirac particles are distinct from their antiparticles of the same helicity, Majorana particles are essentially their own antiparticles in lepton flavor conserving processes like double beta decay. One way to find the nature of the neutrino masses is by observing reactions like neutrino less double beta decay ($0\nu\beta\beta$) which are unique to Majorana neutrinos.

1.4 Neutrino Sources

Neutrinos being the most abundant massive particle, are produced everywhere, like in the earth, the atmosphere, stars, supernovae, and even during the big Bang. The radioactive elements (like Thorium, Uranium) decay while producing neutrinos. Neutrinos are also produced through interactions between cosmic rays and the

Earth's atmosphere. Nuclear fusion reactions that power the Sun also produce neutrinos. The neutrinos produced during the Big Bang persist to this day, in the energy range of micro electron-volts. We can generate neutrinos in nuclear reactors and accelerators. Supernova explosions are a source of extremely high energy neutrinos. One can distinguish between different sources of neutrinos by their energy ranges. Figure 1.3 shows the predicted neutrino flux plotted against neutrino energy for a range of different neutrino sources [22].

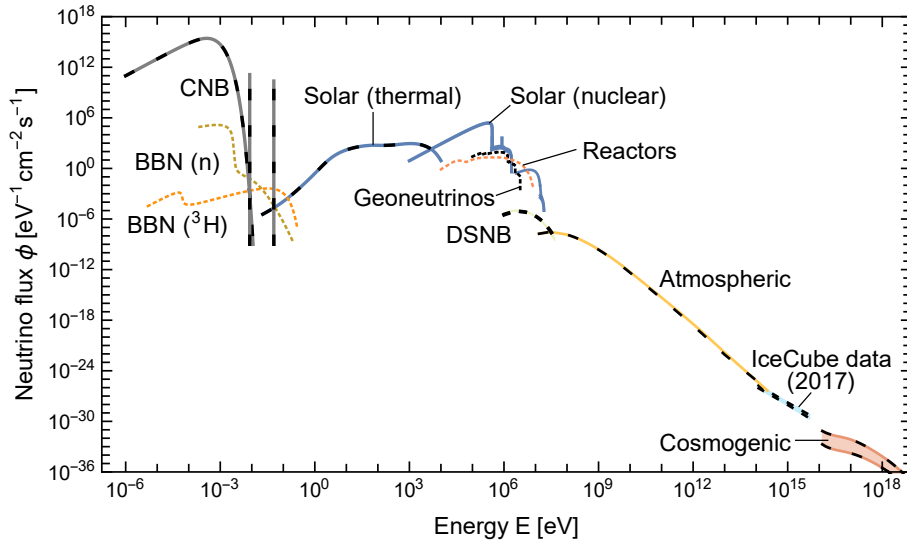


FIGURE 1.3: Predicted neutrino flux from different natural sources [22].

In sun, neutrinos are created abundantly, and these solar neutrinos are the best means of observing the inner workings of the Sun's core directly. Sun produces predominantly electron neutrinos resulting from the fusion processes. We know that the Sun generates energy through two fusion chains - CNO and the proton-proton chains emit neutrinos, as shown in Fig 1.4[23].

Cosmic rays are mostly made up of protons (95%), alpha particles (5%), a small percentage of electrons ($< 1\%$), and heavier nuclei. They constantly bombard the Earth's atmosphere. The collision of primary cosmic rays with the atmospheric nuclei results in a cascade of hadrons, predominantly kaons and pions. These hadrons then produce atmospheric neutrinos as they decay during flight through the atmosphere.

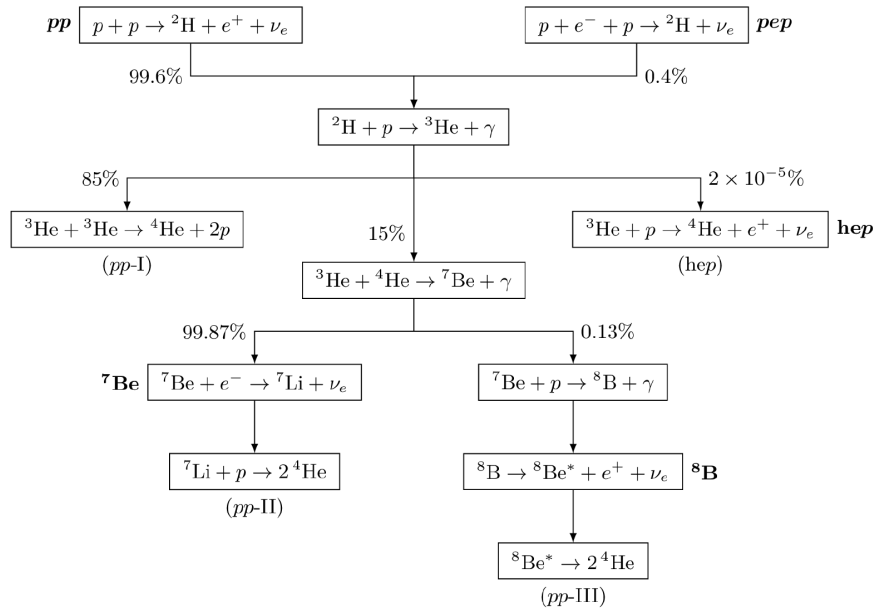


FIGURE 1.4: The proton - proton chain reactions which produce neutrinos in stars, taken from [23]

Artificial neutrino sources like nuclear reactors and accelerators are also available. NOvA is one of those experiments which uses a artificial neutrino beam produced at Fermilab.

1.5 Neutrino interactions

In SM, neutrinos interact with matter only through weak interactions. Because of neutrino's low interaction cross section, their detection is a significant hurdle in experimental particle physics. Even Pauli acknowledged that detecting neutrinos, which are electrically neutral and highly lightweight, would be impossible. Even now, in particle physics experiments, neutrinos are identified by missing energy in known decays or interactions.

It is to be noted that neutrino interactions conserve lepton flavour number. This means that the total lepton number from each generation before and after the interaction should be equal. Each generation of leptons possesses a distinct lepton flavour number. The antiparticles possess a negative lepton number compared to

their corresponding particles. As a result, a neutrino can either interact with a same generation charged lepton or with a neutrino that shares the same flavour.

Neutrino interactions can be broadly classified based on the boson that was exchanged. The interactions mediated through Z^0 boson are called Neutral current (NC) interactions. In this type of interaction, incoming neutrino scatters off the nucleus while maintaining its lepton flavour while going out. NC interactions are identical among all the flavours as Z^0 boson couples with the three flavours equally. A typical representation of NC interaction is shown below:

$$\nu_\alpha + N \rightarrow \nu_\alpha + X \quad (1.11)$$

$$\bar{\nu}_\alpha + N \rightarrow \bar{\nu}_\alpha + X \quad (1.12)$$

Here ν_α ($\bar{\nu}_\alpha$) is incoming neutrino (antineutrino). Here α can be any one of the lepton flavours e, μ or τ . N represents interaction nucleon, while X is a group of hadrons, whose composition depends on the characteristics of the interaction .

The remaining neutrino interactions which mediate through charged $W^\pm$ bosons are called Charge current (CC) interactions. The incoming neutrino scatters off the nucleus and converts into the corresponding lepton of the same generation. A typical representation of CC interaction is shown below:

$$\nu_\alpha + N \rightarrow l_\alpha^- + X \quad (1.13)$$

$$\bar{\nu}_\alpha + N \rightarrow l_\alpha^+ + X \quad (1.14)$$

Here l_α is charged lepton. Here α can be any one of the lepton flavours e, μ or τ . and other notation remains the same.

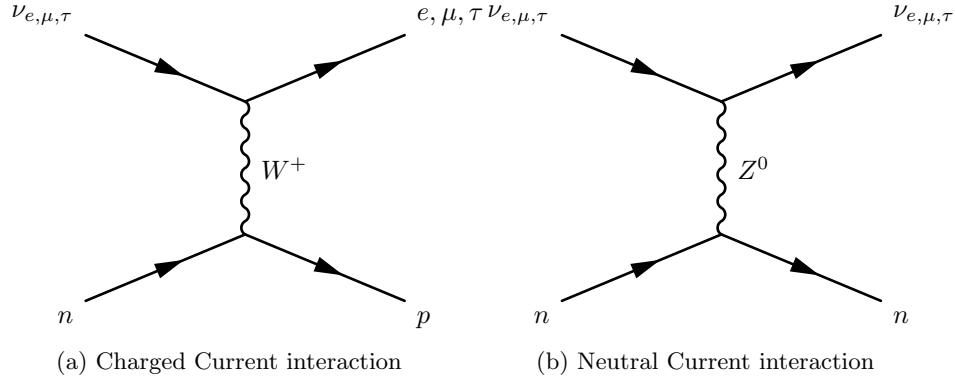


FIGURE 1.5: CC and NC neutrino interactions are depicted through Feynman diagrams.

The neutrino interactions can be further classified with respect to the final nucleus state. The most common interaction types are described below:

- Quasi-Elastic (QE) scattering can happen in both CC and NC processes. The incoming neutrino interacts with nucleons without causing significant change in the nature of nucleons. In CC interactions, the final state consists of a single charged lepton, while in NC interactions case, neutrino is the final state. Both NC and CC interactions have a single nucleon as the final state.
- Resonant Pion Production (RES) process happens at neutrino energies of 2 GeV. The incident neutrino interacts with the nucleon and produces a pion. This process occurs when neutrino sends the nucleus into a resonant baryon state. This resonant state subsequently decays into a pion and the remaining nucleon.
- Deep Inelastic Scattering (DIS) involves an high energy neutrino interacting with the nucleus. The incoming neutrino scatters on a single quark inside a nucleon, displacing it from the nucleus. This typically results in hadronization in the final state. DIS provides valuable insights into the partonic structure of nucleons, revealing information about quarks and gluons within the nucleon when the energy of neutrinos is exceedingly high, typically above 100 GeV. DIS interactions become overwhelmingly dominant and becomes directly proportional to the neutrino energy as seen in fig. 1.6.

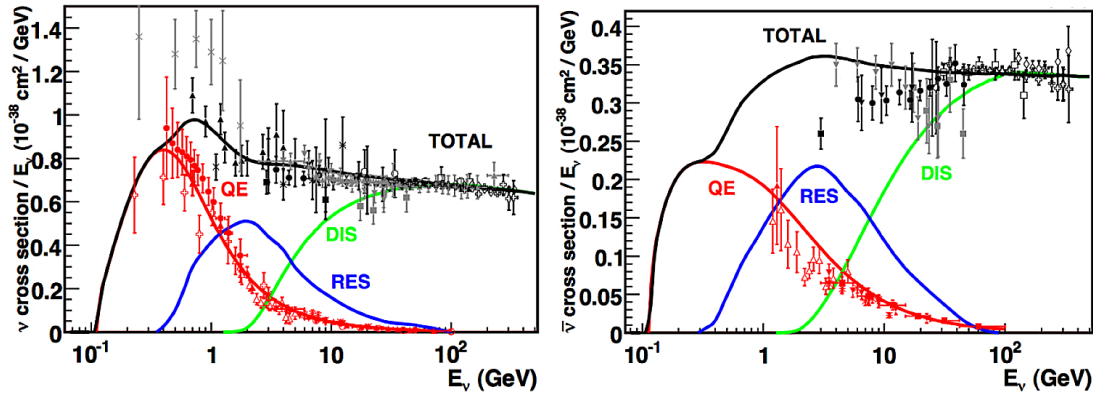


FIGURE 1.6: Neutrino Cross sections. The neutrino(antineutrino)-nucleon cross sections plotted as a function of neutrino(antineutrino) energy are shown in the left (right) panel [24].

In the NOvA energy range (1-4 GeV), Neutrino Cross sections are small for all types of interactions, i.e., about $\sigma \approx 10^{-38} \text{ cm}^2$, as shown in figure 1.6. However, it is important to note that these cross sections do not have good constraints and have large uncertainties of around 20%. These uncertainties are a serious problem to current and future neutrino experiments, as they introduce considerable systematic uncertainties in precision neutrino measurements. Understanding and accounting for this uncertainty is crucial for accurate data analysis and interpretation in neutrino experiments.

1.6 Neutrino Oscillations

Neutrinos exist in a unique flavour state $\langle \nu_\alpha |$ only during their production and when they interact with matter. Each flavour state is a superposition of neutrino mass eigenstates $|\nu_k\rangle$ and vice versa as shown in eqn.1.15.

$$|\nu_\alpha\rangle = \sum_k U_{\alpha k}^* |\nu_k\rangle, \quad |\nu_k\rangle = \sum_\alpha U_{k\alpha} |\nu_\alpha\rangle \quad (1.15)$$

Here α represents the flavour states e, μ and τ , k is the mass eigenstates, and U , a unitary rotation matrix is called the ‘‘Pontecorvo-Maki-Nakagawa-Sakata’’

(PMNS) matrix. The mass and flavour eigenstate basis of neutrinos can be transformed between each other using PMNS matrix as shown in eqn.1.15. The matrix is named after Bruno Pontecorvo, who was the first to propose the idea of neutrino oscillations in 1957 [10]. He proposed neutrino oscillation analogous to neutral kaon mixing. The only viable oscillation was from neutrino to antineutrino, given that only one type of neutrino had been discovered by that time. After discovery of muon neutrino, Pontecorvo developed his initial idea to two flavour neutrino oscillation in 1967 [25]. But two flavour neutrino oscillation was already proposed by Z. Maki, M. Nakagawa, and S. Sakata in 1962 [26]. Thus, matrix also includes the names of these three scientists to honor their roles in solidifying the concept of neutrino oscillations .

$$|\nu_k(t)\rangle = e^{-iE_k t} |\nu_k(0)\rangle = e^{-iE_k t - i\vec{p}_i \cdot \vec{x}_i} |\nu_k(0)\rangle \quad (1.16)$$

$$|\nu_\alpha(t)\rangle = \sum_k e^{-iE_k t} U_{\alpha k}^* |\nu_k\rangle \quad (1.17)$$

Here $e^{-iE_k t}$ represents the time evolution operator corresponding to the free particle Hamiltonian. in natural units.

So, neutrinos travel through space in a superposition of mass eigenstates. We see from eqn.1.17 that the phase of the wave varies by mass eigenstate. This gives rise to neutrino flavour mixing. This means there is a possibility that a neutrino, initially created as one flavour eigenstate, will then interact with matter as a different flavour eigenstate. As this mixing has a periodicity, it is referred to as neutrino oscillation. It is this oscillatory behaviour that allows for the intriguing and significant possibility of neutrinos changing their flavours as they travel. Eqn.1.17 can be further simplified because the neutrinos are ultra relativistic particles, i.e. $|\vec{p}_i| \gg m_i$. In this ultra relativistic region, the energy term can be simplified using Taylor expansion while assuming $p_i = p_k = p$

$$E = \sqrt{p^2 + m_i^2} \approx p + \frac{m_i^2}{2p} \approx E + \frac{m_i^2}{2E} \quad (1.18)$$

Using this approximation and natural units, where ($t \approx L$), the full simplification of eqn.1.16 becomes

$$|\nu_k(t)\rangle = e^{-iL(E-p)} |\nu_k(0)\rangle = e^{-i(\frac{m_i^2 L}{2E})} |\nu_k(0)\rangle \quad (1.19)$$

following eqn.1.19, eqn.1.17 becomes

$$|\nu_\alpha(t)\rangle = \sum_k e^{-i(\frac{m_i^2 L}{2E})} U_{\alpha k}^* |\nu_k\rangle \quad (1.20)$$

This form describes the evolution of neutrino flavour eigenstates with respect to both distance traveled and neutrino energy for neutrino traveling through a vacuum. The probability of flavour oscillation between ν_α to ν_β can be calculated by taking a projection between the initial and propagated neutrino flavour state and squaring it.

$$P_{\alpha \rightarrow \beta} = |\langle \nu_\beta(L) | \nu_\alpha(0) \rangle|^2 \quad (1.21)$$

$$= \left| \sum_i U_{\alpha i}^* U_{\beta i} e^{-i(\frac{m_i^2 L}{2E})} \right|^2 \quad (1.22)$$

Because U is a unitary matrix, eqn. 1.22 can be expanded as shown.

$$\begin{aligned} P_{\alpha \rightarrow \beta} = \delta_{\alpha\beta} & - 4 \sum_{i>j} \Re[U_{\alpha i}^* U_{\alpha j} U_{\beta i} U_{\beta j}^*] \sin^2\left(\frac{\Delta m_{ij}^2 L}{4E}\right) \\ & + 2 \sum_{i>j} \Im[U_{\alpha i}^* U_{\alpha j} U_{\beta i} U_{\beta j}^*] \sin^2\left(\frac{\Delta m_{ij}^2 L}{4E}\right) \end{aligned} \quad (1.23)$$

We can see that neutrino oscillation probability is dependent on parameters like mass squared differences ($\Delta m_{ij}^2 = (m_i^2 - m_j^2)$), distance traveled (L) and neutrino energy (E). We must note that neutrino oscillations are sensitive to differences in masses eigenvalues but not the absolute neutrino mass. Here, it is seen that neutrino oscillations need neutrinos to possess no zero mass, which goes against the standard model, which states that neutrinos are massless. In fact, neutrinos are the only known particles which break the standard model. This makes the study of neutrinos and their oscillations very interesting.

While studying neutrino oscillations, expressing the PMNS matrix in terms of the mixing angles proves to be beneficial, as shown.

$$U = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \quad (1.24)$$

Where $c_{ij} = \cos \theta_{ij}$ and $s_{ij} = \sin \theta_{ij}$, δ_{CP} is the CP-violating phase, which as the name suggests behaves differently for neutrinos and antineutrinos. The PMNS matrix can be further expanded by separating mixing angle parameters.

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{CP}} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta_{CP}} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (1.25)$$

The expanded form is helpful as it splits the PMNS matrix into three components, each depending on the three mixing angles (θ_{12} , θ_{13} and θ_{23}) that drive the oscillations in corresponding oscillation channels.

1.6.1 Neutrino oscillations in matter

The neutrino oscillation probabilities mentioned in eqn.1.24 are derived for neutrino propagation in a vacuum. Although neutrinos have low interaction cross sections, their propagation is affected when they travel through high density matter. This is called the “Mikheyev-Smirnov-Wolfenstein” effect or the matter effect [27, 28]. Coherent forward scattering takes place during the neutrino propagation through matter. A subset of these are NC interactions which are equally likely for all neutrino flavours. Therefore, they have no net effect on the oscillation probabilities. Only $\nu_e(\bar{\nu}_e)$ can interact in remaining CC interactions as the electron is the only lepton present in the matter. Similarly, ν_μ and ν_τ can not engage in the CC scattering as there is a lack of corresponding charged leptons. This unique property of the ν_e requires us to add additional potential, which influences only ν_e part while traveling through the matter. This additional potential term can be understood as a variation in the effective mass of the neutrinos. So, Any neutrino oscillations involving ν_e must be adjusted according to effective mass change, when propagating through matter. This additional potential from the matter effect must be added to the vacuum Hamiltonian.

The derivation of matter effect for two neutrino states case is shown below. The mixing matrix for two neutrino case can be parameterized using a single mixing angle as:

$$U = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \quad (1.26)$$

Adding additional potential from matter effect to the vacuum Hamiltonian:

$$H_M = U \begin{pmatrix} \frac{m_1^2}{2E} & 0 \\ 0 & \frac{m_2^2}{2E} \end{pmatrix} U^\dagger + \begin{pmatrix} \pm\sqrt{2}G_F N_e & 0 \\ 0 & 0 \end{pmatrix} \quad (1.27)$$

$$= \begin{pmatrix} \frac{m_1^2}{2E} \pm \sqrt{2}G_F N_e \cos^2 \theta & \pm\sqrt{2}G_F N_e \cos \theta \sin \theta \\ \pm\sqrt{2}G_F N_e \cos \theta \sin \theta & \frac{m_2^2}{2E} \pm \sqrt{2}G_F N_e \sin^2 \theta \end{pmatrix} \quad (1.28)$$

Here, N_e represents the electron number density within the medium and G_F denotes the Fermi constant. Neutrinos (+) and antineutrinos (-) are differentiated by the sign of the potential term. The modified Hamiltonian is then re-diagonalized, in order to get the new neutrino mixing angles (θ_M) and mass eigenstates (m_M) [23].

$$\Delta m_M^2 = \sqrt{(\Delta m^2 \sin 2\theta)^2 + (\Delta m^2 \cos 2\theta \mp \sqrt{2}G_F N_e)^2} \quad (1.29)$$

$$\tan 2\theta_M = \frac{\tan 2\theta}{1 \mp \frac{\sqrt{2}G_F N_e}{\Delta m^2 \cos 2\theta}} \quad (1.30)$$

The vacuum oscillation probability for two neutrino states is:

$$P_{\alpha \rightarrow \beta} = \sin^2 2\theta \sin^2 \left(\frac{\Delta m^2 L}{4E} \right) \quad (1.31)$$

Now, neutrino oscillations in matter is given by the same equation by replacing $\theta \rightarrow \theta_M$ and $\Delta m^2 \rightarrow \Delta m_M^2$.

The same process can be repeated for three flavour neutrino case by adding additional potential as shown in eqn.1.32.

$$H_M = U \begin{pmatrix} \frac{m_1^2}{2E} & 0 & 0 \\ 0 & \frac{m_2^2}{2E} & 0 \\ 0 & 0 & \frac{m_3^2}{2E} \end{pmatrix} U^\dagger + \begin{pmatrix} \pm\sqrt{2}G_F N_e & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (1.32)$$

The diagonalization for three flavour case is quite complex to show in this thesis. The influence of matter effects on neutrino oscillations varies with the parameters values as seen in eqn.1.24. The matter effects affect the oscillation probabilities of neutrinos and antineutrinos in opposite ways due to the sign difference in eqn.1.30.

1.7 Neutrino experiments

Neutrino experiments are mainly categorized based on three factors: the baseline over which neutrinos travel, the source of neutrinos and the energy of neutrinos. A crucial experimental property that decides the sensitivity of an experiment to specific oscillation parameters is the ratio of the baseline (L) of the experiment to the neutrino energy (E), typically expressed in units of km/GeV as seen in eqn.1.24. There are reactor, solar and accelerator or atmospheric neutrino experiments, each sensitive to different mixing angles.

Reactor experiments are specifically designed to be sensitive towards the disappearance of $\bar{\nu}_e$ s produced through the β -decay process within nuclear reactors. These experiments are sensitive primarily to θ_{13} . For neutrino sector CP violation to exist, we require all three mixing angles to have non zero values. Accurately determining the value of θ_{13} is significant because it was the final mixing angle to be precisely measured. Reactor experiments typically operate at relatively short baselines, in the order of approximately 1 km ($L \approx 1$ km). The energies of the

detected neutrinos in reactor experiments are typically in the range of 1 to 10 MeV. The most noteworthy reactor experiments are Daya Bay [29], Double Chooz [30], and RENO [31].

The electron neutrinos produced within the Sun are the main neutrino source for solar neutrino experiments. The detectable solar neutrinos typically have energies within the range of 0.2–15 MeV, and the propagation length is 1.5×10^{11} km. From the range of L/E , we determine that the solar neutrino experiments are sensitive to $\Delta m^2 \gtrsim 10^{-12} \text{eV}^2$. These experiments are sensitive to θ_{12} . Some noteworthy solar neutrino experiments are Kamiokande [32], SAGE [33], GALLEX [34] and SNO [35] experiment. The SNO experiment played a pivotal role in resolving the solar neutrino problem. It was capable of detecting both the CC interactions, specific to electron neutrinos, and the NC interactions, which could involve neutrinos of any flavour, this provided crucial information about the total neutrino flux reaching the detector.

Atmospheric neutrino experiments are designed for measuring neutrino flux produced through the decay of pions and kaons created through the interactions of cosmic rays with the Earth's atmosphere. The distance between the neutrino source and the detector in these experiments vary based on neutrino direction. For neutrinos coming from above, the baseline is typically around 10 kilometres ($L \approx 10$ km). In contrast, for neutrinos that have traversed the Earth, the baseline can be on the order of thousands of kilometres ($L \approx 10^4$ km). Super-Kamiokande (SuperK) experiment [17, 36] in Japan was the most noteworthy atmospheric neutrino experiment. Atmospheric experiments are primarily sensitive to Δm_{32}^2 and θ_{23} .

Accelerator experiments are also sensitive to θ_{23} and Δm_{32}^2 . They use a well-controlled beam of neutrinos produced using particle accelerators, offering more flexibility and precision in terms of neutrino baseline, energy, and flavour composition compared to other experimental methods. The neutrino beam energy range

typically varies from 10^{-1} to 10 GeV. These experiments are categorized into two types based on the baseline. Experiments with a baseline of $L \approx 1$ km are called Short baseline (SBL) experiments, whereas experiments ranging from 10^2 to 10^3 km are called long baseline (LBL) experiments. The SBL experiments are configured to detect high-frequency neutrino oscillations characterized by large valued Δm^2 . Prominent examples of short baseline experiments include LSND (Liquid Scintillator Neutrino Detector) [37] and the upcoming SBN (Short-Baseline Neutrino) [38] program. LBL experiments have L/E ratio, this makes them sensitive to atmospheric oscillation parameters (θ_{23} and Δm_{32}^2). One of the currently running LBL experiment is NOvA, whose detailed description is provided in chapter 2. Other noteworthy experiments include MINOS [39], T2K [40], and DUNE [41] experiment.

1.8 Oscillation Parameters - Current status

Today we have accurate measurements of most of the oscillation parameters. The mass squared difference and the three mixing angles have been measured up to 10% precision. The current global fit values of oscillation parameters are shown in Table 1.1.

Oscillation parameters	Best Fit Value
$\sin^2 \theta_{12}$	$0.303^{+0.012}_{-0.012}$
$\sin^2 \theta_{13}$	$0.02225^{+0.00056}_{-0.00059}$
$\sin^2 \theta_{23}$	$0.451^{+0.019}_{-0.016}$ (NH)
$\sin^2 \theta_{23}$	$0.569^{+0.016}_{-0.021}$ (IH)
Δm_{21}^2	$7.41^{+0.21}_{-0.20} \times 10^{-5} \text{ eV}^2$
Δm_{32}^2	$+2.507^{+0.026}_{-0.027} \times 10^{-3} \text{ eV}^2$ (NH)
Δm_{32}^2	$-2.486^{+0.025}_{-0.028} \times 10^{-3} \text{ eV}^2$ (IH)

TABLE 1.1: Current global fit of three-flavor neutrino oscillations [42].

By convention, the mass squared difference Δm_{12}^2 is defined to be positive. However, the sign of Δm_{32}^2 , is currently unknown. There are two possible scenarios: When the sign of Δm_{32}^2 is positive, it is Normal Hierarchy (NH) and the Inverted

Hierarchy (IH) is when the sign is negative. Both are shown in fig.1.7. In the Normal Hierarchy case, the mass eigenstates have $m_1 < m_2 < m_3$ order. Conversely, in case of Inverted Hierarchy, the mass eigenstates are arranged in the opposite order ($m_3 < m_2 < m_1$). The naming of the NH and IH is based on the analogy with the masses of charged leptons, where $m_\tau > m_\mu > m_e$.

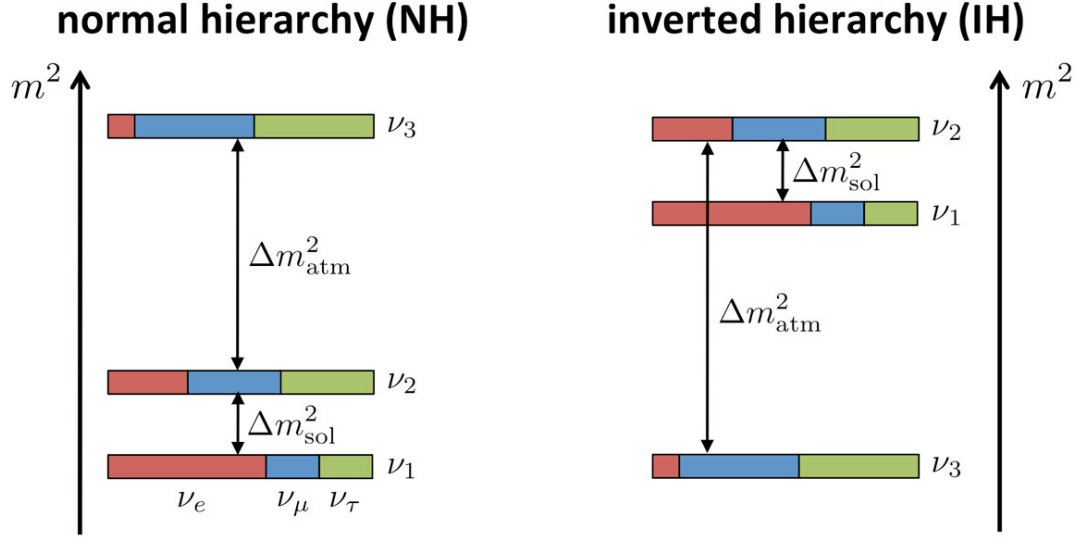


FIGURE 1.7: The two hierarchies of the neutrino mass states. On the left, the normal ordering ($m_1 < m_2 < m_3$) is shown, and the inverted ordering ($m_3 < m_2 < m_1$) is on the right.

The δ_{CP} represents a phase associated with in the neutrino sector. Existing evidence from T2K [43] suggests that values of 0 or π (CP conserving) can be ruled out to some extent, indicating the presence of CP violation in neutrino oscillations. Due to limited experimental measurements, the precise value of δ_{CP} is currently not well determined. However, the best fit value of δ_{CP} from NOvA (NH: 0.82 rad, IH: 1.52 rad) [44] and T2K (NH: -1.97 rad, IH: -1.44 rad) [45] are found to be significantly different. In addition, 1σ allowed regions on the $\sin^2 \theta_{23} - \delta_{CP}$ planes for NOvA and T2K show almost zero overlap for NH case [44]. However in IH case, allowed regions of both experiments coincide. NOvA-T2K Joint Analysis is under progress to address this tension between results of NOvA and T2K. While experimental efforts are ongoing to precisely measure δ_{CP} and improve our understanding of CP violation in the neutrino sector. Future LBL experiment DUNE

which has been designed to be sensitive CP phase, might provide us an accurate value of δ_{CP} [41].

1.9 Outline of Thesis

Chapter 2 of this thesis gives a detailed description of NOvA experiment. It goes through the detector technology used and efforts taken to reduce systematic uncertainties through detector design. We go over Data acquisition system and timing system of NOvA experiment briefly. In Chapter 3, We go over various simulation techniques and software used by NOvA to simulate its detector response, event rates of signal and background. The step by step process of event reconstruction from raw data is also explained. In chapter 4, We go through the application of machine learning in NOvA and how prong CNN which can classify the final particles of neutrino event was developed and validated. Chapter 5 describes the framework for 3+1 sterile neutrino model. We describe does a phenomenological study in which we check the effect of sterile neutrino on degeneracy resolution capabilities of NOvA and DUNE experiments. Finally, the summary and conclusions of the entire thesis work and the future scope of research in this direction are discussed in chapter 6.

Chapter 2

The NOvA Experiment

The NuMI beam Off-Axis Electron neutrino (ν_e) Appearance (NOvA) experiment is a two detector, long baseline neutrino oscillation experiment. The NOvA collaboration comprises of about 266 scientists and engineers from 49 universities spread across eight countries. The NOvA collaboration seeks to make precise measurements of neutrino oscillation parameters like δ_{CP} , θ_{23} , and sign of Δm_{32}^2 . NovA uses an accelerator produced neutrino beam, NuMI (“Neutrinos at the Main Injector”) beam. NuMI beam predominantly contains ν_μ or $\bar{\nu}_\mu$ based on the beam mode. The two detectors are functionally identical and positioned 14.6 mrad off-axis to the neutrino beam trajectory. This off-axis positioning narrows the neutrino energy spectrum to the energy range we are interested in, thereby reducing background. The NOvA Near Detector is 100 meters underground at Fermilab, roughly 1 km from the neutrino source, while the NOvA Far Detector (FD) is on the surface at Ash River, Minnesota, ~ 810 km from away from the neutrino source. The baseline is selected such that the first maxima of the $\nu_\mu \rightarrow \nu_e$ appearance oscillation probability is at the Far detector. Similarly, the detector design was chosen to maximize the detection of ν_e charged current events.

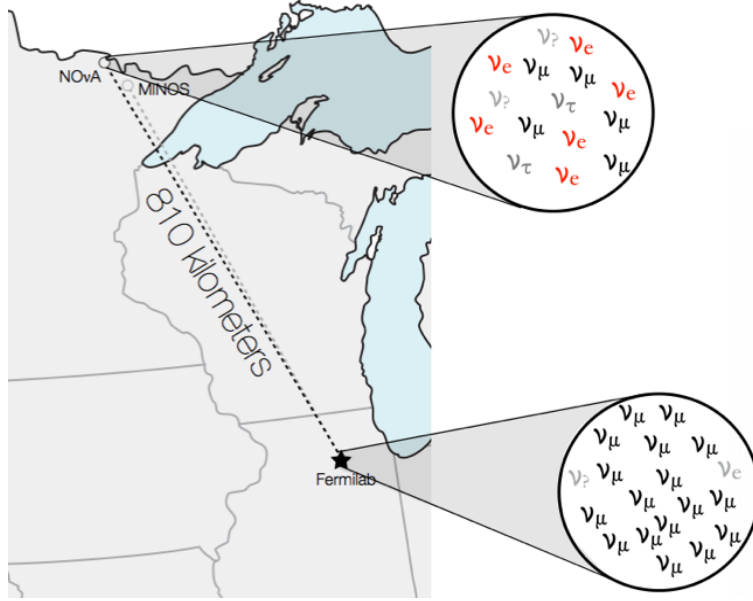


FIGURE 2.1: Map depicting the off-axis beamline of NOvA originating at Fermilab. Courtesy of [46].

2.1 The NuMI Beam

The NuMI beam available at the Fermilab is used as a neutrino source for NOvA. The NuMI beam was initially developed to serve the objectives of the MINOS experiment[47], which ran from 2005 to 2016. As seen in Fig. 2.2, generation of a neutrino beam involves multiple steps. A proton beam is produced and accelerated up to 120 GeV. Then, graphite target is collided with this 120 GeV proton beam. In this stage, secondary mesons are produced and then focused using a magnetic horn. Then, they are allowed to decay in flight to produce neutrinos. This Section provides detailed description of how a neutrino beam is produced at FermiLab.

The beam production process starts from the ion sources at the accelerator complex. Here, hydrogen gas is ionized to create negative hydrogen ions. These hydrogen ions (H^-) are then successively accelerated by radio-frequency (RF) quadrupoles, first to 35 KeV and then to 750 KeV. From there, hydrogen ions are sent to the linear accelerator (LINAC). In LINAC, the hydrogen ions are accelerated to 400 MeV by using RF cavities in addition to drift tubes. The hydrogen ions (H^-) are made to pass down a carbon foil, which is attached to the end of the LINAC.

Fermilab Accelerator Complex

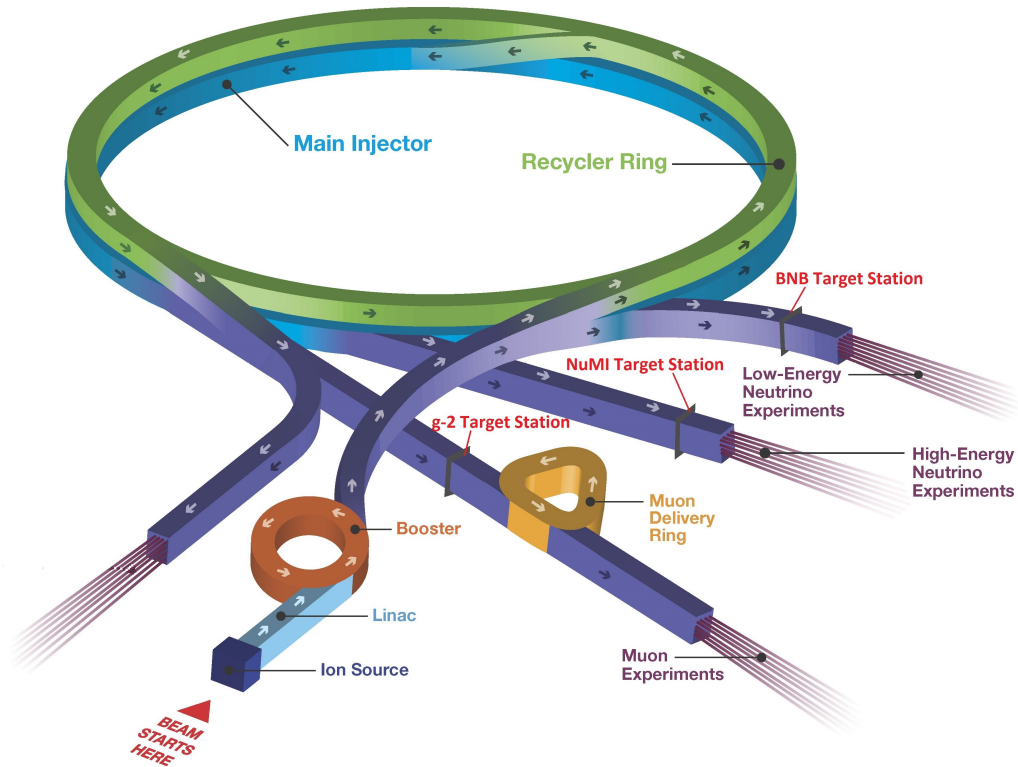


FIGURE 2.2: Overview of the Fermilab accelerator complex in schematic form[48].

The two electrons from H^- are stripped by carbon foil, producing a proton beam. These protons are then packed into packets called bunches, each containing 4×10^{12} protons. Subsequently, these bunches undergo acceleration to reach 8 GeV in the Booster ring. The Booster ring has the capacity to accommodate 84 bunches (1 batch) of protons simultaneously. These proton batches are sent into the Recycler. The original role of the Recycler was to store and cool anti-protons. But now, it is used to pass booster ring protons into the main injector. This is accomplished by a technique known as slipstacking, which by superimposing pairs of bunches, essentially enables the doubling of the beam intensity. Slipstacking is accomplished by injecting two proton batches consecutively while slightly slowing down the first batch. They are combined into one batch with twice as many protons when the two proton batches overlap. These higher intensity batches are then passed to the main injector, here these batches of protons undergo acceleration up to 120

GeV. These main injector proton batches are grouped into spills. A spill is a 12 proton bunches provided in a ten microsecond interval. A spill is supplied approximately once every second. These Protons are then supplied to other experiments as needed, either for direct use or to create other particles like neutrinos.

For the NOvA case, These 120 GeV protons are made to collide with the target. The total beam exposure of experiments in the NuMI beam is calculated by the total number of protons bombarded with the target (Protons on target (POT)). These high energy proton-carbon interactions produce a great variety of secondary particles. Of these secondary particles, the charged mesons (kaons($K^\pm$), pions($\pi^\pm$)) are most abundant. These charged mesons are then diverted to form a focused using magnetic horn pair positioned directly behind the graphite target. We can choose which type of charged mesons to focus on by swapping the direction of current in magnetic horns. The Forward Horn Current (FHC) mode generates magnetic fields that focus the positively charged mesons along the beamline's direction while deflecting negatively charged mesons away from the beamline and inverting the current direction results in the Reverse Horn Current (RHC) mode. In RHC mode, negatively charged mesons are focused opposed to FHC mode. These focused mesons enter into the decay pipe, which has a diameter of two meters and a length of 675 meters. To stop the leftover hadronic components of the beam which travel beyond the decay pipe, an absorber is placed. A series of muon detectors are interspersed between rocks to monitor beam quality, as shown in figure 2.3. More specific details on the NuMI beamline can be found at [49].

As mentioned earlier, The NOvA detectors are positioned 14.6 mrad off-axis to the NuMI neutrino beam. This was done to get a narrow neutrino energy spectrum at the detectors. The logic behind this can be explained by observing the equations 2.1, 2.2 which give the energy and the flux of neutrinos produced from meson decay after traveling z distance in the detector of area A respectively.

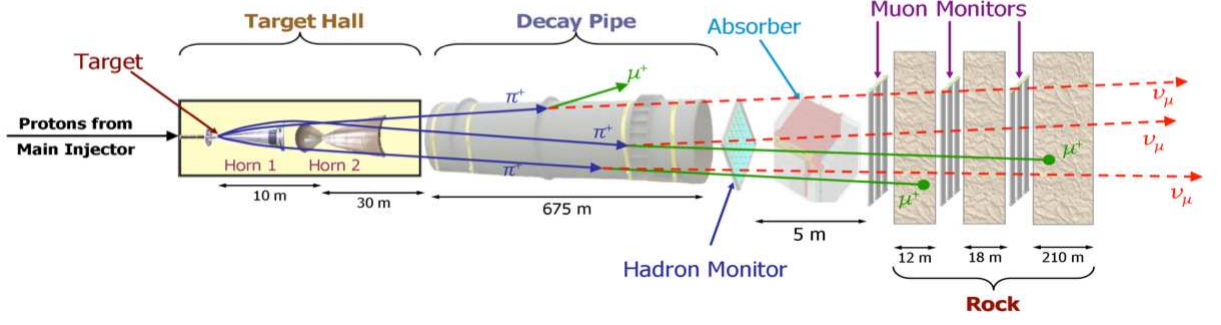


FIGURE 2.3: A Schematic illustrating the production of a neutrino beam for NOvA. It depicts 120 GeV protons colliding the target on the left side, while produced neutrinos exit to the right [49].

$$E_\nu = E_{(\pi,K)} \left(\frac{1 - \frac{m_\mu^2}{m_{(\pi,K)}^2}}{1 + \gamma^2 \theta^2} \right) \quad (2.1)$$

$$\phi = \left(\frac{2\gamma}{1 + \gamma^2 \theta^2} \right)^2 \frac{A}{4\pi z^2} \quad (2.2)$$

Here, θ represents the angle between the outgoing neutrino and the decaying meson trajectory and $\gamma = \frac{E_{(\pi,K)}}{m_{(\pi,K)}}$. Here, we see that the neutrino energy spectrum gets narrow as the angle between meson and neutrino increases. This behaviour is plotted in figure 2.4 [48]. With the increase in the off-axis angle, the neutrino flux decreases. Thus, a choice must be made by weighing between a narrower energy spectrum and preferably high neutrino flux. NOvA was designed off-axis, taking this into account to constrain the neutrino energy spectrum with peak energy at 2 GeV.

Narrowing the neutrino flux reduces the background to the ν_e appearance signal. The Neutral Current (NC) events are the most notable backgrounds. The NC event where neutrino scatters off the nucleus has an event topology similar to that of ν_e CC events. Most of these NC background events are observed in the lower energy ranges, as the neutrino leaves the detector without depositing all of its energy. However, the narrow spectrum ensures that energies of most of these background events stay outside the signal region. The intrinsic ν_e present

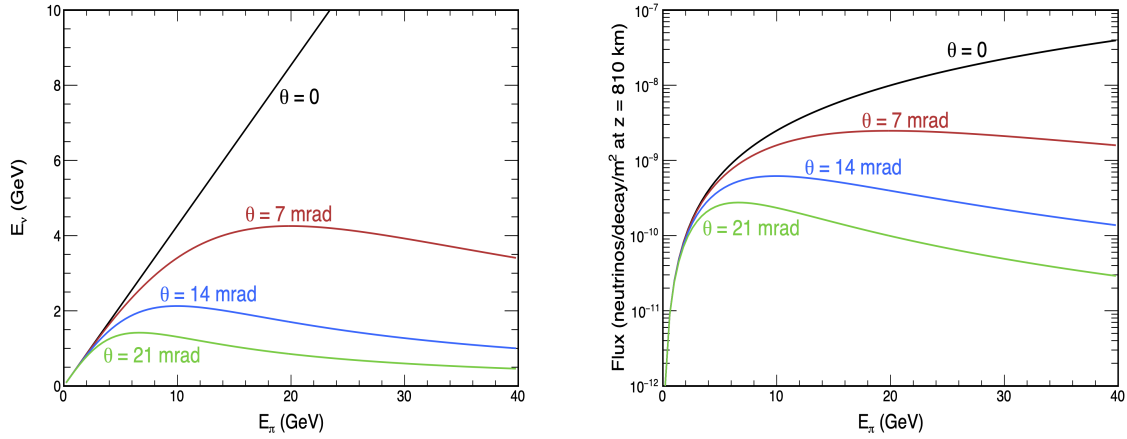


FIGURE 2.4: The neutrino energy(left) and neutrino flux(right) as a function of pion energy. These are shown for various angles of θ in relation to the pion trajectory. [48].

in the beam is another source of background. The intrinsic ν_e events have a wider spectrum than the normal ν_e appeared thorough oscillation. This shape difference in the energy spectrum, can be used to reject a considerable portion of background events.

The charged mesons that decay in the decay pipe produce the neutrinos we seek. The prominent channels of meson decays and their branching ratio are shown in the table 2.1.

Decay Channels	Branching Ratio(%)
$\mu^\pm \rightarrow e^\pm + \nu_e(\bar{\nu}_e) + \bar{\nu}_\mu(\nu_\mu)$	100.0
$K^\pm \rightarrow \mu^\pm + \nu_\mu(\bar{\nu}_\mu)$	63.55
$K^\pm \rightarrow \pi^0 + e^\pm + \nu_e(\bar{\nu}_e)$	5.07
$K^\pm \rightarrow \pi^0 + \mu^\pm + \nu_\mu(\bar{\nu}_\mu)$	3.35
$\pi^\pm \rightarrow \mu^\pm + \nu_\mu(\bar{\nu}_\mu)$	99.98
$\pi^\pm \rightarrow e^\pm + \nu_e(\bar{\nu}_e)$	0.02

TABLE 2.1: Prominent decay modes of charged mesons to neutrinos and their branching fractions.

As we can see from the table 2.1, muon neutrinos dominate the beam's composition in FHC mode. In FHC mode, the contamination from $\bar{\nu}_\mu$ stands at approximately 1.7%, while background from the intrinsic ν_e and $\bar{\nu}_e$ in the FHC beam is around 0.6%. The muon antineutrinos are the predominant flavour in RHC mode and the contamination from ν_μ is about 11.3%. This substantial increase in wrong sign

contamination in RHC mode primarily arises because the antineutrinos have relatively smaller cross section compared to neutrinos. The expected decomposition of neutrino (antineutrino) beam simulated assuming no oscillations is shown in figure 2.5.

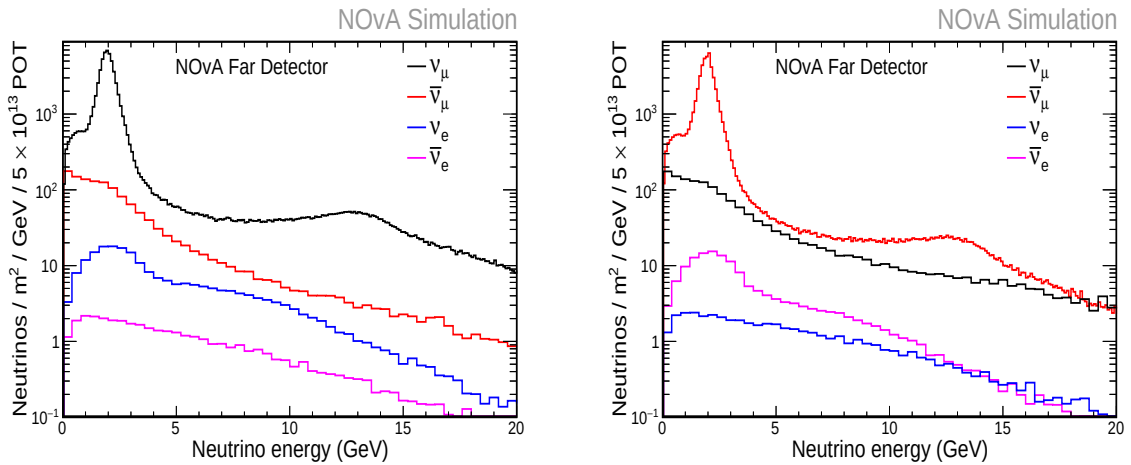


FIGURE 2.5: Expected Neutrino events at FD in FHC mode are shown in the left plot and right plot is for RHC mode. The event rates shown here are corrected for neutrino cross sections in the detector and the flux of the beam, without applying oscillations [50].

2.2 NOvA Detectors

In most LBL neutrino experiments, a common approach involves constructing a minimum of two detectors: one positioned near the neutrino source and the other positioned at a considerable distance, often at the first or second oscillation maxima. Both of these detectors are generally devised to have matching functionality, i.e., similar materials and electronics. Two identical detectors help nullify the effects of their correlated systematic uncertainties, as these effects tend to cancel each other out. NOvA uses the same idea.

The Near Detector (ND) of the NOvA experiment is located 105 meters underground within the MINOS underground area at Fermilab. This underground location of ND significantly reduces the background rate of cosmic muons. The various studies that can be done using ND, include conducting cross section analyses, analyzing the neutrino flux directed towards the far detector and short base

oscillations for new physics. The NuMI beam is angled by 58 milliradians (3.3°) downward through the Earth to allow our beam to travel from the ND to the FD. To make sure that muons of high energy are contained within the detector, a “muon catcher” consisting of 10 steel plates interlaced with detector planes is added at the end of the ND. This muon catcher section is about two-thirds the height (2.5 m) of the other sections of ND. This height limit was imposed by the dimensions of steel plates at hand during the construction of ND. This height difference has minimal effect on muon containment. Since, the NuMI beam is tilted by 3.3° , the neutrinos coming through the top front of the detector will come to a height of 3.1 m when they reach the border of the main detector. Thus, only a small portion of muons will escape muon catcher. Finally, ND has dimensions of $3.8 \text{ m} \times 3.8 \text{ m} \times 16 \text{ m}$, weighing 300 tons.

The Far Detector (FD) is located at Ash River in Minnesota, 810 kilometres from Fermilab. It is relatively larger and heavier than the Near Detector, weighing up to 14 kilotonnes with dimensions $15.8 \text{ m} \times 15.8 \text{ m} \times 60 \text{ m}$. This increased size is essential as we get lower neutrino flux at a larger distance from the NuMI Beam. The FD is situated on the surface with a rock overburden of 3 meters water equivalent (mwe). The positioning of FD makes it susceptible to background in the form of cosmic particles at a frequency of $\sim 130 \text{ kHz}$. The limited rock overburden acts as a shield, providing approximately 12 radiation lengths¹ of protection. This shielding helps diminish the influx of cosmogenic photons, which might otherwise mimic a ν_e signal. The primary reduction of cosmic background is achieved by utilizing the NuMI spill window of 10 milliseconds, which imposes a strict timing cut as neutrinos are only delivered from the beam to the detectors during this brief window. Additional rejection cuts can be applied based on the distinct experimental signatures of down going cosmic muons compared to neutrino events. A scale models of FD and ND detectors are shown in figure 2.6.

¹The radiation length is the mean length, transversed by a photon before undergoing pair production.

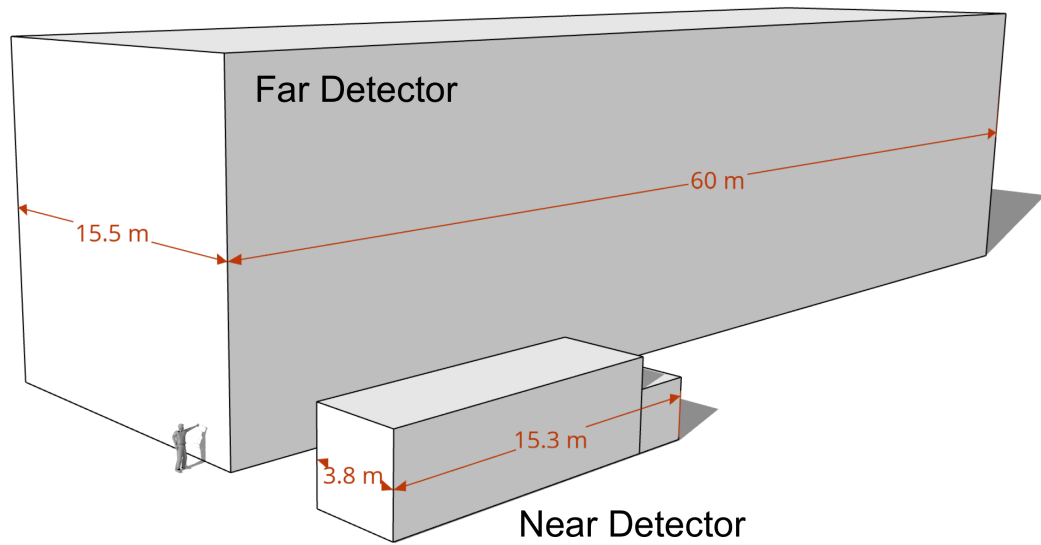


FIGURE 2.6: Model of Far and Near Detectors of NOvA to scale[48].

The building block of both detectors is a cell. Each cell is an extruded PVC tube containing liquid scintillation material. The actual cells are 3.9 cm by 6.6 cm and have walls that range in thickness from 2 to 5 mm, depending on the necessity for structural stability. The length of the extruded cells for the FD (ND) is 15.8 m (3.8 m). The PVC cells that serve as the detector's skeleton must be sturdy enough to bear the entire detector's weight. The corners of the cells are rounded to increase structural stability. Charged particles in the liquid scintillator produce scintillation light, which is the primary particle detection mechanism.

The solution contains 94.6% mineral oil, 5.2% pseudocumene which act as scintillator. Additionally, it contains 0.1% PPO, 0.002% bis-MSB for wavelength shifting and about 0.001% stadis-425 for anti-static protection, and 0.001% vitamin-E used as antioxidant by mass [51]. The solution produces light in the ultraviolet spectrum and subsequently transforms it into violet light, which is visible and ranges in wavelength from 400 to 450 nm. In addition, the PVC surface is coated with TiO_2 to improve the reflectivity of the light being emitted. A single NOvA cell can be seen in figure 2.7

Each cell has a wavelength-shifting (WLS) fiber loop with two ends at the open

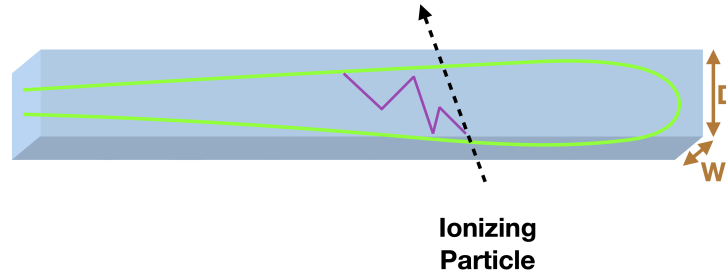


FIGURE 2.7: A single NOvA cell with wavelength-shifting fiber looped inside it[52].

end of the cell. The diameter of a WLS fiber is 0.7 mm. WLS fiber absorbs the light emitted and changes its wavelength to the blue-green range (450–600 nm). This light travels along the length of the fiber through multiple internal reflections. This is not a complete lossless process, especially in FD case with fiber length greater than 28 m, significant attenuation occurs. This attenuation of the scintillation light results in primarily green light surviving to the readout. The two ends of the WLS fiber are then passed into two photodetectors on an Avalanche Photodiode (APD).

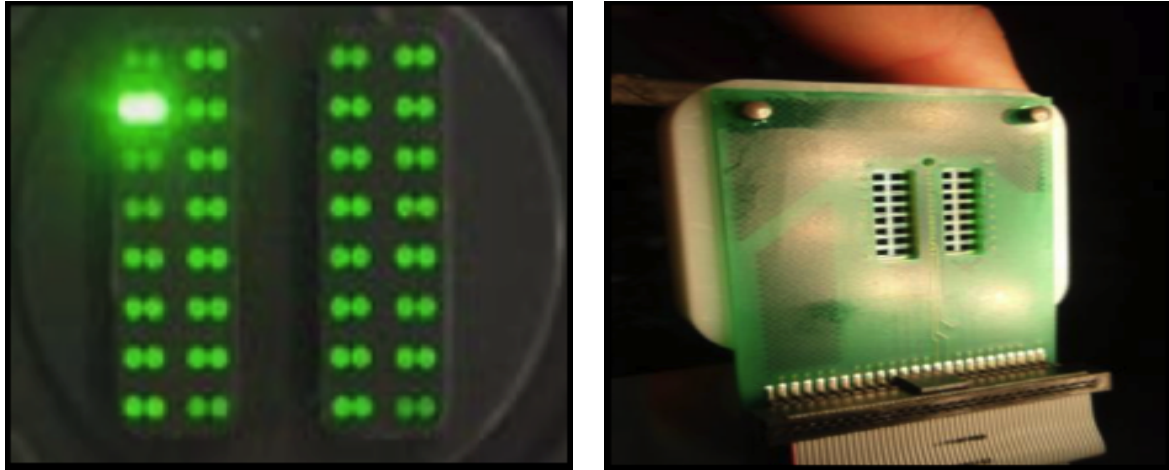


FIGURE 2.8: Left: The end points of WLS fibers from each cell, from total 32 cells. Right: The front face of an APD with 32 inputs. This will be pushed against the fiber ends[52].

The APD used in NOvA has 32 inputs, as seen in fig. 2.8, where each cell forms a pixel. So, 32 cells of NOvA are combined as a single module. A single plane of the detectors is created by aligning multiple modules. Then, the planes are laid in a manner where the cell lengths alternate between vertical and horizontal orientations, thus generating two distinct detector views. Combining the side and

top views allows for the generation of three-dimensional events. Figure 2.9 shows the 3D schematic of the NOvA Detector with cells as its building block. Lastly, the ND is constructed to have 20192 cells grouped into 192 planes, where each plane is comprised of 94 cells. The FD is constructed to have 344064 cells grouped into 896 planes, where each plane contains 384 cells. To keep track of the large number of planes, 64 planes in FD are grouped as a diblock.

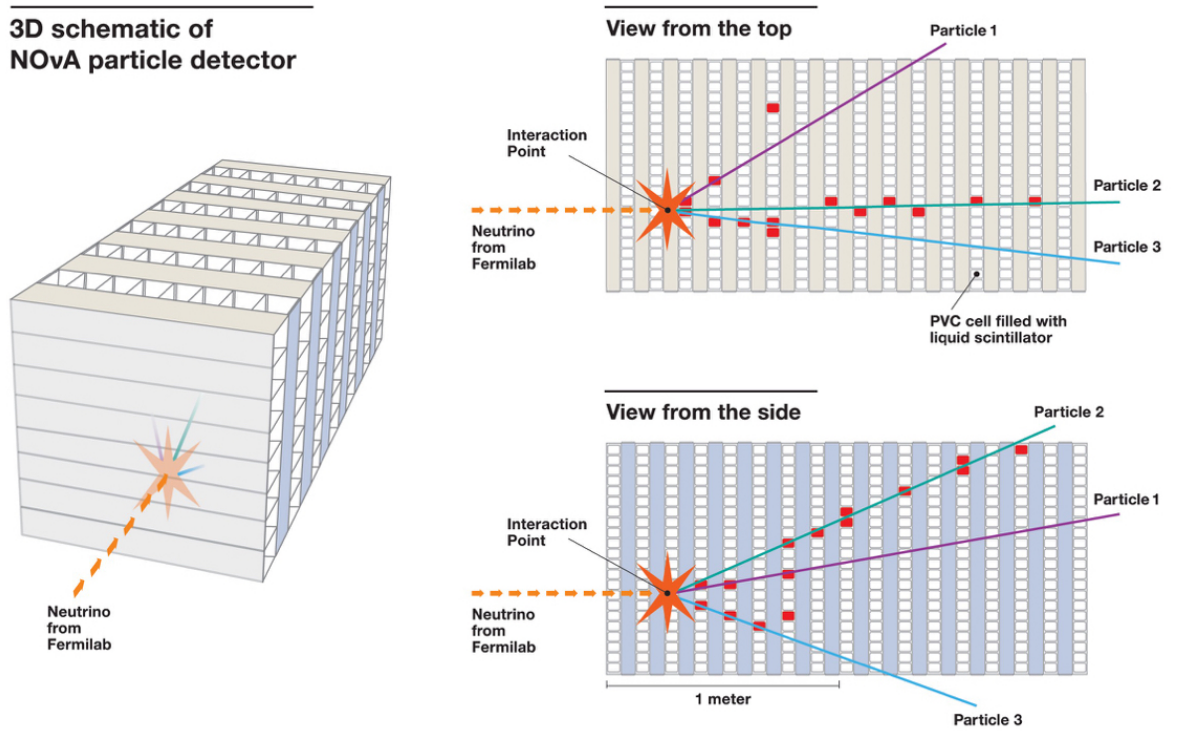


FIGURE 2.9: A 3D representation of the NOvA Detector in schematic form displaying the alternating vertical and horizontal placement of planes to provide us with three-dimensional events[52].

The detector is composed of 35% PVC and 65% active scintillator, measured by mass. Low Z materials are used for constructing the detector, where Z represents the atomic number indicating the protons in the atomic nucleus. The low Z materials have a longer radiation length, a property of the material that gives the mean distance traveled by an electron before its energy is decreased by a $1/e$ factor. This increased radiation length enables the observation of multiple points along a particle's trajectory due to its electromagnetic interactions with the material. The spacing between these observed points depends on the granularity of

the detector. Higher granularity or longer particle paths can enhance the differentiation of various experimental characteristics. Both mineral oil and PVC are carbon-based. These low Z materials produce a Molière radius² of about 11 cm. As a result, the EM showers tend to span across 2 to 3 cell widths, enabling a clear distinction from tracks left behind by more massive particles.

2.3 Data Acquisition System

Notable physics events within the NOvA detectors are read out and saved by the data acquisition (DAQ) system. Components of this DAQ system include avalanche photodiodes (APDs), front end boards (FEBs), data concentrator modules (DCMs), and timing distribution units (TDUs). There are also power distribution boxes and a large number of servers for managing the system's components and handling data. Given that the main objective of NOvA is studying neutrinos, the DAQ system was developed to have a continuous readout with 100% livetime, in which data is continuously gathered and later the decision to keep or drop the data is taken in separate servers.

2.3.1 Avalanche Photodiodes

As mentioned earlier, custom Hamamatsu avalanche photodiodes (APDs) consisting of 32 channels were specifically manufactured and employed to read out all cells within a single module. Each end of the WLS fiber is optically linked to a single APD pixel. These custom APDs used in NOvA have a quantum efficiency of 85% for the 500–550 nm light emitted by the fibers. The energy deposited by photons is recorded in the form of “Hits” when it exceeds the preset ADC (analog to digital conversion) threshold. Thresholds are specified for each APD pixel separately depending on the measured electronic noise of them. This thermal noise

²The Molière radius is the radius of a cylinder containing on average 90% of the shower's energy deposition.

can be minimized by maintaining the APDs at -15°C . Dry nitrogen gas is sent into the APD enclosure to prevent condensation over the cold APDs. The upper limit of bias voltage that can be applied to the APDs is about 425V. This bias voltage can later be adjusted to fine tune the gain up to highest possible gain of 150. Setting gain to the highest increases the hit efficiency to greater than 98% for hits closer to APDs and about 90% for the hits near the far end of an FD cell.

2.3.2 Front End Boards

Every individual APD is linked to a dedicated FEB. These FEBs carry out tasks like integration, shaping, digitizing, and organizing APD signals following the predetermined pixel thresholds for the APDs. The signals from the APDs are subsequently processed by a customized Application-Specific Integrated Circuit (ASIC) which is designed to handle 32 input channels, each corresponding to an individual APD pixel. Each of these channels is then amplified, integrated, and shaped. Following this, several channels are bunched up and transmitted sequentially to the ADC. Multiplexing (MUX) frequency is set to be 16MHz, i.e., Every channel within the bunch is forwarded to the ADC at intervals of 62.5 ns.

A 8:1 MUX is used in the FD while a 2:1 MUX is used at ND, both operating at 16 MHz. The two-fold MUX allows the ND to handle a data rate four times as high. Finally, signal from the MUX is directed into a Field Programmable Gate Array (FPGA). An FPGA is a reprogrammable integrated circuit mainly containing logic gates. The individual pixel thresholds are loaded in FPGA. The FPGA determines whether to retain or discard a signal. This is done by implementing Dual Correlated Sampling (DCS) on the signal. In this method, we compare the signal values of the last three clock ticks. The slow changes are then filtered out, while a fast rise in values is used as an indication of real hits. After filtering, the output of the FPGAs is now ready to be recorded.

2.3.3 Data Concentrator Modules

The processed continuous signals from the FEBs are sent to the Data Concentrator Modules (DCM) located on the detectors. Each DCM can take in signals from 64 FEBs, which is 2048 cells. The FD has 168 DCMs, and the ND has only 14 DCMs. In the FD case, a group of 12 DCMs are called a diblock.

Each DCM collects signals from the connected FPGA for $50\ \mu\text{s}$, called a microslice. The DCM stores 100 microslices and then stitches them together into a millislice containing 5 ms worth of data internally. After accumulating a millislice, the Data Collection Module (DCM) initiates the transfer of data to the buffer farm. This buffer farm acts as a temporary storage facility capable of holding approximately 20 to 30 minutes worth of raw data. This time window gives us time to decide on retaining or discarding the data at an early point in the experiment. The specific duration of this time frame hinges on the ongoing activity and the number of active triggers.

If it is decided to keep the microslice, all the coincident DCM microslices are combined to form an event. Subsequently, this event is transferred to a permanent storage location. Utilizing the buffer farm in this manner prevents the necessity of transferring the entirety of the data to permanent storage, which would overwhelm the system. Even if the permanent storage systems could accommodate the massive volume of recorded data, most of it would not be pertinent to the analysis. It would eventually be discarded during later phases of the analysis. Thus, It is better to discard dispensable data produced by the experiment at the earliest. A representation of FD DAQ system in schematic form is shown in fig. 2.10.

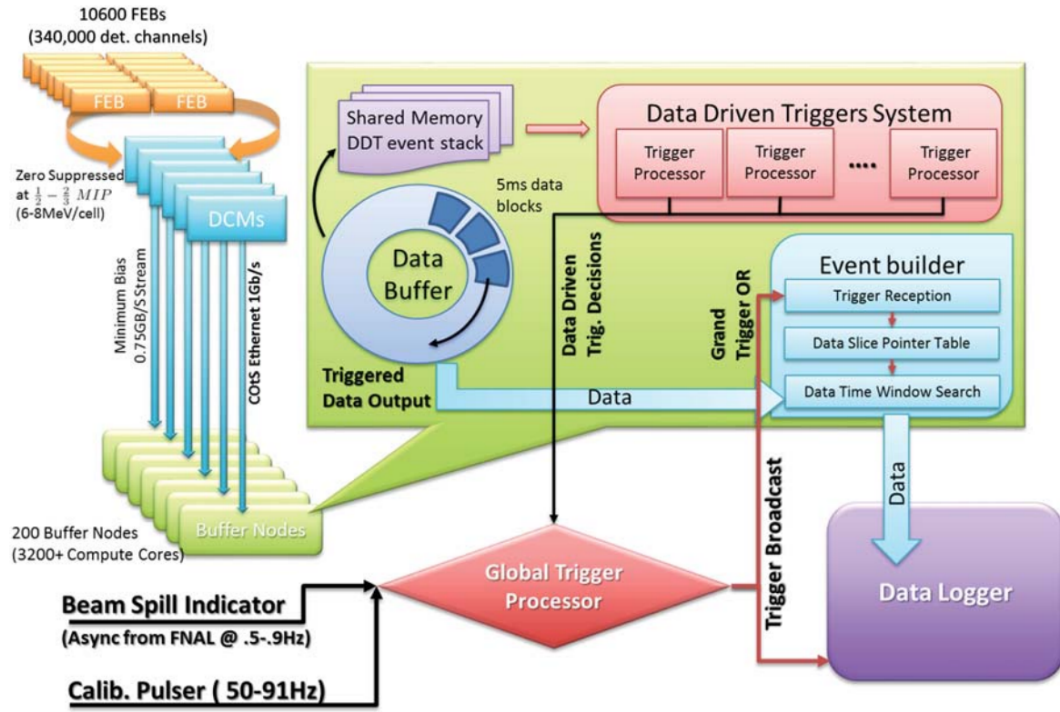


FIGURE 2.10: Data Acquisition (DAQ) system of NOvA FD.[48]

2.4 Timing System

The NOvA DAQ system greatly relies on timing information for readout, processing and saving raw data. To reliably synchronize all the different components of the DAQ system across the two detectors, the NOvA experiment created a dual timing chain system for its detectors. To correctly process and readout for different triggers, the individual detectors must be internally consistent with timing and also have good synchronization between other detector and Fermilab. All detector components of both ND and FD need to be synced within a threshold of 10 ns.

An ensemble of Timing Distribution Units (TDUs) make up the timing system. One of these is connected to GPS, which provides the primary timekeeping. This TDU is designated as the primary(master) unit, and the others in the chain are referred to as secondary(slave) units. The secondary TDUs each connect to 12 DCMs while splitting into 2, connecting 6 DCMs each, as shown in figure 2.11. At the end of the chains, the timing signal is looped back onto itself for calibration

purposes. The primary TDU of ND could not be connected to GPS while placing it near the detector as ND, is 105 m underground. Instead, the primary TDU of ND is placed on the surface and connected to the underground ND using a fiber-optic connection. The NOvA timekeeping starts from 00:00:00 on January 1st 2010 GMT. The base time unit on the NOvA experiment is 15.625 ns as internal clocks of most detector components run at a 64 MHz cycle. The GPS signal running at 10 MHz is used to synchronize the primary TDU clock. The 10 MHz clock is very stable with only a small drift of < 2 per billion per day when crosschecking with the satellite signal can not be performed. A phase-locked loop (PLL) is used to achieve synchronization across frequencies. 56-bit registers are used to encode timestamps, which allows for timestamp validity of 35.7 years from January 1st 2010.

A timing calibration reference (TCR) is added as a safety check at each detector. The TCR unit has its own GPS antenna and receiver and can produce stable trigger outputs each second. A comparison between the primary TDU timing signal and TCR trigger pulses can be used to check for drift and clock errors.

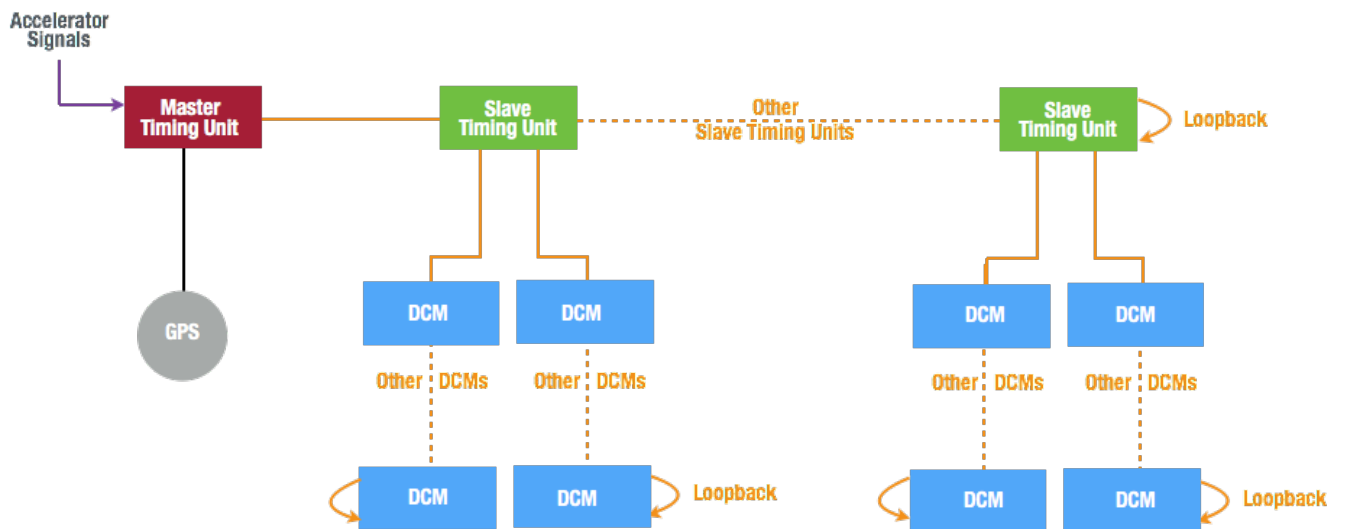


FIGURE 2.11: Timing system schematic of NOvA.[48]

The primary TDU transmits a signal containing the future timestamp. Later, when the sync signal is sent down the line and each component starts counting

down to their respective delay times. When the countdown completes, each component begins counting from the timestamp that was sent and now all components are in sync.

Now that the individual components of timing chains are in sync with each other, raw data can be readout with a time stamp from FEB, which monitors those channels. As mentioned earlier, due to the MUXs the actual readout frequency is considerably lower compared to FEB, which operates at 64 MHz. For the FD(ND), this frequency is lowered to 2(8) MHz. The APDs in the FD are read out with a rise time and a fall time of 460 ns and 7000 ns respectively. Since the ND sample rate is faster than the FD sample rate, the rise time and the fall time are reduced to 140 ns and 4500 ns respectively in the ND.

The APD channels perform readouts using the Duel Correlated Sampling (DCS) algorithm. There are two types of readout methods in DCS: single and multi-point timing. In the single-point timing method, the comparison is done between the current ADC value (s_i) and its counterpart from three samples earlier (s_{i-3}). If the difference between these samples exceeds the threshold of that channel, the current sample is read out by the FEB and it records the difference in ADC as the charge, with the timestamp taken from the TDC value of s_{i-3} . Until August 2014, this single-point timing method was employed exclusively in the FD. It has good resolution, but occasionally assigns two dissimilar pulse shapes with identical TDC and ADC values. The readout is triggered similarly in the multi-point timing method, but all the 4 ADC values are recorded and used to pulse fit. This method is more accurate and has a higher timing resolution. Both methods of the DCS algorithm are illustrated in figure 2.12 [53].

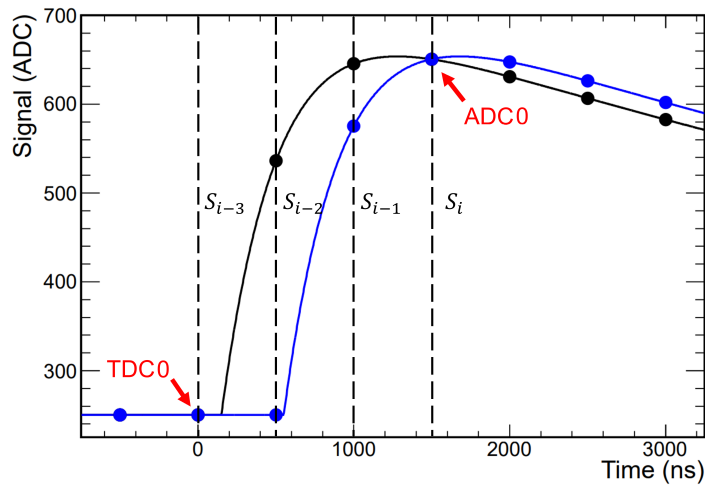


FIGURE 2.12: DCS algorithm. Two pulses of different shapes are shown. In the singlepoint timing method, both pulses will have the same ADC value ($s_i - s_{i-3}$) and the TDC value is time at s_{i-3} . We must use a multipoint timing method to differentiate between these two pulses.[53]

2.4.1 Triggers

Triggers are used to decide which data is to be saved and which data can be discarded. The triggers decide which microslices are to be retained. Triggers save data, which will be important for analysis while discarding less pertinent. Following are the three main types of triggers used in NOvA.

- Data driven triggers (DDT): These Triggers are launched when specific conditions are met in the detectors.
- Signal triggers: These Triggers are launched when DAQ receives signals external to the detectors to readout data.
- Clock triggers: These triggers launch at recurring specific time intervals.

Following are some of the basic triggers used by the NOvA DAQ system.

- DDEnergy Trigger: This a DDT used in the FD. This trigger gets activated when the total ADC reading within the detector exceeds the established threshold. This trigger is used to seek out high-energy cosmic events.

- **DDActivity Trigger:** Whenever activity occurs in the ND, this DDT stores $50\ \mu\text{s}$ of data. This data is utilized for calibrating the ND and monitoring operational conditions.
- **NuMI Trigger:** This trigger launches when the detector gets the NuMI beam spill from Fermilab. The $550\ \mu\text{s}$ slice centered around the $10\ \mu\text{s}$ beam spill is saved.
- **Supernova Trigger:** This specific DDT is crafted to seek supernova neutrino candidates in both FD and ND. The Supernova Early Warning System (SNEWS)[54] can also launch this trigger by sending a signal.

Chapter 3

NOvA Simulation and Event Reconstruction

Current high energy physics experiments like NOvA are complex and collect huge amounts of data over a long periods of time. Simulation of experiments help us correctly interpret the data collected. Latest theories and models used for simulations allow us to make accurate predictions. In NOvA, different Monte-Carlo generators are used to simulate various stages of the experiment. These are step by step, FLUKA and FLUGG for NuMI beam flux simulation, CRY for cosmic ray generation, GENIE for the neutrino interactions, GEANT4 for detector simulation, PhotonTransport for photon propagation in the cells and ReadOutSim to simulate the front end electronics. This chapter will describe each of these methods and show how realistically the total simulation reproduces and reconstructs NOvA data.

3.1 Monte-Carlo simulations

Monte-Carlo (MC) simulations are computational techniques used extensively in physics and engineering to model complex systems that cannot be solved analytically [55]. MC simulations use random sampling to produce numerical simulations

of complex systems. Monte Carlo simulations which have intrinsic randomness integral to them, derive their name from the Monte Carlo Casino in Monaco, renowned for its games of chance. This randomness mimics the inherent uncertainty often found in real-world physical systems, making MC simulations a valuable tool in physics and other fields. By repeatedly sampling random values for uncertain parameters and running simulations with these values, MC simulations provide a way to make predictions and understand the behavior of complex physical systems.

MC simulation depends on the fact that any complex system can be modeled using a probability distribution function (p.d.f). MC simulation requires a random variable distributed according to a desired x . A simple MC algorithm to generate a random variable is as follows:

1. A sequence of random values, denoted as $r_1, r_2, \dots$, is produced following a uniform distribution within the interval $0 < r < 1$.
2. This sequence $r_1, r_2, \dots$ is used to produce another sequence $x_1, x_2, \dots$ such that the x values are distributed according to a probability density $f(x)$ in which one is interested. There are many algorithms for random number generation for any arbitrary function like transformation method, acceptance rejection method and composite method.
3. These values of x can be regarded as simulated measurements, allowing estimation of the probabilities for x to assume values in a given region.
4. In this way one can effectively compute the cumulative measure of the experimental quantity whose probability distribution function is $f(x)$.

There are two main types of uniform random number generators (RNGs):

1. Hardware RNGs generate random numbers from a physical process, rather than by means of an algorithm.

2. Pseudo RNGs employ an algorithm to generate a sequence of numbers that mimic the properties of random number sequences. However, the sequence produced by a PRNG is not genuinely random, as it is entirely determined by an initial value known as the PRNG's seed.

As discussed, MC simulation is very effective in systems with high parameter uncertainty and large degrees of freedom. Monte Carlo simulations offer a powerful and flexible tool for physicists to gain insights into complex systems and phenomena by running a multitude of experiments in a digital environment, all while controlling and manipulating various parameters and conditions as needed. In a complex experiment like NOvA, various MC simulations are used to bridge the gap between the theoretical predictions and the practical experiment.

3.2 Simulation chain of NOvA

The NOvA simulation chain and the packages used at each step are shown in figure 3.1.

3.2.1 NuMI Beamline simulation

The first step of NOvA simulation chain is beam simulation. The beam simulation starts from modelling hadron production at target due to collision between the protons from the accelerator and the graphite target. FLUKA[57] and FLUGG[58] packages were used to simulate this step. The FLUKA is a general-purpose particle physics simulation software. It is primarily used for simulating the passage of particles (including both hadrons and leptons) through matter. The FLUGG stands for “FLUka for GEANT4”. It is a specialized interface that combines FLUKA with GEANT4 [59, 60]. FLUGG serves as a bridge between these two software packages. In the newer analysis, a completely GEANT4 based G4NuMI

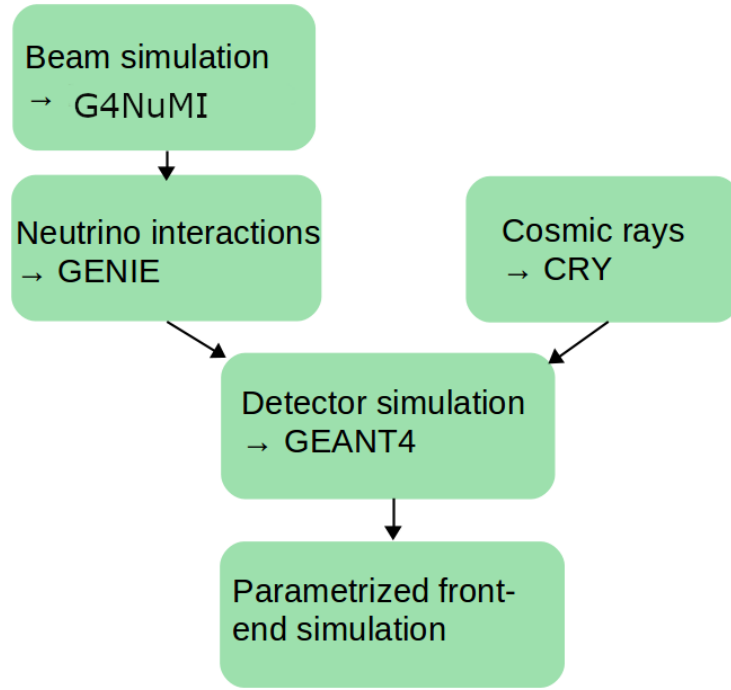
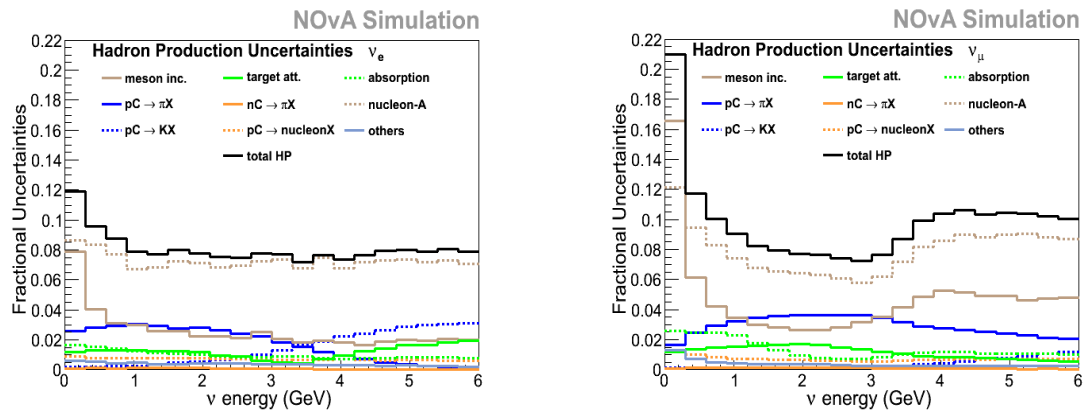


FIGURE 3.1: The simulation chain of NOvA [56].

is used for hadron production simulation. The G4NuMI package uses GEANT4 v10.4p02 with FTFP_BERT [61] hadronic list. These produced hadrons are then propagated through the graphite target, the focusing horns and the decay pipe along NuMI beamline using a GEANT4 simulation.

FIGURE 3.2: The fractional uncertainties of the flux due to hadron production uncertainties as a function of neutrino energy at NOvA ND detector for the ν_e (left) and ν_μ (right) flux. [62].

Hadronic modeling is hard to simulate correctly due to cluttered strong interactions. Hadron production cross section uncertainties are known to be the major contributors to the neutrino flux uncertainties as shown in fig 3.2. To correct for the hadron production mismodeling, NOvA uses the Package to Predict the Flux (PPFX) [63], designed for the MINERvA experiment. The PPFX uses hadron production data from external experiments like NA49 to correct for the hadronic modeling uncertainties between 12-120 GeV. The PPFX is used to weight the hadron production and fine tune the simulated neutrino flux. The PPFX corrected simulated NuMI beam neutrino flux and anti-neutrino flux for the ND can be seen in the figure 3.3.

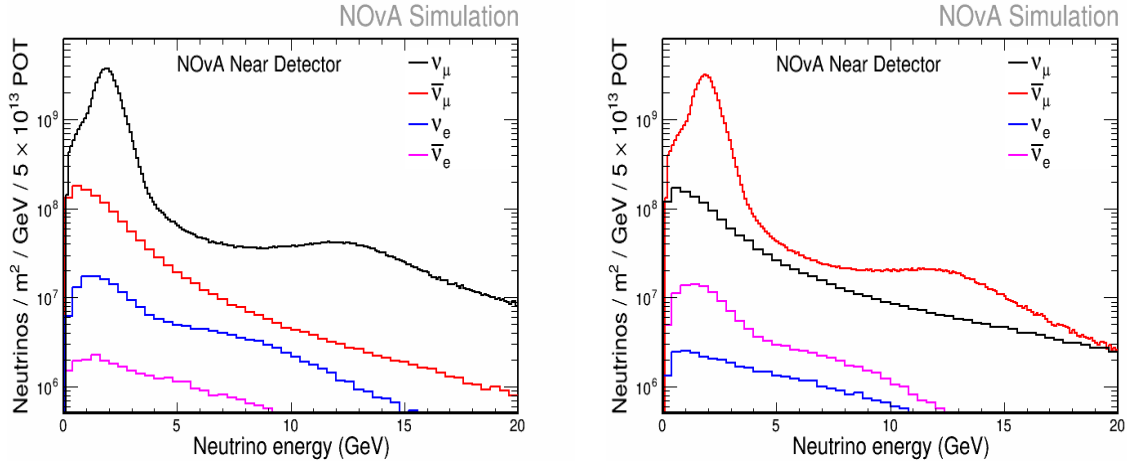


FIGURE 3.3: The PPFX corrected simulated NuMI beam flux at NOvA ND detector for neutrino mode in the left plot and antineutrino mode in the right plot [64].

This corrected neutrino flux is then propagated to both ND and FD in forward and reverse horn current modes assuming no oscillations. In case of FD, multiple set of flux files are generated:

- Nonswap: The original NuMI beam flux propagated to FD.
- Fluxswap: Every ν_e in the original simulated NuMI beam is switched to ν_μ and every ν_μ is switched to ν_e . This amounts to a 0% survival probability for $\nu_\mu \rightarrow \nu_\mu$ and $\nu_e \rightarrow \nu_e$ will have 100% oscillation probability .

- Tauswap: All flavours of neutrinos are switched to ν_τ . The tauswap files allow for the examination of high energy backgrounds originating from ν_τ and neutral current (NC) interactions.

When all the above flux files are combined, any composition of neutrino flux that might be produced through oscillations using any particular set of oscillation parameters can be constructed.

Simulated neutrino flux at the detector, nuclei of the the matter present in the detector and it's surrounding area are given as input to GENIE.

3.2.2 Simulation of Neutrino Interactions

The second step of the simulation chain involve simulating neutrino interactions using neutrino event generator called GENIE. GENIE[65] (“Generates Events for Neutrino Interaction Experiment”) is a software toolkit and event generator designed to simulate and model neutrino interactions with various target materials. In NOvA, GENIE is used to simulate neutrino interactions within the detector. Simulated neutrino flux at the detector, nuclei of the the matter present in the detector and its surrounding area are given as input to GENIE. The initial state of the neutrino interaction with the target nucleon is simulated with a model such as relativistic Fermi gas model and various other models in GENIE. GENIE outputs the 4-vectors of the secondary particles after the primary interaction takes place. Simulated neutrino event rates and respective event kinematics can be produced by combining output of GENIE and previous simulated neutrino flux. The default models of GENIE have been found to be inconsistent with NOvA data. So, those models were tuned based on NOvA's ND data in form of cross section measurements, external measurements and observed data discrepancies to get NOvA specific reweighting [66]. These changes are called tuning.

Cosmic ray simulation is required at NOvA, as the FD is vulnerable to high rate of cosmic rays. Cosmic rays are a significant source of background, even though most of them can be filtered out by timing cuts and analysis techniques. As shown in the figure 3.1, NOvA simulates cosmic ray interactions separate from the neutrino interactions using the Cosmic RaY shower generator (CRY) [67] and overlays it on the output of GENIE.

3.2.3 Detector simulation

The simulated neutrinos, their daughter particles and the cosmic rays are taken as input for the next step of simulation chain, which is the detector simulation. In NOvA, GEANT4 is used to build the exact geometry of both the detectors and their surroundings for detector simulation. The inputs are propagated through the detector using GEANT4 package to simulate their propagation, energy loss and probable decays. The physics processes are modeled in GEANT4 using physics lists. NOvA uses a set of three physics lists called the QGSP_BERT_HP list. For high energy hadrons, Quark Gluon String Precompound (QGSP) physics list which uses QGS model is applied. Bertini cascade model [68] is used BERT list. It models the hadronic processes with energy less than 10 GeV and the thermal neutrons (< 20 MeV) can be tracked using the high precision (HP) neutron [69] model in the HP list. The rock in front of the ND is simulated for limited interactions as it is computationally expensive. These rock simulations are overlaid at a later stage. The GEANT4 geometry simulation, when combined with the simulated particle fluxes, enables the calculation of the energy deposited by individual particles in various sections of the detector. The final output from the GEANT4 contains the list of particles interacted with the detector, the energy deposited and the hit positions.

3.2.4 Photon Transport

NOvA uses custom built ray tracing package, PhotonTransport [70] to simulate the scintillation photon propagation through WLS fiber. Even though GEANT4 can simulate these processes, we use PhotonTransport as it is computationally less expensive. PhotonTransport uses scintillation spectrum of the scintillator, cell wall reflectivity and absorption spectrum of WLS fiber to simulate the photon propagation. The output of this step is then given as input to electronic simulator ReadoutSim. Since, Photon Transport requires scintillation photons its simulation cannot begin until the energy deposited by the charged particle is converted to the scintillation light. After considering the liquid scintillator properties, an estimate of 3360 photons/MeV is used to convert the energy deposited to no of photons.

ReadoutSim package takes the photons from the WLS fiber and propagates it in form of a digitized waveform through the front end electronics of the detector and produce the RawDigits. The resulting file contains information structured similar to the data obtained from NOvA's DAQ system. This simulated files have additional truth level information, which is required for developing analysis tools before we run it on the unboxed data.

3.3 Detector Calibration

The data collected from the physical detector and the data generated through MC simulations are now stored as disconnected hits in the detector. Now, we apply calibration and reconstruction on our data to extract valuable information for physics analyses. Calibration primarily focuses on transforming the recorded ADC (Analog-to-Digital Converter) values associated with the disconnected hits into a reasonable approximation of the actual energy deposited by charged particles in the detector. This step is crucial for achieving precise measurements of the

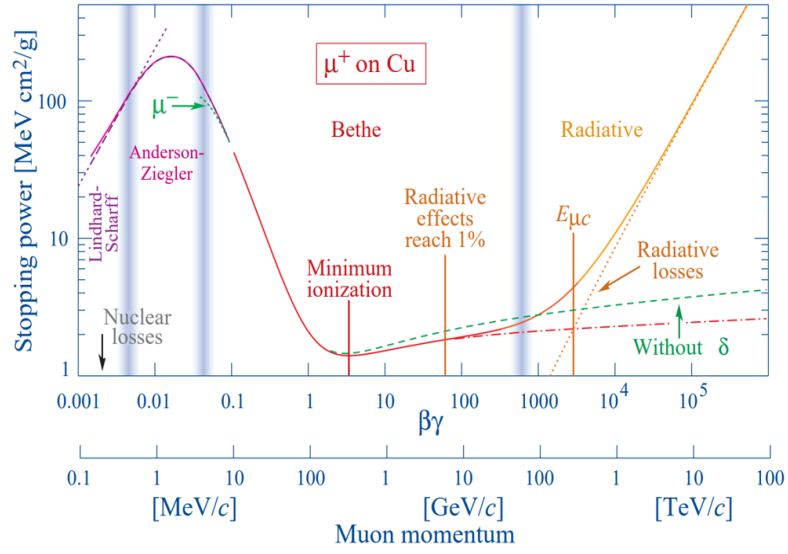
neutrino's energy. The shape and magnitude of the energy spectrum of neutrinos are essential for probing various relevant oscillation parameters.

So, before we can analyze the data from the detector, each recorded ADC value in every cell must undergo individual calibration to establish an absolute energy scale. This calibration process needs to account for several factors, which include the light yield of the liquid scintillator, light attenuation as it travels through the WLS fiber and the quantum efficiency of the APDs. In NOvA, cosmic muons which have almost uniform flux over the detector space are used to attain accurate calibration as they have distinct experimental tracks which can be easily identified. The energy loss of muon traveling through material is theoretically well modeled by Bethe-Bloch equation given in 3.1 and it is accurate in predicting energy deposited in range $0.1 < \beta\gamma < 1000$ as plotted in fig 3.4. Bethe-Bloch equation provides a standard candle for calibration.

$$-\frac{dE}{dx} = 4\pi N_A r_e^2 m_e c^2 z^2 \frac{Z}{A} \frac{1}{2\beta^2} \left[\ln\left(\frac{2m_e c^2 \beta^2 \gamma^2 W_{max}}{I^2}\right) - 2\beta^2 - \delta(\gamma\beta) \right] \quad (3.1)$$

Here, N_A is Avogadro's number, r_e is electron radius, z is the charge, $m_e c^2$ is rest energy of electron, A is the mass number and Z is the atomic number of the detector material, I is the average excitation energy, δ is correction applied ionization energy loss in the material due to density effect. The maximum energy transfer feasible in a single collision is W_{max} , β is the velocity and γ is the gamma factor of muon propagating through the material.

For calibration purposes, the muons tracks need to have good quality. So, containment and tri-cell cuts are made over the data. The tri-cell cut mandates that the muon track must traverse both the cell above and below the current hit as shown in fig 3.5. This requirement serves to prevent the track from passing directly through the corner of a cell, reducing the occurrence of poorly reconstructed

FIGURE 3.4: Mass stopping power (dE/dx) for positive muons in copper [71].

tracks in the dataset. Additionally, it facilitates a trigonometric calculation of the track's path length within a single cell.

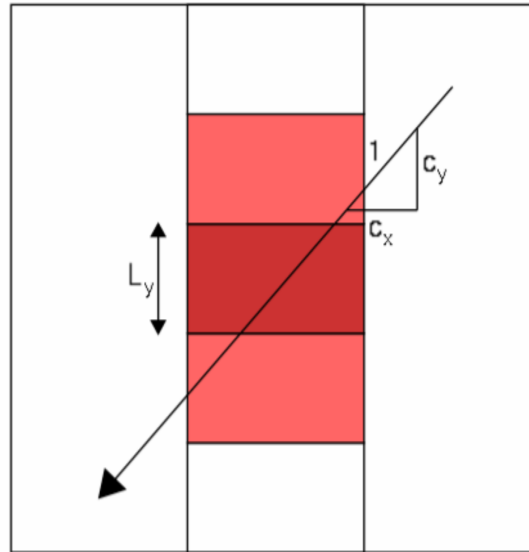


FIGURE 3.5: The tri-cell cut used for detector calibration.

3.3.1 Relative calibration

The recorded ADC values are converted to PECorr, which represents a corrected version of the Photoelectron (PE) values that have been normalized across the

entire detector. It's termed "relative" because its primary function is to correct the differences between individual detector cells. This correction is based on the assumption that the energy distribution of cosmic muons is consistent and uniform throughout the detector. The distance from the readout is labeled as W . W is valued as 0 at center of the cell and the locations closer to the readout have positive W values. For each cell, separate histograms of PE/cm versus W are created. Finally, the relative calibration process is employed to apply corrections to these histograms. The correction factor is

$$T = \frac{PE}{\lambda} \frac{E_{True}}{E_{MIP}} \quad (3.2)$$

Here PE refers to the count of simulated photons recorded by the electronics, T stands for the correction factor, λ represents the count of photons detected without fluctuations, E_{True} represents the true energy deposited in the scintillator and E_{MIP} signifies the energy that would be deposited based solely on the path length through the cell. PE/λ is used to account for the threshold correction, as λ is dependent solely on the path length.

After these corrections, the cosmic muon data should have a uniform energy spectrum across the detector. Now, individual cell variations are corrected for by applying a double exponential fit to the corrected PE/cm vs W plots.

$$y = C + A(e^{x/X_A} + e^{-(\frac{3L}{2} + x/X_A)}) \quad (3.3)$$

Here, x is the distance from the cell center, which is designated as $W = 0$. X_A represents the attenuation length of fiber, while L corresponds to the total cell length. The parameters A and C are free-fitting parameters.

The residuals derived from the above exponential fit undergo a Locally Weighted Scatter plot Smoothing (LOWESS) fit. This smoothing process helps correct and stabilize the cell response. This gives out corrected equivalent photo electrons

called PECorr which is uniform across the detector. PECorr is normalized to the cell response at $W = 0$.

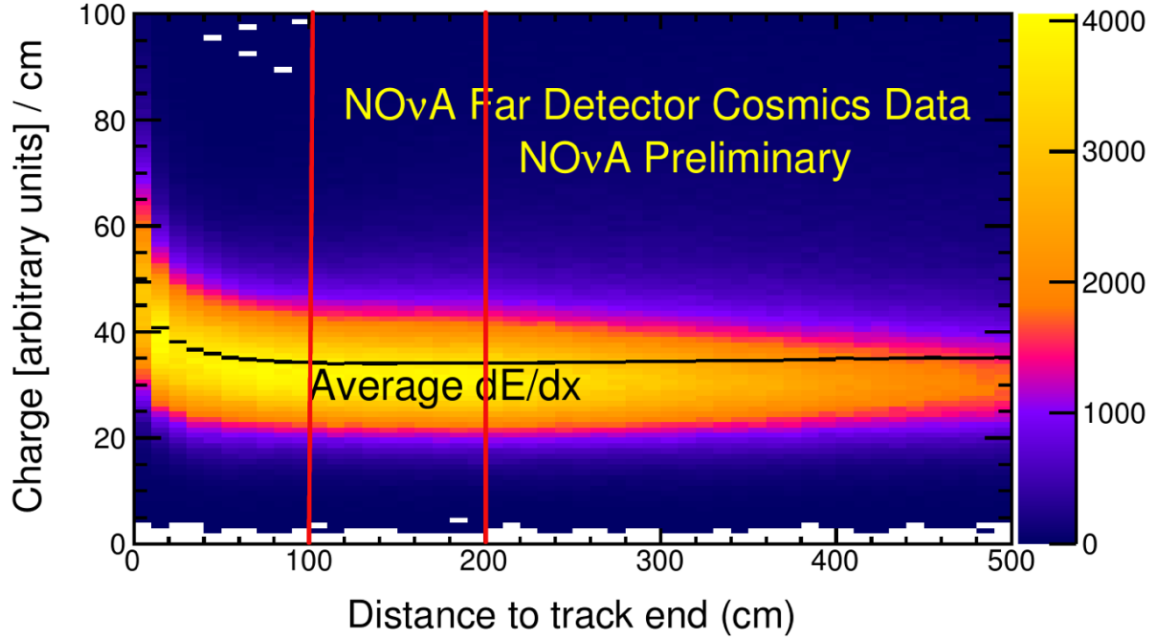


FIGURE 3.6: The dE/dx of stopping muons as function of distance from the end point of track.

3.3.2 Absolute calibration

After normalizing the cell response across the detector, the resulting PECorr values need to be calibrated to an absolute energy scale. While through-going muons are utilized for relative calibration, stopping muons are employed for absolute calibration purposes. The dE/dx behaviour of stopping muons at end of their tracks is compared with the Bethe-Bloch curve in fig 3.4, for an accurate estimate of energy deposition. To measure the detector response, a window of 1 m length along the track, positioned 1 m from the stopping point is used. In this region, the dE/dx (rate of energy loss) is approximately flat. Now, Attenuation corrected dE/dx is measured for both data and MC as shown in fig 3.7a. Then, we obtain the scale factor to transform PECorr to GeV using the truth information from MC. This scale factor is subsequently utilized for transforming PECorr/cm values

of our data to MeV/cm, enabling an absolute energy scale calibration as shown in the fig 3.7b.

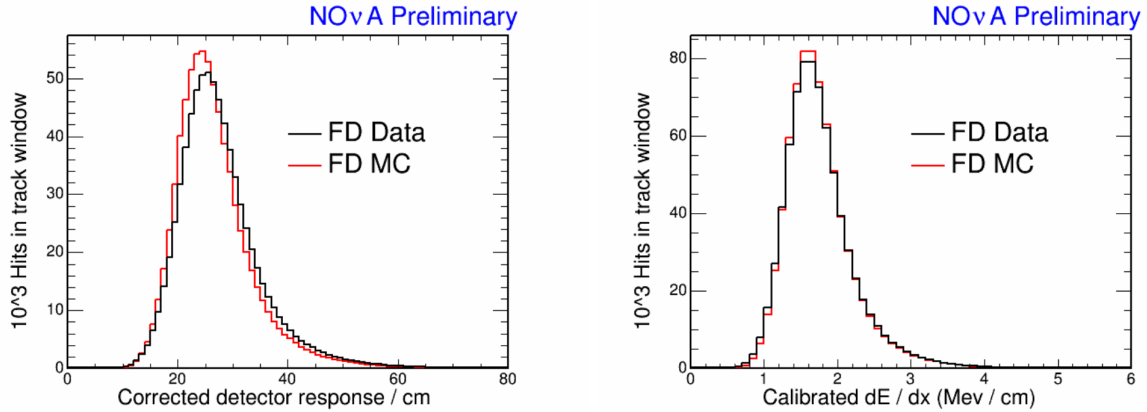


FIGURE 3.7: (a) Attenuation corrected dE/dx. (b) Absolute calibrated dE/dx.

3.4 Event Reconstruction

We want to get physics quantities like particle identities, four-momenta from the calibrated data hits and times from each detector cell for a physics analysis. This is done through event reconstruction. While many of the reconstruction tasks in NOvA are now done using machine learning techniques, it is important to understand the conventional reconstructions methods. This section will concentrate on conventional reconstruction algorithms, which make use of fundamental physics principles governing particle interactions and propagation and the space-time connection between detector data hits.

NOvA reconstruction chain has three steps:

1. Clustering
2. Vertexing
3. Prong Reconstruction

3.4.1 Clustering

NOvA DAQ reads out all hits within $550 \mu\text{s}$ of NuMI or cosmic triggers. In that $550 \mu\text{s}$ window, there will be numerous overlaying particle interactions as seen in fig 3.8. In the FD, background from cosmic rays must be separated from the neutrino interactions as shown in fig 3.9. We see that the NOvA timing is reasonable good at separating the interactions, yet timing criteria is not sufficient as the particle interactions might occur at the same time. Like in ND, we get several neutrino interactions per each $10\mu\text{s}$ beam spill because of the high neutrino flux as shown in fig 3.10. So, the first step in reconstruction is to isolate these interactions from each other and the surrounding detector activity using both spatial and temporal information. This done using clustering algorithms, in which all the hits from a single event are identified and then grouped together into a slice. Accurate clustering will also lead to the recognition of noise hits.

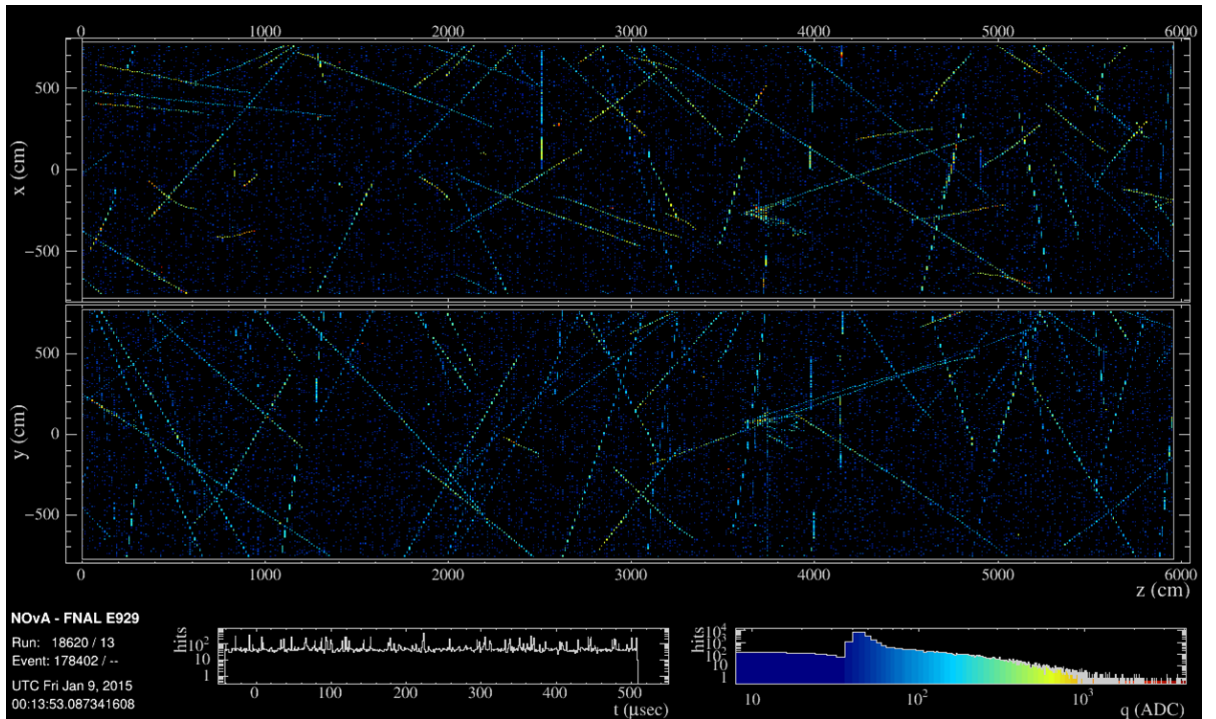


FIGURE 3.8: The event display showing readout of $550 \mu\text{s}$ of NuMI trigger at FD. The top part of the event display shows the x-view and the bottom pane shows the y-view. Hit colors correspond to ADC values.

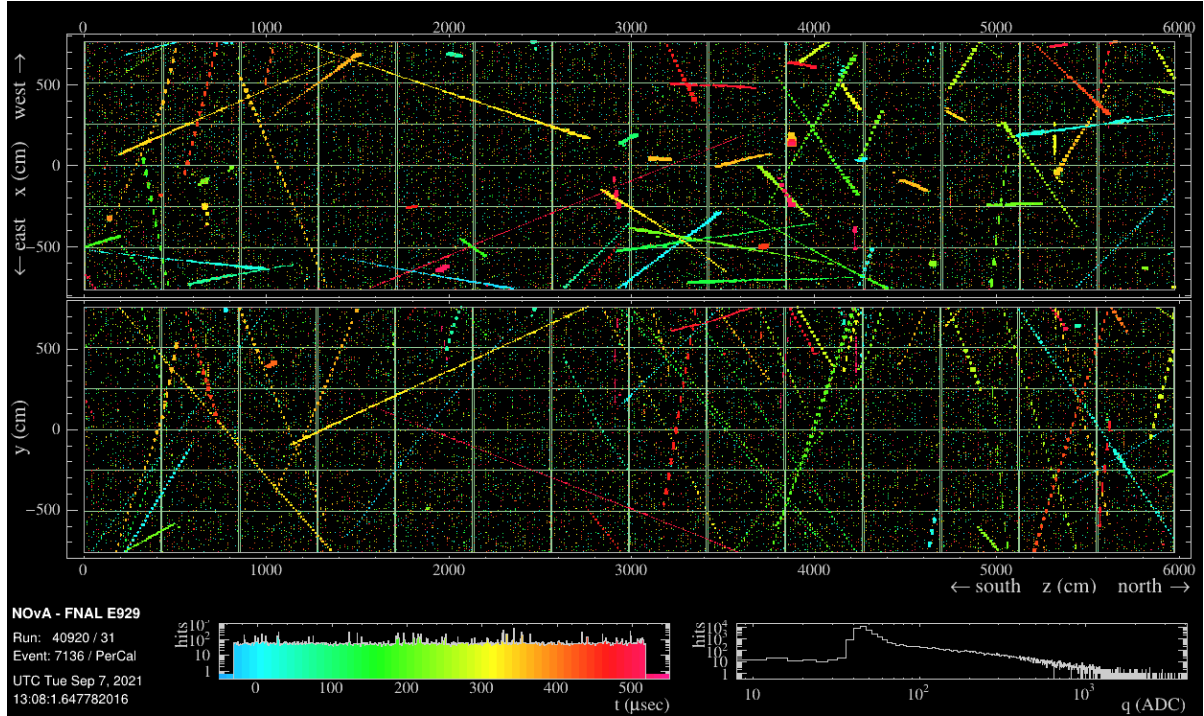


FIGURE 3.9: The event display showing readout of FD Cosmic trigger at FD. Hit colors correspond to readout timing.

Initial NOvA analyses used 4DSlicer, which was built upon “Density-Based Spatial Clustering for Applications with Noise” [72] (DBSCAN) clustering algorithm. DBSCAN depends on the number of neighboring hits to define clusters and distinguish signal hits from noise. 4DSlicer was shown to be successful but exhibited several common failure scenarios. The most prominent scenario is the potential merging of two clearly distinct slices when there are hits bridging them. This is particularly troublesome in the ND case, because we expect several neutrino interactions to take place in the beam spill window, leading to a pile-up particularly during periods of high beam intensity.

Later, Time Density Slicer (TDSlicer) [73] is implemented to improve upon the 4DSlicer. The TDSlicer uses a different clustering algorithm [74, 75] compared to the 4DSlicer, resulting in a more forceful approach to clustering. This reduces the likelihood of merging clusters incorrectly but may lead to the occasional erroneous splitting of clusters. When a bridge could potentially form between two clusters, and TDSlicer divides the cluster containing the bridge into two separate clusters,

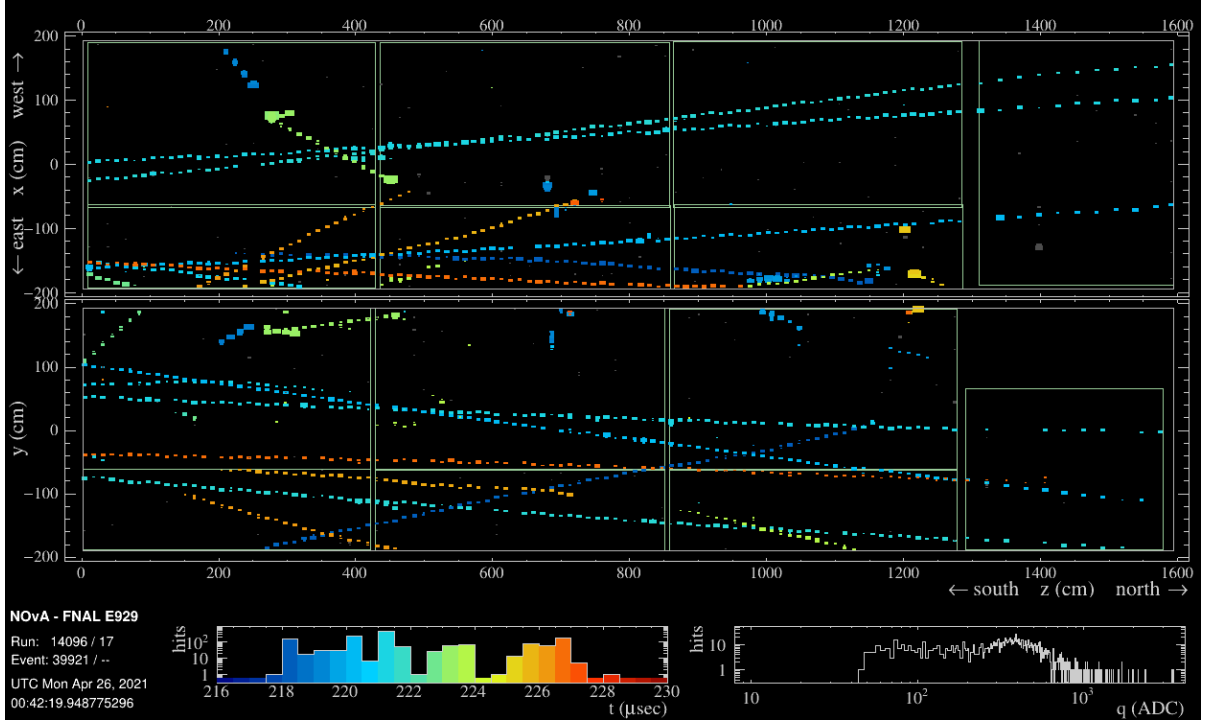


FIGURE 3.10: The event display showing readout of 10 μ s of NuMI beam spill at ND. Hit colors correspond to readout timing.

the hits forming the bridge are then assigned to the individual clusters. Some of these hits may not be accurately grouped by the Slicer. However, it has been observed that the misgrouping of these hits does not notably affect the energy estimation and event reconstruction overall [73].

The TDSlicer algorithm mainly contains three steps. The first is to find the centroid hits using density clustering algorithm Rodriguez and Liao [74]. Centroid hits are determined through location of the local maxima in density for the given group of hits. The algorithm prescribes finding the density of points (ρ_i) around each point at first, given by

$$\rho_i = \sum_j \exp\left(\frac{-d_{ij}^2}{\tau^2}\right) \quad (3.4)$$

$$d_{ij} = \left| |\Delta t| - \frac{R}{c} \right| / \tau \quad (3.5)$$

Here d_{ij} is the euclidean distance between two hits i and j in a detector view, c is speed of light, Δt is temporal distance, R is the distance between two hits in spatial dimension. τ is the scaling factor associated with the timing resolution for each hit, with a value of 80 ns in the FD and 16 ns in the ND. The difference in τ for ND and FD is because the sampling frequency of FD DAQ is four times lower than ND DAQ.

After the density of each hit is calculated, isolation score (δ_i) of each hits is calculated. δ_i is an euclidean distance between i and the nearest point in space with a greater density, as shown

$$\delta_i = \min_{j|\rho_j > \rho_i} (d_{ij}) \quad (3.6)$$

Finally, To create a well defined selection, value of the maximum hit isolation is set to a number greater than the maximum isolation possible between two hits occurring within the data spill. The hits with higher density and isolation are picked as the centroids of the cluster in each view. But, there is a minimum threshold set to finish time density slicing. The threshold is $\rho > 3$ and $\delta > 8$ for the ND, while it is $\rho > 10$ and $\delta > 6$ for the FD. The hits are then grouped with their closest centroid if their time of flight time difference is less than 10τ .

The second step of the TDSlicer uses centroids to build 3D clusters in xzt and yzt space using Prim's algorithm [75]. From the centroid, the algorithm builds the cluster by iteratively adding the hit outside the cluster which is closest to cluster's edge. The points with $d_{ij} > 8(5)\tau$ for the ND(FD) are not added to the cluster.

In the third step, the 3D clusters in each view are combined across both views to create the final collection of 4D slices. This process involves calculating the average zt values for every conceivable cluster pairs originating from two views. These pairs are iteratively merged, starting with the best pair with smallest zt, until all paired

clusters have been consolidated. Any clusters that remain unmatched are classified as noise.

Two key metrics: purity and completeness (also known as efficiency) are used to assess the quality of the generated slices. Completeness provides a measure of how much of a particle's energy is included in the slice. Purity indicates the amount of energy in the slice that originates from sources other than the event. These metrics are evaluated using MC datasets, where the true information about a particle passing through a portion of the detector is known.

$$\text{Purity} = \frac{\text{Energy deposited by True particle in the slice.}}{\text{Total Energy deposited in the slice}} \quad (3.7)$$

$$\text{Completeness} = \frac{\text{Energy deposited by True particle in the slice.}}{\text{Total Energy from True particle}} \quad (3.8)$$

The TDSlicer and 4DSlicer can be compared using the above metrics. The TDSlicer attains a mean completeness of 0.849 and 0.970 for ND and FD MC data respectively. It achieves a mean purity of 0.994 and 0.987 for ND and FD MC data. On the other hand, the 4DSlicer exhibits a mean completeness of 0.930 and 0.964 for ND and FD MC data and a mean purity of 0.984 and 0.986 for ND and FD MC simulated data respectively. It is to be noted that the completeness at ND for TDSlicer is lower than 4Dslicer. But, the TDSlicer produces a greater number of good slices (slices with efficiency > 0.9 and purity > 0.9) than 4DSlicer, and scales more effectively with increasing data intensity as shown in fig 3.11.

3.4.2 Vertexing

The next step of reconstruction is to determine the interaction vertex, the point where neutrino interaction took place inside the detector. In NOvA, it is assumed that all energy deposited during an event originates from a single point, a premise

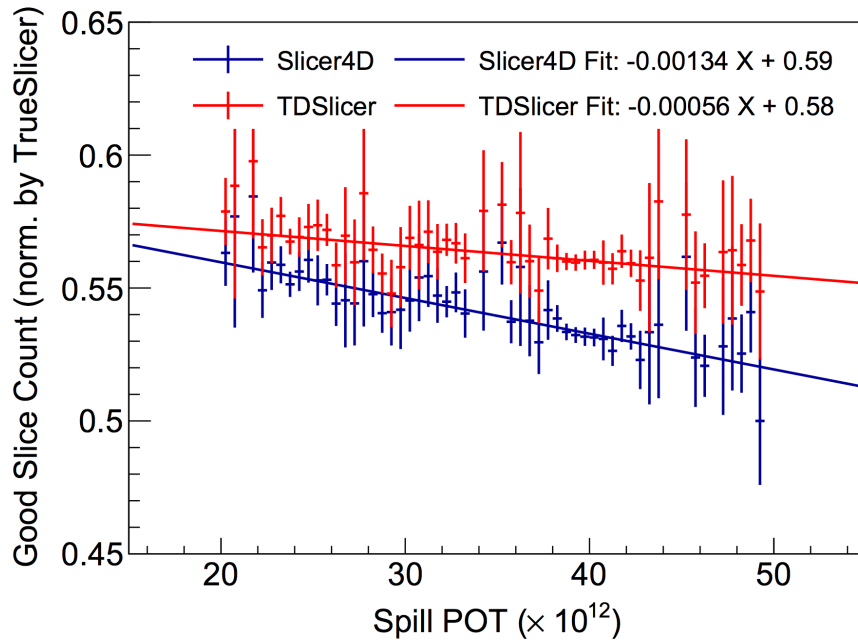


FIGURE 3.11: Comparison between 4DSlicer and TDSlicer by plotting The fraction of good slices as a function of intensity.

that holds true for the majority of neutrino interactions. However, this assumption does not account for any secondary interactions like particle decays or hard scatters, that might occur after the initial neutrino interaction.

The algorithm used in NOvA for vertexing is called "elastic arms". It operates under the assumption that all particle tracks (arms) originate from a common vertex point. These arms are produced by applying a multi point Hough transform on all hits within a slice for both views. The arms are produced by detecting the linear features in the event using the Hough transformation. It accomplishes this by fitting each pair of hits within a view with a straight line and parameterizing them in a space defined by the coordinates (ρ, θ) .

$$\rho = x \cos \theta + y \sin \theta \quad (3.9)$$

Here, ρ is the closet distance from origin, and θ represents the angle formed by ρ with the positive x-axis. Any given pair of hits can be mapped to the Hough space as a single set of (ρ, θ) values as shown in fig 3.12.

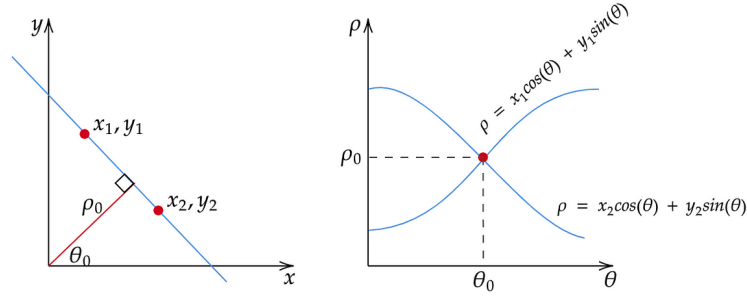


FIGURE 3.12: A schematic diagram illustrating the coordinate transformation of the Hough transform. The left plot depicts the line in Cartesian coordinate space; The right plot transforms the line into the Hough space.

To decrease computational expense, limits are set on which pair of hits to be passed through the Hough transform. If there is large gap ($\sqrt{15000}$ cm) between a pair in XZ or YZ space, they are not selected. This limits also reduce the tendency of the transform to overweight distant hit pairs. The hits with same x or y in xz or yz space respectively are ignored unless they are far apart, to reduce over production of horizontal lines.

Each pair of hits is given a weight, which is determined using Gaussian smearing. The weight assigned to a pair decreases when the hits are located far apart or have significantly different angles relative to the neutrino momentum. This approach reduces the impact of hit pairs associated with other particle tracks. Once all the hit pairs are mapped to the binned Hough space, collinear hits pairs will pop up as peaks, with value above the calculated threshold. The threshold value is specific to each view. To enhance the accuracy of the produced lines, a deviation from the bin center is permitted through an adjustment to the (ρ, ϕ) parameters. To achieve this, a weighted average is computed over a 7×7 bin area positioned at the center of the peak bin. The weights used for this averaging calculation are determined based on the content of each bin and its distance from the peak bin.

If all bins in the parameter space which are above the threshold are considered as representatives of lines in the data, the number of lines get overestimated. To

address this issue, the Hough transform is applied multiple times. After each application, only the highest peak is chosen, and the hit pairs associated with that peak are removed from the slice. The Hough transform is then reapplied to the remaining hits within the slice, identifying additional lines. If the new lines are found to be very similar to previous line in the list, the hits related to the new line are removed from the slice. This is continued until there are no more peaks crossing the threshold or when the number of lines is reached to maximum, typically limited to 10, and added to the list. The application of multi point Hough transform on a event can be seen in fig 3.13.

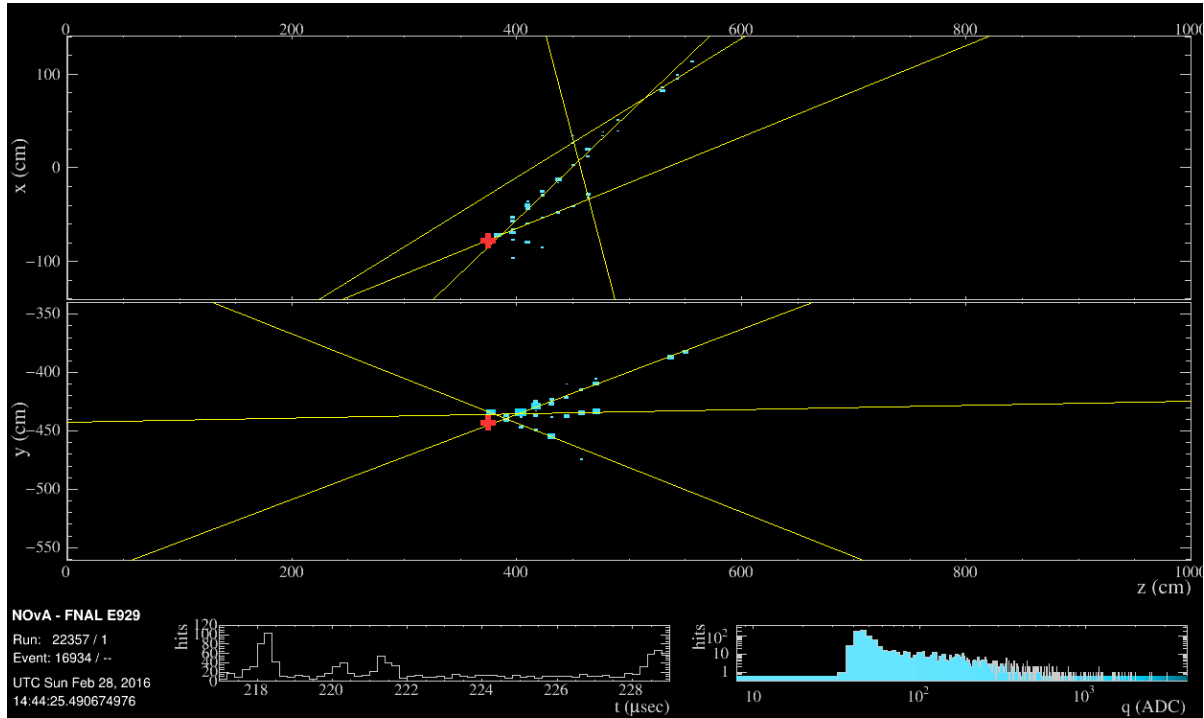


FIGURE 3.13: The multi point Hough transform on a event in FD. The red cross is position of interaction vertex produced by elastic arms algorithm

The lines obtained from the Multi-Point Hough Transform can be utilized to determine the interaction vertex of the given slice. This is done using Elastic Arms vertexing algorithm. The first step of the algorithm is to minimize the energy function

$$E = \sum_{i=1}^N \sum_{j=1}^M V_{ij} M_{ij} + \lambda \sum_{i=1}^N \left(\sum_{j=1}^M V_{ij} - 1 \right)^2 + \frac{2}{\lambda_\nu} \sum_{j=1}^M D_j \quad (3.10)$$

$$V_{ij} = \frac{e^{-\beta M_{ij}}}{e^{-\beta \lambda} + \sum_{k=1}^M e^{-\beta M_{ik}}} \quad (3.11)$$

where M_{ij} is the distance from the hit i and the arm j , V_{ij} is the strength of the association, D_j is the distance between the interaction vertex and the earliest hit along the arm j , and the λ terms control the strength of each term. The energy function 3.11 can be broke down into three individual terms, each with a specific function. The goodness of fit between the arms and hits is measured in the first term. The second term is the penalty for the hits which are not connected to any arm. The third term penalizes the distance between the first hit and the interaction vertex. For V_{ij} , $e^{-\beta M_{ij}}$, $e^{-\beta \lambda}$ give the likelihood of a given hit i to belong to the arm j and likelihood that the hit i is a noise respectively. β gives the value for the influence of an arm in terms of distance to the hits. Finally using the “simulated annealing” [76], the β value which governs the associations with far hits to a given arm is slowly decreased, which ensures that we avoid the local minima, and the fit is guided towards the global minimum, ultimately leading to the optimal vertex location.

3.4.3 Prong reconstruction

Once the interaction vertex has been determined, the next stage of reconstruction chain involves organizing the hits within the event into prongs, which are defined as directed final state particle clusters. Prongs serve as the foundation for subsequent analyses, which include energy estimation and particle identification.

The algorithm used for the Prong clustering is Fuzzy-K Means (FKM) clustering algorithm [77]. Here, Fuzzy-K means that a given hit may be connected to k number of clusters. FKM operates under the assumption that there is only one

vertex within the event, and all the hits within the cells are linked or associated with this single vertex. NOvA required modification of FKM as cluster center value k is not known in advance because the number of prongs can vary from one event to another and normalization done during FKM require connecting noise hits to at least one cluster. So, Possibilistic Fuzzy Clustering (PFC) algorithm [78] is used to solve these issues. The normalization requirement of FKM algorithm is not present in PFC [79] and it is able to find total number of clusters without us giving it the k value.

This algorithm is seeded by the angles that correspond to clusters of high cell activity. These angles are placed in a density matrix with 360 bins spanning from $-\pi$ to π with the interaction vertex at its origin. The beam path (positive z direction) is collinear to the 0 angle. The density matrix is given as follows

$$w_k = \sum_{h=1}^H e^{-\left(\frac{\theta_k - \theta_h}{\sigma_h}\right)^2} \quad (3.12)$$

$$\text{where } \theta_k = -\pi + \pi k / 180, 0 \leq k < 360 \quad (3.13)$$

The initial angles are calculated from the highest bins in the density matrix. Now, given a cell hit, its corresponding cluster is assigned by calculating cell membership (μ_{ij}):

$$\mu_{ij} = e^{-\frac{m d_{ij} \sqrt{a}}{\beta}} \quad (3.14)$$

$$\text{where } d_{ij} = \left(\frac{x_j - a_i}{\sigma_j}\right)^2 \quad (3.15)$$

Here a is the number of clusters, m is the degree of fuzziness of clusters in the slice, is set at 2 for NOvA, β is the spread of hits around cluster centers, is set at 4 for NOvA. The algorithm is iterated and cluster centers are updated until deviation between the old and new centers are less than 10^{-7} radians. The prong reconstruction of a event is shown in fig 3.14. Once the prongs are determined by

this algorithm, the prongs in the two views are matched to make 3D prongs which can now be used for analysis.

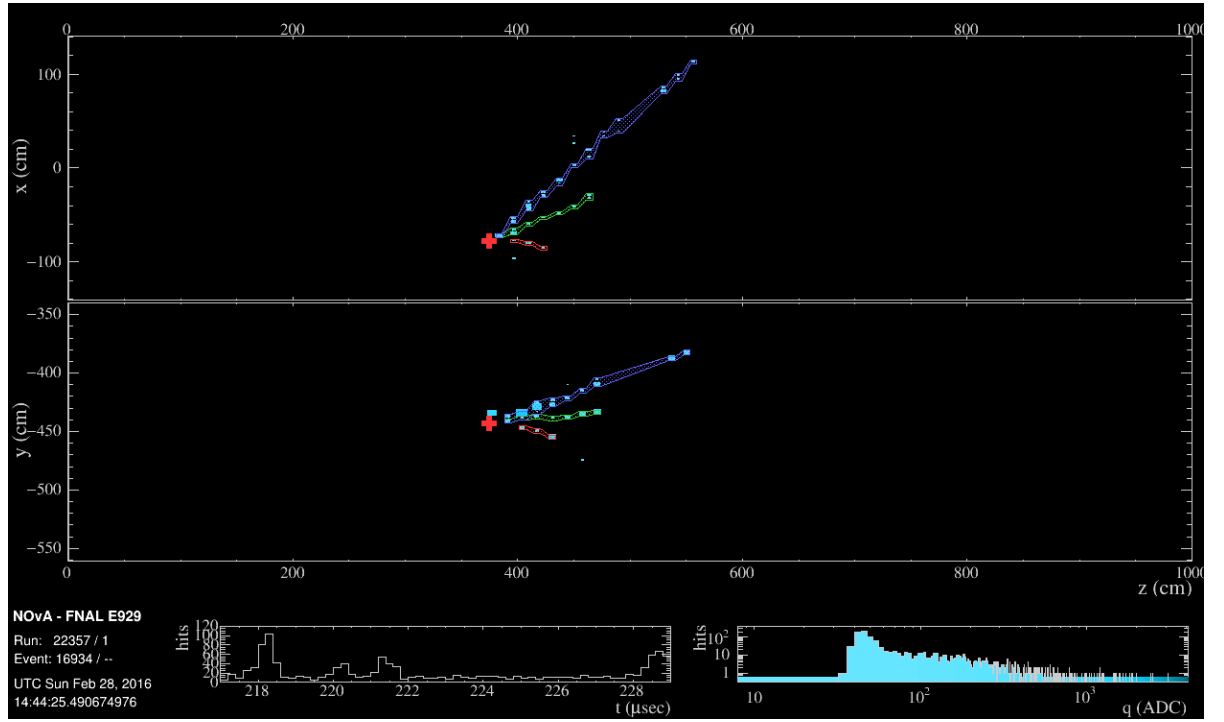


FIGURE 3.14: The prong reconstruction of event in fig.3.13 using FuzzyK algorithm. Each of the prong identified by FuzzyK algorithm are shown in different colors.

Chapter 4

Prong Convolutional Neural Network

The primary physics goal of NOvA is study of neutrino oscillations and to make precise measurements of θ_{23}, δ_{13} and Δm_{32}^2 . NOvA can also be used to study neutrino cross-section measurements and exotic physics. To achieve its goals, NOvA needs to have good neutrino interaction classification and neutrino energy measurement. Accurately categorizing the particle interactions as either signal or background is very important. The categorization was commonly done by reconstructing high-level components such as clusters, tracks associated with particle interactions recorded in the detector using traditional reconstruction methods. After reconstruction, we can summarize the directions, shapes, and energies of these particle interactions with a fewer number of quantities. In NOvA, we have to deal with a huge amount of data to classify the signal and background. This is where machine learning is very helpful.

This chapter gives a brief introduction to machine learning. Then, a detailed description of machine learning applications in the NOvA experiment is given, focusing mainly on the event and particle classifier.

4.1 Introduction to Machine learning

Machine learning, a sub-field of artificial intelligence, uses computational algorithms to enable computers to learn from data and make predictions or decisions without explicit programming. ML algorithms shine when huge volumes of training data is available as their performance is proportional to the size of data up to some extent. ML algorithms give better performance over the traditional methods in most of the cases as they extract the features in the data which cannot be gleaned through regular methods. ML algorithms are complex and require specialized hardware to speed up their learning. The internal workings of ML algorithms are complex and not easily interpretable, making it challenging to understand why a particular prediction or outcome was generated. Researchers are actively working on making machine learning more transparent and interpretable [80, 81].

4.1.1 Artificial neural networks

Artificial Neural Network (ANN) is a basic and popular ML algorithm. ANNs are computational models which draw inspiration from the workings of the human brain. ANNs learn from data through a process called training, where ANNs modify the strength of connections (weights) between neurons to optimize the network's ability to make predictions or classifications. The power of ANNs lies in their capacity to capture complex, non-linear relationships in data, making them well-suited for tasks such as classification and pattern recognition. These networks comprise interconnected nodes called neurons, arranged into layers, typically an input layer, one or more hidden layers, and an output layer as shown in the fig 4.1. These networks are fully connected, implying that every neuron in each layer is linked to every neuron in the subsequent layer.

The various components of an ANN are described here:

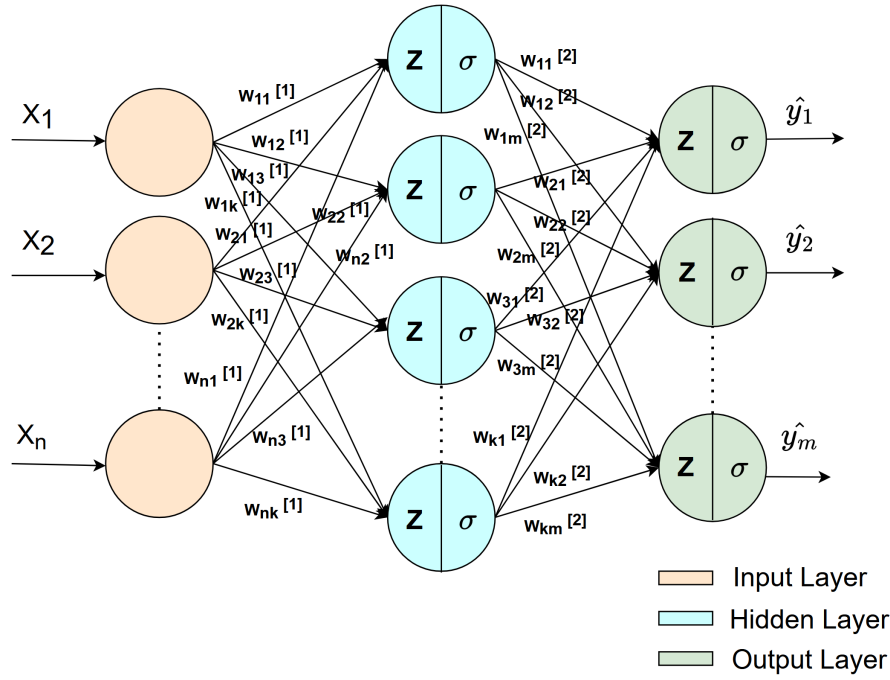


FIGURE 4.1: Schematic overview of an Artificial Neural Network.

- **Neuron:** This is the building block of an ANN. A neuron takes a set of inputs (x_i) and converts them into an output (y). This process involves multiplying each input by an internal adjustable weight (w_i) and summing these products with a learnable bias. This sum (Z) is sent to an activation function (σ), and the neuron's output is determined by the result of this activation function.

$$y = \sigma(Z) = \sigma\left(\sum_{i=1}^N w_i x_i + b\right) \quad (4.1)$$

- **Input Layer:** It is the initial layer of the ANN. Neurons in this layer receive input data features (x_i) and pass them on to the subsequent layers. The dimensions of input data dictates the number of neurons in the input layer.
- **Hidden Layers:** These are the intermediate layers positioned between the input and output layers. Every neuron in this layer processes information from the neurons of the previous layer and its output is passed to the next layer. The term “hidden” reflects that these layers are not directly connected to the input or output of the network. ANNs can have multiple hidden layers.

- **Output Layer:** The output layer serves as the final layer within the neural network. It generates the network's outputs or predictions. It synthesizes this information based on the processed information from the preceding layers. The number of neurons in the output layer varies depending on the task, such as regression, binary classification or multi-class classification, or regression.
- **Weights and Bias:** Every connection between neurons in adjacent layers carries an associated weight. Weights determine the strength of the connection and are adjusted during training to learn the relationships in the data. Learning the correct weights is a crucial part of training a neural network. In addition to weights, each neuron may have a bias term. Bias helps shift the activation function and allows the network to model more complex functions by controlling the threshold for activation.
- **Activation Functions:** They introduce non-linearity to the neural network. Non-linearity is important because most of the physics problems are not linearly separable and cannot be solved using linear functions. Some popular activation functions like sigmoid, softmax, rectified linear unit (ReLU), and hyperbolic tangent (tanh) are shown in fig 4.2.

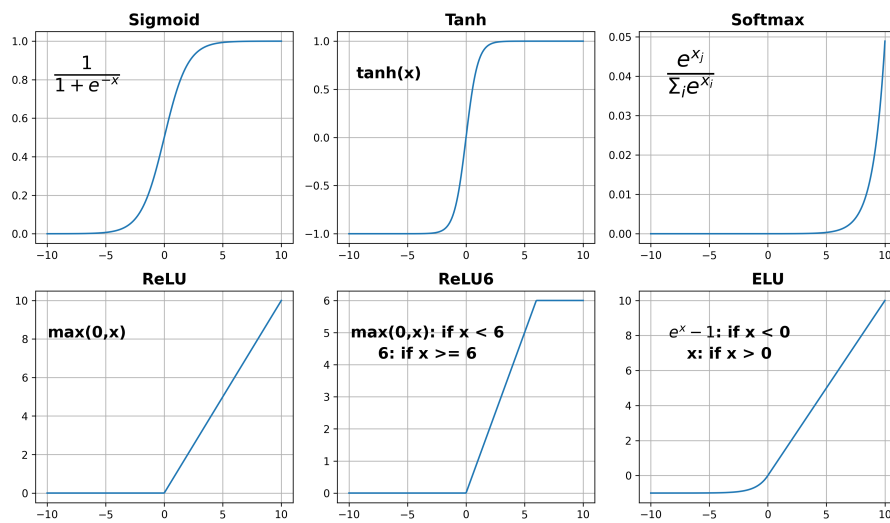


FIGURE 4.2: Activation Functions.

4.1.2 Training Procedure

Once the ANN architecture is decided, it is trained over the data. Before starting the training, all the values of parameters (weights and biases) in the model need to be initialized. Bad initialization can lead to issues like vanishing or exploding gradients, which hinder training. While proper initialization begets faster convergence and better generalization. Xavier/Glorot Initialization [82], He Initialization [83] and Orthogonal Initialization [84] are few popular initialization algorithms. In this thesis, Xavier/Glorot Initialization is the default initializer unless stated otherwise. Xavier/Glorot initialization sets the weights using a distribution that takes into account the number of input and output units for each layer. It helps maintain consistent variance in activations across layers, promoting stable training.

After initialization, a batch of input data is passed to the input layer and its output is passed forward layer by layer to the final layer. Neurons in every layer apply the weights on their inputs and the result is sent to an activation function, and send the output to the next layer. This is called Forward propagation. The initial output, we get from the forward propagation is quite random. We want the model to learn from the input data and give expected output. For this, we try to reduce the difference between the output of the network and the expected output through the loss functions.

The loss function (cost function) calculates the error between the network predictions and the expected values. The aim of training is to minimize the error, making the network's predictions as accurate as possible. Selection of a loss function hinges on the task at hand, type of data, and the model architecture. The model which is required to output a numerical value is called regression type. Mean Squared Error (MSE) Loss is well suited for regression.

$$\text{MSE loss} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (4.2)$$

Here, y_i and $\hat{y}_i$ are the true and predicted values respectively. n represents the number of samples in the batch of input data. MSE penalizes large errors heavily.

Different kind of loss functions are defined for classification problems. Binary Cross Entropy (BCE) loss shown in eqn 4.3 is used in binary classification problems.

$$\text{BCE loss} = -\frac{1}{n} \sum_{i=1}^n [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (4.3)$$

The first term of eqn 4.3 penalizes deviations when $y_i = 1$, and the second term does the same when $y_i = 0$. Minimizing this loss during training aligns predicted probabilities with true labels.

For multiple class classification problem, Categorical Cross Entropy (CCE) is used as the loss function. This loss function is generally used in tandem with One-hot encoding. One-hot encoding transforms categorical data into a binary format. Each class in the data is represented by a unique binary vector, where one element is set to 1 to indicate the category, and all other elements are set to 0. This ensures that categorical data can be used as input for ML training and prevents model from preferring the class with larger y_i values.

$$\text{CCE loss} = -\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^C y_{ij} \log(\hat{y}_{ij}) \quad (4.4)$$

Here, C is output classes. CCE loss penalizes deviations between predicted and actual class probabilities. The goal is to minimize this loss during training to align predicted probabilities with true class labels, enabling accurate multi-class classification.

There are a multitude of loss functions [85], each possessing its unique advantages. Custom build loss function can also be defined based on the task at hand.

Once the loss is calculated, it is utilized for updating weights and parameters in a way that the loss will be reduced. This is called back propagation. In back propagation, gradients of the loss function are computed with respect to the network's parameters, so that these parameters can be adjusted to minimize the loss during training.

The minimization of loss is done through optimizers. Optimizers are used to iteratively adjust model parameters toward minimizing the loss. Without an optimizer, it would be hard to find the optimal set of parameters manually, especially in complex models. Stochastic Gradient Descent (SGD) is a popular and most commonly used optimizer to train models. SGD is an extension of the regular gradient descent but it processes data in mini batches rather than the entire dataset at once. The mini batches are selected randomly from the total data, introducing the stochastic component. This stochastic nature introduces randomness into the parameter updates, which can help escape local minima and explore the parameter space more effectively. For each mini-batch, SGD computes the gradient of the loss function with respect to the model parameters. It then updates the model parameters in the opposite direction of the gradient to minimize the loss. The update rule for a parameter θ is:

$$\theta = \theta - \alpha \cdot \nabla L(\theta) \quad (4.5)$$

Here, $\nabla L(\theta)$ is the gradient of the loss function with respect to θ and α is the learning rate, a hyperparameter that controls the step size while updating parameters. A hyperparameter is a configuration setting external to the model that influences its training process and performance, often set before the training and not learned from data.

SGD with Nesterov Accelerated Gradient optimizer was used while training the Prong CNN described in this thesis. It is an extension of the SGD algorithm that incorporates the concept of momentum for faster convergence. The standard momentum term helps model parameters continue moving in their current direction, allowing them to build up speed and overcome small fluctuations in the training data. In standard case, the momentum term is added after computing the gradient based on the current position, which can lead to overshooting the minimum. Nesterov's approach [86] computes the gradient slightly ahead of the current position, allowing the momentum term to adjust the position in a more accurate direction. This anticipatory update can prevent overshooting. The update rule for the momentum term is as follows:

$$v_t = \mu.v_{t-1} + \alpha.\nabla L(\alpha t) \quad (4.6)$$

Here, v_t the momentum term at time step. μ is the momentum coefficient, a hyperparameter typically set between 0 and 1. It determines the fraction of the previous momentum term that is retained and added to the current update. $\nabla L(\alpha t)$ is the gradient of the loss function with respect to the model parameters θ_t at time step t . The iterative process of feed forward and back propagation while optimizing the model parameters is called training. This can be repeated to get better performance from the model.

4.1.3 Convolutional neural networks

NOvA uses tracking calorimeters as detector, record energy deposited in the PVC cells throughout the detector area. These records can be viewed as images which shows the physics of the interaction as seen in fig.2.9. We can classify the events through these images using ML algorithms, bypassing the need for reconstruction. While some individual variations may exist between events, ML algorithms will be able to extract shared similarities between the events and use them to identify

various types of events and particles. But, It is not advisable to use the standard NNs described above. This is because NNs treat each input pixel as independent point, which doesn't capture the spatial relationships and hierarchies present in images. The solution to this issue comes in form of the Convolutional neural networks (CNNs).

CNNs were developed as part of computer vision task of making computers to see and understand the world around them. CNNs were specifically designed to address the challenges associated with processing and understanding visual data, such as images and videos. The development of CNNs marked a significant breakthrough in computer vision and has since revolutionized the field.

A basic CNN model as shown in the fig 4.3 contains layers like the Convolutional layer, Pooling layer, fully connected layers, each with specific role for feature extraction and training the model.

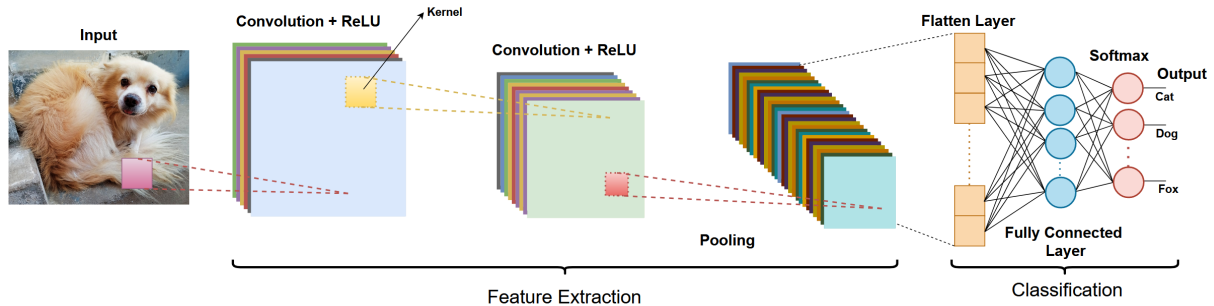


FIGURE 4.3: Schematic overview of a basic CNN architecture.

Convolutional layers are the most important blocks of a CNN. Convolutional layer applies convolution on its inputs. Convolution involves applying a set of learnable filters known as kernels to the input data. The kernels are matrices of weights. The kernels strides across the input data by taking dot products at different positions. The result of each dot product produces a single value in the output, which forms a feature map. The stride of filter is a hyperparameter whose value can be customized, it determines how many pixels the filter moves when sliding across the input as shown in the fig 4.4.

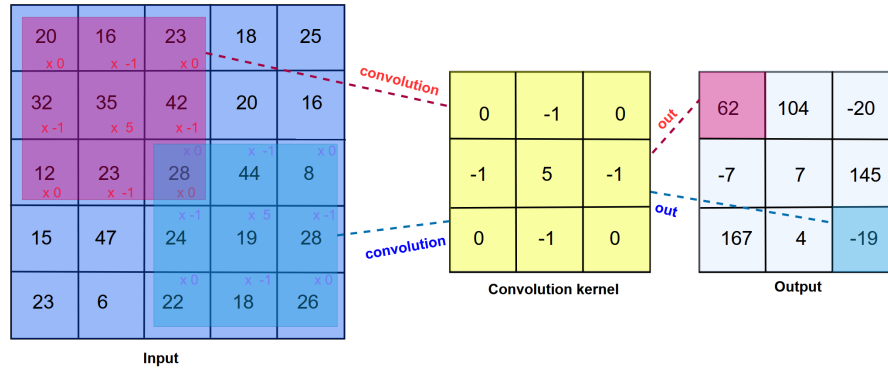


FIGURE 4.4: Operation process of a Convolutional layer.

A feature map is a grid of values arranged in two dimensions. Each value in the feature map represents the degree of similarity between the filter and a local region of the input data. The feature maps capture local patterns and features in the input, such as edges, textures, or specific shapes. Convolutional layers generally consist of multiple filters, each specialized in detecting different features. These filters are learned during the training process, allowing the network to adapt to the specific patterns in the data. The output of a convolutional layer is a set of multiple feature maps, one for each filter.

Main advantage of convolutional layers is parameter sharing. Each filter is used across the entire input, which reduces the number of learnable parameters compared to fully connected layers. Parameter sharing makes CNNs more efficient and helps capture translation-invariant features. General CNN architectures typically consist of multiple convolutional layers stacked on top of each other. In the initial layers, these networks identify basic features like edges or high-frequency patterns, capturing fundamental visual elements. As layers progress deeper into the network, they capture more abstract and high-level features. These features often do not correspond to attributes identified by traditional reconstruction techniques. Instead, they extract abstract features which carry a higher level of classification potential. This hierarchical extraction allows the CNN to gather the essential information required to efficiently tackle a specific task, adapting its understanding

of the input data as it moves through the layers. This improves the network's performance.

The pooling layer is an important layer employed in CNNs. The pooling operation can be applied independently to each feature map generated by the convolutional layers. It operates on small windows of the feature map. Two common types of pooling are:

- **Max Pooling:** The maximum value within each pooling window is retained in the output feature map. Max pooling is effective at preserving prominent features and edges.
- **Average Pooling:** The average value of the values within each pooling window is computed and placed in the output feature map. Average pooling can help reduce noise in the data.

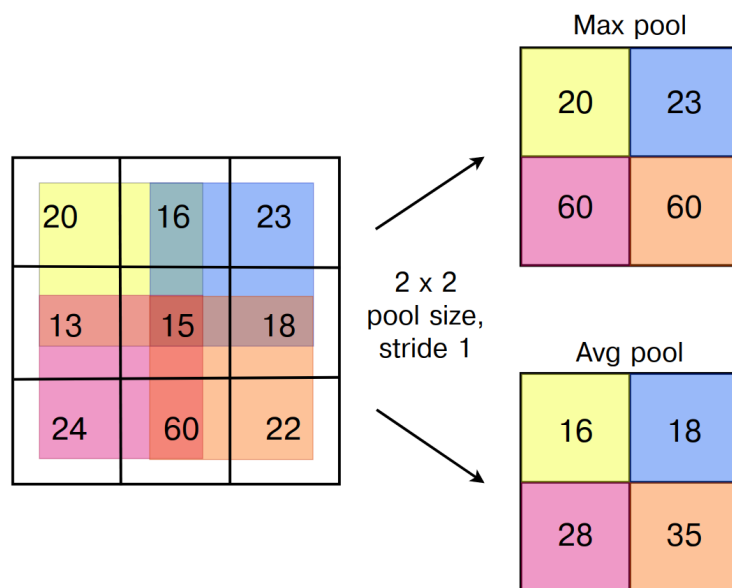


FIGURE 4.5: Operation process of Max Pooling and Average Pooling layers.

The hyperparameters, pooling window size and the strides can be varied according to the needs of the model. Pooling can help reduce the computational cost of the network by down scaling the feature maps. Low-level features are captured by early convolutional layers, while higher-level features are captured by later layers

with reduced spatial dimensions after pooling. The pooling layers also provide translation-invariance ,i.e., that the specific location of features in the input data becomes less important. This property helps CNNs recognize patterns regardless of their precise position in the input.

A batch normalization (BatchNorm) layer used in deep neural networks(DNNs) to stabilize and accelerate the training of DNNs. This layer normalizes the input of a layer by adjusting and scaling, it based on the mean and variance learnt across a batch of data. It ensures that the inputs to each layer in a neural network are centered and have a consistent variance, which aids in faster convergence during training.

The fully connected (Dense) Layer is used as the final layer of the CNN. The primary reason for using dense layers is to capture global patterns and relationships across the entire feature space. The input to the dense layer must be flattened from the two-dimensional feature maps into a one-dimensional vector. This flattening helps us connect each neuron in the dense layer to all neurons in the previous layer. This dense layer is responsible for making predictions. In image classification tasks, it produces class probabilities for each class. The activation function used in the final layer is decided based on the problem at hand. For example, in multi-class classification, the softmax function is often used to compute class probabilities.

Normal CNNs with a limited number of layers struggle to capture intricate details and patterns in images. This can be countered by adding depth to the model,i.e., more layers. This type of models are called Deep CNNs, excel at representing images by learning a hierarchy of features, enabling superior image analysis and understanding. NOvA uses Deep CNNs for classification of events and particle tracks.

4.2 Machine learning in NOvA

In 2016, NOvA was the first HEP experiment to use a Convolutional Neural Network (CNN) in a physics measurement to classify candidate neutrino interactions. This decision resulted in a significant 30% rise of effective exposure over the conventional classification methods [87].

4.2.1 Event CNN

NOvA uses Event CNN which takes full event pixel map of both x and y views as input and the output is an event classifier. Pixel maps represent 2-dimensional arrays where each detector cell corresponds to the element in array. These maps encompass a total of 200 planes, with each view (x,y) containing 100 planes, and each view has a height of 80 cells. The plane containing the first hit of the event is taken as the first plane within the pixel map, and the subsequent 8 planes should register more than 4 hits. This safeguards against premature initiation of the pixel map caused by outlier hits located upstream of the event's core. While beginning the pixel map with the reconstructed vertex seems like a good idea, it is not advisable as it introduces unwanted dependencies from vertex reconstruction into the network.

The array values are computed based on the calibrated energy associated with each hit. These energy values are scaled such that they are able to be represented as 8 bit unsigned characters with values ranging from 0 to 255, to enhance data storage and transfer efficiency during the training phase. Hits surpassing 278 MeV are capped to a maximum value of 255. This approach shrinks the pixel map's memory size by eight times compared to using floating-point representation [88], all the while retaining an acceptable precision level as shown in fig 4.6.

Event CNN can classify the events into 4 classes:

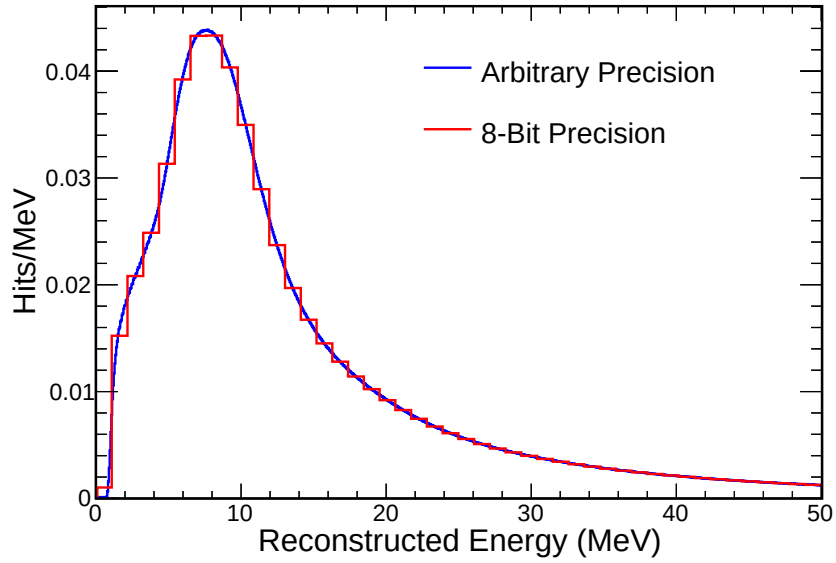


FIGURE 4.6: The plot shows the comparison between the energy spectrum of individual hits employing 8 bit precision binning with only 256 bins and an arbitrary fine binning.

1. ν_μ CC interaction
2. ν_e CC interaction
3. NC interaction
4. Cosmics

While Event CNN can classify events, identification of final state particles of an event is needed to better our energy reconstruction and enable cross-section measurements of final states. Context enriched Prong CNN [89] is one of the networks being trained with a goal to identify all the final state particles ($e^\pm, p^+, \mu, \pi^\pm, \gamma$) of a given neutrino event.

4.3 Prong CNN

Prong CNN uses a four-tower Siamese-type MobileNetV2 architecture to include context information i.e it takes both event (Context) views and prong (Independent) views. The position of hits within the slice determines the pixel map's

boundaries, while the hit energy values are scaled and filled into 100×80 arrays. These arrays are filled with hits of the specific prong of interest for both x and y views. Furthermore, for the entire event's X and Y-views, the same pixel maps utilized in event classification are employed. This context enables the network to grasp the context of particle interactions. Context provided to the network assists it to learn from the relative topology of the prong inside the event pixel map and implied conservation laws from physical quantities.

The earlier iteration of Prong CNN used by NOvA trained on neutrino (FHC) and anti-neutrino(RHC) dataset separately. The new Prong CNN, we developed and discussed in this thesis, is trained on a combined (FRHC) neutrino and anti-neutrino dataset. We trained these networks using TensorFlow 1.12.0, Keras 2.2.4.

4.4 Training Datasets:

The dataset we used for training the network is taken from the FD MC simulated data. Preprocessing data is a must before starting training. The prong and the event selected must be fully contained within detector boundaries so that no energy is lost outside the detector. A purity cut was applied to ensure that we get clusters with identifiable prongs that are realistic in appearance, while not removing too much of the training sample. These cuts are unchanged from the previous version of prong CNN: 40% for electrons, 50 % for photons and muons and 35% for protons and pions. Prongs with less than 2 hits in any x or y view are removed. Finally, prongs exceeding 5 m in length were excluded as they are predominantly muons to reduce overall computation time.

After making the above mentioned cuts we have following number of prongs for each dataset:FHC had 6.80 M, RHC had 5.97 M prongs, Combined: 12.8 M.

This training dataset is unbalanced, because the chosen sample tends to lack representation of particles generated less frequently in neutrino interactions, specifically

Selection cuts	Description
Containment Cut	Selects the prong and the event contained within the detector boundaries
Cosmic Veto	Removes cosmic events
Cell Hits	No of Cell hits > 2 for both x and y views
Purity Cut	Realistic looking cluster with prongs $\mu^\pm, \gamma$: 50 %, $e^\pm$: 40 %, $\pi^\pm, p^+$: 35 %
Prong length	Cut prongs with prong length more than 5 m

TABLE 4.1: Various Cuts used for preprocessing the data for training.

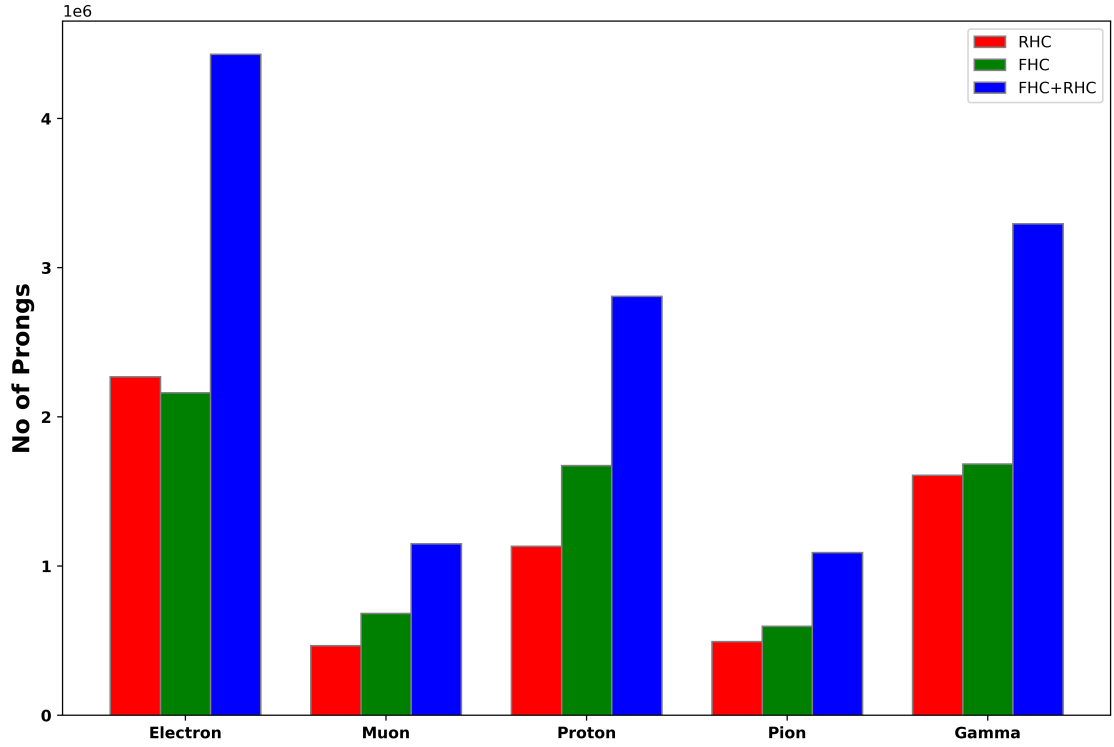


FIGURE 4.7: No: of prongs belonging to different classes after initial data preprocessing cuts are applied.

charged pions. one needs to train the network on a balanced dataset to make sure that network does not learn bias towards the class with more samples. Since, our training dataset is unbalanced, it is later balanced to contain an equal number of each type of particle and Horn current type. After balancing dataset, we have a sample of 4.6 M prongs for training and 0.4 M for evaluation.

4.5 Architecture

Previous version of Prong CNN[90] used MobileNetv2 [91] architecture. This architecture details can be found at TensorFlowTrainingUtils repository in NOvA GitHub.

Mobilenets are family of efficient CNN architectures which are known for being computationally lightweight. MobileNets achieve their efficiency by utilizing depth-wise separable convolutions [92]. Instead of the traditional approach where a standard convolution operates on both the depth and spatial dimensions of the input tensor simultaneously, depth-wise separable convolutions split this into two separate layers: depth-wise convolutions and point-wise convolutions. This separation significantly reduces the computational cost while preserving accuracy to a large extent.

The depth-wise convolutions perform convolution independently for each input channel as shown in fig 4.8. The output and input of this convolution have the same size. Then, a batch norm layer and activation function is applied. This is passed to the point-wise convolution in which, a 1×1 convolution is used to convolve all the channels as shown in fig 4.8. Again, a batch norm layer is applied before passing information to next layer.

The new Prong CNN developed by us, uses a modified MobileNetv2 architecture adding in new features from MobileNetv3 [93]. The overall number of layers of Prong CNN were reduced and the composition and quantity of convolutional layers in each layer was optimized to minimize run time on CPUs without loss of performance. When this new architecture was adapted for Prong CNN case, the number of parameters of the network went down by a quarter ($\sim 4\text{M}$ to $\sim 1\text{M}$). This new modified Prong CNN is computationally cheaper and is able to get to higher classification efficiency in fewer epochs of training.

New features introduced in new Prong CNN network are listed below:

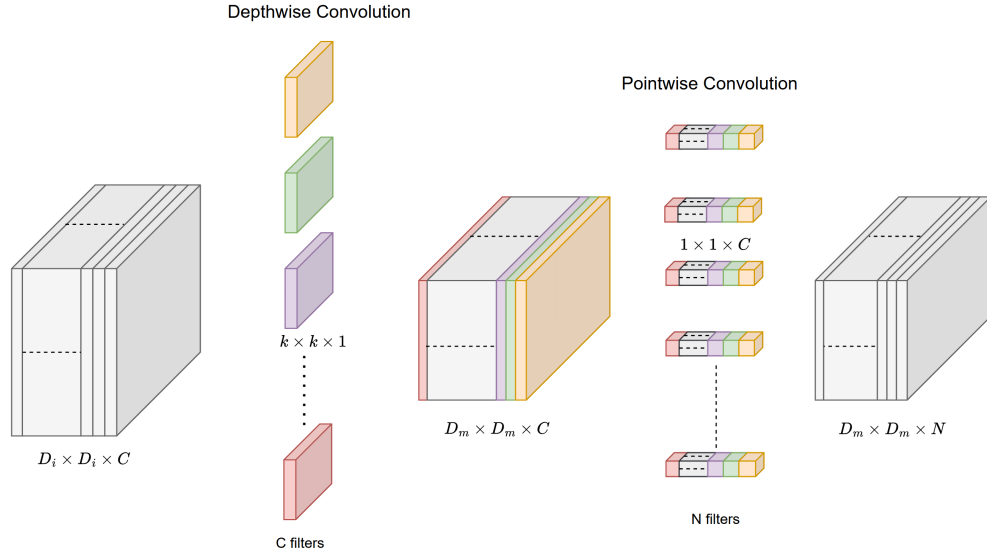


FIGURE 4.8: Depth-wise separable Convolution.

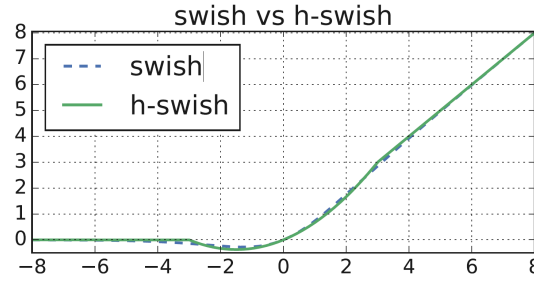


FIGURE 4.9: Comparison between Swish and Hard Swish.

1. New type of Activation - hard swish
2. Squeeze excite block

A activation function called Swish introduced in [94] is used as a replacement for ReLU, it is found to significantly improves the accuracy of neural networks.

$$swish(x) = x \cdot sigmoid(x) \quad (4.7)$$

Since, sigmoid function which is expensive to compute, it is replaced with with its piece-wise linear hard analog ReLU shown below.

$$hswish(x) = \frac{ReLU6(x + 3)}{6} \quad (4.8)$$

Squeeze and excite (SE) block was introduced in Squeeze-and-Excitation Networks[95]. SE block learns the channel interdependencies and then scales the output of the layer proportional to the channel importance. This gives more weight to channels with greater importance. Many SE blocks are later proposed trying to improve channel attention while reducing extra computation required. We found Efficient Channel Attention (ECA) block a good replacement to standard SE block. An ECA block [96] is just like an SE block. It includes a squeeze module for collecting global spatial information and an efficient excitation module for learning cross-channel interaction. Unlike SE block which considers correspondence between all the channels, excitation module in an ECA block emphasizes on correspondence between each channel and its k-nearest neighbors. This approach is adopted to regulate model complexity. The value of k is proportional to number of channels and it is calculated as:

$$k = \psi(C) = \left\lfloor \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right\rfloor_{\text{odd}} \quad (4.9)$$

In our case, γ and b are taken as 2 and 1 respectively. The networks were trained on the The Fermilab Wilson cluster Institutional Cluster at Argonne National Lab. SHERPA was used to optimize the hyperparameters. The training is carried out using framework from TensorFlowTrainingUtils repository.

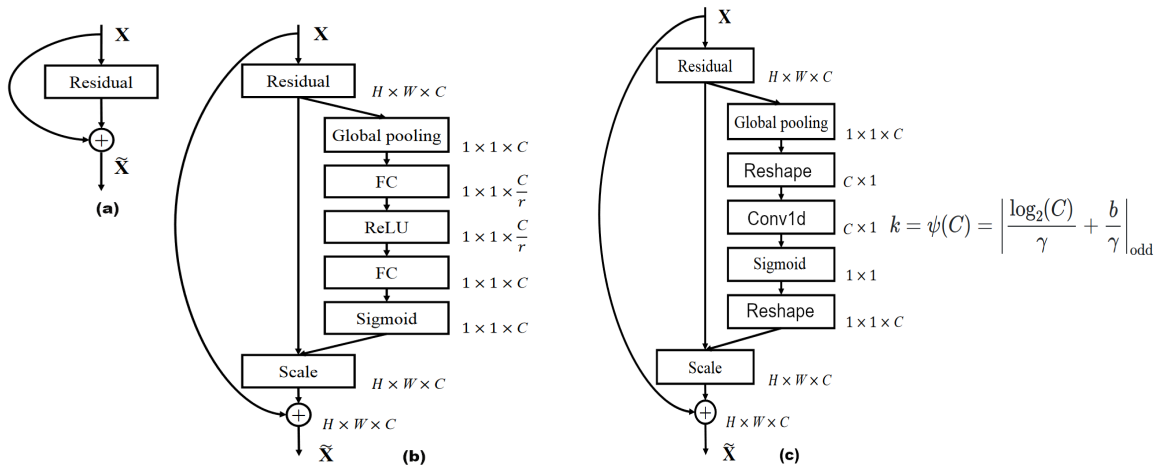


FIGURE 4.10: (a) Residual connection, (b) Residual connection with SE Block, (c) Residual connection with ECA Block

Architecture of this new Prong CNN is shown in block diagram form in figure 4.11.

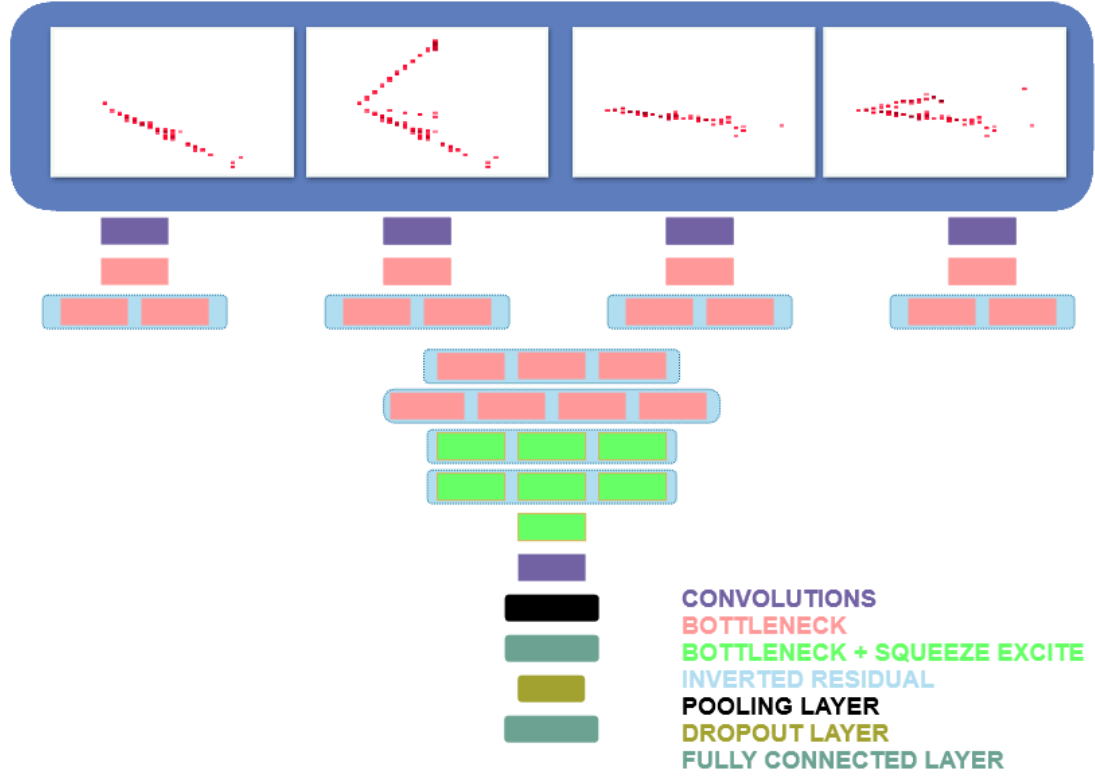


FIGURE 4.11: Block diagram of New Prong CNN Architecture based on modified MobileNetV2 architecture, adding Efficient channel attention block for improved channel attention.

4.6 Performance Metrics

First in the agenda was to compare network trained on FHC+RHC combined to separately trained networks. The preliminary study showed that double FHC + RHC network is comparatively better than RHC and FHC networks trained separately. Thus, we continued training network on combined dataset.

The comparison between performance of new prong CNN and old model on FD MC dataset is shown below using the confusion matrix of efficiency and purity in figures 4.12 and 4.13 respectively. ROC curves (receiver operating characteristic curve) of New Prong CNN and old model are shown in figure 4.14, in it AUC stands for Area under the curve, greater it is better the performance. The new network is up to 82% efficiency for all particles while having 86% total efficiency.

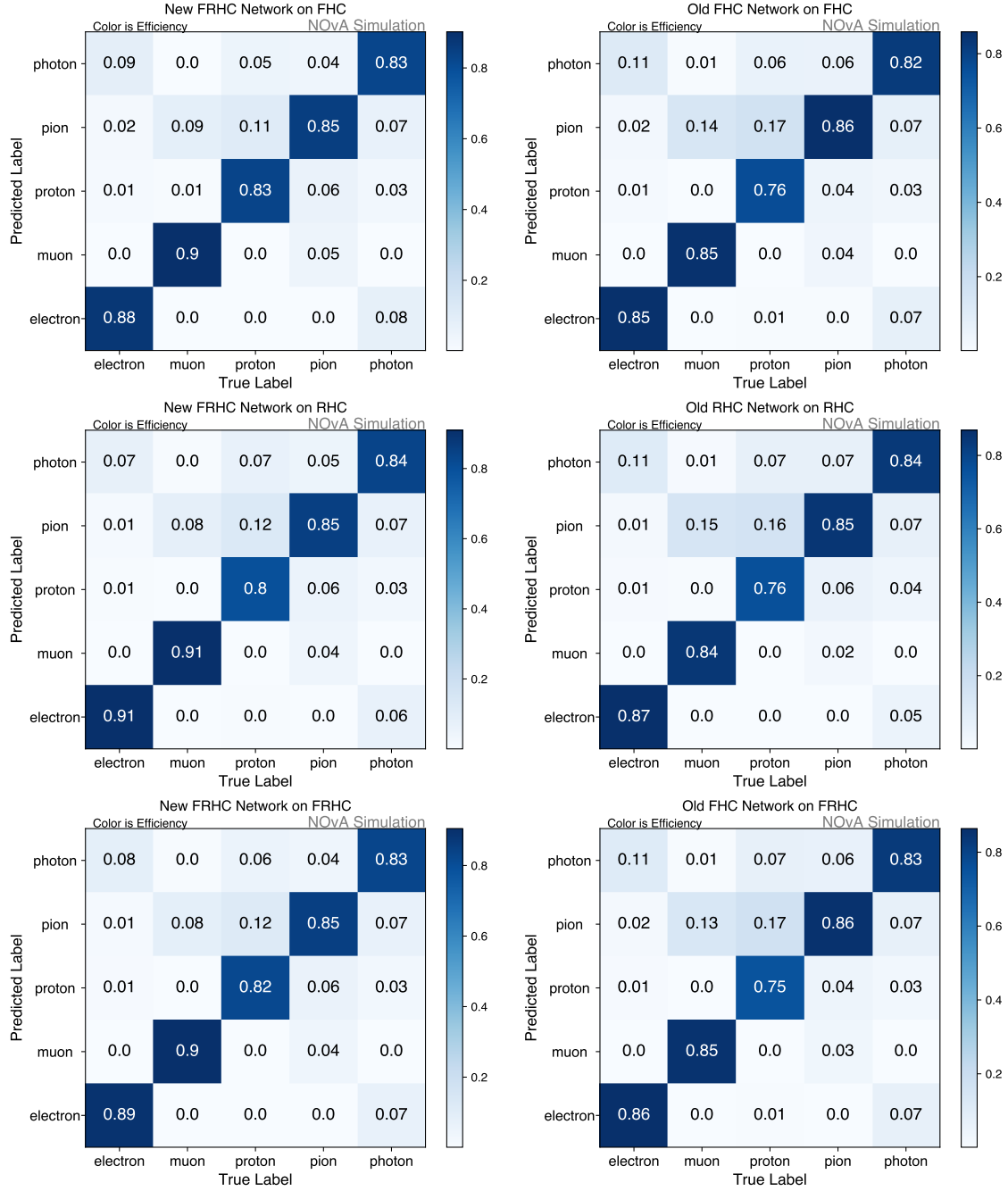


FIGURE 4.12: Classification matrices showing the comparison between performance of new prong CNN and old model on FD MC. They are normalized to show the efficiency of each category along the diagonal. The off-diagonal shows how each category is misclassified.

In conclusion, New version of Prong CNN is trained on FD MC dataset with modified MobileNetV2 architecture adding Efficient channel attention block for improved channel attention. New prong network is quarter in size compared to previous model but performs better. New Prong Network improved the classification efficiency by 3% compared to previous Prong CNN.

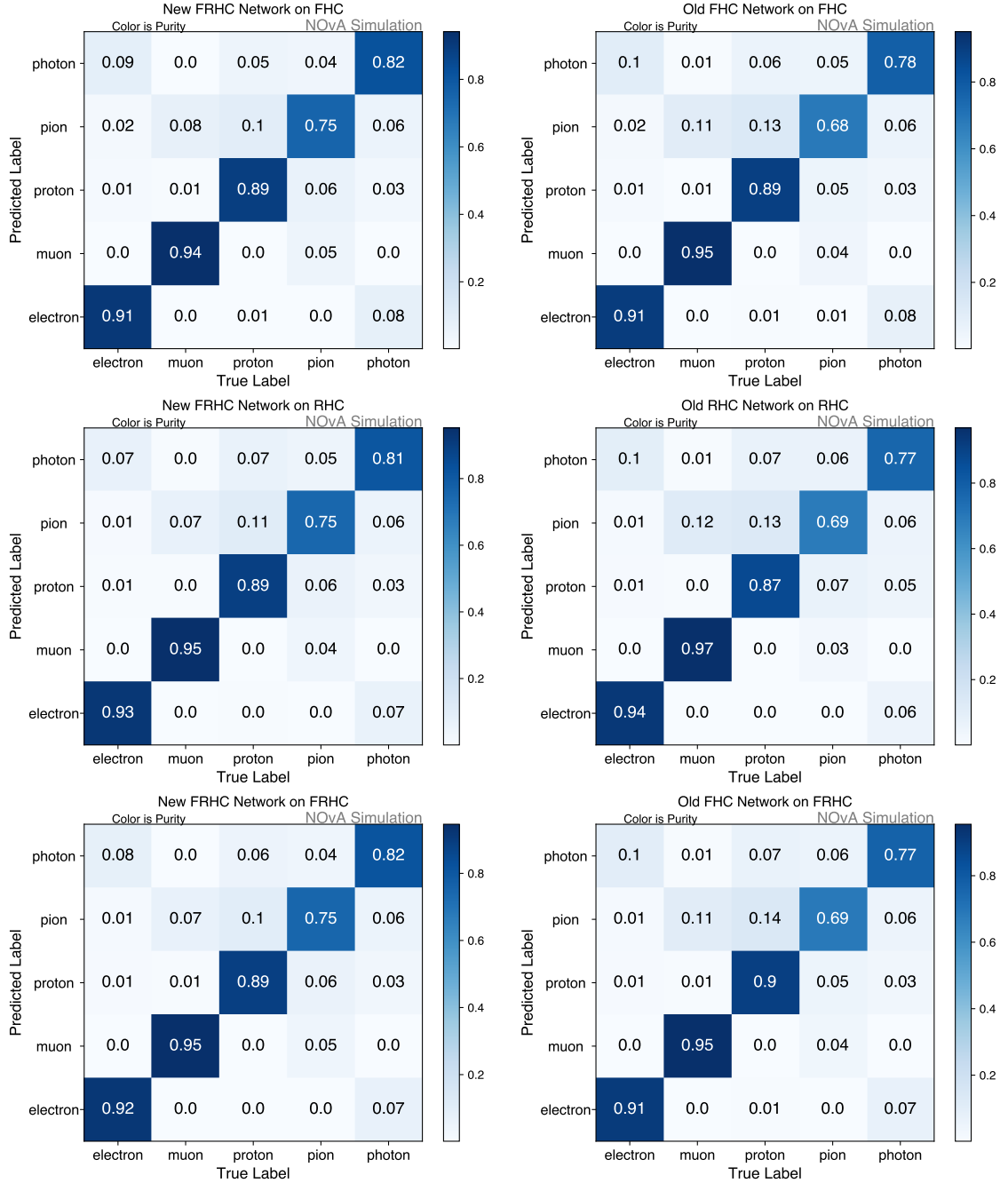


FIGURE 4.13: Classification matrices showing the comparison between performance of new prong CNN and old model on FD MC. They are normalized to show the purity of each category along the diagonal. The off-diagonals show the common backgrounds to each category.

4.7 Application of Prong CNN

4.7.1 Calibration of π^0 mass peak

ML algorithms trained on specific datasets might sometimes memorize patterns within that dataset rather than learning the underlying physics. Performance of

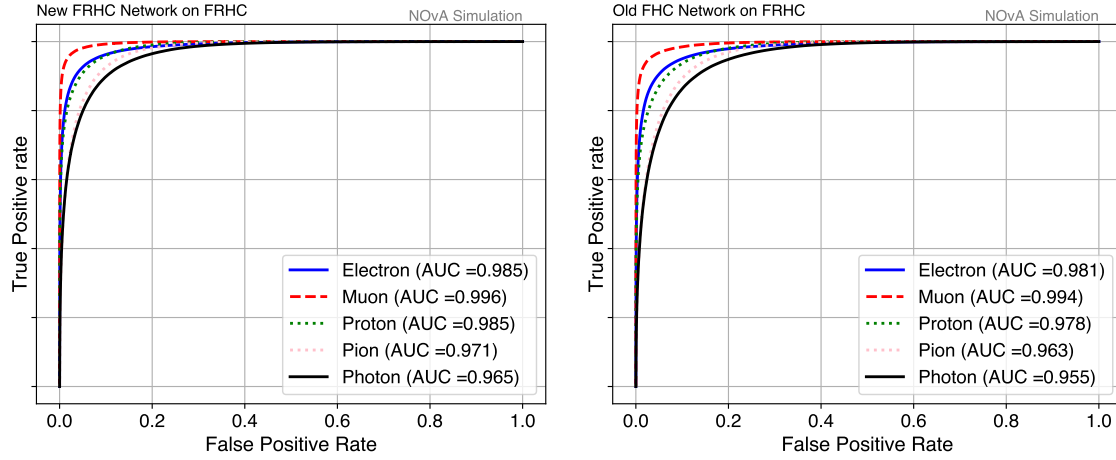


FIGURE 4.14: The ROC curves of each class showing the comparison between performance of new prong CNN and old model on FD MC.

Model	Efficiency(%)	Purity(%)
Eval on FRHC	on FD MC	FD MC
FRHC	86.0	86.4
Old Network	83.0	84.3

TABLE 4.2: Comparison of efficiency and purity metrics between new and old prong CNN.

Prong CNN like any ML algorithms must be compared between actual data and MC dataset to make sure that it is not learning features distinct to MC dataset it trained on. The π^0 mass peak serves as the standard candle in particle physics to assess detector response. In the primary decay channel of π^0 , it decays into two photons, resulting an invariant mass peak at 135 MeV, which is the mass of π^0 . This, photon-derived mass peak is employed within NOvA as a validation measure for both the absolute energy calibration and event reconstruction processes.

In NOvA experiment, π^0 originates at the neutrino interaction vertex, swiftly decaying, roughly 8.4×10^{-17} seconds later, into two photons in close proximity to the vertex. Subsequently, these photons travel an average distance of one radiation length before initiating an electromagnetic shower through pair production.

The invariant mass $M_{\gamma\gamma}$ of π^0 can be calculated from the energy from the photons it decays into using

$$M_{\gamma\gamma} = \sqrt{2E_{\gamma 1}E_{\gamma 2}(1 - \cos \theta)} \quad (4.10)$$

Here, θ is the angle between the photons and $E_{\gamma 1}, E_{\gamma 2}$ are the energies of the first and second photons.

The detector response study is done using π^0 events from the ND. The events selected must be fully contained within ND boundaries, except muon catcher so that no energy is lost outside the detector. The events must have only two reconstructed prongs, each with photon score from Prong CNN greater than 0.75. The two prongs must not have more than 4 consecutive planes of hits.

The process to determine a photon's energy involves summing up the total calibrated energy of all hits within the prong which gives us visible energy. But, the whole photon's energy is not converted to visible energy as some of the energy is deposited in the detector plastic. The average dead material within the detector is estimated at estimated at 35.8% based on a simulation of photon showers. So, in order to correct for the energy deposited in the dead material, the measured visible energy must be adjusted by a factor of 1.56. Meanwhile, the cosine of the angle between the prongs can be calculated by taking the dot product of the unit vectors that represent the directions of the two prongs.

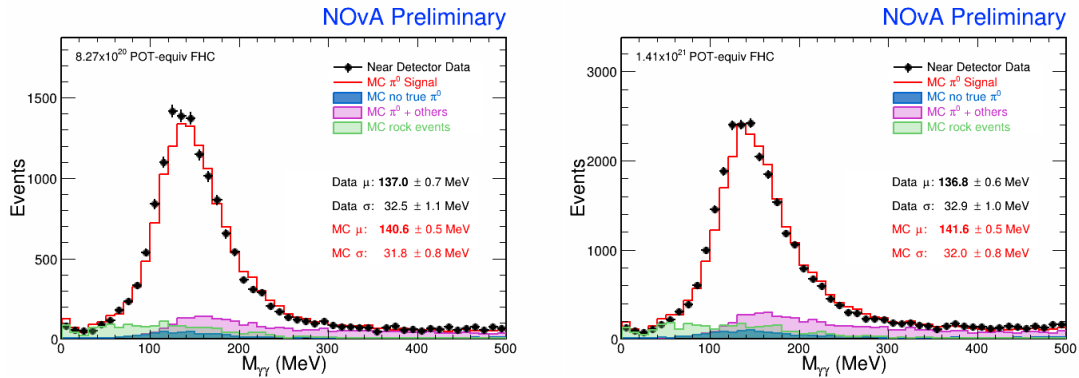


FIGURE 4.15: The reconstructed π^0 invariant mass using selection based on new(old) Prong CNN is on the left(right). The invariant mass is reconstructed using ND FHC. The MC is scaled to the data POT for both plots.

The reconstructed invariant mass of π^0 can be seen in fig 4.15, 4.16. Here, the events with $M_{\gamma\gamma} > 175$ MeV are mainly because of the events with hadronic activity clustered into the photon prongs.

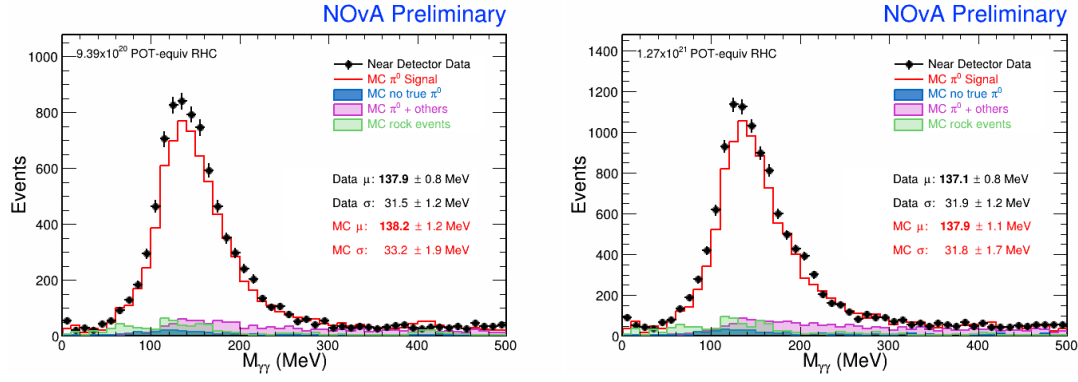


FIGURE 4.16: The reconstructed π^0 invariant mass using selection based on new(old) Prong CNN is on the left(right). The invariant mass is reconstructed using ND RHC. The MC is scaled to the data POT for both plots.

In all four plots above 4.15, 4.16, a small discrepancy between data and MC mass peak can be noticed. This discrepancy is about 2.6% in the new Prong CNN selection using ND FHC and about 3.5% in the old Prong CNN selection using ND FHC. The same can be seen in ND RHC case 4.16, discrepancy is about 2.1% in the new Prong CNN selection and about 5.8% in the old Prong CNN selection. We can see that the new Prong CNN have better agreement with data and MC when compared to old Prong CNN.

Chapter 5

Degeneracy resolution capabilities of NOvA and DUNE in the presence of light sterile neutrino

In addition to developing prong CNN for the NOvA experiment, a phenomenological study in the framework of 3+1 sterile neutrino model was done to check the effect of sterile neutrino on degeneracy resolution capabilities of NOvA and DUNE experiments. This chapter describes that study.

5.1 Introduction to sterile neutrinos

Sterile neutrinos are hypothetical particles that do not interact via any of the fundamental interactions other than gravity. The term sterile is used to distinguish them from active neutrinos, which are charged under weak interaction. The theoretical motivation for sterile neutrino explains the active neutrino mass after spontaneous symmetry breaking, by adding a gauge singlet term (sterile neutrino) to the Lagrangian under $SU(3)_c \otimes SU(2)_L \otimes U(1)_r$ where the Dirac term appears through the Higgs mechanism, and Majorana mass term is a gauge singlet, and

hence appears as a bare mass term[97]. The diagonalization of the mass matrix gives masses to all neutrinos due to the See-Saw mechanism.

Some experimental anomalies also point towards the existence of sterile neutrinos. Liquid Scintillator Neutrino Detector(LSND) detected $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$ transitions indicating $\Delta m^2 \approx 1eV^2$ which is inconsistent with $\Delta m_{32}^2, \Delta m_{21}^2$ (LSND anomaly)[98]. Measurement of the width of Z boson by LEP gave number of active neutrinos to be 2.984 ± 0.008 [99]. Thus the new neutrino introduced to explain the anomaly has to be a sterile neutrino. MiniBooNE, designed to verify the LSND anomaly, observed an unexplained excess of events in low-energy region of $\bar{\nu}_e, \nu_e$ spectra, consistent with LSND [100]. SAGE and GALLEX observed lower event rate than expected, explained by the oscillations of ν_e due to $\Delta m^2 \geq 1eV^2$ (Gallium anomaly)[101–103]. Recent precise predictions of reactor anti-neutrino flux has increased the expected flux by 3% over old predictions. With the new flux evaluation, the ratio of observed and predicted flux deviates at 98.6 % C.L(Confidence level) from unity, this is called “Reactor anti-neutrino Anomaly”[104]. This anomaly can also be explained using sterile neutrino model.

Short-baseline(SBL) experiments are running to search for sterile neutrinos. SBL experiments are the best place to look for sterile neutrino, as they are sensitive to new expected mass-squared splitting $\Delta m^2 \simeq 1eV^2$. However, SBL experiments cannot study all the properties of sterile neutrinos, mainly new CP phases introduced by sterile neutrino models. These new CP phases need long distances to become measurable[105, 106], thus can be measured using Long baseline(LBL) experiments. With the discovery of relatively large value for θ_{13} by Daya Bay[107], the sensitivity of LBL experiments towards neutrino mass hierarchy and CP phases increased significantly. In this context, some phenomenological studies regarding the sensitivity of LBL experiments can be found in recent works[108–112]. Using recent global fits of oscillation parameters in the 3+1 scenario[113], current LBL experiments can extract two out of three CP phases (one of them being standard

δ_{13})[106]. The phenomenological studies of LBL experiments in presence of sterile neutrino is studied by several groups[114–119]. Now, the sensitivity of LBL experiments towards their original goals decreases due to sterile neutrinos. It is seen in case of the CPV measurement; new CP-phases will decrease the sensitivity towards standard CP phase (δ_{13}). This will reduce degeneracy resolution capacities of LBL experiments. In this chapter, we study hierarchy- θ_{23} - δ_{13} degeneracies using contours in θ_{23} - δ_{13} plane and how they are affected by the introduction of sterile neutrinos. We attempt to find the extent to which these degeneracies can be resolved in future runs of NOvA and DUNE.

5.2 3+1 Sterile Neutrino Framework

In a 3+1 sterile neutrino model, the flavour and mass eigenstates are connected through a 4×4 mixing matrix. A convenient parameterization of the mixing matrix is [120]

$$U = R_{34}\tilde{R}_{24}\tilde{R}_{14}R_{23}\tilde{R}_{13}R_{12}. \quad (5.1)$$

where R_{ij} and $\tilde{R}_{ij}$ represent real and complex 4×4 rotation in the plane containing the 2×2 sub-block in (i,j) sub-block

$$R_{ij}^{2 \times 2} = \begin{pmatrix} c_{ij} & s_{ij} \\ -s_{ij} & c_{ij} \end{pmatrix} \quad \tilde{R}_{ij}^{2 \times 2} = \begin{pmatrix} c_{ij} & \tilde{s}_{ij} \\ -\tilde{s}_{ij}^* & c_{ij} \end{pmatrix} \quad (5.2)$$

Where, $c_{ij} = \cos \theta_{ij}$, $s_{ij} = \sin \theta_{ij}$, $\tilde{s}_{ij} = s_{ij}e^{-i\delta_{ij}}$ and δ_{ij} are the CP phases.

There are three mass squared difference terms in 3+1 model- Δm_{21}^2 (solar) $\simeq 7.5 \times 10^{-5} \text{eV}^2$, Δm_{31}^2 (atmospheric) $\simeq 2.4 \times 10^{-3} \text{eV}^2$ and Δm_{41}^2 (sterile) $\simeq 1 \text{eV}^2$. The mass-squared difference term towards which the experiment is sensitive depends

on L/E of the experiment. Since SBL experiments have small a very small L/E, $\sin^2(\Delta m_{ij}^2 L/4E) \simeq 0$ for Δm_{21}^2 and Δm_{31}^2 . Δm_{41}^2 term survives. Hence, SBL experiments depend only on sterile mixing angles and are insensitive to the CP phases. The oscillation probability, $P_{\mu e}$ for LBL experiments in 3+1 model, after averaging Δm_{41}^2 oscillations and neglecting MSW effects,[121] is expressed as sum of the four terms

$$P_{\mu e}^{4\nu} \simeq P_1 + P_2(\delta_{13}) + P_3(\delta_{14} - \delta_{24}) + P_4(\delta_{13} - (\delta_{14} - \delta_{24})). \quad (5.3)$$

These terms can be approximately expressed as follows:

$$\begin{aligned} P_1 = & \frac{1}{2} \sin^2 2\theta_{\mu e}^{4\nu} + [a^2 \sin^2 2\theta_{\mu e}^{3\nu} - \frac{1}{4} \sin^2 2\theta_{13} \sin^2 2\theta_{\mu e}^{4\nu}] \sin^2 \Delta_{31} \\ & + [a^2 b^2 - \frac{1}{4} \sin^2 2\theta_{12} (\cos^4 \theta_{13} \sin^2 2\theta_{\mu e}^{4\nu} + a^2 \sin^2 2\theta_{\mu e}^{3\nu})] \sin^2 \Delta_{21}, \end{aligned} \quad (5.4)$$

$$P_2(\delta_{13}) = a^2 b \sin 2\theta_{\mu e}^{3\nu} (\cos 2\theta_{12} \cos \delta_{13} \sin^2 \Delta_{21} - \frac{1}{2} \sin \delta_{13} \sin 2\Delta_{21}), \quad (5.5)$$

$$\begin{aligned} P_3(\delta_{14} - \delta_{24}) = & ab \sin 2\theta_{\mu e}^{4\nu} \cos^2 \theta_{13} [\cos 2\theta_{12} \cos(\delta_{14} - \delta_{24}) \sin^2 \Delta_{21} \\ & - \frac{1}{2} \sin(\delta_{14} - \delta_{24}) \sin 2\Delta_{21}], \end{aligned} \quad (5.6)$$

$$\begin{aligned} P_4(\delta_{13} - (\delta_{14} - \delta_{24})) = & a \sin 2\theta_{\mu e}^{3\nu} \sin 2\theta_{\mu e}^{4\nu} [\cos 2\theta_{13} \cos(\delta_{13} - (\delta_{14} - \delta_{24})) \sin^2 \Delta_{31} \\ & + \frac{1}{2} \sin(\delta_{13} - (\delta_{14} - \delta_{24})) \sin 2\Delta_{31} - \frac{1}{4} \sin^2 2\theta_{12} \cos^2 \theta_{13} \cos(\delta_{13} - (\delta_{14} - \delta_{24})) \sin^2 \Delta_{21}], \end{aligned} \quad (5.7)$$

With the parameters defined as

$$\begin{aligned}
\Delta_{ij} &\equiv \Delta m_{ij}^2 L/4E, \text{ a function of baseline}(L) \text{ and neutrino energy}(E) \\
a &= \cos \theta_{14} \cos \theta_{24}, \\
b &= \cos \theta_{13} \cos \theta_{23} \sin 2\theta_{12}, \\
\sin 2\theta_{\mu e}^{3\nu} &= \sin 2\theta_{13} \sin \theta_{23}, \\
\sin 2\theta_{\mu e}^{4\nu} &= \sin 2\theta_{14} \sin \theta_{24}.
\end{aligned} \tag{5.8}$$

The CP phases introduced due to sterile neutrinos persist in the $P_{\mu e}$ even after averaging out Δm_{41}^2 lead oscillations. Last two terms of equation 5.3, give the sterile CP phase dependence terms. The term $P_3(\delta_{14} - \delta_{24})$ depends on the sterile CP phases δ_{14} and δ_{24} , while P_4 depends on a combination of δ_{13} and $\delta_{14} - \delta_{24}$. Thus, we expect LBL experiments to be sensitive to sterile phases. We note that the probability $P_{\mu e}$ is independent θ_{34} . One can see that θ_{34} will effect $P_{\mu e}$ if we consider earth mass effects. Since matter effects are relatively small for NOvA and DUNE, their sensitivity towards θ_{34} is negligible. The amplitudes of atmospheric-sterile interference term (eq.5.7) and solar-atmospheric interference term (eq.5.5), are of the same order. This new interference term reduces the sensitivity of experiments to the standard CP phase(δ_{13}). We also note that the sterile phases δ_{14} and δ_{24} are always together as $\delta_{14} - \delta_{24}$ in a linear dependent form. From this linear dependency, we can deduce that only the difference between the sterile phases ($\delta_{14} - \delta_{24}$) effect the oscillation probability in vacuum case.

5.3 Experiment simulation specifications

We used GLoBES (General Long Baseline Experiment simulator) [122, 123] to simulate the data for different LBL experiments including NOvA and DUNE. The neutrino oscillation probabilities for the 3+1 model are calculated using the new physics engine available from Ref.[124].

NOvA[125, 126] is an LBL experiment which started its full operation from October 2014. NOvA has two detectors, the near detector is located at Fermilab (300 ton, 1 km from NuMI beam target) while the far detector(14 Kt) is located at Northern Minnesota 14.6 mrad off the NuMI beam axis at 810 km from NuMI beam target, justifying “Off-Axis” in the name. This off-axis orientation gives us a narrow beam of flux, peak at 2 GeV[127]. For simulations, we used NOvA setup from Ref.[128]. We used the full projected exposure of 3.6×10^{21} POT (protons on target) expected after six years of runtime at 700kW beam power. Assuming the same runtime for neutrino and anti-neutrino modes, we get 1.8×10^{21} POT for each mode. Following [129] we considered 5% normalization error for the signal, 10 % error for the background for appearance and disappearance channels.

DUNE (Deep Underground Neutrino Experiment)[130, 131] is the next generation LBL experiment. Long Base Neutrino Facility(LBNF) of Fermilab is the source for DUNE. Near detector of DUNE will be at Fermilab. Liquid Argon detector of 40 kt to be constructed at Sanford Underground Research Facility situated 1300 km from the beam target, will act as the far detector. DUNE uses the same source as of NOvA; we will observe beam flux peak at 2.5GeV. We used DUNE setup give in Ref.[132] for our simulations. Since DUNE is still in its early stages, we used simplified systematic treatment, i.e., 5% normalization error on signal, 10 % error on the background for both appearance and disappearance spectra. We give experimental details described above in tabular form in tables 5.1 and 5.2.

Oscillation parameters are estimated from the data by comparing observed and predicted ν_e and ν_μ interaction rates and energy spectra. GLoBES calculates event rates of neutrinos for energy bins taking systematic errors, detector resolutions, MSW effect due to earth’s crust etc into account. The event rates generated for true and test values are used to plot χ^2 contours. GLoBES uses its inbuilt algorithm to calculate χ^2 values numerically considering parameter correlations as well as systematic errors. In our calculations we used χ^2 as:

$$\chi^2 = \sum_{i=1}^{\text{\#ofbins}} \sum_{E_n=E_1, E_2..} \frac{(O_{E_n,i} - (1 + a_F + a_{E_n})T_{E,i})^2}{O_{E_n,i}} + \frac{a_F^2}{\sigma_F^2} + \frac{a_{E_n}^2}{\sigma_{E_n}^2} \quad (5.9)$$

Here $O_{E_n,i}$ and $T_{E_n,i}$ are computed using the true and test oscillation parameters respectively. They represent the event rates for the i^{th} bin in detectors of various experiments being used in simulation using true and test parameter values respectively; uncertainties related to the neutrino flux is given by a_F and uncertainties related to the detector mass is given by a_{E_n} ; σ_F, σ_{E_n} are the standard deviations of a_F, a_{E_n} . The final calculated χ^2 value gives the confidence level at which the test parameter values can be excluded with reference to true parameter values used in simulation. It provides an excellent preliminary evaluation model to estimate the experiment performance.

Name of Experiment	NOvA	DUNE
Detector type	Tracking Calorimeters	LArTPCs
Baseline(Far/Near)	812 km/1 km	1300 km/500 m
Location	Minnesota	South Dakota
Target mass(Far/Near)	14 kt/290 t	40 kt/8 t
POT(yr^{-1})	6.0×10^{20}	1.1×10^{21}
Exposure(years)	6	10

TABLE 5.1: Details of experiments used for simulation using GLOBES.

Name Of Exp	Rule	Normalization error	
		signal(%)	background(%)
NOvA	ν_μ disappearance	2	10
	$\bar{\nu}_\mu$ disappearance	2	10
	ν_e appearance	5	10
	$\bar{\nu}_e$ appearance	5	10
DUNE	ν_μ disappearance	5	10
	$\bar{\nu}_\mu$ disappearance	5	10
	ν_e appearance	5	10
	$\bar{\nu}_e$ appearance	5	10

TABLE 5.2: Systematic errors associated with NOvA and DUNE

In figure 5.1, we plot the oscillation probability($P_{\mu e}$) as a function of energy while varying δ_{14} (-180° to 180°) and keeping $\delta_{24} = 0$ for the three best fit values of

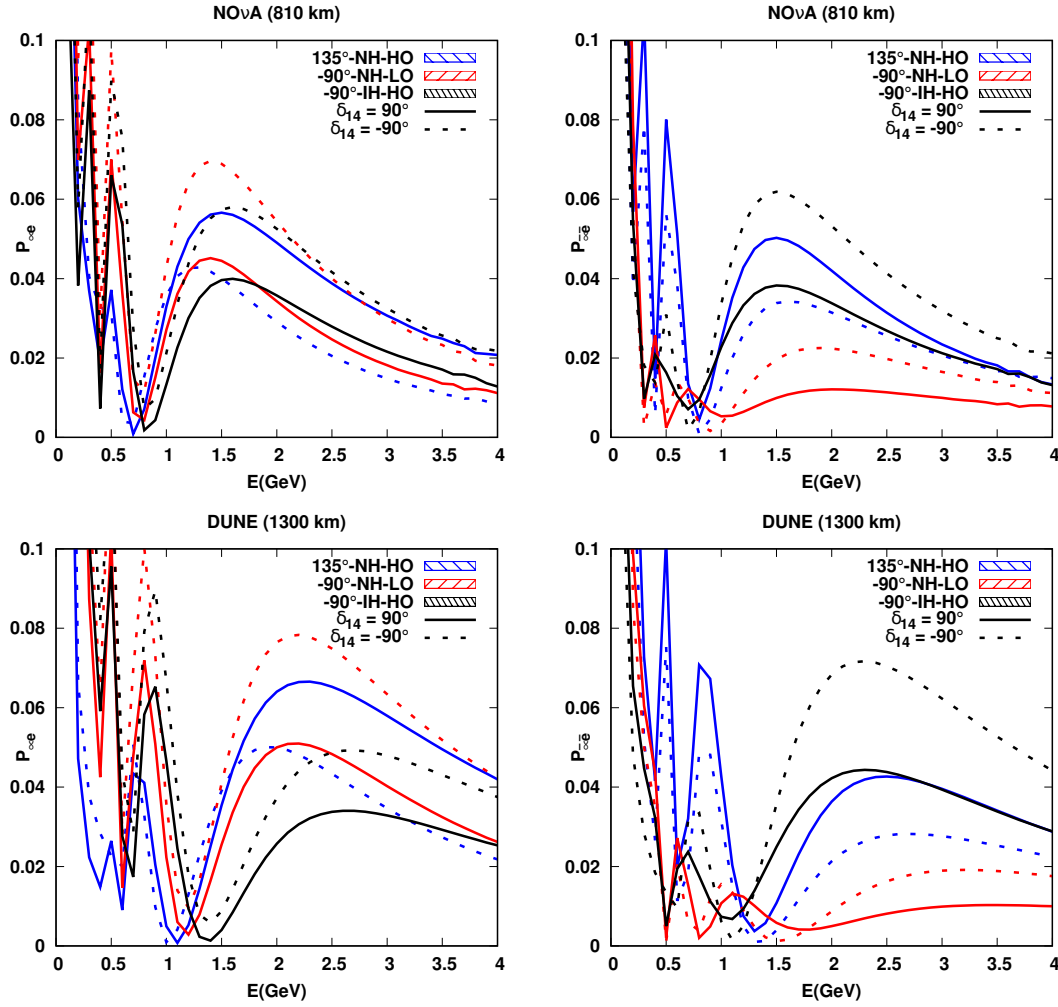


FIGURE 5.1: The oscillation probability $P_{\mu e}$ as a function of energy. The Top(bottom) panel is NOvA(DUNE). The bands correspond to different values of δ_{14} , ranging from -180° to 180° when $\delta_{24} = 0^\circ$. Inside each band, the probability for $\delta_{14} = 90^\circ$ ($\delta_{14} = -90^\circ$) case is shown as the solid (dashed) line. The left(right) panel corresponds to neutrinos(anti-neutrinos).

latest NOvA results[133] i.e; NH-LO- $1.48\pi[\delta_{13}]$, NH-HO- 0.74π and IH-HO- 1.48π . Where, HO implies $\sin^2 \theta_{23} = 0.62$ and LO implies $\sin^2 \theta_{23} = 0.40$. For the flux peak of NOvA, $E \approx 2\text{GeV}$, We observe a degeneracy between all best-fit values due to the presence of δ_{14} band for neutrino case. While, only NH-HO and IH-HO bands overlap in anti-neutrino case. We see that δ_{14} phase decreases both octant and hierarchy resolution capacity for neutrino case and only mass hierarchy resolution capacity for anti-neutrino case. The second row plots $P_{\mu e}$ for DUNE at baseline 1300 km. We observe smaller overlap between bands compared to NOvA. Thus, the decrease of degeneracy resolution capacity for DUNE is less than NOvA. Similarly we plot $P_{\mu e}$ while varying δ_{24} (-180° to 180°) in figure 5.2

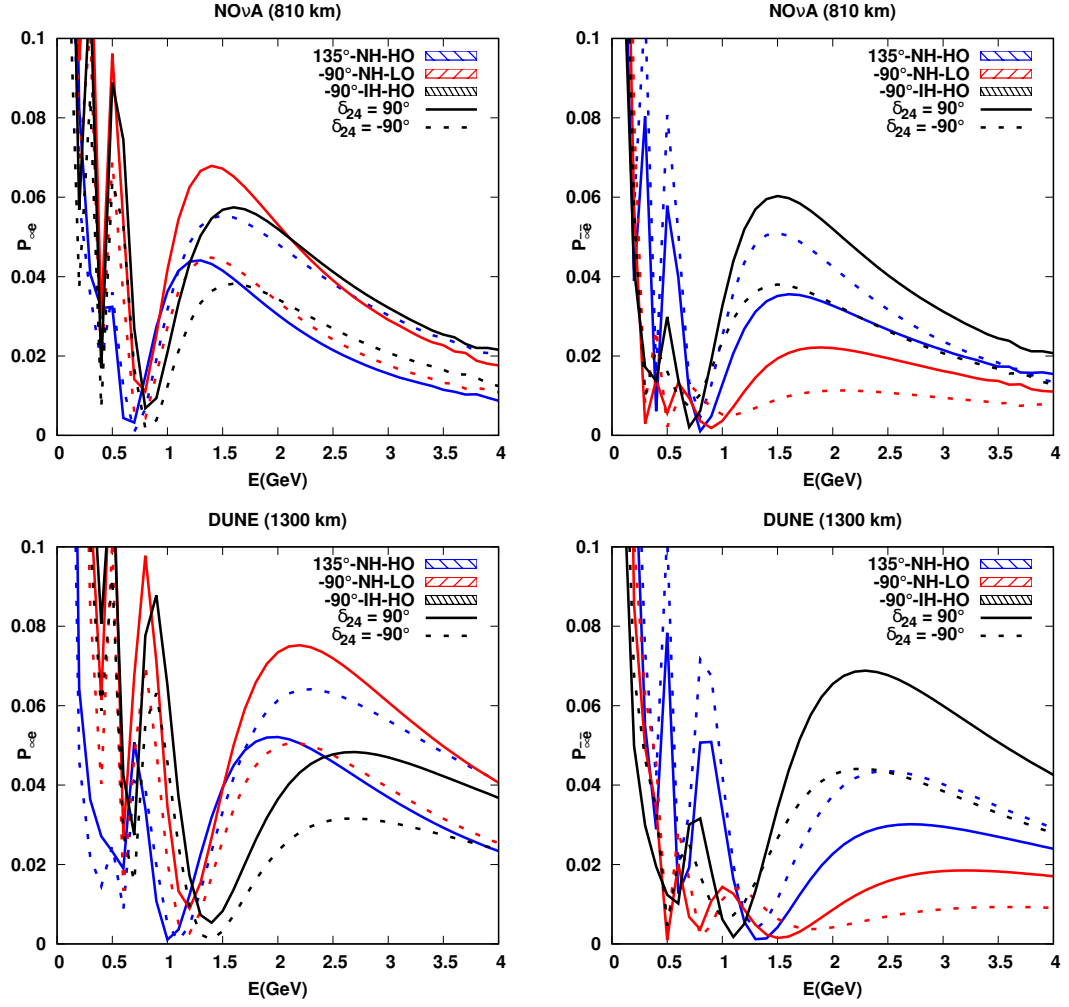


FIGURE 5.2: The oscillation probability $P_{\mu e}$ as a function of energy. The Top(bottom) panel is NOvA(DUNE). The bands correspond to different values of δ_{24} , ranging from -180° to 180° when $\delta_{14} = 0^\circ$. Inside each band, the probability for $\delta_{24} = 90^\circ$ ($\delta_{24} = -90^\circ$) case is shown as solid (dashed) line. The left(right) panel is for neutrinos(anti-neutrinos).

and keeping $\delta_{14} = 0^\circ$. We see that δ_{24} has similar effect as of δ_{14} only change is reversal of δ_{24} band extrema i.e; $\delta_{24} = -90^\circ$ gives same result as $\delta_{14} = 90^\circ$ and vice versa. This can be explained using equations 5.3 in which we see δ_{14} and δ_{24} are always together with opposite signs. Overall from the probability plots, we observe that the addition of new CP phases decrease octant and mass hierarchy resolution capacities.

In Fig:5.3(a), we plot $P_{\mu e}$ for the NOvA experiment with the inclusion of matter effects using GLoBES. The various bands of $(\delta_{14} - \delta_{24})$ are plotted while varying individual values of δ_{14}, δ_{24} inside the band. In Fig:5.3(b), we plot the same type

Parameter	True value	Marginalization Range
$\sin^2\theta_{12}$	0.304	No Marginalization
$\sin^2\theta_{13}$	0.085	[0.075,0.095]
$\sin^2\theta_{23}$	0.623(HO),0.404(LO)	[0.32,0.67]
$\sin^2\theta_{14}$	0.025	No Marginalization
$\sin^2\theta_{24}$	0.025	No Marginalization
$\sin^2\theta_{34}$	0.025	No Marginalization
δ_{13}	135(NH-HO),-90(NH-LO,IH)	[-180,180]
δ_{14}	[-180,180]	[-180,180]
δ_{24}	[-180,180]	[-180,180]
Δm_{21}^2	$7.50 \times 10^{-5} \text{ eV}^2$	No Marginalization
$\Delta m_{31}^2 \text{ (NH)}$	$2.40 \times 10^{-3} \text{ eV}^2$	No Marginalization
$\Delta m_{31}^2 \text{ (IH)}$	$-2.33 \times 10^{-3} \text{ eV}^2$	No Marginalization
Δm_{41}^2	1 eV^2	No Marginalization

TABLE 5.3: Oscillation parameters considered in numerical analysis. The $\sin^2\theta_{23}$ and δ_{13} are taken from latest NOvA results[133].

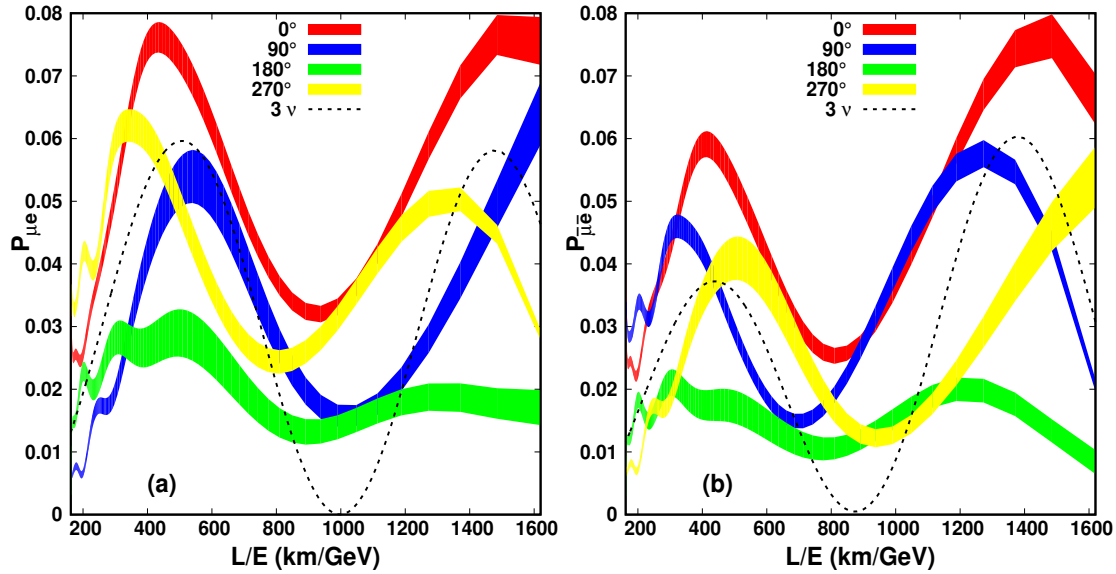


FIGURE 5.3: (a)The electron neutrino(b)The electron anti-neutrino) appearance probability $P_{\mu e}(P_{\mu \bar{e}})$ for NOvA as function of $L/E(\text{km/GeV})$ for various bands of $(\delta_{14} - \delta_{24})$ while varying individual values of δ_{14}, δ_{24} inside the band.

of plot for electron anti-neutrino appearance. We note that even after considering matter effects, individual values of δ_{14}, δ_{24} cause only small variations in $P_{\mu e}(P_{\mu \bar{e}})$, while the difference $\delta_{14} - \delta_{24}$ still plays a major role. In Fig:5.4(a,b), we plot similar plots for DUNE. We see that for the DUNE case also, only the difference $\delta_{14} - \delta_{24}$ is important. Thus, we can say that the difference between the sterile phases has

a more dominating effect on oscillation probability compared to small variations due to changes in individual values.

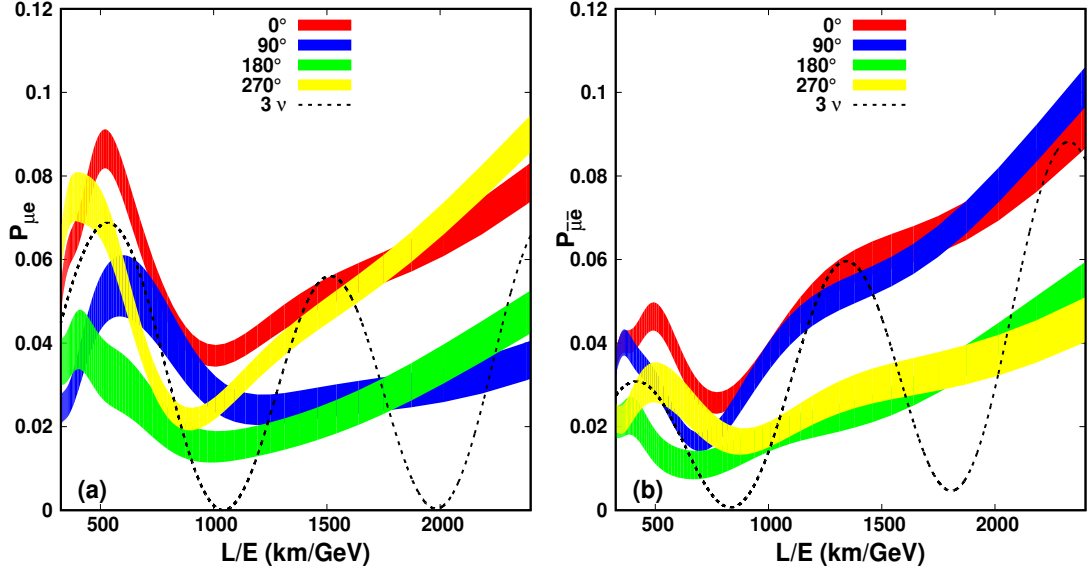


FIGURE 5.4: (a)The electron neutrino((b)The electron anti-neutrino) appearance probability $P_{\mu e}(P_{\bar{\mu e}})$ for DUNE as function of $L/E(\text{km/GeV})$ for various bands of $(\delta_{14} - \delta_{24})$ while varying individual values of δ_{14}, δ_{24} inside the band.

In the next section, we explore how parameter degeneracies are affected in the 3+1 model and the extent to which these degeneracies can be resolved in future runs of NOvA and DUNE. We also find which value of $(\delta_{14} - \delta_{24})$, least affects on the parameter degeneracy of resolution of δ_{13}, θ_{23} for NOvA and DUNE.

5.4 Results

5.4.1 Effect of sterile neutrino on $\sin^2\theta_{23}$ and δ_{13}

We explore allowed regions in $\sin^2\theta_{23}-\delta_{13}$ plane from NOvA and DUNE simulation data with different runtimes, considering latest NOvA results as true values. Using combined analysis of the disappearance and appearance data, NOvA reported preferred solutions[133] at normal hierarchy (NH) with two degenerate best-fit points, one in the lower octant (LO) and $\delta_{13} = 1.48\pi$, the other in higher octant (HO) and $\delta_{13} = 0.74\pi$. Another solution of inverted hierarchy (IH), 0.46σ away

from best fit is also reported. Table 5.3 shows true values of oscillation parameters and their marginalization ranges we used in our simulation. By studying the allowed regions, we understand the extent to which future runs of NOvA and DUNE will resolve these degeneracies, if the best fit values are true values.

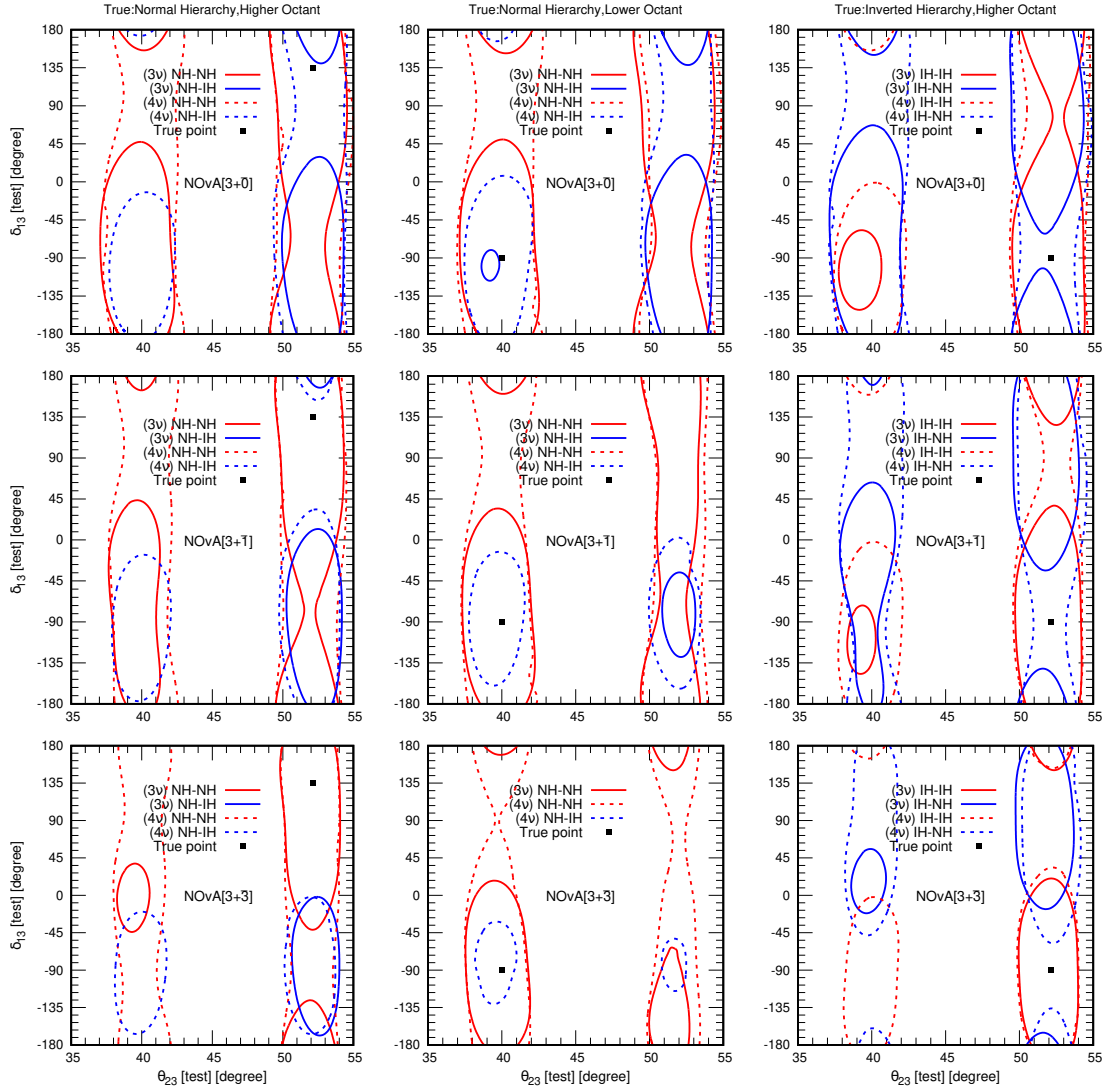


FIGURE 5.5: Contour plots of allowed regions in the test plane, θ_{23} vs δ_{13} , at 90% C.I with top, middle and bottom rows for NOvA runs of $3 + \bar{0}$, $3 + \bar{1}$ and $3 + \bar{3}$ years respectively.

In the first row of figure 5.5, we show allowed areas for NOvA[$3 + \bar{0}$]. In first plot of first row, we show 90% C.L allowed regions for true values of $\delta_{13} = 135^\circ$ and $\theta_{23} = 52^\circ$ and normal hierarchy. We plot test values for both NH and IH, of 3 and 3+1 neutrino models. We observe that introducing sterile neutrino largely decreases the precision of θ_{23} . The WO-RH region, for 3ν case confined between

45° to -180° of δ_{13} , confines the whole δ_{13} region for 4ν case. The WH-RO region of 3ν case doubles, covering the entire region of δ_{13} for 4ν case. The 3+1 model also introduces a small WH-WO region, that was absent in 3ν model. In the second plot of first row (true value $\delta_{13} = -90^\circ$, $\theta_{23} = 40^\circ$ and normal hierarchy), for the 3ν case, we see RH-RO region excluding 45° to 150° of δ_{13} , while RH-WO region covers whole of the δ_{13} region. In 3+1 model, both RH-RO, RH-WO regions cover whole of the δ_{13} region. WH-RO solution occupies a small region for 3ν case, covering half of δ_{13} region for 4ν case. WH-WO region covers whole of the δ_{13} region for 4ν case. In the third plot of first row, true values are taken as, $\delta_{13} = -90^\circ$, $\theta_{23} = 52^\circ$ and inverted hierarchy. The RH-RO region covers the entire range of δ_{13} for both 3ν and 4ν case, where as, RH-WO region almost doubles from 3ν case to 4ν case. A small range of δ_{13} excluded from WH-RO for 3ν case are covered in 4ν case. WH-WO region of 3ν case excludes 60° to 150° of δ_{13} while full δ_{13} range is covered for 4ν case.

In the second row of the figure, we plot allowed regions for NOvA[3+ $\bar{1}$]. We take true values as best fit points obtained by NOvA. We observe an increase in precision of parameter measurement, due to an increase in statistics, from added 1 yr of anti-neutrino run. In the first plot of the second row, the RH-RO octant region covers entire δ_{13} range for both 3ν and 4ν case. RH-WO region includes -180° to 45° of δ_{13} for 3ν case, while whole range of δ_{13} is covered in 4ν case. A slight increase in the area of WH-RO is observed from 3ν to 4ν case. 4ν introduces WH-WO region which was resolved for 3ν case. In the second plot, RH-RO region allows full range of δ_{13} for 4ν case, while it was restricted to lower half of CP range in 3ν case. We see WH-RO solution which was resolved in 3ν case, is reintroduced in 4ν case. We also see a slight increase in the size of WH-WO solution from 3ν to 4ν . In third plot, RH-RO region covers whole CP range for 4ν while 35° to 125° of δ_{13} are excluded in 3ν case. The almost resolved RH-WO solution for 3ν doubles for 4ν case. WH-RO, WH-WO cover entire region of δ_{13} for 4ν case.

In the third row, we show allowed regions for NOvA[3+ $\bar{3}$]. In the first plot, it can be seen that small area of RH-WO in case of 3ν case now covers the whole of δ_{13} region for 4ν case. While the 3ν case has WH-W δ_{13} degeneracy, 4ν case introduces equal sized WH-WO-W δ_{13} degeneracy. In second plot, for 3ν case: most of δ_{13} values above 0° are excluded, but for 4ν case we see contour covers whole of δ_{13} range. Already present small area of RH-WO of 3ν is also increased for 4ν case. 4ν case also introduces a small region of WH solutions which were not present in 3ν case. In the third plot, we see that 4ν introduces RH-WO region of the almost equal size of RH-RO region of 3ν case. We observed a slight increase in WH-RO region for 4ν over 3ν case, while the WH-WO region almost triples for 4ν case.

In the figure 5.6, we show allowed parameter regions for DUNE experiment for different run-times. DUNE being the next generation LBL experiment it is expected to have excellent statistics. Hence, We plot 99% C.L regions for DUNE. In the first row of figure 5.6, We show 99% C.L for DUNE[1+ $\bar{0}$]. In the first plot, RH-RO region covers entire of δ_{13} range for both 3ν and 4ν case. The RH-WO region which covers only lower half of δ_{13} region for 3ν case covers the whole range for 4ν case. A small region of WH is also observed. The second plot we see that all WH solutions are resolved. RH-WO covers the whole range of δ_{13} for both 3ν and 4ν case. RH-RO solutions exclude 0° to 155° of δ_{13} for 3ν case, while 20° to 100° of δ_{13} is excluded for 4ν case. In third plot, we see that 4ν case extends RH-RO to whole range of δ_{13} while 30° to 140° of δ_{13} were excluded for 3ν case. We can see that DUNE clearly has better precision than NOvA experiment. In the second row, we show allowed regions for DUNE[1+ $\bar{1}$]. We see the WH solutions are resolved for both 3ν and 4ν cases for all the best-fit values. In the first plot, 4ν case introduces RH-WO solution of similar size as RH-RO region of 3ν case. In the second plot, there is no considerable change in 4ν , compared to 3ν case for RH-RO region, while RH-WO octant is approximately doubled for 4ν case compared to 3ν case. In the third plot, 4ν case introduces small region of

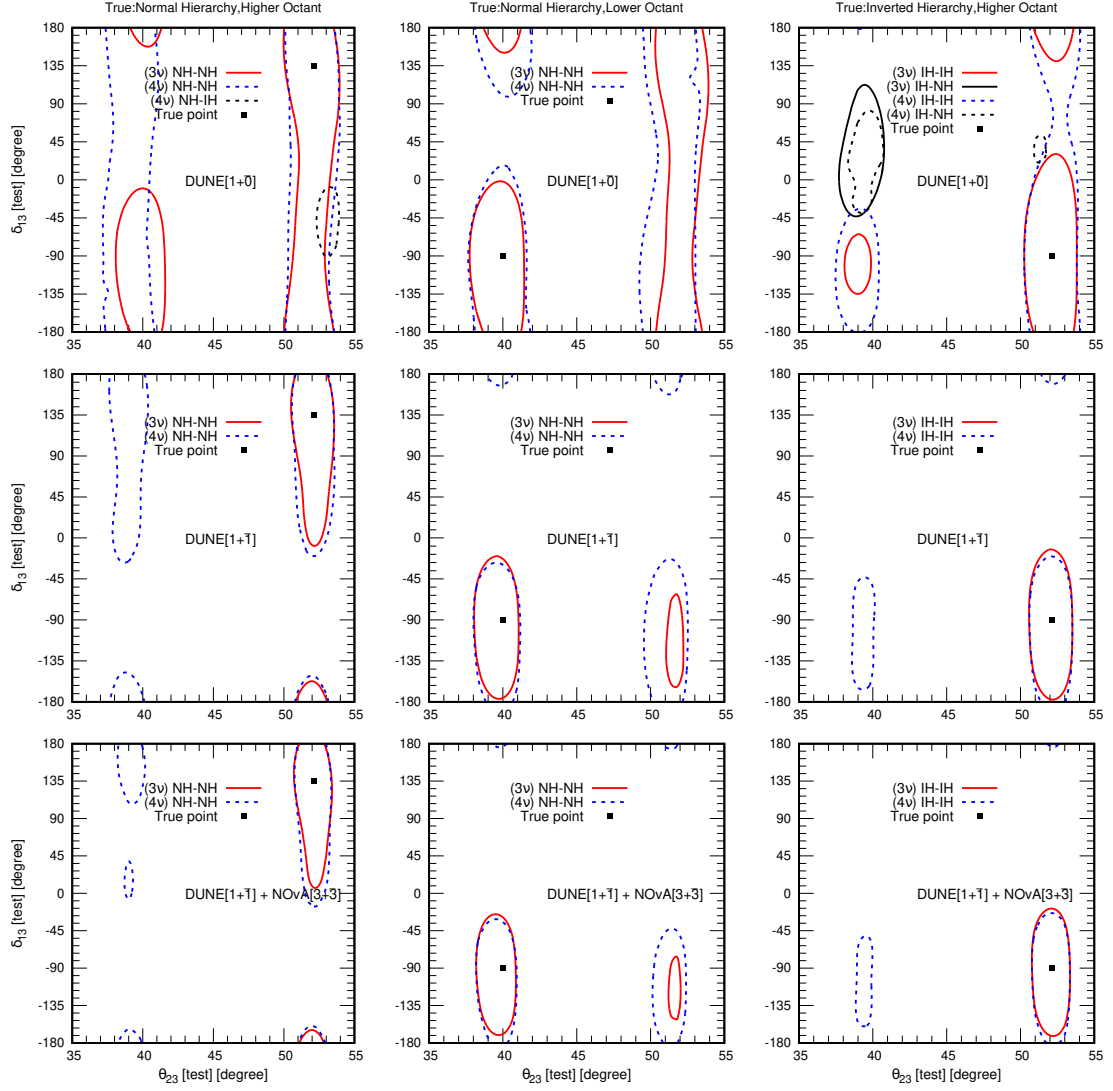


FIGURE 5.6: Contour plots of allowed regions in the test plane θ_{23} vs δ_{13} at 99% C.L. with top, middle and bottom rows for DUNE runs of $1 + \bar{0}$, $1 + \bar{1}$ years and DUNE[$1 + \bar{1}$]+NOvA[$3 + \bar{3}$] respectively.

RH-WO which covers -45° to -170° of δ_{13} . In third row, we combine statistics of DUNE[$1 + \bar{1}$] and NOvA[$3 + \bar{3}$]. There is a small improvement in precision from the combined result over the result from DUNE[$1 + \bar{1}$] alone. In the first plot, we see a small RH-WO region is introduced by 4ν case. In the second plot, there is no considerable change between 3ν and 4ν case for RH-RO region, while RH-WO octant almost doubles over 3ν case for 4ν case. In the third plot, 4ν case introduces small region of RH-WO which covers -35° to -160° of δ_{13} .

In the next figure 5.7, we show allowed parameter regions for DUNE experiment,

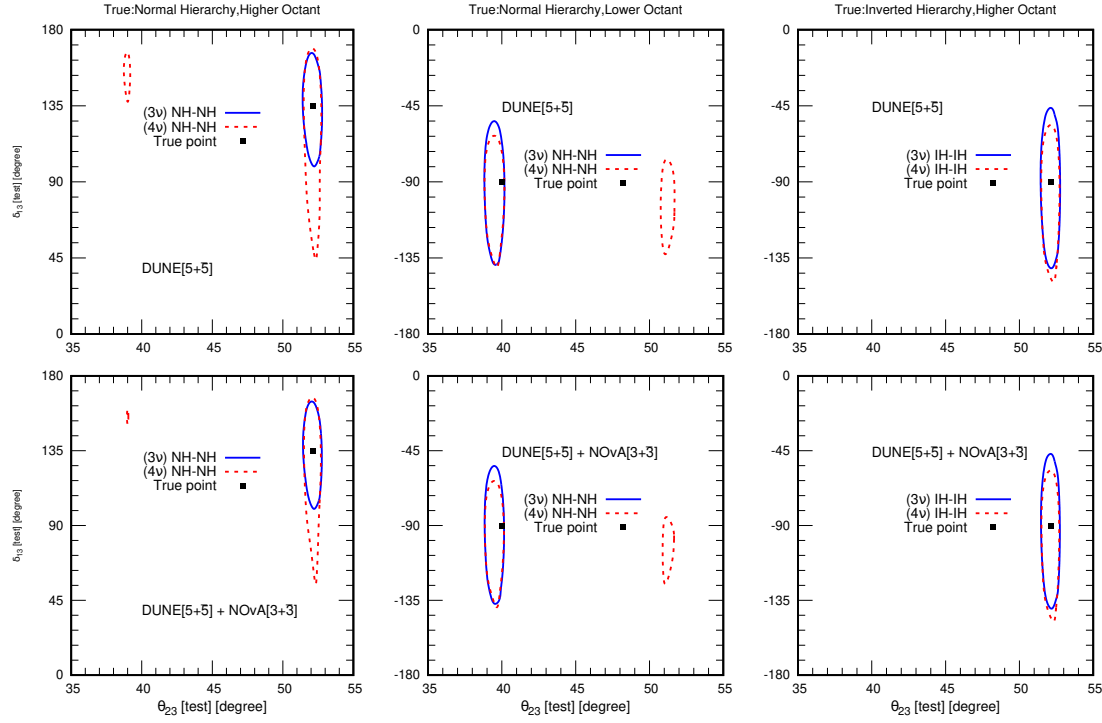


FIGURE 5.7: Contour plots of allowed regions in the test plane θ_{23} vs δ_{13} at 99% C.L with top and bottom rows for DUNE[5 + 5] and NOvA[3 + 3] + DUNE[5 + 5] respectively.

at 99% C.L for DUNE[5+ $\bar{5}$]. We see that WH regions completely disappear for all the true value assumptions. In the first plot, RH-RO region covers a small δ_{13} range for both 3ν and 4ν case indicating high precision measurement capacity of DUNE. We see that δ_{13} range for 4ν case is approximately doubled as compared to the 3ν case. A small region of RH-WO is observed for 4ν case. In the second plot, RH-RO region covers small δ_{13} range of equal area for both 3ν and 4ν case. A small region of RH-WO is observed for 4ν case. In the third plot, the RH-WO solution is resolved. There is an increase in precision due to an increase in statistics. DUNE[5+ $\bar{5}$] clearly has a better precision compared to the NOvA[3+ $\bar{3}$] experiment. In the second row, we combine full run of NOvA and DUNE to check their degeneracy resolution capacity. The WH solutions are resolved for both 3ν and 4ν cases for all the best-fit values. In the first plot, RH-WO solution is almost resolved for 4ν case. In the second plot, RH-RO region covers small δ_{13} range of equal area for both 3ν and 4ν case. A small region of RH-WO is observed for 4ν case. We observe a slight improvement in degeneracy resolution, on consideration

of combined statistics of full run DUNE and NOvA, over DUNE[5+ $\bar{5}$].

5.4.2 Effect of sterile phases on $\sin^2\theta_{23}$ and δ_{13}

We show allowed regions in $\sin^2\theta_{23}$ - δ_{13} plane from NOvA and DUNE simulation data for different $(\delta_{14} - \delta_{24})$ values taking 2019 NOvA results[134] as true values. We marginalise over all the sterile mixing angles and Δm_{31}^2 in our analysis.

In figure 5.8, we show allowed areas at 90% C.I for NOvA[3+ $\bar{3}$] i.e., 3 years of neutrino and 3 years of anti-neutrino run time for NH-HO and NH-LO true values. The figure 5.8a, in which NH-HO is taken as true values, we observe that the parameter resolution sensitivity is highest when value of $\delta_{14} - \delta_{24} = 180^\circ$. Same case follows for figure 5.8b, in which NH-LO is taken as true values.

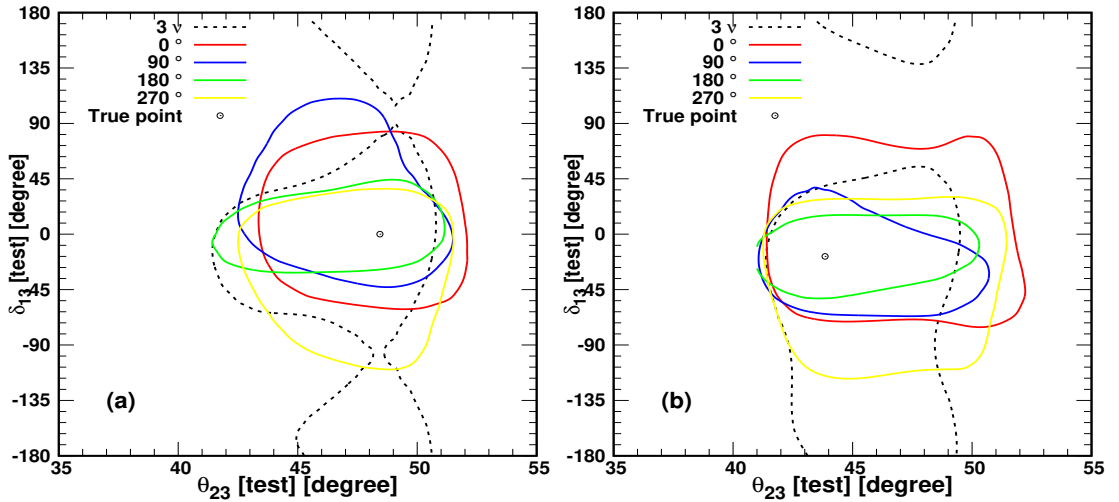


FIGURE 5.8: Contour plots of allowed regions in the test plane, θ_{23} vs δ_{13} for different $(\delta_{14} - \delta_{24})$ values at 90% C.I regions for NOvA[3+ $\bar{3}$]. We take NH-HO(NH-LO) as true values for a(b) case.

In figure 5.9, we show allowed areas at 90% C.I for DUNE[3+ $\bar{3}$] i.e., 3 years of neutrino and 3 years of anti-neutrino run time for NH-HO and NH-LO true values. The figure 5.9a, in which NH-HO is taken as true values, we observe that the parameter resolution sensitivity is highest when value of $\delta_{14} - \delta_{24} = 90^\circ$. Same case follows for figure 5.9b, in which NH-LO is taken as true values.

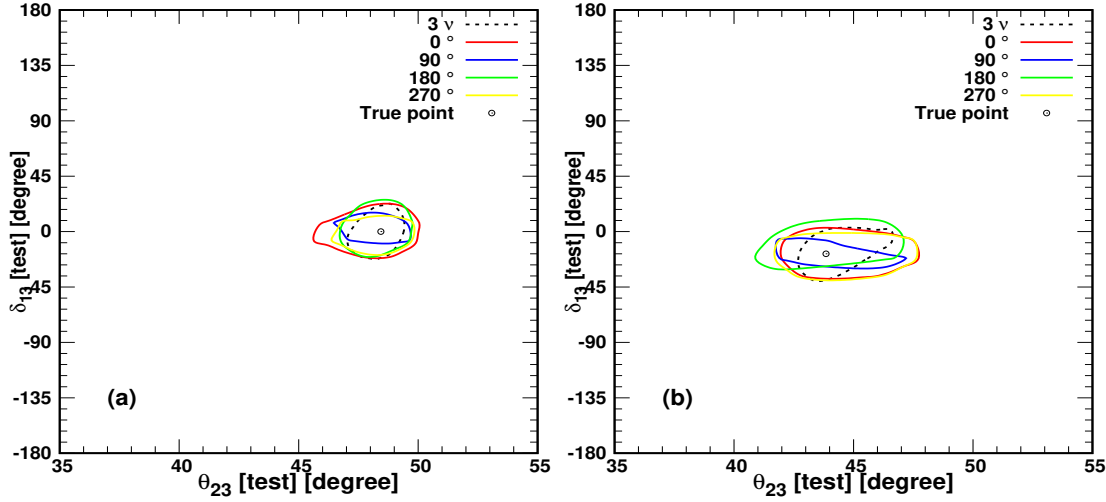


FIGURE 5.9: Contour plots of allowed regions in the test plane, θ_{23} vs δ_{13} for different $(\delta_{14} - \delta_{24})$ values at 90% C.I regions for DUNE[3+3]. We take NH-HO(NH-LO) as true values for a(b) case.

5.5 Conclusions

We have discussed how the presence of a sterile neutrino will affect, the physics potential of the proposed experiment DUNE and future runs of NOvA, in the light of latest NOvA results[133]. The best-fit parameters reported by NOvA still contain degenerate solutions. We attempt to see the extent to which these degeneracies could be resolved in future runs for the 3+1 model. Latest NOvA best-fit values are taken as our true values. First, we show the degeneracy resolution capacity, for future runs of NOvA. We conclude that NOvA[3+ $\bar{3}$] could resolve WH-WO solutions for first two true value cases, at 90% C.L for 3ν case, but not for 4ν case. DUNE[1+ $\bar{1}$] could resolve WH and RH-W δ_{13} solutions for both 3ν and 4ν case. WO degeneracy is resolved for 3ν case at 99% C.L except for small RH-WO region for the second case of true values. DUNE[1+ $\bar{1}$] combined with NOvA[3+ $\bar{3}$] shows increased sensitivity towards degeneracy resolution. Finally, for the full planned run of DUNE[5+ $\bar{5}$], all the degeneracies are resolved at 99% C.L for 3ν case while a tiny region of WO linger on for 4ν case. For combined statistics of DUNE[5+ $\bar{5}$] and NOvA[3+ $\bar{3}$], we observe that all the degeneracies are resolved at 99% C.L for both 3ν and 4ν case except for the NH-LO case. Thus, we conclude that NOvA

and DUNE experiments together can resolve all the degeneracies at 99% C.L even in the presence of sterile neutrino, if one of the current best-fit values of NOvA, is the true value.

We find that only the difference between sterile phases ($\delta_{14} - \delta_{24}$) is important, where as the individual values of δ_{14}, δ_{24} play a negligible role in case of neutrino oscillation at LBL experiments. We note the NOvA and DUNE have different value of $(\delta_{14} - \delta_{24})$, at 180° for NOVA and 90° for DUNE at which they get good parameter resolution. This difference in the result maybe due to difference in matter effects and statistics between NOvA and DUNE.

Chapter 6

Conclusion and Future work

The NOvA collaboration aims to precisely determine neutrino oscillation parameters, namely the values of δ_{CP} , θ_{23} , and sign of Δm_{32}^2 . Achieving these objectives relies heavily on accurately classifying neutrino interactions and measuring neutrino energy. Accurately categorizing the particle interactions as either signal or background is very important. We studied the traditional reconstruction methods used for reconstructing high-level components such as clusters, tracks associated with particle interactions recorded in the detector. After reconstruction, we can summarize the directions, shapes, and energies of these particle interactions with a fewer number of quantities. In the context of a large-scale experiment like NOvA, managing vast amounts of data to differentiate between signal and background becomes challenging. This is where Machine Learning proves highly beneficial.

NOvA uses a Convolutional Neural Network (CNN) to classify candidate neutrino interactions. While Event CNN can classify events, identification of final state particles of an event is needed to better our energy reconstruction and enable cross-section measurements of final states. Context enriched Prong CNN is one of the networks being trained with a goal to identify all the final-state particles ($e^\pm, p^\pm, \mu, \pi^\pm, \gamma$) of a given neutrino event.

In this thesis, we described the development of Prong CNN, and modifications made compared to previous version. We trained single Prong CNN from both neutrino and anti-neutrino events. Our updated Prong CNN incorporates a modified MobileNetv2 architecture infused with additional features from MobileNetv3 [93]. We decreased the overall number of layers in the network and fine-tuned the composition and quantity of convolutional layers within each layer to optimize runtime on CPUs without compromising performance. Upon adapting this revised architecture to the Prong CNN scenario, the network's parameter count decreased by approximately a quarter. This modified Prong CNN is now more computationally efficient, enabling higher classification efficiency within fewer training epochs. The newer prong CNN has a global efficiency greater than 86% with improved the classification efficiency by 3% compared to previous Prong CNN. In the next NOvA analysis run, the new Prong CNN will adopted to give better particle identification and energy resolution. In the future, I hope to incorporate the ML techniques learnt in NOvA to upcoming DUNE experiment envisaged at Fermilab later in this decade.

Besides developing the prong CNN for the NOvA experiment, a phenomenological study was conducted within the framework of the 3+1 sterile neutrino model. This study aimed to assess the impact of sterile neutrinos on the degeneracy resolution capabilities of both the NOvA and DUNE experiments. We await results from DUNE to verify our study. We hope that this work aids the future analysis of DUNE.

In future, more experiments will start depending on machine learning tools. But, most of the current ML tools being used in HEP behave like a black boxes. This is not optimal condition for HEP experiments because it is crucial to understand the underlying physics and the reasoning behind the model's decisions. Interpretability is essential for building trust in the machine learning models and their applications. So, there has been a rising interest in interpretable ML tools. We would like to

work towards increasing the interpretability of ML tools which will enable us to extract meaningful physical insights from data and maintain transparency in experimental results in future.

Bibliography

- [1] A. Ianni, “*Solar neutrinos and the solar model*,” Physics of the Dark Universe **4** (2014) 44–49. <https://www.sciencedirect.com/science/article/pii/S2212686414000211>. DARK TAUP2013.
- [2] A. H. Bécquerel, “*Sur les radiations invisibles émises par les corps phosphorescents*,” C. R. Acad. Sci. Paris **122** (1896) 501. <http://cds.cern.ch/record/261888>.
- [3] A. H. Bécquerel, “*Sur quelques propriétés nouvelles des radiations invisibles émises par divers corps phosphorescents*,” C. R. Acad. Sci. Paris **122** (1896) 559. <http://cds.cern.ch/record/261892>.
- [4] A. H. Bécquerel, “*Sur les propriétés différentes des radiations invisibles émises par les sels d’uranium, et du rayonnement de la paroi anticathodique d’un tube de Crookes*,” C. R. Acad. Sci. Paris **122** (1896) 762. <http://cds.cern.ch/record/261893>.
- [5] J. Chadwick, “*Intensitätsverteilung im magnetischen Spectrum der β -Strahlen von radium B + C*,” Verhandl. Dtsc. Phys. Ges. **16** (1914) 383. <http://cds.cern.ch/record/262756>.
- [6] C. D. Ellis and W. A. Wooster, “*The average energy of disintegration of radium E*,” Proc. Roy. Soc. Lond. A **117** (1927) no. 776, 109–123.
- [7] W. Pauli, “*Letter to the physical society of Tübingen*,”.

- [8] E. Fermi, “*An attempt of a theory of beta radiation. 1.*,” Z. Phys. **88** (1934) 161–177.
- [9] C. L. Cowan, F. Reines, F. B. Harrison, H. W. Kruse, and A. D. McGuire, “*Detection of the free neutrino: A Confirmation*,” Science **124** (1956) 103–104.
- [10] B. Pontecorvo, “*Inverse beta processes and nonconservation of lepton charge*,” Zh. Eksp. Teor. Fiz. **34** (1957) 247.
- [11] **OPAL**, M. Z. Akrawy *et al.*, “*Measurement of the Z^0 Mass and Width with the OPAL Detector at LEP*,” Phys. Lett. B **231** (1989) 530–538.
- [12] **ALEPH**, D. Decamp *et al.*, “*Determination of the Number of Light Neutrino Species*,” Phys. Lett. B **231** (1989) 519–529.
- [13] G. Danby, J. M. Gaillard, K. A. Goulianos, L. M. Lederman, N. B. Mistry, M. Schwartz, and J. Steinberger, “*Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos*,” Phys. Rev. Lett. **9** (1962) 36–44.
- [14] **DONUT**, K. Kodama *et al.*, “*Observation of tau neutrino interactions*,” Phys. Lett. B **504** (2001) 218–224, [arXiv:hep-ex/0012035](#).
- [15] J. N. Bahcall, N. A. Bahcall, and G. Shaviv, “*Present status of the theoretical predictions for the Cl-36 solar neutrino experiment*,” Phys. Rev. Lett. **20** (1968) 1209–1212.
- [16] R. Davis, Jr., D. S. Harmer, and K. C. Hoffman, “*Search for neutrinos from the sun*,” Phys. Rev. Lett. **20** (1968) 1205–1209.
- [17] **Super-Kamiokande**, Y. Fukuda *et al.*, “*Evidence for oscillation of atmospheric neutrinos*,” Phys. Rev. Lett. **81** (1998) 1562–1567, [arXiv:hep-ex/9807003](#).

- [18] N. Jelley, A. B. McDonald, and R. G. H. Robertson, “*The Sudbury Neutrino Observatory*,” *Ann. Rev. Nucl. Part. Sci.* **59** (2009) 431–465.
- [19] **ATLAS**, G. Aad *et al.*, “*Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*,” *Phys. Lett. B* **716** (2012) 1–29, [arXiv:1207.7214](#).
- [20] N. Palanque-Delabrouille *et al.*, “*Neutrino masses and cosmology with Lyman-alpha forest power spectrum*,” *JCAP* **11** (2015) 011, [arXiv:1506.05976](#).
- [21] **KATRIN**, M. Aker *et al.*, “*Direct neutrino-mass measurement with sub-electronvolt sensitivity*,” *Nature Phys.* **18** (2022) no. 2, 160–166, [arXiv:2105.08533](#).
- [22] E. Vitagliano, I. Tamborra, and G. Raffelt, “*Grand Unified Neutrino Spectrum at Earth: Sources and Spectral Components*,” *Rev. Mod. Phys.* **92** (2020) 45006, [arXiv:1910.11878](#).
- [23] C. Giunti and C. W. Kim, *Fundamentals of Neutrino Physics and Astrophysics*. 2007.
- [24] *Fundamental Physics at the Intensity Frontier*. 5, 2012. [arXiv:1205.2671](#).
- [25] B. Pontecorvo, “*Neutrino Experiments and the Problem of Conservation of Leptonic Charge*,” *Zh. Eksp. Teor. Fiz.* **53** (1967) 1717–1725.
- [26] Z. Maki, M. Nakagawa, and S. Sakata, “*Remarks on the unified model of elementary particles*,” *Prog. Theor. Phys.* **28** (1962) 870–880.
- [27] S. P. Mikheev and A. Y. Smirnov, “*Resonant amplification of neutrino oscillations in matter and solar neutrino spectroscopy*,” *Nuovo Cim. C* **9** (1986) 17–26.
- [28] L. Wolfenstein, “*Neutrino Oscillations in Matter*,” *Phys. Rev. D* **17** (1978) 2369–2374.

- [29] **Daya Bay**, F. P. An *et al.*, “*Observation of electron-antineutrino disappearance at Daya Bay*,” Phys. Rev. Lett. **108** (2012) 171803, arXiv:1203.1669.
- [30] **Double Chooz**, Y. Abe *et al.*, “*Indication of Reactor $\bar{\nu}_e$ Disappearance in the Double Chooz Experiment*,” Phys. Rev. Lett. **108** (2012) 131801, arXiv:1112.6353.
- [31] **RENO**, J. K. Ahn *et al.*, “*Observation of Reactor Electron Antineutrino Disappearance in the RENO Experiment*,” Phys. Rev. Lett. **108** (2012) 191802, arXiv:1204.0626.
- [32] **Kamiokande-II**, K. S. Hirata *et al.*, “*Observation of B-8 Solar Neutrinos in the Kamiokande-II Detector*,” Phys. Rev. Lett. **63** (1989) 16.
- [33] **SAGE**, J. N. Abdurashitov *et al.*, “*Solar neutrino flux measurements by the Soviet-American Gallium Experiment (SAGE) for half the 22 year solar cycle*,” J. Exp. Theor. Phys. **95** (2002) 181–193, arXiv:astro-ph/0204245.
- [34] **GALLEX**, P. Anselmann *et al.*, “*Solar neutrinos observed by GALLEX at Gran Sasso*,” Phys. Lett. B **285** (1992) 376–389.
- [35] **SNO**, A. Bellerive, J. R. Klein, A. B. McDonald, A. J. Noble, and A. W. P. Poon, “*The Sudbury Neutrino Observatory*,” Nucl. Phys. B **908** (2016) 30–51, arXiv:1602.02469.
- [36] **Super-Kamiokande**, Y. Fukuda *et al.*, “*The Super-Kamiokande detector*,” Nucl. Instrum. Meth. A **501** (2003) 418–462.
- [37] **LSND**, G. B. Mills, “*Neutrino oscillation results from LSND*,” Nucl. Phys. B Proc. Suppl. **91** (2001) 198–202.
- [38] **SBND**, **MicroBooNE**, S. Tufanli, “*The Short Baseline Neutrino Program at Fermilab*,” PoS **EPS-HEP2017** (2017) 141.

- [39] **MINOS**, P. Adamson *et al.*, “*Improved search for muon-neutrino to electron-neutrino oscillations in MINOS*,” Phys. Rev. Lett. **107** (2011) 181802, [arXiv:1108.0015](#).
- [40] **T2K**, K. Abe *et al.*, “*The T2K Experiment*,” Nucl. Instrum. Meth. A **659** (2011) 106–135, [arXiv:1106.1238](#).
- [41] **DUNE**, B. Abi *et al.*, “*Long-baseline neutrino oscillation physics potential of the DUNE experiment*,” Eur. Phys. J. C **80** (2020) no. 10, 978, [arXiv:2006.16043](#).
- [42] I. Esteban, M. C. Gonzalez-Garcia, M. Maltoni, T. Schwetz, and A. Zhou, “*The fate of hints: updated global analysis of three-flavor neutrino oscillations*,” JHEP **09** (2020) 178, [arXiv:2007.14792](#).
- [43] **T2K**, K. Abe *et al.*, “*Constraint on the matter–antimatter symmetry-violating phase in neutrino oscillations*,” Nature **580** (2020) no. 7803, 339–344, [arXiv:1910.03887](#). [Erratum: Nature 583, E16 (2020)].
- [44] **NOvA**, M. A. Acero *et al.*, “*Improved measurement of neutrino oscillation parameters by the NOvA experiment*,” Phys. Rev. D **106** (2022) no. 3, 032004, [arXiv:2108.08219](#).
- [45] **T2K**, K. Abe *et al.*, “*Improved constraints on neutrino mixing from the T2K experiment with 3.13×10^{21} protons on target*,” Phys. Rev. D **103** (2021) no. 11, 112008, [arXiv:2101.03779](#).
- [46] L. Cremonesi, “*Cross-section measurements with NOvA*,” <https://doi.org/10.5281/zenodo.4155399>.
- [47] **MINOS**, I. Ambats *et al.*, “*The MINOS Detectors Technical Design Report*,”.
- [48] **NOvA**, D. S. Ayres *et al.*, “*The NOvA Technical Design Report*,”.

- [49] P. Adamson *et al.*, “*The NuMI Neutrino Beam*,” Nucl. Instrum. Meth. A **806** (2016) 279–306, [arXiv:1507.06690](#).
- [50] L. Soplin., “*2017-2018 beam plots*,” NOvA internal document DocDb 20843-v12 (2018) .
- [51] S. Mufson *et al.*, “*Liquid scintillator production for the NOvA experiment*,” Nucl. Instrum. Meth. A **799** (2015) 1–9, [arXiv:1504.04035](#).
- [52] S. C. Tognini, *Observation of multiple-muon seasonal variations in the NOvA Near Detector*. PhD thesis, Goias U., 2018.
- [53] C. Backhouse., “*Improving timing resolution*,” NOvA internal document DocDb 6567-v1 (2014) .
- [54] P. Antonioli *et al.*, “*SNEWS: The Supernova Early Warning System*,” New J. Phys. **6** (2004) 114, [arXiv:astro-ph/0406214](#).
- [55] C. Robert and G. Casella, *Monte Carlo statistical methods*. 2004.
- [56] **NOvA**, A. Aurisano *et al.*, “*The NOvA simulation chain*,” J. Phys. Conf. Ser. **664** (2015) no. 7, 072002.
- [57] A. Fasso *et al.*, “*The FLUKA code: Present applications and future developments*,” eConf **C0303241** (2003) MOMT004, [arXiv:physics/0306162](#).
- [58] M. Campanella, A. Ferrari, P. R. Sala, and S. Vanini, “*First Calorimeter Simulation with the FLUGG Prototype*,”.
- [59] J. Allison *et al.*, “*Geant4 developments and applications*,” IEEE Trans. Nucl. Sci. **53** (2006) 270.
- [60] **GEANT4**, S. Agostinelli *et al.*, “*GEANT4—a simulation toolkit*,” Nucl. Instrum. Meth. A **506** (2003) 250–303.

- [61] J. Allison *et al.*, “Recent developments in *Geant4*,” Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **835** (2016) 186–225.
- [62] **EMPHATIC**, T. Akaishi *et al.*, “*EMPHATIC: A Proposed Experiment to Measure Hadron Scattering and Production Cross Sections for Improved Neutrino Flux Predictions*,” [arXiv:1912.08841](#).
- [63] **MINERvA**, L. Aliaga *et al.*, “Neutrino Flux Predictions for the NuMI Beam,” Phys. Rev. D **94** (2016) no. 9, 092005, [arXiv:1607.00704](#).
[Addendum: Phys.Rev.D 95, 039903 (2017)].
- [64] A. T. C. Sutton, *Domain Generalization with Machine Learning in the NOvA Experiment*. PhD thesis, Virginia U., 2022.
- [65] C. Andreopoulos *et al.*, “The GENIE Neutrino Monte Carlo Generator,” Nucl. Instrum. Meth. A **614** (2010) 87–104, [arXiv:0905.2517](#).
- [66] **NOvA**, M. A. Acero *et al.*, “Measurement of the double-differential muon-neutrino charged-current inclusive cross section in the NOvA near detector,” Phys. Rev. D **107** (2023) no. 5, 052011, [arXiv:2109.12220](#).
- [67] C. Hagmann, D. Lange, and D. Wright, “Cosmic-ray shower generator (CRY) for Monte Carlo transport codes,” in *2007 IEEE Nuclear Science Symposium Conference Record*, vol. 2, pp. 1143–1146. 2007.
- [68] D. Wright and M. Kelsey, “The *Geant4* Bertini Cascade,” Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **804** (2015) 175–188.
<https://www.sciencedirect.com/science/article/pii/S0168900215011134>.
- [69] L. Thulliez, C. Jouanne, and E. Dumonteil, “Improvement of *Geant4* Neutron-HP package: From methodology to evaluated nuclear data library,”

- Nuclear Instruments and Methods in Physics Research Section A:
Accelerators, Spectrometers, Detectors and Associated Equipment **1027**
(2022) 166187. <https://www.sciencedirect.com/science/article/pii/S016890022101055X>.
- [70] A. Aurisano, “*The NOvA Detector Simulation*,” NOvA internal document DocDb 13577-v3 (2015) .
- [71] **Particle Data Group**, R. L. Workman and Others, “*Review of Particle Physics*,” PTEP **2022** (2022) 083C01.
- [72] M. Ester, H. P. Kriegel, J. Sander, and X. Xiaowei, “*A density-based algorithm for discovering clusters in large spatial databases with noise*,” <https://www.osti.gov/biblio/421283>.
- [73] J. H. Dan Pershey and M. Judah., “*TDSlicer Technote*,” NOvA internal document DocDb 27689-v5 (2019) .
- [74] A. Rodriguez and A. Laio, “*Clustering by fast search and find of density peaks*,” American Association for the Advancement of Science **344** (2014) no. 6191, 1492–1496.
- [75] R. C. Prim, “*Shortest connection networks and some generalizations*,” Bell System Technical Journal **36** (1957) 1389–1401.
<https://api.semanticscholar.org/CorpusID:120649066>.
- [76] C. G. Scott Kirkpatrick and M. Vecchi, “*Optimization by Simulated Annealing*,”
- [77] J. Bezdek, *Pattern Recognition With Fuzzy Objective Function Algorithms*. 01, 1981.
- [78] R. Krishnapuram and J. Keller, “*A possibilistic approach to clustering*,” IEEE Transactions on Fuzzy Systems **1** (1993) no. 2, 98–110.
- [79] M.-S. Yang and K.-L. Wu, “*Unsupervised possibilistic clustering*,”

- [80] C. Rudin, C. Chen, Z. Chen, H. Huang, L. Semenova, and C. Zhong, “*Interpretable machine learning: Fundamental principles and 10 grand challenges*,” *Statistic Surveys* **16** (2022) 1–85.
- [81] C. Molnar, G. Casalicchio, and B. Bischl, “*Interpretable machine learning—a brief history, state-of-the-art and challenges*,” in *Joint European conference on machine learning and knowledge discovery in databases*, pp. 417–431, Springer. 2020.
- [82] X. Glorot and Y. Bengio, “*Understanding the difficulty of training deep feedforward neural networks*,” in *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, Y. W. Teh and M. Titterton, eds., vol. 9 of *Proceedings of Machine Learning Research*, pp. 249–256. PMLR, Chia Laguna Resort, Sardinia, Italy, 13–15 May, 2010. <https://proceedings.mlr.press/v9/glorot10a.html>.
- [83] K. He, X. Zhang, S. Ren, and J. Sun, “*Delving deep into rectifiers: Surpassing human-level performance on imagenet classification*,” in *Proceedings of the IEEE international conference on computer vision*, pp. 1026–1034. 2015.
- [84] W. Hu, L. Xiao, and J. Pennington, “*Provable benefit of orthogonal initialization in optimizing deep linear networks*,” arXiv preprint arXiv:2001.05992 (2020) .
- [85] K. Janocha and W. M. Czarnecki, “*On loss functions for deep neural networks in classification*,” arXiv preprint arXiv:1702.05659 (2017) .
- [86] I. Sutskever, J. Martens, G. Dahl, and G. Hinton, “*On the importance of initialization and momentum in deep learning*,” in *Proceedings of the 30th International Conference on Machine Learning*, vol. 28 of *Proceedings of Machine Learning Research*, pp. 1139–1147. PMLR, Atlanta, Georgia, USA, 17–19 Jun, 2013.

- [87] C. Backhouse and R. B. Patterson, “*Library Event Matching event classification algorithm for electron neutrino interactions in the NOvA detectors*,” Nucl. Instrum. Meth. A **778** (2015) 31–39, [arXiv:1501.00968](#).
- [88] A. Aurisano, A. Radovic, D. Rocco, A. Himmel, M. D. Messier, E. Niner, G. Pawloski, F. Psihas, A. Sousa, and P. Vahle, “*A Convolutional Neural Network Neutrino Event Classifier*,” JINST **11** (2016) no. 09, P09001, [arXiv:1604.01444](#).
- [89] F. Psihas, E. Niner, M. Groh, R. Murphy, A. Aurisano, A. Himmel, K. Lang, M. D. Messier, A. Radovic, and A. Sousa, “*Context-Enriched Identification of Particles with a Convolutional Network for Neutrino Events*,” Phys. Rev. D **100** (2019) no. 7, 073005, [arXiv:1906.00713](#).
- [90] T. W. Micah Groh, “*CVN Training for the 2020 Analysis*,” NOvA internal document DocDb 42897 (2020) .
- [91] M. S. et al., “*MobileNetV2: Inverted Residuals and Linear Bottlenecks*,” [arXiv:arXiv:1801.04381 \[cs.CV\]](#).
- [92] F. Chollet, “*Xception: Deep Learning with Depthwise Separable Convolutions*,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1800–1807. IEEE Computer Society, Los Alamitos, CA, USA, jul, 2017.
<https://doi.ieeecomputersociety.org/10.1109/CVPR.2017.195>.
- [93] A. H. et al., “*Searching for MobileNetV3*,” [arXiv:arXiv:1905.02244](#).
- [94] P. R. et al., “*Searching for Activation Functions*,” [arXiv:arXiv:1710.05941](#).
- [95] J. H. et al., “*Squeeze-and-Excitation Networks*,” [arXiv:arXiv:1709.01507](#).
- [96] Q. W. et al., “*ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks*,” [arXiv:arXiv:1910.03151](#).

- [97] R. Volkas, “*Introduction to sterile neutrinos*,” Progress in Particle and Nuclear Physics **48** (2002) 161, [hep-ph/0111326](#).
- [98] C. Athanassopoulos *et al.*, “*Candidate events in a search for anti-muon-neutrino — $\bar{\nu}_\mu$ — anti-electron-neutrino oscillations*,” Physical Review Letters **75** (1995) 2650, [nucl-ex/9504002](#).
- [99] S. Schael *et al.*, “*Precision electroweak measurements on the Z resonance*,” Physics Reports **427** (2006) 257, [hep-ex/0509008](#).
- [100] A. Aguilar-Arevalo *et al.*, “*Unexplained Excess of Electron-Like Events From a 1-GeV Neutrino Beam*,” Physical Review Letters **102** (2009) 101802, [arXiv:0812.2243 \[hep-ex\]](#).
- [101] J. Abdurashitov *et al.*, “*Measurement of the response of a Ga solar neutrino experiment to neutrinos from an Ar-37 source*,” Physical Review C **73** (2006) 045805, [nucl-ex/0512041](#).
- [102] M. Acero, C. Giunti, and M. Laveder, “*Limits on $\nu(e)$ and anti- $\nu(e)$ disappearance from Gallium and reactor experiments*,” Physical Review D **78** (2008) 073009, [arXiv:0711.4222 \[hep-ph\]](#).
- [103] C. Giunti and M. Laveder, “*Statistical Significance of the Gallium Anomaly*,” Physical Review C **83** (2011) 065504, [arXiv:1006.3244 \[hep-ph\]](#).
- [104] G. Mention, M. Fechner, T. Lasserre, T. Mueller, D. Lhuillier, M. Cribier, and A. Letourneau, “*The Reactor Antineutrino Anomaly*,” Physical Review D **83** (2011) 073006, [arXiv:1101.2755 \[hep-ex\]](#).
- [105] N. Klop and A. Palazzo, “*Imprints of CP violation induced by sterile neutrinos in T2K data*,” Physical Review D **91** (2015) no. 7, 073017, [arXiv:1412.7524 \[hep-ph\]](#).

- [106] A. Palazzo, “*3-flavor and 4-flavor implications of the latest T2K and NO ν A electron (anti-)neutrino appearance results*,” Physics Letters B **757** (2016) 142, arXiv:1509.03148 [hep-ph].
- [107] **Daya Bay Collaboration**, F. An *et al.*, “*Independent measurement of the neutrino mixing angle θ_{13} via neutron capture on hydrogen at Daya Bay*,” Physical Review D **90** (2014) no. 7, 071101, arXiv:1406.6468 [hep-ex].
- [108] S. C and R. Mohanta, “*Towards extracting the best possible results from NO ν A*,” European Physical Journal C **76** (2016) no. 6, 302, arXiv:1605.00523 [hep-ph].
- [109] S. C, K. Deepthi, and R. Mohanta, “*A comprehensive study of the discovery potential of NO ν A, T2K and T2HK experiments*,” Advances in High Energy Physics **2016** (2016) 9139402, arXiv:1408.6071 [hep-ph].
- [110] K. Deepthi, S. C, and R. Mohanta, “*Revisiting the sensitivity studies for leptonic CP-violation and mass hierarchy with T2K, NO ν A and LBNE experiments*,” New Journal of Physics **17** (2015) no. 2, 023035, arXiv:1409.2343 [hep-ph].
- [111] M. Ghosh, P. Ghoshal, S. Goswami, N. Nath, and S. Raut, “*New look at the degeneracies in the neutrino oscillation parameters, and their resolution by T2K, NO ν A and ICAL*,” Physical Review D **93** (2016) no. 1, 013013, arXiv:1504.06283 [hep-ph].
- [112] S. Goswami and N. Nath, “*Implications of the latest NO ν A results*,” arXiv:1705.01274 [hep-ph].
- [113] J. Kopp, P. Machado, M. Maltoni, and T. Schwetz, “*Sterile Neutrino Oscillations: The Global Picture*,” Journal of High Energy Physics **2013** (2013) 050, arXiv:1303.3011 [hep-ph].
- [114] M. Ghosh, S. Gupta, Z. Matthews, P. Sharma, and A. Williams, “*Study of parameter degeneracy and hierarchy sensitivity of NO ν A in presence of*

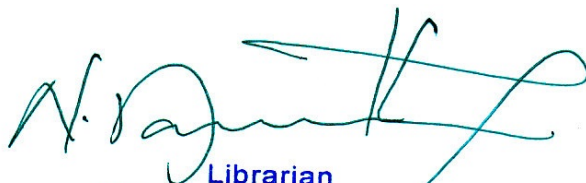
- sterile neutrino*,” Physical Review D **96** (2017) no. 7, 075018, arXiv:1704.04771 [hep-ph].
- [115] S. Agarwalla, S. Chatterjee, and A. Palazzo, “*Physics Reach of DUNE with a Light Sterile Neutrino*,” Journal of High Energy Physics **2016** (2016) 016, arXiv:1603.03759 [hep-ph].
- [116] D. Dutta, R. Gandhi, B. Kayser, M. Masud, and S. Prakash, “*Capabilities of long-baseline experiments in the presence of a sterile neutrino*,” Journal of High Energy Physics **2016** (2016) 122, arXiv:1607.02152 [hep-ph].
- [117] S. Agarwalla, S. Chatterjee, and A. Palazzo, “*Signatures of a Light Sterile Neutrino in T2HK*,” Journal of High Energy Physics **2018** (2018) 091, arXiv:1801.04855 [hep-ph].
- [118] S. Agarwalla, S. Chatterjee, and A. Palazzo, “*Octant of θ_{23} in danger with a light sterile neutrino*,” Physical Review Letters **118** (2017) no. 3, 031804, arXiv:1605.04299 [hep-ph].
- [119] S. Choubey, D. Dutta, and D. Pramanik, “*Measuring the Sterile Neutrino CP Phase at DUNE and T2HK*,” European Physical Journal C **78** (2018) no. 4, 339, arXiv:1711.07464 [hep-ph].
- [120] P. Adamson and et al. [NOvA Collaboration], “*Search for active-sterile neutrino mixing using neutral-current interactions in NOvA*,” Phys. Rev. D **96** (2017) 072006, arXiv:1706.04592 [hep-ex].
- [121] A. Y. Smirnov, “*The MSW effect and matter effects in neutrino oscillations*,” Phys. Scripta T **121** (2005) 57, hep-ph/0412391.
- [122] P. Huber, M. Lindner, and W. Winter, “*Simulation of long-baseline neutrino oscillation experiments with GLoBES (General Long Baseline Experiment Simulator)*,” Computer Physics Communications **167** (2005) 195, hep-ph/0407333.

- [123] P. Huber, J. Kopp, M. Lindner, M. Rolinec, and W. Winter, “*New features in the simulation of neutrino oscillation experiments with GLoBES 3.0: General Long Baseline Experiment Simulator*,” Computer Physics Communications **177** (2007) 432, [hep-ph/0701187](#).
- [124] J. Kopp, “*New physics engine for the inclusion sterile neutrinos and non-standard interactions in GLoBES*,” Available at:
<https://www.mpi-hd.mpg.de/personalhomes/globes/tools.html>.
- [125] P. Adamson and et al. [NOvA Collaboration], “*First measurement of muon-neutrino disappearance in NOvA*,” Physical Review D **93** (2016) no. 5, 051104, [arXiv:1601.05037 \[hep-ex\]](#).
- [126] P. Adamson and et al. [NOvA Collaboration], “*First measurement of electron neutrino appearance in NOvA*,” Physical Review Letters **116** (2016) no. 15, 151806, [arXiv:1601.05022 \[hep-ex\]](#).
- [127] D. S. Ayres and et al. [NOvA Collaboration], “*The NOvA Technical Design Report*,”.
- [128] S. K. Agarwalla, S. Prakash, S. K. Raut, and S. U. Sankar, “*Potential of optimized NOvA for large θ_{13} & combined performance with a LArTPC & T2K*,” Journal of High Energy Physics **1212** (2012) 075, [arXiv:1208.3644 \[hep-ph\]](#).
- [129] D. S. Ayres and et al. [NOvA Collaboration], “*NOvA: Proposal to Build a 30 Kiloton Off-Axis Detector to Study $\nu_\mu \rightarrow \nu_e$ Oscillations in the NuMI Beamline*,” [hep-ex/0503053](#).
- [130] R. Acciarri and et al. [DUNE Collaboration], “*Long-Baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE): Conceptual Design Report, Volume 2: The Physics Program for DUNE at LBNF*,” [arXiv:1512.06148 \[physics.ins-det\]](#).

- [131] R. Acciarri and et al. [DUNE Collaboration], “*Long-Baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE): Conceptual Design Report, Volume 1: The LBNF and DUNE Projects*,” arXiv:1601.05471 [physics.ins-det].
- [132] T. Alion and et al. [DUNE Collaboration], “*Experiment Simulation Configurations Used in DUNE CDR*,” arXiv:1606.09550 [physics.ins-det].
- [133] P. Adamson and et al. [NOvA Collaboration], “*Constraints on Oscillation Parameters from ν_e Appearance and ν_μ Disappearance in NOvA*,” Phys. Rev. Lett. **118** (2017) 231801, arXiv:1703.03328 [hep-ex].
- [134] **NOvA**, M. A. Acero *et al.*, “*First Measurement of Neutrino Oscillation Parameters using Neutrinos and Antineutrinos by NOvA*,” Phys. Rev. Lett. **123** (2019) no. 15, 151803, arXiv:1906.04907.

Classification of particles with a convolutional neural network for neutrino and anti neutrino events in the NOvA experiment

by Akshay Chatla



Librarian
Indira Gandhi Memorial Library
UNIVERSITY OF HYDERABAD
Central University P.O.
HYDERABAD-500 046.

Submission date: 19-Jan-2024 12:24PM (UTC+0530)

Submission ID: 2273759737

File name: Akshay_Chatla_16PHPH08_PhD_Thesis.pdf (18.07M)

Word count: 31293

Character count: 156058

Rundin & Bambah



Classification of particles with a convolutional neural network for neutrino and anti neutrino events in the NOvA experiment

ORIGINALITY REPORT

24%

SIMILARITY INDEX

23%

INTERNET SOURCES

17%

PUBLICATIONS

4%

STUDENT PAPERS

PRIMARY SOURCES

1

export.arxiv.org

Internet Source

7%

2

arxiv.org

Internet Source

3%

3

preview.hindawi.com

Internet Source

2%

4

www.hindawi.com

Internet Source

1%

5

www.arxiv-vanity.com

Internet Source

1%

6

downloads.hindawi.com

Internet Source

1%

7

indico.fnal.gov

Internet Source

1%

8

scholarworks.iu.edu

Internet Source

1%

indico.cern.ch

Bindu A. Bapat



9	Internet Source	<1 %
10	lss.fnal.gov Internet Source	<1 %
11	"XXIII DAE High Energy Physics Symposium", Springer Science and Business Media LLC, 2021 Publication	<1 %
12	dspace.uohyd.ac.in Internet Source	<1 %
13	inspirehep.net Internet Source	<1 %
14	Submitted to University of Hyderabad, Hyderabad Student Paper	<1 %
15	digital.library.adelaide.edu.au Internet Source	<1 %
16	resolver.caltech.edu Internet Source	<1 %
17	Submitted to University College London Student Paper	<1 %
18	escholarship.org Internet Source	<1 %
19	tel.archives-ouvertes.fr Internet Source	<1 %

Bindu A. Bambah



This is to certify that the thesis entitled “Classification of particles with a convolutional neural network for neutrino and antineutrino events in the NOvA experiment” has been screened by the Turnitin software at the library of University of Hyderabad. The software shows 24% similarity index out of which, 16% came from the candidate’s research articles related to this thesis . That leaves only 8% similarity from other sources which is lower than maximum similarity (10%) allowed by University of Hyderabad . Therefore, this thesis is free from plagiarism.

Bindu A. Bambah

Prof. Bindu A Bambah

Supervisor

School of Physics

University of Hyderabad



Date: 23.01.2024

Rukmani Mohanta

Prof. Rukmani Mohanta

Co-Supervisor

School of Physics

University of Hyderabad

Dr. Rukmani Mohanta

Professor

School of Physics

UNIVERSITY OF HYDERABAD

Hyderabad-500 046.