Electromagnetic Shower from Cosmic Muons in NOνA Detector

Nitin Yadav

A thesis
submitted for the degree of
Doctor of Philosophy

Department of Physics
Indian Institute of Technology Guwahati
Guwahati 781039, India

March 2018

Electromagnetic Shower from Cosmic Muons in NOνA Detector

Nitin Yadav

A thesis
submitted for the degree of
Doctor of Philosophy

Supervisor:
Dr. Bipul Bhuyan

Department of Physics
Indian Institute of Technology Guwahati
Guwahati 781039, India

March 2018

Declaration

The work in this thesis is based on research carried out at the Department of Physics, Indian Institute of Technology Guwahati, India and Fermi National Accelerator Laboratory (FNAL), USA under the supervision of Dr. Bipul Bhuyan, Dr. Peter Shanahan (Fermilab) and Prof. Sanjib Mishra (University of South Carolina). No part of this thesis has been submitted elsewhere for any other degree or qualification. Works presented in the thesis are all my own unless referenced to the contrary in the text.

Signed: _____

Date: _____

Certificate

It is certified that the work contained in the thesis entitled "*Electromagnetic Shower from Cosmic Muons in NOνA Detector*" by Mr. Nitin Yadav, a PhD student at the Department of Physics, IIT Guwahati was carried out under my supervision and has not been submitted elsewhere for award of any degree.

Dr. Bipul Bhuyan
Indian Institute of Technology Guwahati, India

Dr. Peter Shanahan
Fermi National Accelerator Laboratory, USA

Prof. Sanjib Mishra
University of South Carolina, USA

To my grandparents and parents...

Acknowledgements

I feel very humble in writing the current page. It has taken a numerous number of people help and contributions in my life before I could earn this moment. My current status of learning, while achieving a Ph.D. degree, is a sum of uncountable small and big learnings which I have gathered throughout my life and therefore I would like to take this opportunity, before I could start writing a formal apparent acknowledgment, to deeply express my gratitude toward all, known and unknown, evident or non evident, people who made my journey of this learned experience so far possible.

I can not begin, without the mention of my beloved institution which provided me one of the most prestigious affiliation, the Indian Institute of Technology Guwahati, a naturally beautiful campus, in the foothills of Great Himalayas on the banks of majestic Brahmaputra, with its lakes, hills and flowers serve a splendid feat to one's eyes and it has contributed immensely to my learning process as this all has inculcated in me, slowly and silently, a thoughtful phenomenon to know more about the nature and its functioning, and its laws.

Having said that, now firstly and foremost, I would like to express my sincere gratitude to my advisor Dr. Bipul Bhuyan for the continuous support during my research work, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D study. Alongwith, I would also like to thank Dr. Pahi Saikia for her support.

Besides my advisor, I would like to thank the rest of my thesis committee members: Prof. P. Poulose, Dr. Udit Raha, and Dr. M. K. Bhuyan, for reviewing my yearly progress and for giving me their insightful comments and encouragement, but also for the hard question which incited me to widen my research from various perspectives. I would also like to thank, all former and current, Head of Department of Physics, teaching and non-teaching staff at IITG.

My sincere thanks also goes to Dr. Peter Shanahan and Prof Sanjib R. Mishra who provided me an opportunity to learn under their able supervision. I am thankful to Fermilab administration and US Govt. for giving an access to the Fermilab laboratory and research facilities. Without their precious support it would not have been possible to conduct this research. I am thankful to Albert Stebbins and Sonya Stebbins for their support. I am also thankful to all the members of Housing Office at Fermilab, specially Jack Hawkins, Cheryl Bentham Jacqueline Cyko, Jeane Lasusa, Thomas Eggleston and Daniel Lee, for making my stay much comfortable and memorable at Fermilab.

I greatly owe a lot in my learnings to Dr. Hongyue Duyang and I sincerely thank him as with him I learned the basics of the Brem work and as a result of his help throughout, I could also do the DiF work. I would also like to thank Prof. Stanley Wojcicki for giving me an opportunity to have multiple fruitful discussion with him on DiF work. He was very kind to give me his time and his valuable suggestions on my work.

Thanks to the India-Fermilab Neutrino Collaboration and its members for providing me the opportunity to work at Fermilab and supporting my stay. I thank Dr. Rajendran Raja, the founder of the collaboration, Prof. Brajesh Chandra Chaudhary and Dr. Shekhar Mishra. I also thank Prof. Raj Gandhi for arranging a special lecture series at Fermilab to help us in learning the theoretical framework of neutrinos. A special thank to Sonam for helping me out for several things during my stay at Fermilab.

I would like to say thanks to all the students, post docs and members of NOνA collaboration, particularly Chris Backhouse, Patricia Vahle, Gavin Davies, Evan Niner, Dominick Rocco, Keith Matera, Xuebing Wu, Michael Baird, Martin Frank, Mathew Muethar, Louise Suter, Pengfei Ding, Shaokai Yang, José Sepúlveda, Erika Catano-Mur,

Barnali and Kanika Sachdev, for helping me in providing solutions to all the technical, computing or Physics related issues.

Thanks to all my teachers, from my school days till date, whose efforts and blessings have helped me in reaching to the current stage.

I thank my fellow labmates Deepanwita Dutta, Kamal Jyoti Nath, Maharnab, Dhiraj, Dibyajyoti, Jyotirmoi and Devender for their help and discussions. A special thanks to Navaneeth, Deepthi, Pavan, Ram, Anuj, Prahlad, Niharika, Biju, Rahul, Sujay and Chandita for their friendship and help all throughout.

This acknowledgement would remain incomplete without the mention of the love, the care, the hard work and the sacrifice of my ancestors, my grandparents, my mother, my father, my relatives and Guruji in bringing me up and also the utmost affection, given to me by my sisters and my dear ones Neelima Di, Teshma Di and Anamika, Tommy, Krishna, Kanha, Dolcy and Deepanjali, that helped me in keep moving blissfully and to them any degree of acknowledgment would somewhere, definitely, leave in me a void of ungratefulness.

-Nitin Yadav

Abstract

NOνA (NuMI Off-Axis ν_e Appearance) experiment is a neutrino $\nu_\mu \rightarrow \nu_e$ oscillation experiment with a baseline of 810 km. The signature for a $\nu_\mu \rightarrow \nu_e$ oscillation in the far detector is the production of an electromagnetic (EM) shower due to the charge current (CC) interaction of ν_e with the detector material. Cosmic muon induced electromagnetic showers can serve as a source of EM shower sample that can be used as a data-driven method for validating the EM shower modelling and EM shower identification algorithms at NOνA. The work reported in the thesis is the first ever implementation of cosmic muon induced EM shower for such a purpose in the NOνA experiment. We extracted EM shower sample from two sources - muon undergoing bremsstrahlung (Brem) and muon decay in flight to electron (DiF). Brem sample, extracted from cosmic muon, turns out to be a very useful data-driven method for testing EM shower modelling and the particle identification (PID) algorithms for the detection of ν_e-CC interactions. A new muon removal (MR) technique is developed and applied to remove the muon hits inside the bremsstrahlung shower. However, a complete removal of the muon energy in the event is not possible with high efficiency. DiF induced EM shower, on the other hand, is exactly similar to ν_e-CC EM shower because DiF is a pure electron initiated EM shower. Therefore, DiF EM shower, once extracted, mimicks the shower of a ν_e-CC interaction event. We present the extraction methods for both brem and DiF events in MC and cosmic data in far and near detector of NOνA and its use in data driven validation of ν_e EM shower modelling and ν_e signal selection efficiency across the detector. A systematic uncertainty is also assigned on ν_e appearance signal selection based on this study to take into account the difference in EM shower modelling in data and MC.

Preface

Neutrinos are the most abundant particles in the universe after photon. Billions of neutrinos are passing through 1 cm^2 of earth every second. Despite of having such a huge abundance across the universe, the neutrino remained elusive and unknown till it was discovered during early nineteenth century. In 1911, Lise Meitner and Otto Hahn found that β-decay energy spectrum is continuous rather than discrete. The expected discrete β decay energy spectrum was consistent with the two particles coming out of the decay. On the contrary, the continuous nature of decay was an indication of a third particle being emitted which carries away the part of energy along with e^- and in 1930, it led Wolfgang Pauli to postulate a new particle which has the characteristics of being a spin 1/2, massless, chargeless and undetectable particle. Soon after, in 1931, Fermi put forward the theory of β decay with the inclusion of neutrino.

Neutrino is one of the fundamental particle in the Standard Model (SM) of particle physics. It is a chargeless fermion (spin 1/2 particle) with a very tiny mass. It interacts via weak force and gravity, though gravitational force is negligible due to its tiny mass [1][2]. The cosmological upper limit on mass of neutrino is of the order of 1 eV or less [3][4].

Neutrino was first propounded by Wolfgang Pauli in 1930 as a remedy to explain the continuous nature of β-decay spectrum [5]. Soon after in 1932, Fermi put forward the theory of beta decay with the inclusion of neutrino in it [6]. Bethe and Peierls obtained the cross section of neutrino interaction in the matter to be very small ($\sigma < 10^{-44} cm^2$) and concluded that "*there is no practically possible way of observing the neutrino*" [7]. In 1946, Pontecarvo [8] took the challenge and proposed a radiochemical method of detecting neutrino based on Cl-Ar reaction as shown in Eq.1.

$$\nu + {}^{37}Cl \rightarrow e^- + {}^{37}Ar \tag{1}$$

In 1956, Fred Reines and Clyder Cowan detected the neutrinos in a nuclear reactor. First atmospheric neutrinos were detected in Kolar mines in 1960 in India [9]. In 1956

Chein-Shiung Wu discovered parity violations in beta decay [10] and in 1957 Goldhaber et al. established that neutrinos are left handed. Few years later in 1960, R Davis observed the solar neutrinos in an experiment based on Cl-Ar (Eq.1) reaction [11]. Following that, successive experiments were carried out to detect and study neutrinos. In 1985, Kamiokande and IMB experiments observed less amount of solar neutrinos than predicted by Standard Solar model [12]. The problem was termed as the solar neutrino deficit problem [13].

The results indicated the deficit in numbers of neutrinos coming from Sun. The deficit was explained by SNO and Super Kamiokande (Super-K) experiments, as due to oscillation process among different kind of neutrinos (μ, e, τ) and bagged the Nobel prize in Physics in 2015 [14][15]. Neutrino oscillation imply that neutrinos have mass, which otherwise were considered massless. Current generation of neutrino experiments such as NOνA and T2K use a ν_μ beam produced at the accelerators to study the phenomenon of neutrino oscillation and for the precision measurement of the neutrino oscillation parameters [16], [17]

NOνA is a long baseline two detector neutrino experiment designed to measure the mass hierarchy and δ_{CP} associated with neutrino oscillation [18] with a baseline of 810 km. The near detector (ND) is a 290 ton detector and located 1 km downstream of the beam target whereas the far detector (FD) is 14 kt and located in Ash river, Minnesota, downstream of beam line at a distance of 810 km from the beam target. Both the detectors are made up of finely segmented plastic cells and filled with liquid scintillator. The cell width is 3.9 cm by 6.6 cm and its length is 15.5 m (3.9 m) long in far (near) detector. NOνA studies neutrino oscillations in both $\nu_\mu \rightarrow \nu_\mu$ ($\bar{\nu_\mu} \rightarrow \bar{\nu_\mu}$) disappearance and $\nu_\mu \rightarrow \nu_e$ ($\bar{\nu_\mu} \rightarrow \bar{\nu_e}$) appearance mode. In $\nu_\mu \rightarrow \nu_e$ ($\bar{\nu_\mu} \rightarrow \bar{\nu_e}$) appearance mode, the signal is an electromagnetic (EM) shower initiated by an electron produced in ν_e charge current interaction in the detector. The success of the experiment is based on how accurately the simulations are done and how good is the efficiency of signal (EM shower generated in ν_e-CC interaction in FD) detection along with how well the systematic uncertainties are controlled in the detector.

Parameters	Best fit	3σ
$\Delta m^2_{21}[10^{-5}eV^2]$	7.37	6.93 - 7.97
$\lvert\Delta m^2_{32}\rvert$ $[10^{-3}eV^2]$	2.50	2.37 - 2.63 (normal mass hierarchy
$\lvert\Delta m^2_{32}\rvert$ $[10^{-3}eV^2]$	2.46	2.33 - 2.60 (inverted mass hierarchy
$\sin^2\theta_{12}$	0.297	0.250 - 0.354
$\sin^2\theta_{23}$	0.437	0.379 - 0.616 (normal hierarchy)
$\sin^2\theta_{23}$	0.567	0.383 - 0.637 (inverted hierarchy)
$\sin^2\theta_{13}$	0.0214	0.0185 - 0.0246 (normal hierarchy)
$\sin^2\theta_{13}$	0.0218	0.0186 - 0.0248 (inverted hierarchy)
δ/π	1.35	0.83 - 1.99 (normal hierarchy)
δ/π	1.32	0.92 - 1.99 (inverted hierarchy)

TABLE 1: Best fit values for neutrino oscillation parameters [27].

The main goals of current generation of neutrino experiment are to measure the six parameters that govern the neutrino oscillations $\Delta m^2_{21}, \Delta m^2_{32}, \theta_{12}, \theta_{23}, \theta_{13}$ and δ_{CP}. The first precision measurement of θ_{12} and Δm^2_{21} have been done by SuperK [19] , SNO [20] and KamLAND [21] experiment. Reactor exeperiment like Daya Bay [22], RENO [23] and Double Chooz [24] have measured θ_{13}. Along with NOνA, the other contemporary accelerator based neutrino oscillation experiment are T2K [25] and MINOS [26]. The mass hierarchy and δ_{CP} are yet to be measured. The octant of θ_{23} is also not known. Table 1 summarizes the current knowledge of the parameters.

In $\nu_\mu \rightarrow \nu_e$ appearance mode, ν_e interacts in the far detector through a quasi elastic charged current interaction and produces an electromagnetic shower. Fig.1 shows the ν_e charged current and neutral current interactions. The outgoing electron (in case of ν_e charged current) produces an electron which creates an electromagnetic shower in the NOνA detectors.

The topologies of the interaction in far detector is shown in Fig.2. Production of EM shower is the characteristic signal for ν_e appearance during CC interaction and therefore shower modelling must be done accurately. Also the particle identification (PID) algorithms should be able to identify the EM shower event. External source of EM showers from data (other than ν_e-CC) would be very useful to verify the shower modelling and particle-ID at NOνA.

Since the location of far detector is on the surface, the cosmic rate in far detector is $\sim$

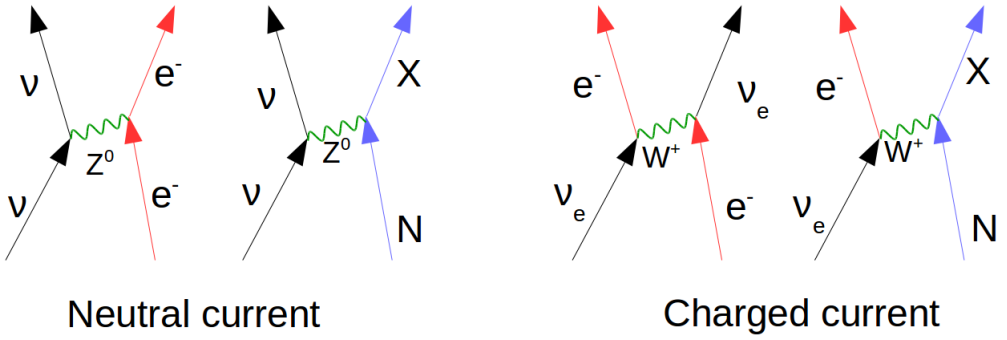

FIGURE 1: Feynmann diagrams for the Neutral Current (NC) and Charged Current (CC) interactions of neutrinos with matter.

148 kHz [18]. The cosmic muon can induce an electromagnetic shower via soft scattering in the form of bremsstrahlung radiation (Brem) and by muon decay in flight (DiF) to a sub-GeV electron. The total cosmic muon induced shower rate, as measured, in far detector is 72 kHz and in ND is 22 Hz [28]. DiF rate is 70 Hz in FD and 1 Hz in ND [29] . In both the cases, EM showers can be extracted[1]. EM showers from bremsstrahlung radiation needs muon to be removed from the shower region, whereas EM shower from muon decay in flight is a pure electron shower. Once the EM shower (bremsstrahlung or decay in flight) is extracted, it can be used as data-driven method for verifying or improving shower modelling and PID at NOνA. We have first studied the bremsstrahlung shower and its usage to validate and benchmark the ν_e-CC shower modelling and PIDs. A far detector event display of bremsstrahlung muon to give an EM shower before muon removal (MR) is shown in Fig.3. Fig.4 shows a similar bremsstrahlung event after removing the muon track from the EM shower. We also extract and study the muon decay in flight EM shower (more similar to ν_e-CC shower) which is a much purer EM shower sample to revalidate the ν_e-CC shower modelling and particle-ID in NOνA. We have also used these samples to study the detector efficiency as a function of EM shower location in the far detector. Fig.5 shows an event display for a muon decay in flight event that produces an EM shower in the far detector.

[1]However, in ND cosmic muons are less in numbers due to its ∼105 m depth and shower is extracted from muons which are produced abundantly in nearby rock, surrounding the near detector, through the beam ν_μ interactions

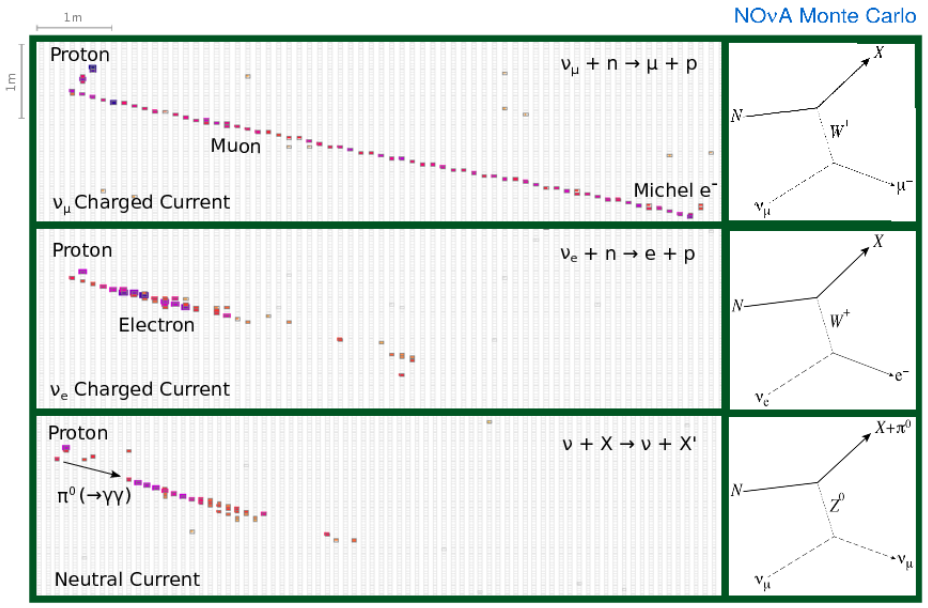

FIGURE 2: Event topologies of different kinds of ν interaction in NOνA far detector. A ν_μ-CC interaction exhibits itself as a long track. ν_e-CC interaction results in an electron which materializes as a shower. π^0 produced in neutral current (NC) interactions in the detector also results in an electromagnetic shower through pair production ($\gamma \rightarrow e^+e^-$).

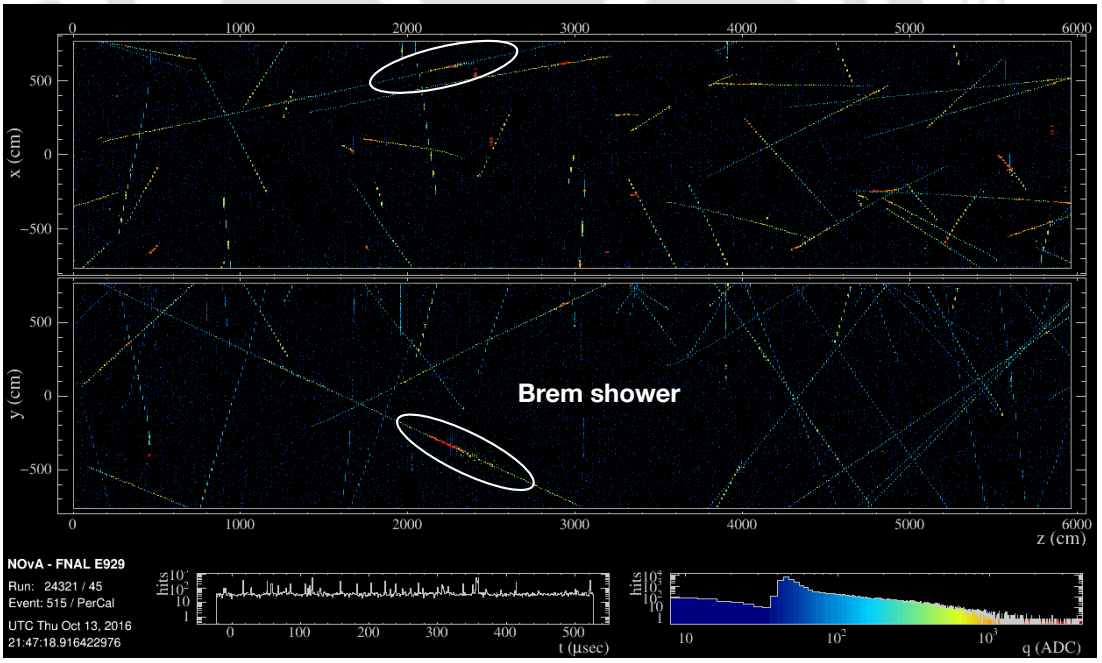

FIGURE 3: A track of a muon undergoing bremsstrahlung EM shower in the far detector. The circled areas show the bremsstrahlung EM shower in two views (X-Z and Y-Z) of the detector. The above image is in a 500μ sec window of a cosmic event in the far detector in X-Y and X-Z view. The bottom left plot shows the timing of the hits in the detector and the bottom right plot shows the charge deposition information in the event.

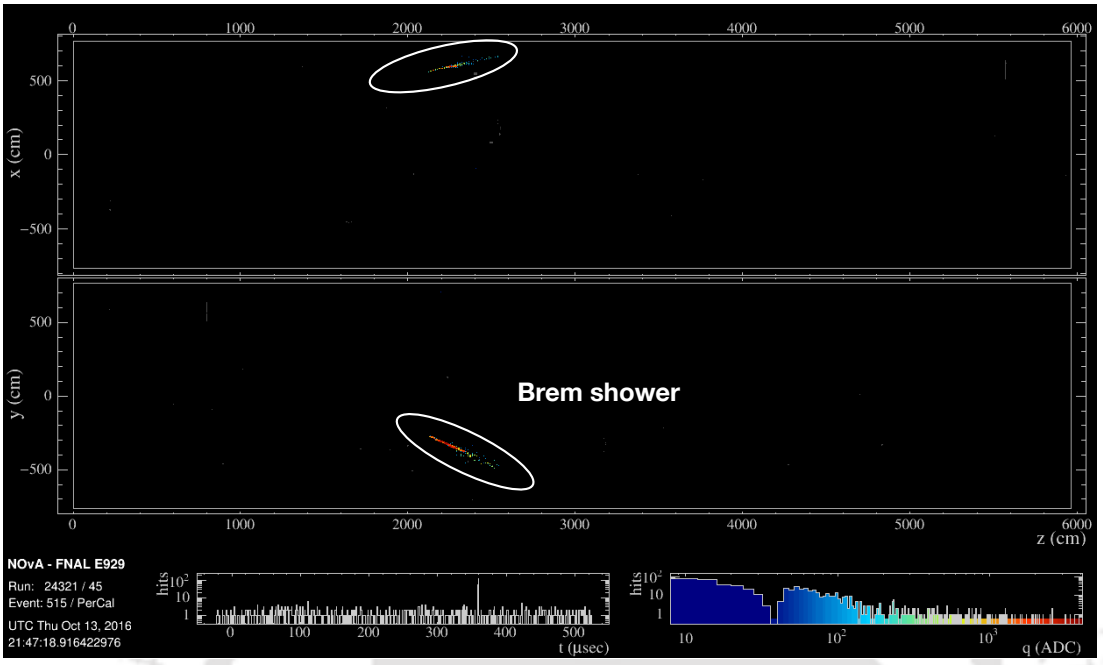

FIGURE 4: The muon hits are removed from the muon track as mentioned in Fig.3. The only hits remains, belong to bremsstrahlung EM shower. The circled areas show the extracted bremsstrahlung EM shower in two views (X-Z and Y-Z) of the detector. The above image is in a 500μ sec window of a cosmic event in the far detector in X-Y and X-Z view. The bottom left plot shows the timing of the hits in the detector and the bottom right plot shows the charge deposition information in the event.

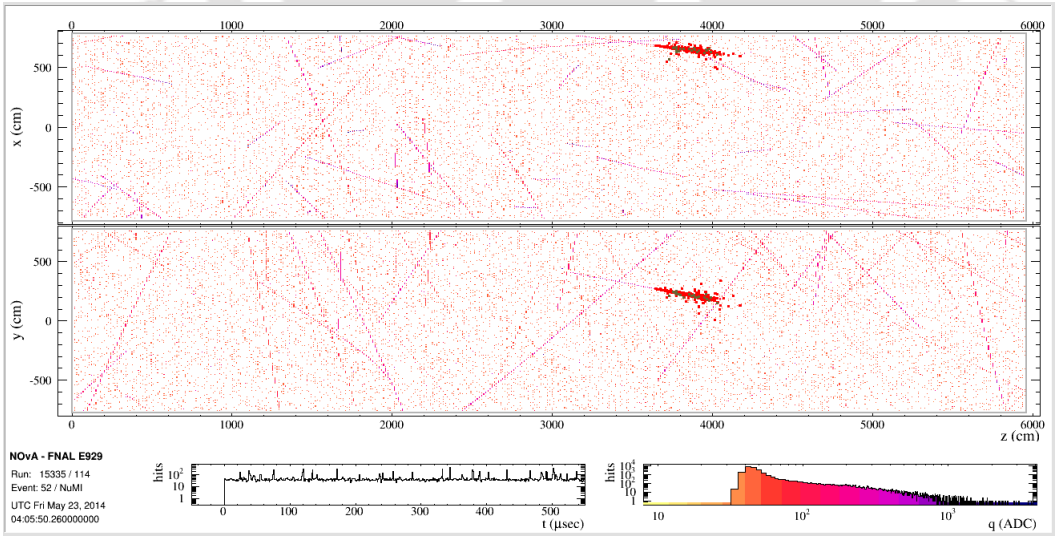

FIGURE 5: A muon undergoing a decay in flight process to give an EM shower in the far detector. The thick red dots correspond to DiF EM shower hits. The above image is in a 500μ sec window of a cosmic event in the far detector in X-Y and X-Z view. The bottom left plot shows the timing of the hits in the detector and the bottom right plot shows the charge deposition information in the event.

The thesis describes the extraction of electromagnetic showers induced by cosmic muons and their usage in detecting (PID) and in the validation of the ν_e-CC generated EM shower modelling in the NOνA detectors. The efficiency of selecting ν_e-CC signal in the far detector can also be studied using EM showers. We have carried out the aforementioned studies based on two kinds of EM showers induced by muons through two processes - bremsstrahlung and muon decay in flight. The work is done in collaboration with groups from Fermilab (Dr. Peter Shanahan), University of South Carolina (Prof. Sanjib R Mishra and Dr. Hongyue Duyang) and Indian Institute of Technology Guwahati (Dr. Bipul Bhuyan). Dr. Hongyue Duyang and Prof. Sanjib R Mishra carried out the preliminary studies of bremsstrahlung EM shower work as described in Dr. Hongyue Duyang thesis work [2]. Based on that intial work, the bremsstrahlung EM shower work was further extended as reported in this thesis.

The details of the study are included in the following chapters :

Chapter 1 of the thesis includes a detailed account of the theory of neutrino oscillations. Since the neutrinos are weakly interacting particles, a review of the weak interactions from the point of view of neutrino interaction will be included in the chapter. The chapter will start with two flavor neutrino oscillation formalism before it is generalized to three flavor neutrino oscillation formalism. Since NuMI (neutrino at the main injector) beam traverses its path inside the earth to reach the far detector, a brief discussion about Mikheyev-Smirnov-Wolfenstein (MSW) effect will also be discussed [30]. NOνA experiment is collecting data with both ν_μ and $\bar{\nu_\mu}$ beam [31], that enables NOνA to explore the possibility of measurement of CP-violation in the neutrino sector. Along with the MSW effect, NOνA can also explore mass-hierarchy of neutrino mass states. The chapter will finish with the discussion of NOνA sensitivity towards the above physics goals.

Chapter 2 includes the description of the NOνA detectors and its DAQ system. NOνA has two detectors - far (FD) and near (ND). Both the detectors are similar built using PVC cells filled with liquid scintillator and have similar architecture except the size. NOνA far detector is 14 kt and near detector is 290 ton. However, the operating

[2] https://scholarcommons.sc.edu/cgi/viewcontent.cgi?article=3622&context=etd

environment of both detectors are different as near detector is placed ∼105 m down the earth surface level and is exposed to high intensity NuMI beam of neutrinos, whereas NOνA far detector is located 810 km downstream of the NuMI neutrino production point and is situated on the surface with an overburden of few radiation lengths. This results in the far detector being exposed to cosmics at a rate of about 148 kHz. Due to the high intensity neutrino beam at ND, the front end board (FEB) sampling ratio of all channels is 8:1 in contrast to 2:1 in far detector. This enables the data to be collected at a faster rate at ND due to more neutrino interaction per beam spill in near detector. Data is collected from FEBs in trigger less mode to allow the NOνA to take non-beam triggered cosmic data for calibration. The spill trigger is applied at the buffer node. This chapter includes a detailed description of the NOνA detector geometry, the data acquisition system, online monitoring for data quality, the monitoring of the DAQ system, NOνA simulations and the detector calibrations.

Chapter 3 includes a discussion on the ν_e-CC EM shower reconstruction and ν_e-CC event identification algorithms for ν_e-appearance mode analysis. NOνA aims to identify ν_μ and ν_e interaction in the detector for its main physics goals. It also aims to identify particle signatures of pions and muons. The reconstruction of events at NOνA consists of various steps starting from hits which contain the information about charge deposition and its location in 2-D form. After this, the hits are identified for spatial and temporal correlation to produce new objects called as clusters. Cosmic tracks are reconstructed using cosmic track algorithm. Hough Transformation algorithm is applied on the hits of the cluster to identify the pattern in the form of straight lines. Hough transformation is followed by identifying the vertex in the event using Elastic Arms algorithm. Once the vertex is identified, the next step is to form prongs which has a starting point and direction and hits associated with it. After this point, particle identification algorithms (PID) are used to classify the particles and the events.

Chapter 4 begins with the description of selection and extraction of bremsstrahlung EM shower from the cosmic muons. The study of bremsstrahlung EM shower as a data driven method to benchmark shower modelling and the use of it as a controlled sample to study detector inefficiency as a function of shower position and energy is also discussed.

Such evaluation provides a strong basis for understanding the systematic uncertainty associated with the $\nu_\mu \rightarrow \nu_e$ oscillations. In the chapter, we include the description of the preselection and the final set of selection criteria and their optimization for selection of muon that produces the bremsstrahlung shower. Once such an event is selected, one needs to remove the muon hits in the event. This is achieved by muon removal (MR) algorithm which removes the muon from inside and outside the bremsstrahlung shower region. Using variables such as shower energy and angle, the bremsstrahlung shower sample is reweighed to match with the ν_e-CC generated shower. Using the reweighed sample, the detector inefficiency has been studied which is presented in this chapter.

Chapter 5 includes the detailed selection of EM shower generated by muon decay in flight. The usage of DiF sample to re-validate the ν_e-CC modelling and to study the detector inefficiency as a function of shower's kinematics are discussed.

Chapter 6 includes the final results followed by a discussion on the usefulness of the two samples to study EM showers in the NOνA detectors. Both the bremsstrahlung and DiF samples are used as data driven method for validation of shower modelling and particle-IDs. The agreement between data and MC is well within 4%. Comparison of results for Brem and DiF are presented in this chapter. Further, we argue the importance of similar studies for other current generation of neutrino experiments such as T2K and possibly also for future neutrino experiments as well.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Neutrinos are the most abundant particles in the Universe after photon. Billions of neutrinos pass through a 1 cm^2 of earth every second. Despite of its such a huge abundance across the Universe, it was only, no earlier than early nineteenth century, it was discovered. In 1911, Lise Meitner and Otto Hahn found that β-decay energy spectrum is continuous rather than discrete [47]. The expected discrete β decay energy spectrum, was consistent with the two particles coming out of the decay. On the contrary, the continuous nature of β-decay spectrum was an indication of a third particle being emitted which carries away a part of energy along with electron. In 1930, this observation led Wolfgang Pauli to postulate a new particle which has the characteristics of being a spin 1/2, massless, chargeless and undetectable particle. He wrote in his famous letter [48]:

> *Dear Radioactive Ladies and Gentlemen,*
>
> *As the bearer of these lines, to whom I graciously ask you to listen, will explain to you in more detail, because of the wrong statistics of the N- and Li-6 nuclei and the continuous beta spectrum, I have hit upon a desperate remedy to save the exchange theorem of statistics and the law of conservation of energy. Namely, the possibility that in the nuclei there could exist electrically neutral particles, which I will call neutrons, that have spin 1/2 and obey the exclusion principle and that further differ from light quanta in that they do not travel with the velocity of light. The mass of the neutrons should be of the same order of magnitude as the electron mass and in any event not larger than 0.01 proton mass. - The continuous beta spectrum would*

then make sense with the assumption that in beta decay, in addition to the electron, a neutron is emitted such that the sum of the energies of neutron and electron is constant. Now it is also a question of which forces act upon neutrons. For me, the most likely model for the neutron seems to be, for wave-mechanical reasons (the bearer of these lines knows more), that the neutron at rest is a magnetic dipole with a certain moment . The experiments seem to require that the ionizing effect of such a neutron can not be bigger than the one of a gamma-ray, and then is probably not allowed to be larger than 1013cm. But so far I do not dare to publish anything about this idea, and trustfully turn first to you, dear radioactive people, with the question of how likely it is to find experimental evidence for such a neutron if it would have the same or perhaps a 10 times larger ability to get through [material] than a gamma-ray. I admit that my remedy may seem almost improbable because one probably would have seen those neutrons, if they exist, for a long time. But nothing ventured, nothing gained, and the seriousness of the situation, due to the continuous structure of the beta spectrum, is illuminated by a remark of my honored predecessor, Mr. Debye, who told me recently in Bruxelles: Oh, Its better not to think about this at all, like new taxes. Therefore one should seriously discuss every way of rescue. Thus, dear radioactive people, scrutinize and judge. - Unfortunately, I cannot personally appear in Tubingen since I am indispensable here in Zurich because of a ball on the night from December 6 to 7. With my best regards to you, and also to Mr. Back, your humble servant

W. Pauli.

After the Pauli's introduction of neutrino[1], the complete equation of β decay is shown in Eq.1.1 in which a neutron converts to a proton in nucleus with the emission of an e^- and $\bar{\nu}$.

$$n \rightarrow p + e^- + \bar{\nu} \tag{1.1}$$

Based on the Pauli's idea, Fermi put forward the theory of β decay with the inclusion of neutrino in 1931 [6], which is discussed in the next section.

[1]What Pauli called as "neutron" is now a days called as "neutrino',' after the discovery of neutron in 1932 by Chadwick [49].

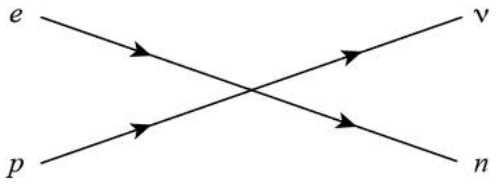

FIGURE 1.1: Fermi four point interaction of β decay.

1.1 The Weak Interaction

Neutrinos interact via the weak force. Fermi described the beta decay in the form of four point interaction which involves e, p, n and ν fields at the same point in space and time as shown in Fig.1.1.

The matrix element for the four point vertex can be written as,

$$M = G_F(\bar{u_n}\gamma_\mu u_p)(\bar{u_\nu}\gamma^\mu u_e) \tag{1.2}$$

where G_F is the coupling constant, u_i's are the Dirac spinors and γ_μ's are the Dirac gamma matrices. This interaction is a vector-vector type and involves no propagator and it conserves parity. Prior to 1950's, parity was considered to be conserved in all the processes, however, in 1956, Wu et al., discovered the parity violation in weak interactions [10]. With the inclusion of parity violation the vector-vector nature of interaction in Eq.1.2, is replaced by $V - A$ current and the matrix element takes the form as,

$$M = \frac{G_F}{\sqrt{2}}[\bar{u_n}\gamma^\mu(1-\gamma^5)u_p][\bar{u_\nu}\gamma_\mu(1-\gamma^5)u_e] \tag{1.3}$$

The theory could explain charge current weak interaction but it was not a complete theory. Later on Sheldon Glashow, Abdus Salam and Steven Weinberg unified electromagnetic and weak interaction and put forward the electroweak theory [50][51][52]. The new theory is based on a gauge group $SU(2) \times U(1)$ that predicts three gauge bosons ($W^\pm$ and Z^0). The four point interaction in the Fermi theory was replaced by ($W^\pm$,Z^0) propagator. The weak interaction mediated by $W^\pm$ (charged current) and Z^0 (neutral current) are shown in Fig.1.2.

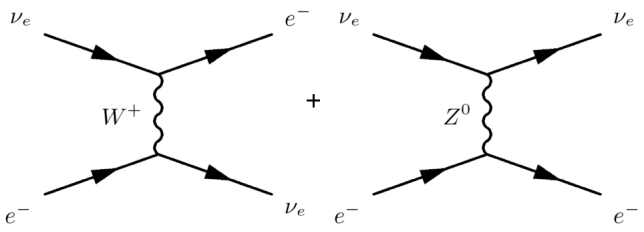

FIGURE 1.2: Two types of weak interaction. The charged current is mediated by $W^{\pm}$ and neutral current is mediated via Z^0

From Eq.1.3 the $e - \nu$ vertex,

$$j_{\bar{\mu}} = \bar{u_\nu}\gamma_\mu\left(\frac{1-\gamma^5}{2}\right)u_e \tag{1.4}$$

can be written as,

$$j_{\bar{\mu}} = \bar{\nu_L}\gamma_\mu e_L \tag{1.5}$$

where, ν_L and e_L are the left handed chiral projections as,

$$\bar{\nu_L} = \bar{u_\nu}\frac{1}{2}\left(1+\gamma^5\right) \; and \; e_L = \frac{1}{2}\left(1-\gamma^5\right)u_e \tag{1.6}$$

As a consequence of $V - A$ theory, the weak charged current interaction allows the coupling to only left-handed chiral fermions and therefore the right handed coupling to charge current interaction vanishes. The coupling of left-handed and right-handed neutrino to Z^0 is $\propto I_z - qsin^2\theta_W$ and $\propto qsin^2\theta_W$, respectively, where θ_W is the Weinberg angle and q is the electric charge. Due to being a neutral particle, only left-handed coupling survives with Z^0. Therefore, the right-handed neutrinos do not couple with either

weak currents. In summary, the left-handed neutrinos with weak isospin z projection $I_z = 1/2$ and corresponding left handed lepton ($I_z = -1/2$) form a weak isospin doublet and right-handed neutrino is singlet ($I_z = 0$).

1.2 Mass of Neutrino

The mass term for fermion in the Dirac Lagrangian can be written as [53],

$$L_m^{Dirac} = -m\bar{\psi}\psi \tag{1.7}$$

where ψ and $\bar{\psi}$ are the Dirac fields. Using chiral projection, it can be written as,

$$L_m^{Dirac} = -m(\bar{\psi_L}\psi_R + \bar{\psi_R}\psi_L) \tag{1.8}$$

where ψ_{LR} are two independent chiral fields, in other words, in case of massive fermion both chiralities should exist. However, in case of neutrinos, only left handed chirality states exist and as a consequence, the neutrinos are massless in the Standard Model.

In the minimal extension of Standard Model, one can write a general Dirac mass term assuming the existence of a right handed neutrinos in the following form,

$$L_m^{Dirac} = -\sum_{a,b} \bar{\nu}_{aL}(x) M^D_{ab} \nu_{bR}(x) + h.c. \tag{1.9}$$

where a, b indices run over e, μ, τ and M^D is a 3x3 complex matrix. The mass written in the above form in Eq.1.9 introduces a right handed sterile neutrino and and the tiny masses of neutrinos will result in a very small Yukawa couplings which are not of the order of Standard Model Yukawa couplings in case of other fermions.

Majorana mass can be included without the inclusion of right handed sterile neutrinos. Since neutrinos are chargeless, the four independent chirality states collapse to two independent chirality states and conjugate fields are the opposite chirality fields. Considering this, the Majorana term can be written in the form,

$$L^M = -\frac{1}{2}\bar{\nu}_L M^M (\nu_L)^c + h.c. \tag{1.10}$$

where superscript c stands for charge conjugation. To generate the definite mass of neutrino through Majorana term, a particle must be its antiparticle. Also through See-Saw mechanism the tiny masses of neutrinos can be explained [54]. We can represent all the mass terms (Dirac and Majorana mass terms) as-

$$L_{mass} = \begin{pmatrix} \bar{\nu_L^c} & \bar{\nu_R^c} \end{pmatrix} \begin{pmatrix} m_L & m_D \\ m_D & m_R \end{pmatrix} \begin{pmatrix} \nu_L^c \\ \nu_R^c \end{pmatrix} \tag{1.11}$$

Diagonalizing the above mass matrix $\begin{pmatrix} m_L & m_D \\ m_D & m_R \end{pmatrix}$, we get the observable masses as eigenvalues-

$$m_{1,2} = \frac{1}{2}\left((m_L + m_R \pm \sqrt{(m_L - m_R)^2 + 4m_D^2} \right) \tag{1.12}$$

Now suppose $m_L = 0$,(otherwise it will require a Higgs triplet which is not available in Standard Model), $m_R >> m_D$, the mass eigenvalues reduces to-

$$m_1 \sim \frac{m_D^2}{m_R} \tag{1.13}$$

and

$$m_2 \sim m_R(1 + \frac{m_D^2}{m_R}). \tag{1.14}$$

It suggests that, if m_R is very large then m_2 is very large and m_1 is very small which corresponds to the low mass of neutrinos. This is called as Seesaw Mechanism[2].

1.3 Neutrino Oscillations

The evidence that neutrinos are not massless, but massive, is given by the experimental discovery of neutrino oscillation which can be explained using quantum formalism [55].

At time t=0, a flavor eigen state can be expressed as follows:

$$|\nu_\alpha(0)\rangle = \sum_i U_{\alpha i}|\nu_i(0)\rangle \tag{1.15}$$

The states evolution with time can be described by Schrdinger equation

$$i\frac{\partial|\psi(t)\rangle}{\partial t} = H|\psi(t)\rangle \tag{1.16}$$

where H is the Hamiltonian. The general solution of above equation can be written as,

$$|\psi(t)\rangle = e^{-iHt}|\psi(0)\rangle \tag{1.17}$$

where $|\psi(0)\rangle$ is the state at time t=0. At time t=0, the flavor states of neutrino can be written as the linear combination of neutrino mass states as,

$$|\nu_\alpha\rangle = \sum_k U^*_{\alpha k}|\nu_k\rangle \tag{1.18}$$

with $\alpha = e, \mu, \tau$ and $U^*_{\alpha k}$ is the unitary matrix . The mass states are eigen states of Hamiltonian in vacuum

[2]Taken from https://warwick.ac.uk/fac/sci/physics/staff/academic/boyd/gradlec/neutrino_lec_3.pdf

$$H_0|\nu_k\rangle = E_k|\nu_k\rangle \tag{1.19}$$

with energy eigenvalues

$$E_k = \sqrt{p^2 + m^2} \tag{1.20}$$

In Schrodinger Eq.1.17, the massive states evolve with time as

$$|\nu_k(t)\rangle = e^{-iHt}|\nu_k\rangle \tag{1.21}$$

Where,

$$H = H_0 + H_m \tag{1.22}$$

where H_0 is given by Eq.1.19 and H_m is given as,

$$H_m|\nu_k\rangle = V_\alpha^m|\nu_k\rangle \tag{1.23}$$

where V_α^m is the matter potential, which arises due to charged current and neutral current interaction of ν_e in the matter.

From Eq.1.15 and Eq.1.17, we can write the time evolution of flavor states as,

$$|\nu_\alpha\rangle = \sum_k U_{\alpha k}^* e^{-iH_k t}|\nu_k\rangle \tag{1.24}$$

where at t=0, $|\nu_\alpha(t = 0)\rangle = |\nu_\alpha\rangle$. Using the inverted form of Eq.1.18 in above equation, we can write

$$|\nu_\alpha\rangle = \sum_\beta \Big(\sum_k U_{\alpha k}^* e^{-iH_k t} U_{\beta k} \Big) |\nu_\beta\rangle \tag{1.25}$$

From above equation, we can see that at time, t the left side combination of mass eigen states do not remain same as at time t=0 and it changes and this leads to a different flavor of eigen state at time t.

The probability of transiting into another flavor can be found out by squaring the transition amplitude $\langle\nu_\beta|\nu_\alpha(t)\rangle$,

$$P_{\nu_\alpha\to\nu_\beta}(t) = |\langle\nu_\beta|\nu_\alpha(t)\rangle|^2 = \sum_{j,k} U^*_{\alpha k}U_{\beta k}e^{-i(H_k-H_j)t}U_{\alpha j}U^*_{\beta j} \tag{1.26}$$

1.3.1 Oscillations in Vacuum

In vacuum,

$$H = H_0 \tag{1.27}$$

and the Hamiltonian H_0 in vacuum is given by $H_0|\nu_k\rangle = H_k|\nu_k\rangle$. We can write the exponential term in Eq.1.26 as,

$$H_k t = E_k t - p_k.x = (E_k - p_k)L \tag{1.28}$$

Assuming neutrinos travel at speed of light, we can approximate x=t=L (where L is length of distance between source and detector) and $E = |p|$,

$$\Big(E_k - p_k\Big)L = \frac{E_k^2 - p_k^2}{E_k + p_k}L = \frac{m_k^2}{E_k + p_k}L \approx \frac{m_k^2}{2E}L \tag{1.29}$$

Therefore,

$$H_k - H_j \approx \frac{m_k^2 - m_j^2}{2E} \approx \frac{\Delta m_{kj}^2}{2E} \tag{1.30}$$

Hence the transition probability in Eq.1.26 becomes,

$$P_{\nu_\alpha \to \nu_\beta}(t) = \sum_{k,j} U^*_{\alpha k} U_{\beta k} U_{\alpha j} U^*_{\beta j} exp\Big(- i \frac{\Delta m^2_{kj} L}{2E} \Big) \tag{1.31}$$

$$\begin{aligned} P_{\nu_\alpha \to \nu_\beta}(t) = \delta_{\alpha\beta} - 4 \sum_{k>j} Re\big[U^*_{\alpha k} U_{\beta k} U_{\alpha j} U^*_{\beta j}\big] sin^2\big(\frac{\Delta m^2_{kj} L}{2E}\big) \\ +2 \sum_{k>j} Im\big[U^*_{\alpha k} U_{\beta k} U_{\alpha j} U^*_{\beta j}\big] sin^2\big(\frac{\Delta m^2_{kj} L}{2E}\big) \end{aligned} \tag{1.32}$$

In case of three active neutrino flavors oscillations, the unitary matrix in $|\nu_\alpha\rangle = \sum_k U^*_{\alpha k} |\nu_k\rangle$, can be parameterized as-

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13} e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13} e^{-i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.33}$$

$$\begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{12}c_{23} \end{pmatrix} \tag{1.34}$$

where $s_{ij} = \sin\theta_i$ and $c_{ij} = \cos\theta_i$ are the mixing angles and δ is the Dirac CP phase. This way, the unitary matrix is described in terms of three mixing angles, $\theta_{12}, \theta_{23}, \theta_{13}$ and a CP-violating phase , δ (Dirac phase)[3]. The above matrix is called as Pontecarvo Maki Nakagawa Sakata (PMNS) matrix [56].

The argument of sinusoidal term in Eq.1.32, can be written in S.I. units as -

$$\frac{\Delta m^2_{kj} c^4 L}{4E\hbar c} \tag{1.35}$$

Expressing L in km and E in GeV and δm^2_{kj} in eV^2, Eq.1.35 can be written as,

[3]We are not writing the Majorana phase as it has no bearing on the oscillation results.

$$\frac{1.27\Delta m^2_{kj}L}{E} \tag{1.36}$$

Therefore, Eq.1.37 can be written as,

$$\begin{aligned} P_{\nu_\alpha\to\nu_\beta}(t) = \delta_{\alpha\beta} - 4\sum_{k>j} Re\big[U^*_{\alpha k}U_{\beta k}U_{\alpha j}U^*_{\beta j}\big] sin^2\Big(\frac{1.27\Delta m^2_{kj}L}{E}\Big) \\ +2\sum_{k>j} Im\big[U^*_{\alpha k}U_{\beta k}U_{\alpha j}U^*_{\beta j}\big] sin^2\Big(\frac{1.27\Delta m^2_{kj}L}{E}\Big) \end{aligned} \tag{1.37}$$

1.3.2 Two Flavor Neutrino Approximation

Neutrino oscillations can be approximated to two neutrino flavor oscillation for most of the ongoing long baseline neutrino experiments because the oscillation amongst the three flavors are decoupled due to the two different mass squared differences. The two flavor approximation must be considered as instructive only with the advancement of our knowledge about the three flavor neutrino oscillations. Ignoring the relatively smaller term in Eq.1.37 for $\alpha = \beta$, we can write

$$P_{\nu_\mu\to\nu_\mu}(t) \approx 1 - 4\sum_{k>j} |U_{\mu k}|^2|U_{\mu j}|^2 \sin^2\Big(\frac{1.27\Delta m^2_{kj}L}{E}\Big) \tag{1.38}$$

Eq.1.38 is the survival probability for a muon neutrino of energy E after traveling a distance of L km.

For two flavor approximation we consider $\theta_{13} = 0$ and for the long baseline experiment of ratio of order L/E $\sim$ 500, we can approximate sinusoidal term with $\Delta m^2_{21} \sim 0$ and $\Delta m^2_{31} \sim \Delta m^2_{32} = \Delta m^2_{atm}$. Then Eq.1.38 takes the form-

$$P_{\nu_\mu\to\nu_\mu} \approx 1 - \sin^2 2\theta_{23} \sin^2\Big(\frac{1.27\Delta m^2_{atm}L}{E}\Big) \tag{1.39}$$

which is a two flavor neutrino oscillation equation. Long baseline experiments like MINOS [25], T2K[57] and NOνA [58] experiments are sensitive to the following appearance and disappearance channels:

$$\begin{aligned} P_{\nu_\mu \to \nu_e} &\approx 1 - 4|U_{\mu 3}|^2 |U_{\mu 3}|^2 \sin^2 \left(\frac{1.27\Delta m^2_{32} L}{E}\right) \\ &\approx \sin^2 2\theta_{13} \sin^2 \theta_{23} \sin^2 \left(\frac{1.27\Delta m^2_{32} L}{E}\right) \end{aligned} \tag{1.40}$$

and they are also sensitive to following disappearance channel-

$$\begin{aligned} P_{\nu_\mu \to \nu_\mu} &\approx 1 - 4|1 - U_{\mu 3}|^2 |U_{\mu 3}|^2 \sin^2 \left(\frac{1.27\Delta m^2_{32} L}{E}\right) \\ &\approx 1 - \cos^2 2\theta_{13} \sin^2 \theta_{23} \sin^2 \left(\frac{1.27\Delta m^2_{32} L}{E}\right) \end{aligned} \tag{1.41}$$

There are Reactor based neutrino oscillation experiments like Daya Bay [59], Double Chooz [60] and RENO [61] experiments, which are sensitive to following disappearance channel:

$$\begin{aligned} P_{\nu_e \to \nu_e} &\approx 1 - 4|U_{e3}|^2 |1 - U_{e3}|^2 \sin^2 \left(\frac{1.27\Delta m^2_{32} L}{E}\right) \\ &\approx 1 - \sin^2 \theta_{13} \sin^2 \left(\frac{1.27\Delta m^2_{32} L}{E}\right) \end{aligned} \tag{1.42}$$

1.3.3 Matter Effect

As described in section 1.1, neutrinos can interact with matter through charge-current (CC) and neutral-current (NC) reaction. Neutrinos suffer coherent and incoherent forward scattering while traveling through the matter. Amount of incoherent scattering is negligible due to very long free mean path of interaction in the Earth, so it can be safely ignored. All the neutrinos (ν_e, ν_μ, ν_τ) interact with matter via neutral currents and not through charge current reaction, as there are only electrons in the matter (muons and taus are not there in the matter). However, only ν_e can interact with medium via charged

currents interactions. This changes (as in Eq.1.22) the effective Hamiltonian. In the case of two-neutrino mixing, the mixing angle in vacuum is replaced by an effective angle in matter. The amount of change in the mixing angle depends on matter density. For certain densities, even a small mixing angle in vacuum, the effective mixing angle can become maximal in matter. This is called as MSW effect on the name of SP Mikheev, Smirnov and Wolfenstein [30].

Due to NC and CC interactions, the vacuum Hamiltonian gets modified by the following terms,

$$\begin{aligned} V_{NC} &= -G_F N_n/\sqrt{2} \\ V_{CC} &= G_F N_e/\sqrt{2} \end{aligned} \tag{1.43}$$

Where N_n and N_e are the neutron and electron density inside the earth respectively.

Considering only the two flavor cases, the evolution equation for neutrinos mass eigen state can we written as,

$$i\frac{d}{dt}\begin{pmatrix}\nu_1(t)\\ \nu_2(t)\end{pmatrix} = H\begin{pmatrix}\nu_1(t)\\ \nu_2(t)\end{pmatrix} \tag{1.44}$$

where, the Hamiltonian in the mass eigen basis can be written as follows, since, $E \approx |p| + \frac{m^2}{2|p|}$,

$$H = \begin{pmatrix}E_1 & 0\\ 0 & E_2\end{pmatrix} \approx |p| + \begin{pmatrix}\frac{m_1^2}{2|p|} & 0\\ 0 & \frac{m_2^2}{2|p|}\end{pmatrix} \tag{1.45}$$

Converting it to the flavor eigen basis by using $H_f = UHU^\dagger$, where U is

$$U = \begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix} \tag{1.46}$$

$$H = |p| + \frac{m_1^2 + m_2^2}{4p} + \frac{\Delta m_{21}^2}{4|p|} \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ -\sin 2\theta & \cos 2\theta \end{pmatrix} \tag{1.47}$$

And the diagonalizing angle is ,

$$\tan 2\theta = \frac{2H_{f12}}{H_{f22} - H_{f11}} \tag{1.48}$$

In the case of MSW effect the above Hamiltonian becomes,

$$H = |p| + \frac{m_1^2 + m_2^2}{4p} - \frac{1}{\sqrt{2}} G_F N_n + \begin{pmatrix} -\frac{\Delta m_{21}^2}{4|p|}\cos 2\theta + \sqrt{2} G_F N_e & \frac{\Delta m_{21}^2}{4|p|}\sin 2\theta \\ \frac{\Delta m_{21}^2}{4|p|}\sin 2\theta & \frac{\Delta m_{21}^2}{4|p|}\cos 2\theta \end{pmatrix} \tag{1.49}$$

Where $\frac{1}{\sqrt{2}} G_F N_n$ and $\sqrt{2} G_F N_e$ are NC and CC contributions to effective Hamiltonian respectively. The NC term is written out of the matrix as it is common to all types of neutrinos. However, due to devoid of μ, τ in matter, CC term is written with M_{11} entry, as only electron can have this interaction. The diagonalizing angle in this case is,

$$\tan 2\theta_M = \frac{\Delta m_{21}^2 \sin 2\theta}{\Delta m_{21}^2 \cos 2\theta - A} \tag{1.50}$$

where $A = 2\sqrt{2} G_F N_e E$, E is neutrino energy. From the above equation if,

$$\begin{aligned} A &= \Delta m_{21}^2 \cos 2\theta \\ \implies 2\sqrt{2} G_F N_e E &= \Delta m_{21}^2 \cos 2\theta \\ \implies N_e &= \frac{\Delta m_{21}^2 \cos 2\theta}{2\sqrt{2} G_F E} \end{aligned} \tag{1.51}$$

Hence for the value of N_e in Eq.1.51, the mixing can become maximal (the effective angle in the matter can become $\pi/2$) which is the MSW effect.

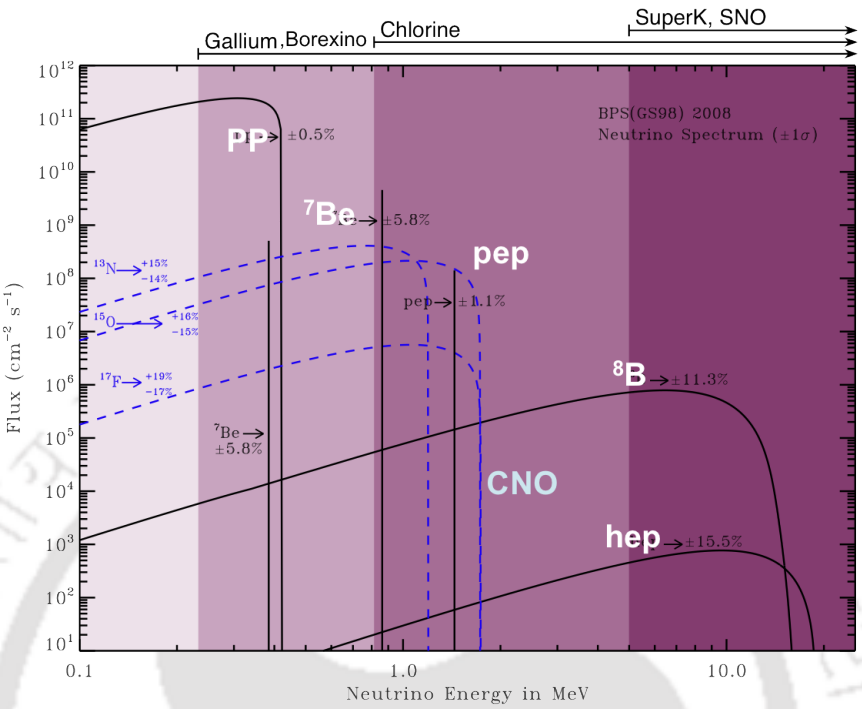

FIGURE 1.3: Neutrino flux from the Sun [32]

1.4 Experimental Evidence for Neutrino Oscillations

1.4.1 Solar Neutrinos

Sun is a great source of neutrinos. It produces the neutrinos in its energy production through solar fusion reactions. Fig.1.3 shows the flux of neutrinos produced inside the Sun through various fusion reactions [32]. Solar neutrinos energy lies in the 0.1 - 10 MeV range [62].

There were several experiments designed to detect neutrinos based on radiochemical technology. In 1960s, Ray Davis performed an experiment at Homestake mines to detect the solar neutrinos [63]. The experiment was based on the following reaction,

$$\nu_e + {}^{37}Cl \rightarrow {}^{37}Ar + e^- \quad (E_{th} = 0.814\ MeV) \tag{1.52}$$

Due to the reaction threshold, the Homestake experiment was sensitive to following neutrinos produced in proton-proton chain in Standard Solar Model,

$$^{8}B \rightarrow^{8} B^{*} + e^{+} + \nu_{e}(E_{th} =\sim 10\ MeV) \tag{1.53}$$

Later on during 1990s, other experiments, like SAGE [64] and GALLEX-GNO [65] measured the rates of neutrinos produced in p-p reaction. The threshold of neutrino detection was lower than that of Homestake.

$$p + p \rightarrow d + e^{+} + \nu_{e} \quad (SAGE\ E_{th}\ =\ 0.3\ MeV) \tag{1.54}$$

$$\nu_{e} +^{71} Ga \rightarrow e^{-} +^{71} Ge \quad (GALLEX\ E_{th}\ =\ 0.233\ MeV) \tag{1.55}$$

All the above radiochemical experiments could only tell the rates of neutrinos based on the energy. However, the information about time and direction of neutrino events can not be obtained using such methods.

In 1996, physicists in Japan performed the Kamiokande [15] experiment based on water cherenkov detectors. With this technology, a real time neutrino can be detected with the time and direction information. It used the water cherenkov detectors to measure the elastic neutrino scattering on electrons,

$$\nu_{e} + e^{-} \rightarrow \nu_{e} + e^{-}(E_{th} \sim 5\ MeV) \tag{1.56}$$

All of the above neutrino experiments consistently found the rate of solar neutrinos to be within 1/3 and 1/2 of the expected rate as predicted by Solar Standard Model (SSM) [66]. The problem was termed as Solar Neutrino Anomaly [67].

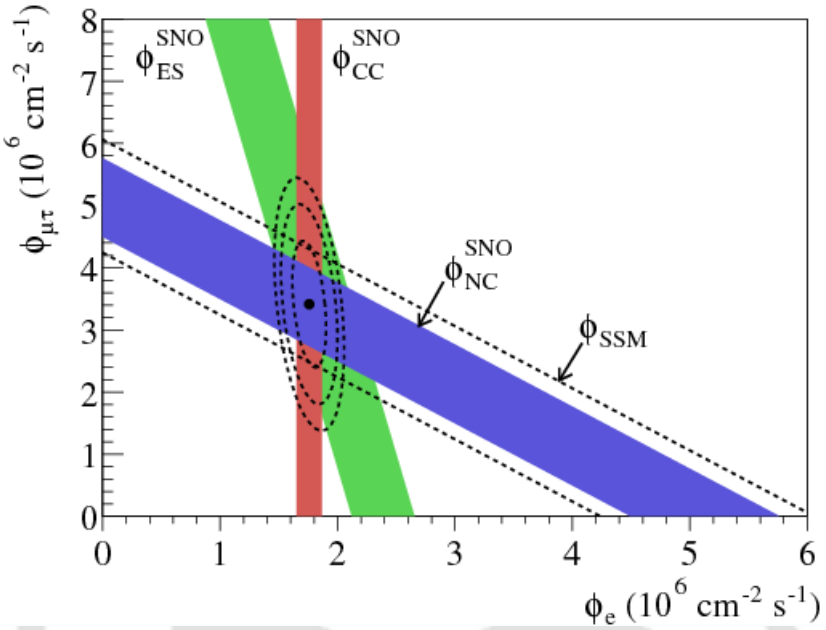

FIGURE 1.4: Flux of $\nu_{\mu,\tau}$ vs ν_e. The dashed line shows the SSM prediction which is in complete agreement with the flux (blue band) measured with the NC [33].

In 2001, the solar neutrino anomaly was resolved by the experiment called, Sudbury Neutrino Observatory (SNO) experiment [68]. SNO has the ability to measure the neutrino flux through three different reactions as given below with sensitivity for all the kind of neutrino fluxes,

$$\begin{aligned} \nu_e + d \rightarrow p + p + e^- \quad (CC,\ E_{th} > 5\ MeV) \\ \nu_x + d \rightarrow p + n + \nu_x \quad x = e, \mu, \tau \quad (NC,\ E_{th} > 2.2\ MeV) \end{aligned} \tag{1.57}$$

CC reaction is only sensitive to electron neutrinos and NC is sensitive to all the flavors. This enables SNO to establish the flavor transition in the solar neutrinos as shown in Fig.1.4

In 2002, the solar neutrino oscillations was confirmed with the reactor neutrino in the KamLAND experiment which measured the reactor neutrinos ($\bar{\nu_e}$) coming from nuclear reactor [34]. The neutrinos were detected via inverse β decy as ,

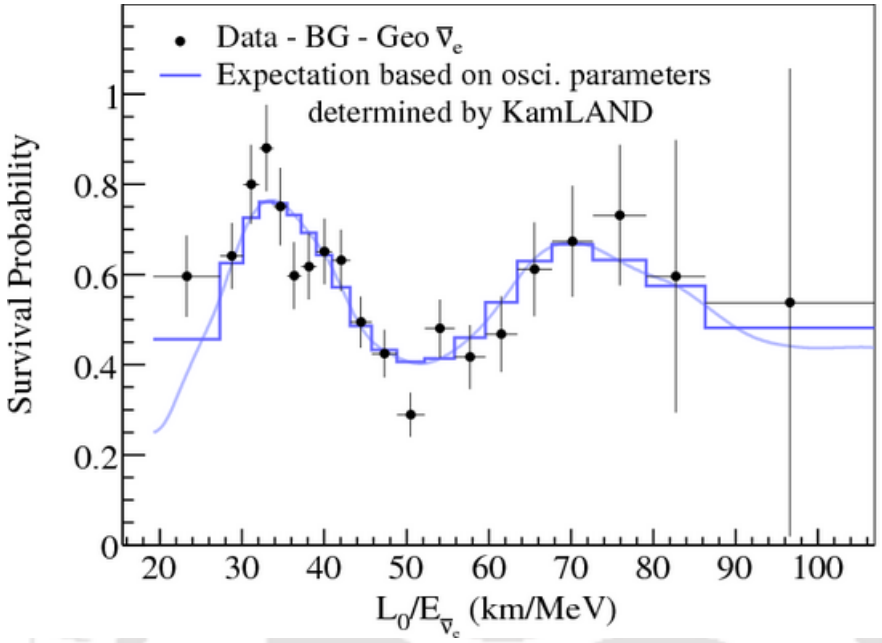

FIGURE 1.5: Survival probability as a function of L/E form the KamLAND experiment. THe experimental data points clearly agree with the theoretical predictions based on neutrino oscillations [34].

$$\bar{\nu}_e + p \rightarrow e^+ + n \quad (\ E_{th} > 2.6\ MeV) \tag{1.58}$$

Fig.1.5 shows the survival probability as a function of ratio E_ν/L. This clearly shows that the neutrinos oscillate.

The combined fit (Fig.1.6) to all the solar neutrino data gives the best fit oscillation parameter with a precision of 3.5 % as,

$$\Delta m^2_{21} = 7.54^{+0.26}_{-0.22} \times 10^{-5} eV^2, \quad \tan^2\theta_{12} = 0.307^{+0.018}_{-0.016}, \quad m_2 > m_1$$

1.4.2 Reactor Neutrinos

Nuclear reactors are good source of anti-neutrinos. The energy range of neutrinos that come from reactors are in few MeVs. The reactor experiments measure the probability,

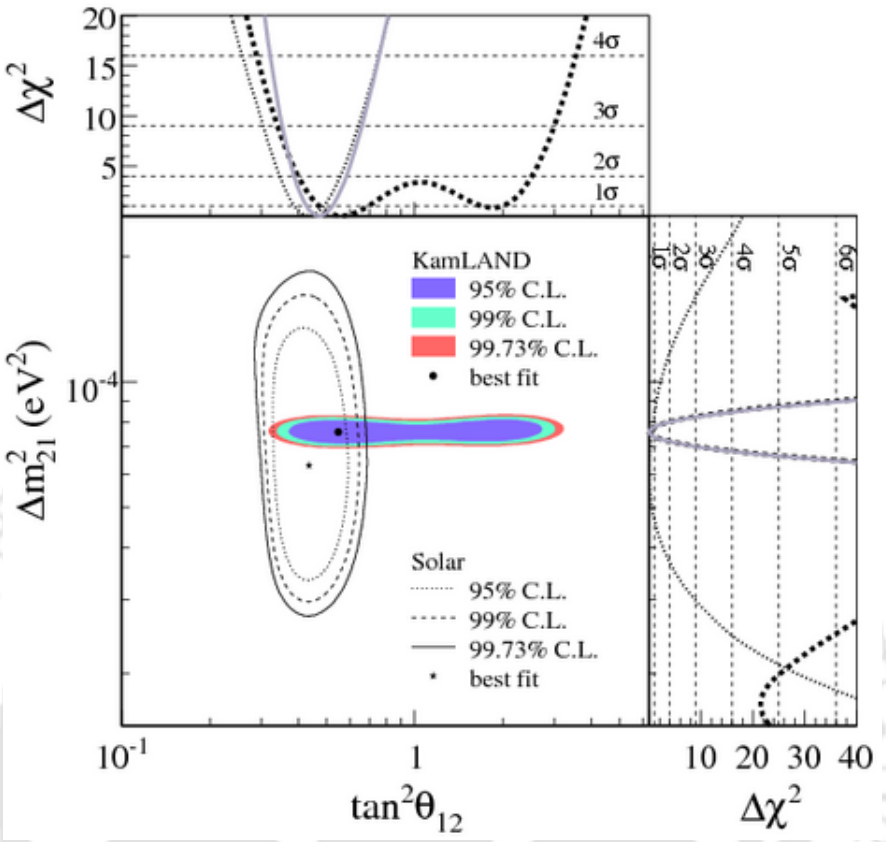

FIGURE 1.6: Best fit results on Δm^2_{21} and θ_{12} from all solar neutrino experiments. Results from KamLAND experiment are also included [34].

$$P(\nu_e \to \nu_e) = P(\bar{\nu}_e \to \bar{\nu}_e) = \sin^2 2\theta_{13} \sin^2 \left(\frac{\Delta m^2_{31} L}{4E_\nu}\right) \tag{1.59}$$

which makes the reactor neutrino to measure θ_{13} very precisely. ChOOZ, Daya Bay, RENO and Double Chooz are some of the current reactor experiments. Current constrains on θ_{13} comes from the reactor experiment such as Daya Bay as shown in Fig.1.7.

1.4.3 Atmospheric Neutrinos

In atmosphere, neutrinos are produced in the collision of primary cosmic rays, primarily protons with nuclei (air molecule) that produces secondary particles, mostly being pions [69]. Pion decay to a muon and a muon neutrino and subsequently the muon decays to

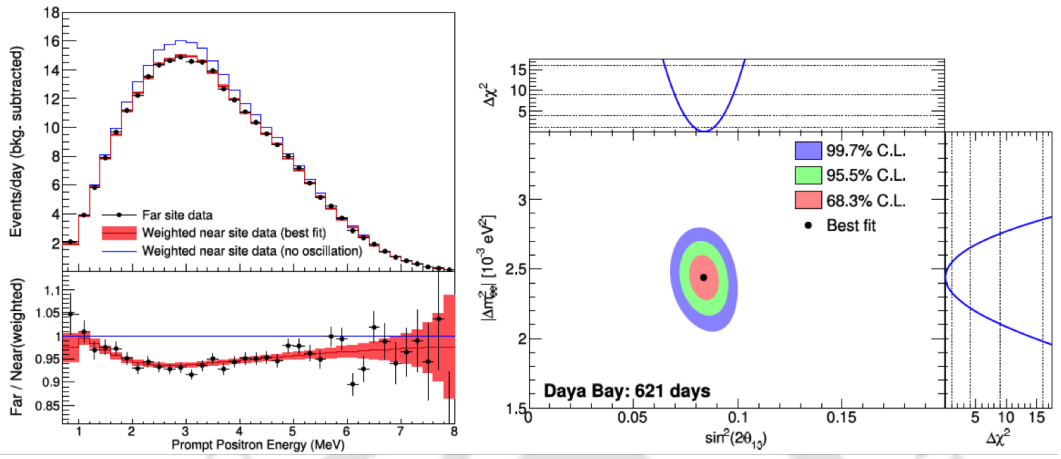

FIGURE 1.7: Right: $\bar{\nu_e}$ energy spectrum from Daya Bay experiment. Left: The parameter space (Δm^2_{ee} and $\sin^2(2\theta_{12})$) constrained from the experiment [35].

a positron, a muon anti-neutrino and an electron neutrino, as shown in the decay chain below. One can clearly see that ratio of ν_μ to ν_e is 2:1.

$$\pi^+ \rightarrow \mu^+ + \nu_\mu \ , \qquad \mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu \tag{1.60}$$

Super Kamiokande (SK) experiment measures the flux of neutrinos in GeV range along with the direction of neutrino coming from Sun. SK experiment could measure the neutrinos energy coming to detector from different directions. It was found, as shown in Fig.1.8 that the neutrinos coming from down side ($\cos_\theta = -1$) are less in numbers than the predictions. However, neutrinos coming from above were in agreement with the prediction. This was understood as the neutrino oscillation while passing through the matter (MSW effect) coming from the other side of the earth.

The other ongoing experiments which measure the neutrino oscillation at atmospheric length are T2K, MINOS and NOνA.

Measurements of atmospheric muon neutrino disappearance are sensitive to $|\Delta m^2_{32}|$ and $\sin^2 2\theta_{23}$. Referred to Fig.1.9 and Fig.1.10 which shows the latest results from NOνA [37], where upper octant is preferred at 0.2σ and the best fit values found as given by Eq.1.61

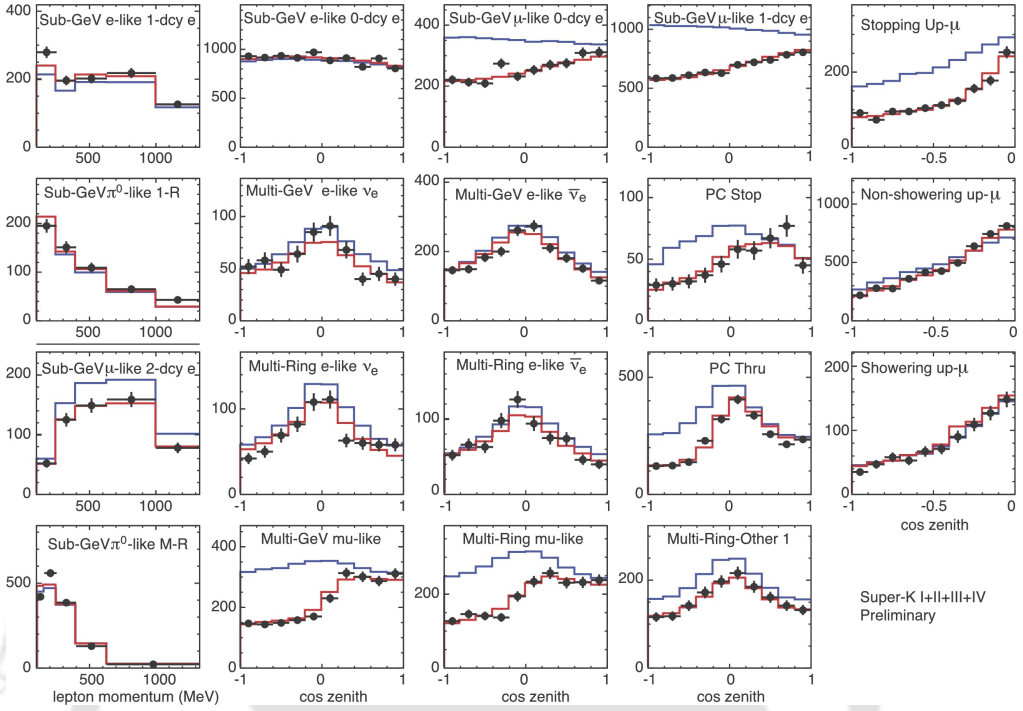

FIGURE 1.8: Up and down neutrinos spectrum in SuperK experiment. Upward going neutrinos are oscillated away [36].

$$
\begin{aligned}
|\Delta m^2_{32}| &= 2.444^{+0.079}_{-0.077} \times 10^{-3} eV^2 \\
\sin^2\theta_{23} &= 0.558^{+0.041}_{-0.033}, \quad \textit{Upper octant} \\
\sin^2\theta_{23} &= 0.475^{+0.036}_{-0.044}, \quad \textit{Lower octant}
\end{aligned}
\tag{1.61}
$$

1.5 Future Experiments

Neutrino experiments have entered into the precision era. The unknowns in the neutrino sector such as mass hierarchy, CP violation in neutrinos sector and which octant does the θ_{23} value occur, are yet to be constrained or yet to be known. The future neutrino experiments are being designed to measure the above unknowns with no ambiguity.

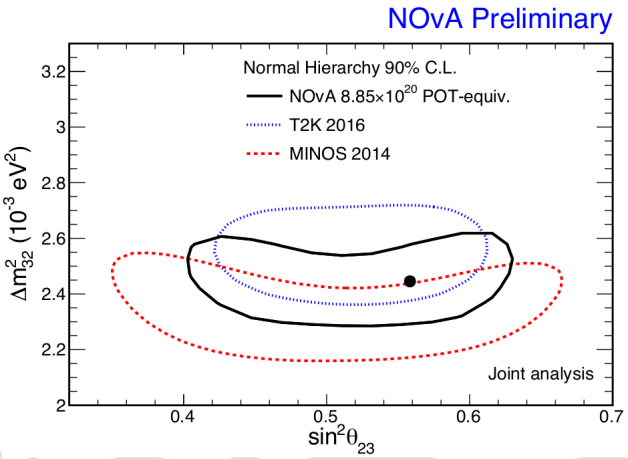

FIGURE 1.9: NOνA joint fit analysis results. The result is consistent with the world expectations. Taken from [37]

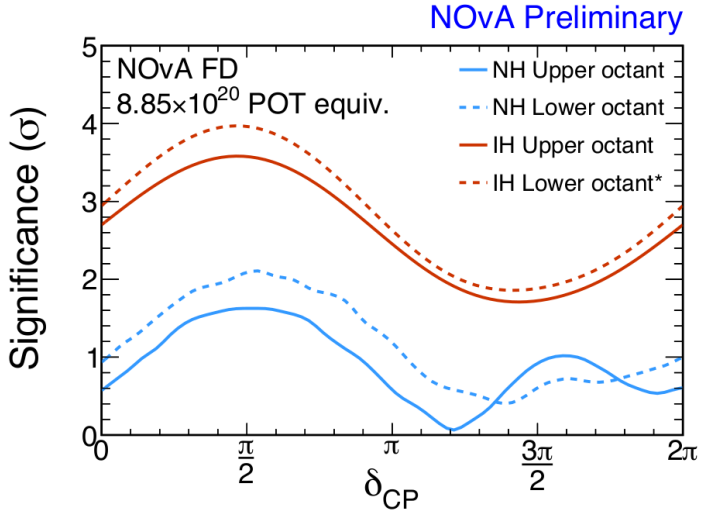

FIGURE 1.10: NOνA joint fit analysis results. IH at $\delta_{CP} = \pi/2$ disfavored at greater than 3σ. Taken from [37]

1.5.1 Mass Hierarchy

The mass hierarchy of neutrino mass eigen states is yet to be determined by the neutrino oscillations experiments. Whether it is normal hierarchy or inverted hierarchy (Fig.1.11) is not known. Sensitivity to MH comes from the matter effect, and that grows with distance. Future experiments like DUNE, INO and HyperK are being designed with main aim to reasonably resolve the mass hierarchy.

[4] http://cerncourier.com/cws/article/cern/49351

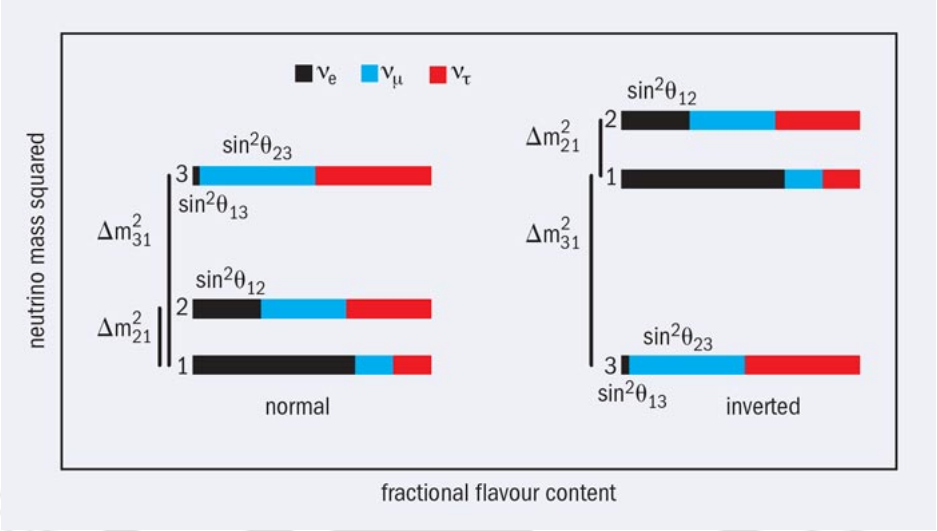

FIGURE 1.11: Possible mass hierarchy scenario in the neutrino sector[4].

1.5.2 CP Symmetry in Lepton Sector

A measure of CP violation can be measured in terms of asymmetry as given by-

$$A_{CP}^{(\alpha,\beta)} = P(\nu_\alpha \to \nu_\beta) - P(\bar{\nu_\alpha} \to \bar{\nu_\beta}) \tag{1.62}$$

Plug in probabilities in Eq.1.62, we get

$$\begin{aligned} P(\nu_\alpha \to \nu_\beta) - P(\bar{\nu}_\alpha - \bar{\nu}_\beta) &= 4\sum_{j>k} Im(U^*_{\beta j}U_{\beta k}U_{\alpha j}U^*_{\beta k}) \sin\left(\frac{\Delta m^2_{jk}L}{2E}\right) \\ &\propto \sin\delta \ (where,\ \delta\ is\ Dirac\ phase\ as\ in\ PMNS\ matrix.) \end{aligned} \tag{1.63}$$

T2k has given a hint of CP violation as shown in Fig.1.13 in the neutrino sector. Future experiments, such as DUNE and HyperK, are designed which have the potential to establish it. In order to observe CP asymmetry, the initial and final flavour states must be different, i.e., $\alpha \neq \beta$, as shown in Eq.1.63, and it should be disentangled from mass hierarchy. Fig.1.12 shows the projected sensitivities to rejection of: maximal mixing,

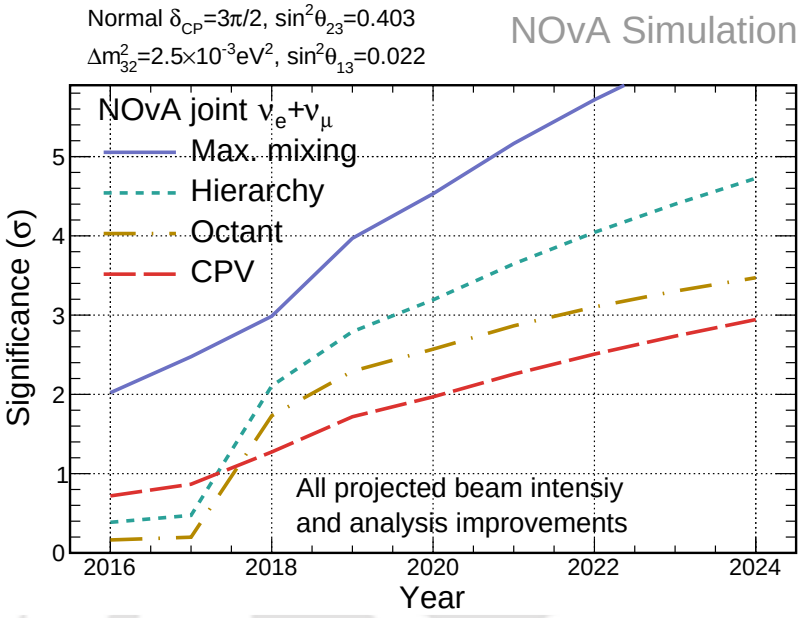

FIGURE 1.12: NOνA experiment projected sensitivities to rejection of: maximal mixing (violet), wrong hierarchy (green), wrong octant (yellow), and CP conservation (red), assuming true oscillation parameters NH $3\pi/2$ and $\sin^2\theta_{23} = 0.403$. Joint fit combines the electron neutrino appearance and muon neutrino disappearance channels, and global reactor constraint $\sin^2\theta_{13} = 0.0805 \pm 0.005$.

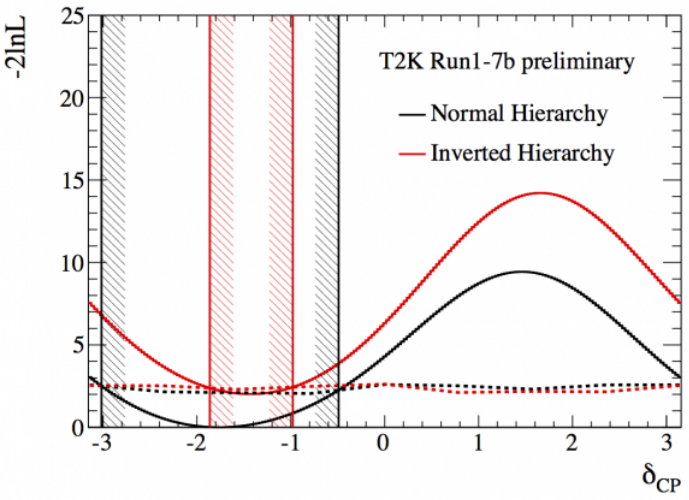

FIGURE 1.13: χ^2 vs CP violation phase as measured at T2K [5].

wrong hierarchy , wrong octant, and CP conservation, by combining the electron neutrino appearance and muon neutrino disappearance channels, and global reactor constraint

[5] http://t2k-experiment.org/2016/07/t2k-presents-first-cp-violation-search-result/

NuFIT 3.2 (2018)

	Normal Ordering (best fit)		Inverted Ordering ($\Delta\chi^2 = 4.14$)		Any Ordering
	bfp $\pm 1\sigma$	3σ range	bfp $\pm 1\sigma$	3σ range	3σ range
$\sin^2\theta_{12}$	$0.307^{+0.013}_{-0.012}$	$0.272 \to 0.346$	$0.307^{+0.013}_{-0.012}$	$0.272 \to 0.346$	$0.272 \to 0.346$
$\theta_{12}/^\circ$	$33.62^{+0.78}_{-0.76}$	$31.42 \to 36.05$	$33.62^{+0.78}_{-0.76}$	$31.43 \to 36.06$	$31.42 \to 36.05$
$\sin^2\theta_{23}$	$0.538^{+0.033}_{-0.069}$	$0.418 \to 0.613$	$0.554^{+0.023}_{-0.033}$	$0.435 \to 0.616$	$0.418 \to 0.613$
$\theta_{23}/^\circ$	$47.2^{+1.9}_{-3.9}$	$40.3 \to 51.5$	$48.1^{+1.4}_{-1.9}$	$41.3 \to 51.7$	$40.3 \to 51.5$
$\sin^2\theta_{13}$	$0.02206^{+0.00075}_{-0.00075}$	$0.01981 \to 0.02436$	$0.02227^{+0.00074}_{-0.00074}$	$0.02006 \to 0.02452$	$0.01981 \to 0.02436$
$\theta_{13}/^\circ$	$8.54^{+0.15}_{-0.15}$	$8.09 \to 8.98$	$8.58^{+0.14}_{-0.14}$	$8.14 \to 9.01$	$8.09 \to 8.98$
$\delta_{CP}/^\circ$	234^{+43}_{-31}	$144 \to 374$	278^{+26}_{-29}	$192 \to 354$	$144 \to 374$
$\frac{\Delta m^2_{21}}{10^{-5}\ \mathrm{eV}^2}$	$7.40^{+0.21}_{-0.20}$	$6.80 \to 8.02$	$7.40^{+0.21}_{-0.20}$	$6.80 \to 8.02$	$6.80 \to 8.02$
$\frac{\Delta m^2_{3\ell}}{10^{-3}\ \mathrm{eV}^2}$	$+2.494^{+0.033}_{-0.031}$	$+2.399 \to +2.593$	$-2.465^{+0.032}_{-0.031}$	$-2.562 \to -2.369$	$\begin{bmatrix}+2.399 \to +2.593\\ -2.536 \to -2.395\end{bmatrix}$

FIGURE 1.14: Three-flavor oscillation parameters from our fit to global data as of November 2017. The numbers in the 1st (2nd) column are obtained assuming NO (IO), i.e., relative to the respective local minimum, whereas in the 3rd column we minimize also with respect to the ordering. Taken from [38].

1.5.3 θ_{23} Octant

Whether θ_{23} is $< \pi/4$ or $> \pi/4$ is also not known. Knowing θ_{23} will help knowing the admixture of flavors in ν_3 mass eigen states. Current measurements are consistent with $\theta \sim 45^\circ$. This means that the ν_3 state contains equal mixture of ν_μ and ν_τ flavor eigen states, apart from ν_e. This might indicate a new symmetry in the neutrino sector [70].

1.6 Current Status of Oscillation Parameters

The main goals of current generation of neutrino experiment are to precisely measure the six parameters that govern the neutrino oscillations ($\Delta m^2_{21}, \Delta m^2_{32}, \theta_{12}, \theta_{23}, \theta_{13}$ and δ_{CP}). The first precision measurement of θ_{12} and Δm^2_{21} have been done by SuperK , SNO and KamLAND experiment. Reactor experiments like Daya Bay , RENO and Double Chooz have measured θ_{13}. Along with NOνA, the other contemporary accelerator based neutrino oscillation experiment are T2K and MINOS . The mass hierarchy and δ_{CP} are yet to be measured. The quadrant of θ_{23} is also not known. Table 1 summarizes the current knowledge of the parameters .

1.7 Relevance of the Thesis Work

Neutrino experiments have already entered into a precision measurement era. More and more precise and long baseline future neutrino experiments are being designed to answer the above questions. Also current systematic uncertainty in neutrino oscillation experiments are up to 10 -15 %, which makes it even more challenging to extract or constrain the oscillation parameters precisely. This boils down to a point of better understanding of relevant simulations in the experiment, accurate understanding of the detector and better identification of signal over backgrounds. In electron neutrino appearance oscillation mode, the signal in neutrino quasi elastic interaction is always an outgoing electron. In the precision era in the neutrino sector, the precision of the measurement depend on how well is the agreement between the simulations and the reality and how well, an experiment can measure the signal etc. The work reported in this thesis, helps in benchmarking the neutrino simulations, the signal identification efficiency using data driven methods in ν_e-appearance mode. To enhance the confidence in the analysis tools and simulations is the main theme of the reported work and is described in the chapters to follow.

1.8 Chapter Summary

Neutrino, since its inception, has drawn significant amount of attention and has also brought excitement in the field. Neutrinos are potentially the key factors in solving many mysteries of universe. Neutrino oscillations are sought after very actively through several ongoing experiments and several experiments are being designed. The future seems full of neutrino research and following are the several important questions that are to be answered by current and ongoing neutrino experiments-

- What is the absolute mass scale of neutrino?
- What is the mass hierarchy?
- What is δ-CP phase and why there exists matter-antimatter asymmetry in Nature?
- What is the octant of θ_{23}? Does it indicate any underlying symmetry ?

- What is the type of neutrinos? Is it Dirac or Majorana?
- Are there any sterile neutrinos?

This chapter introduces the theory of neutrino oscillation, followed by an overview of the status of the experimental measurement of the neutrino oscillation parameters. The chapter also includes a short discussion on the importance of developing new data driven methods to control the systematic uncertainties associated with the neutrino oscillation, which is the main objective of the thesis.

Chapter 2

NOνA Experiment

NOνA stands for NuMI Off-axis ν_e-appearance[1], where NuMI is the neutrino beam produced at Fermilab. NOνA far detector is located at 810 km away downstream of the beam production point, as shown in Fig.2.1. By doing this, NOνA's far detector sensitivity to detect $\nu_\mu \rightarrow \nu_e$ oscillation enhances near the first oscillation maximum. Also due to 810 km of earth, matter effect also plays a role. The near detector is located at 1 km away from the beam production point which is used to characterize the beam before it has a chance to oscillate. There is a large degree of systematic cancellation by using the far detector prediction derived from near detector rate as the two detectors use the same detector technology. Besides this, the other studies that can be done on near detector are short baseline oscillations, cross-section studies etc. Except the size, the near and far detector are made up of same type of materials and design. Detectors are made of low Z material which allows electromagnetic showers from electrons to build up over many planes and cells to provide enough topological information along with energy deposition to differentiate it from other electromagnetic shower producing backgrounds such as $\pi^0 \rightarrow \gamma\gamma$.

[1]The chapter is based on NOνA technical design reports [41] and NOνA internal database of collaboratively blessed plots at https://nusoft.fnal.gov/nova/blessedplots/ and NOνA internal document database https://nova-docdb.fnal.gov/cgi-bin/private/DocumentDatabase

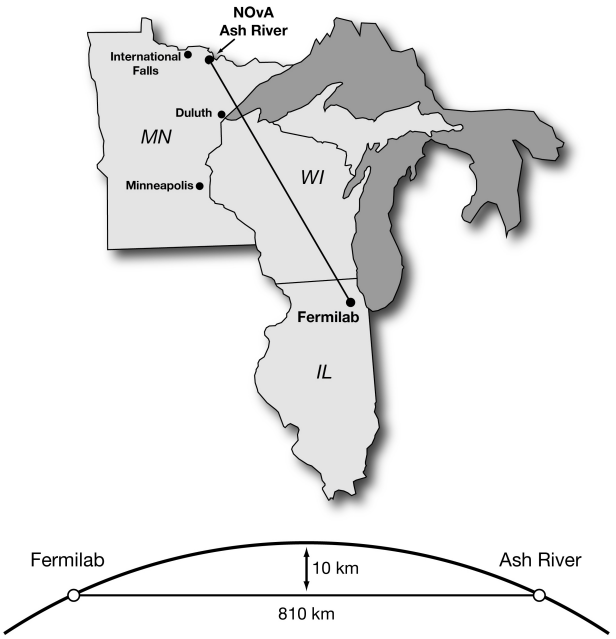

FIGURE 2.1: Geographical location of near and far detector. The beam travels a distance of about 810 km inside the earth from Fermilab, Batavia to Ashriver, Minnesota.

NOνA can measure oscillation in both $\nu_\mu \to \nu_\mu$ ($\bar{\nu}_\mu \to \bar{\nu}_\mu$) disappearance mode and $\nu_\mu \to \nu_e$ ($\bar{\nu}_\mu \to \bar{\nu}_e$) appearance mode. This makes NOνA sensitive to parameters, θ_{13}, θ_{23} and Δm^2_{23}. Due to matter in between the near detector and far detector and because of the availability of both neutrino and anti-neutrino beam, NOνA can measure the sign of atmospheric mass splitting and dirac-CP violating phase, δ_{CP}. NOνA has the potential to significantly constrain the value of θ_{13} although the best measurement will come from reactor experiments. NOνA is optimized for resolving mass hierarchy and is sensitive to the "octant" of θ_{23} (in case if it is not maximal, whether it is $\theta_{23} > \pi/4$ or $\theta_{23} < \pi/4$). NOνA sensitivity to mass hierarchy, δ_{CP} and the octant depend critically on having θ_{13} available from reactor experiments. Fig.2.2 shows the capability of NOνA to measure the aforesaid parameters by measuring the oscillation probabilities in the NOνA experiment for $P_{\bar{\nu_\mu}\to\bar{\nu_e}}$ vs $P_{\nu_\mu\to\nu_e}$ (for 18×10^{20} proton on target (POT) in each neutrino and anti-neutrino mode). The central starred point is for no matter effect, maximal θ_{23}, no CP violation, and $\sin^2 2\theta_{13} = 0.095$. Inclusion of matter effect splits the oscillation in red (inverted hierarchy) and blue (normal hierarchy) ellipses. The different point depicts

different value of δ phase. The probabilities further get changed depending on the θ_{23} octant, with values of $\sin^2 2\theta_{23} = 0.97$. The starred point with 1σ and 2σ contours show the mass ordering, δ and octant that can be measured simultaneously through these two oscillation channels.

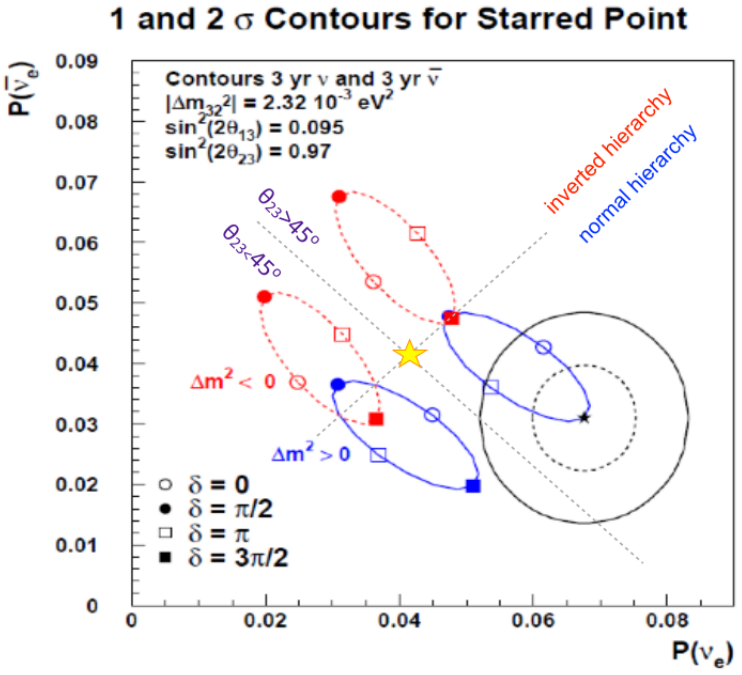

FIGURE 2.2: Neutrino and anti-neutrino electron appearance probabilities as a function of δ for the normal and inverted hierarchies. The star point is a possible measurement scenario that NOνA may observe. The contours are representative of statistical errors on the measurement [39].

2.1 NuMI Beam

NuMI beam is a muon neutrino beam produced at Fermilab. High energy protons (120 GeV/c) are allowed to hit on graphite target to produce large number of charged kaons and pions as shown in Fig.2.3. The graphite target, 1.2 m in length, is of two interaction lengths. The produced hadrons are focussed by the two magnetic "horns" downstream. NuMI uses pulsed magnetic horns which are water-cooled, parabolic aluminum horns, pulsed with up to 200 kA to focus the pions. The horns act like lenses to particles and the focal length is proportional to pion momentum. Particle which are not focused by first

horn are focused by a second horn. One can choose either the positive charge particle or negative charged particles by changing the polarity of the horns. After this, the focused particles are allowed to decay in a decay pipe volume. The configuration which selects neutrinos is called as the Forward Horn Current (FHC) and the configuration which selects the anti-neutrinos is called as the Reverse Horn Current (RHC). The decay pipe volume is 675 meters in length and 1 meter in radius and is maintained at low pressure of about 0.4 torr and is filled with helium. In the decay pipe, pions and kaons decay to ν_μ through the following decay modes:

$$\begin{aligned} \pi^+ &\to \mu^+ + \nu_\mu \qquad \pi^- \to \mu^- + \bar{\nu}_\mu \quad (BR.\ 99.98\%) \\ K^+ &\to \mu^+ + \nu_\mu \qquad K^- \to \mu^- + \bar{\nu}_\mu \quad (BR.\ 63.55\%) \end{aligned} \tag{2.1}$$

NuMI beam also contains the small component of ν_e and wrong sign $\bar{\nu}_\mu$ in neutrino mode, or vice-versa, which are produced through the following processes,

$$\begin{aligned} \pi^+ &\to e^+ + \nu_e \qquad \pi^- \to e^- + \bar{\nu}_e \quad (BR.\ 0.01\%) \\ K^+ &\to \pi^0 + e^+ + \nu_e \qquad K^- \to \pi^0 + e^- + \bar{\nu}_e \quad (BR.\ 5.07\%) \\ K^+ &\to \pi^0 + \mu^+ + \nu_\mu \qquad K^- \to \pi^0 + \mu^- + \bar{\nu}_\mu \quad (BR.\ 3.35\%) \\ K^0_L &\to \pi^+ + e^- + \bar{\nu}_e \qquad K^0_L \to \pi^- + e^+ + \nu_e \quad (BR.\ 40.55\%) \\ K^0_L &\to \pi^+ + \mu^- + \bar{\nu}_\mu \qquad K^0_L \to \pi^- + \mu^+ + \nu_\mu \quad (BR.\ 27.04\%) \\ \mu^+ &\to e^+ + \nu_e + \bar{\nu}_\mu \qquad \mu^- \to e^- + +\bar{\nu}_e + \nu_\mu \quad (BR.\ 100.00\%) \end{aligned} \tag{2.2}$$

Simulation estimates that such contamination is $< 2\%$ in neutrino mode and $\sim 10\%$ in the anti-neutrino mode. After the decay pipe, a water-cooled aluminum beam absorber is positioned followed by hadron absorbers and ionization chambers to monitor the rate of muon passing through it.

The NuMI neutrinos energy spectrum depends on the off-axis angle, θ which is measured from the direction of beam as shown in Fig.2.4.

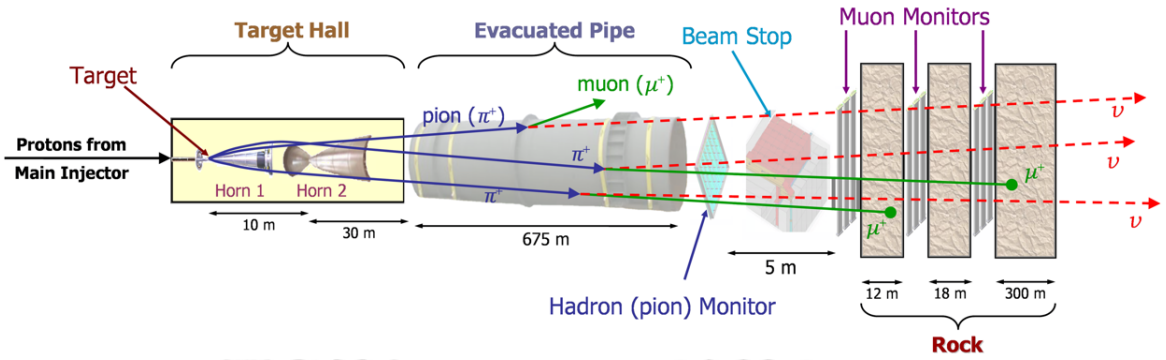

FIGURE 2.3: Schematic diagram of NuMI beam components [40].

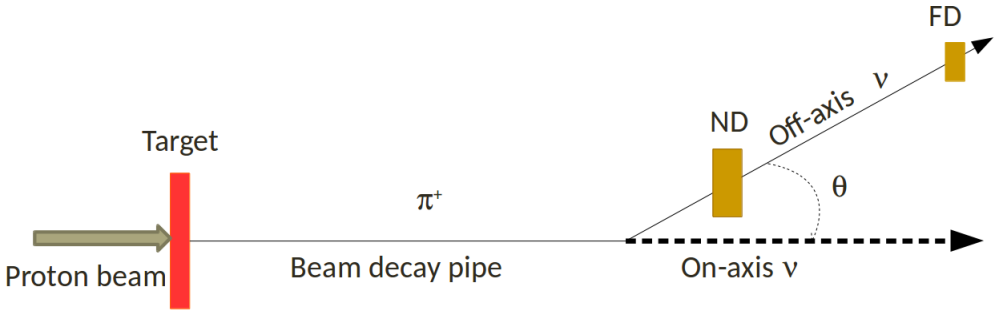

FIGURE 2.4: Off-axis neutrino beam schematics.

The energy spectrum of neutrinos depend on the off-axis angle and given by following equation,

$$E_{\nu_\mu} = \frac{(0.43)E_\pi}{1+\gamma^2\theta^2} \tag{2.3}$$

where E_π is the energy of the parent pion[2].

Fig.2.5 shows the energy spectrum of ν_μ in NuMI beam as a function of angle ($\theta = 0$ corresponds to on axis beam). As can be seen in Fig.2.5, at off-axis, the energy spectrum of ν is almost horizontal with respect to parent pion energy. This make the off-line beam ν energy more or less independent (varying only within a $\sim GeV$ range) after a certain

[2]Replace 0.43 by 0.96 to get the energy vs off-axis angle spectrum of neutrinos coming from kaons.

energy of the energy of parent particles as can be seen in Fig.2.5. The resulting Off-axis neutrino beam is relatively narrower in energy than On-axis neutrino energy spectrum.

NOνA is placed at 14.6 mrad off the central axis and therby NOνA gets a sharp neutrino energy spectra peaking at 2 GeV as shown in Fig.2.6 which is close to the $\nu_\mu \rightarrow \nu_e$ oscillation first maximum for NOνA far detector. The narrow peak at 2 GeV has an advantage of removing the background coming from beam electron neutrinos. These ν_e are produced through three-body decay and tends to have broader energy spectrum. Narrower beam also helps in reducing the high energy neutral current background. Fig.2.7 shows the simulated signal and background events in NOνA using FLUKA.

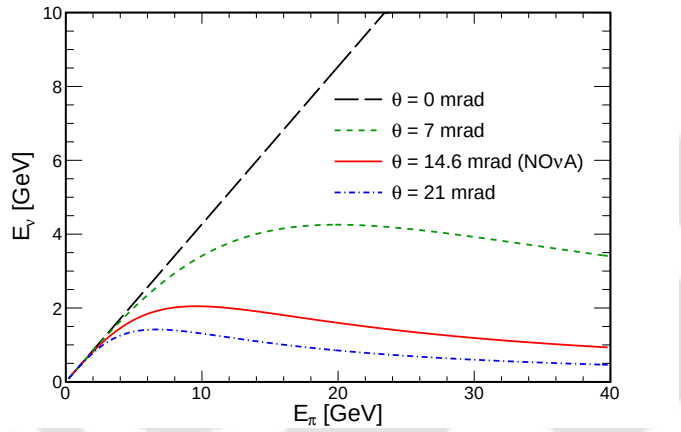

FIGURE 2.5: Neutrino energy distribution as a function of parent pion energy at different off-axis angle.

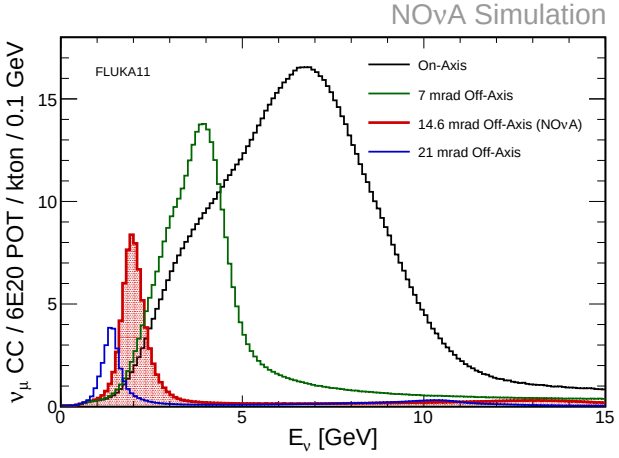

FIGURE 2.6: The off-axis beam results in a narrower energy distribution of neutrinos.

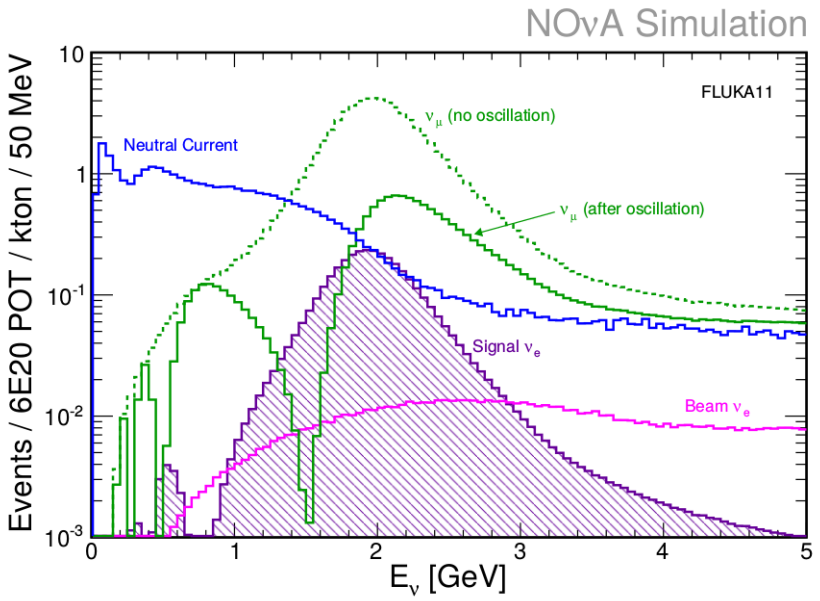

FIGURE 2.7: NuMI beam spectrum with signal (ν_e) and backgrounds. Neutral current events are the main source of backgrounds to ν_e-CC channel [41].

2.2 NOνA Detectors

NOνA experiment comprises two identical detectors separated by 810 km. The near detector is located near the beam production point and is 1 km from the beam target at Fermilab. The second detector is called as far detector and is located at Ash river in Minnesota. Both the detectors are located 14.6 mrad off axis and beam falls perpendicular to the front planes of the detector. Near detector is 290 ton and far detector is 14 kt. Owing to the proximity of near detector to beam production, the near detector is relatively smaller in size than far detector with sufficient transverse and longitudinal size for neutrino event containment. Event rate is smaller at far detector (neutrino beam spreads with distance) and hence it requires far detector to be big.

Near detector greatly reduces the systematic uncertainty in predicting the flux rate at far detector and is used to characterize the beam. It is also used for the several neutrino cross section studies [71][72]. Near detector is situated ∼100 m underground at Fermilab and this makes near detector almost free from cosmogenic events. The far detector is located on the surface and it has a modest concrete and barite overburden

which corresponds to 14 radiation lengths and a significant part of the electromagnetic cosmic rays are stopped. The cosmic rate measured in far detector is 148 kHz.

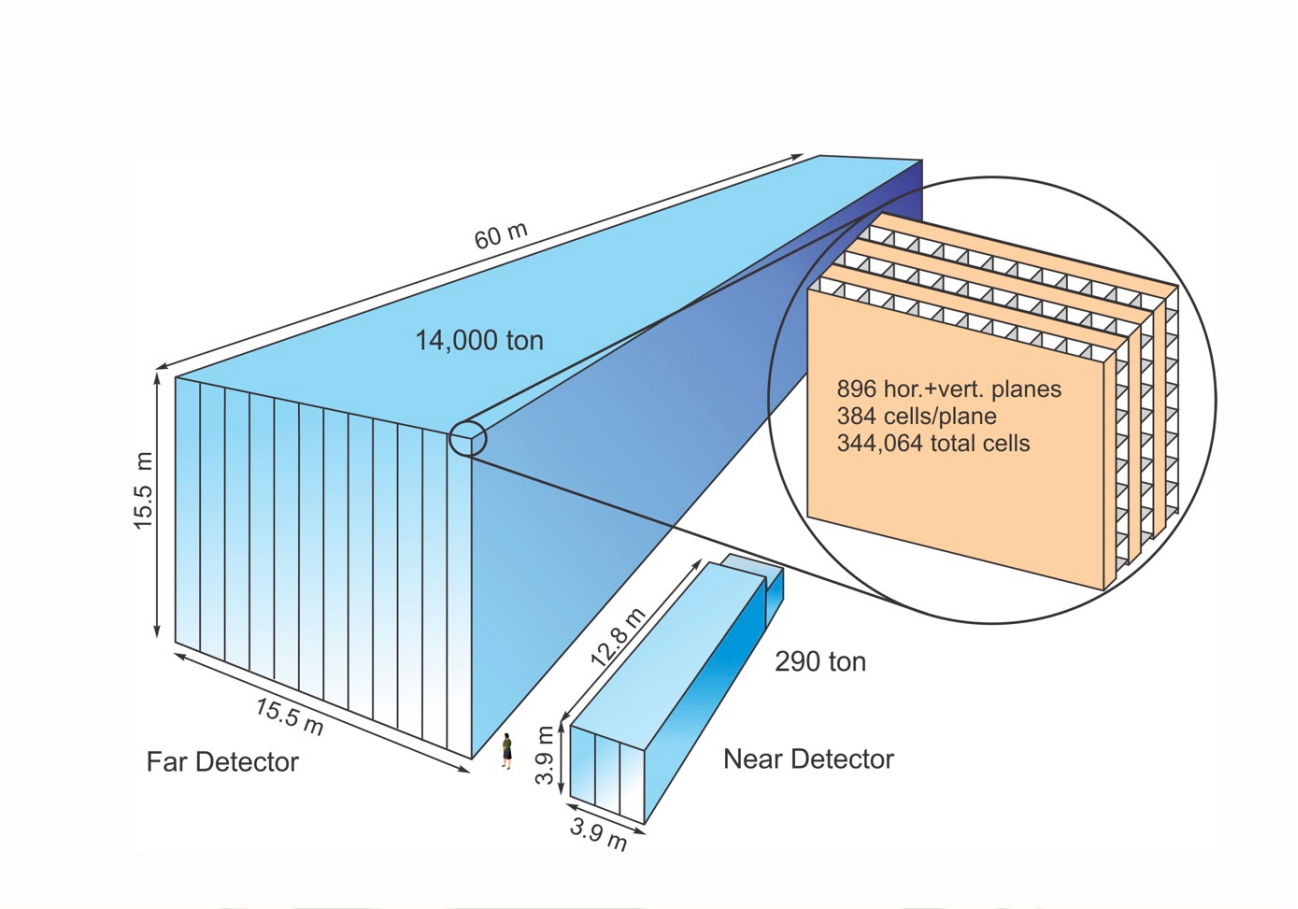

FIGURE 2.8: NOνA detectors- Far detector and near detector [41].

As mentioned earlier, both the detectors are identical except the size. The basic unit of NOνA detector is a cell, which is 3.9 cm × 6.0 cm in dimension and is 4 m (15.5 m) long in near (far) detector as shown in the Fig.2.9. The inner walls of cells are highly reflective and are made reflective by using titanium dioxide TiO_2. The reflectivity is 92% at 430 nm near the peak of the scintillator emission spectrum. The unit cell is filled with scintillator whose chemical composition is 95% pseudocumene (1,2,4- Trimethylbenzene) + 5% mineral oil +∼ 1% WLS agents. The wavelength shifting fiber (WLS) agents used are - (PPO(2,5-Diphenyloxazole)+ bis-MSB (1,4-bis-(o-methyl-styryl-)-benzene)). WLS agents shift the wavelength of photons, which are emitted by a charged particle passing through it, from 270 - 320 nm to 390 - 440 nm.

A fiber of thickness 7 mm loops inside the cell to carry the light to readout. The looped fiber collects more light in the cell in comparison to single strand of fiber in the cell . The fiber picks the light from scintillator and shifts its wavelength to 450-650 nm. This light from a single cell is delivered in a fiber to a readout electronics. The fiber end readout is a single pixel of a 32 pixel avalanche photo diode (APD) array.

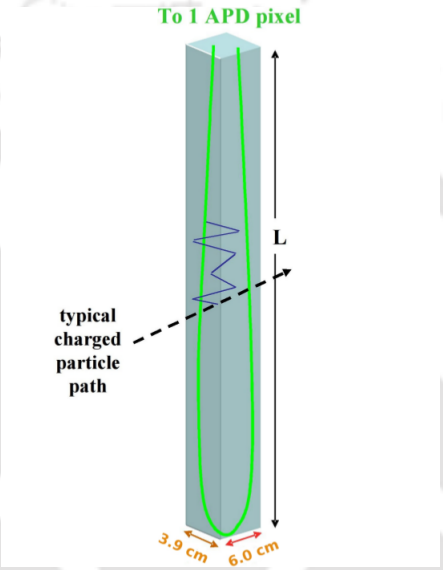

FIGURE 2.9: A unit cell in NOνA detectors.

PVC extrusions are made by 16 unit cell put together side by side, as shown in Fig.2.10. The outer walls of extrusion are 4.8 mm thick and the inner walls between cells are 3.3 mm thick . The corners of the extrusions are rounded to reduce the stress on PVC structure from its weight.

Two extrusions (32 cells) are glued by adhesive epoxies, that are inert to the liquid scintillator, to form a modules. The fibers from each cell in the modules are taken to optical connector using fiber raceways as shown in Fig.2.11. Optical connector has 32 slots which maps to 32 pixels on APD. Each slot contains two fiber ends from each cell. For more details, please see the Technical Design Reports of NOνA [41].

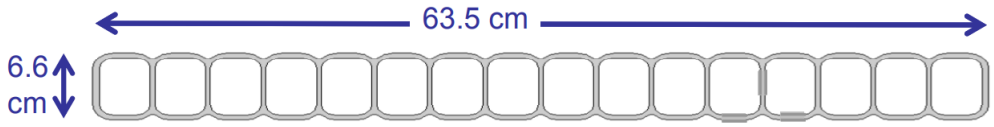

FIGURE 2.10: Cross section view of an extrusion (16 cells) in NOνA with, roughly, an aspect ratio of 1.5:1.

The modules (32 PVC cells) are put together to make planes. Both the detectors are made up such planes arranged in alternate fashion. The alternate planes helps in reconstruction of 3-D view of the event. Both the detectors' shape is rectangular parallelepiped as can be seen from Fig.2.8.

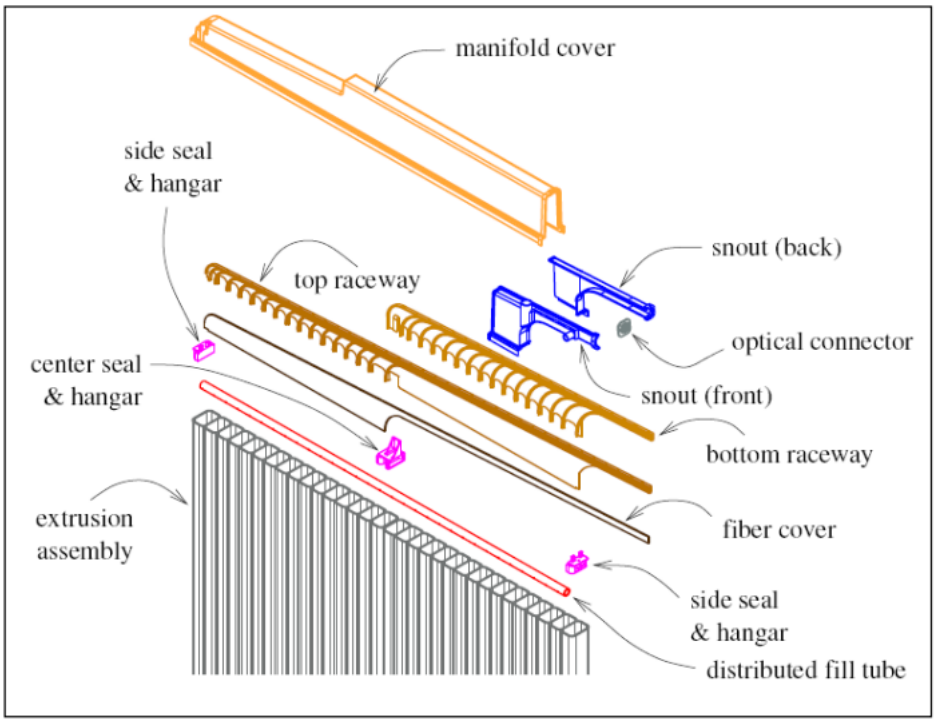

FIGURE 2.11: Various component of a NOνA plane assembly.

As mentioned earlier, the alternate planes (cells perpendicular to each other) enables to create a 3-D view in the detector. In near detector, a plane is made up of three modules. Twenty four planes constitute a block. Near detector is made up of total eight blocks. Near detector also contains a muon catcher to range out muons. Muon catcher contains

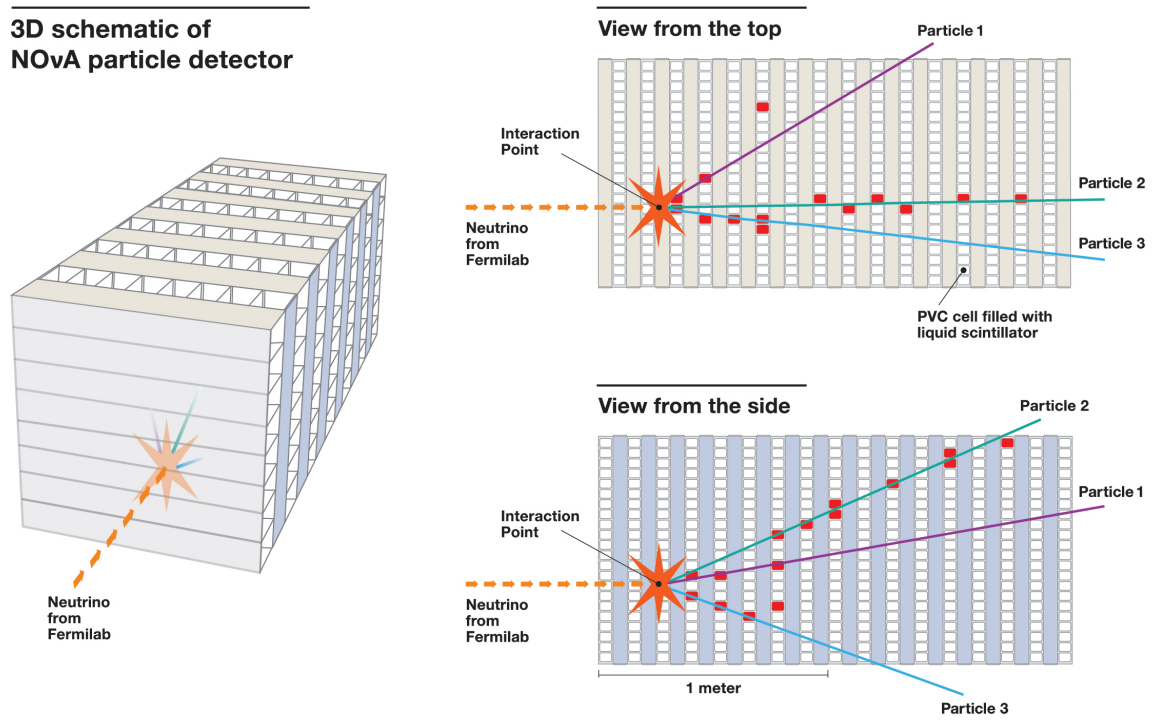

FIGURE 2.12: NOνA detectors are made up of planes, made of cells, put in alternate fashion. Two 2-D views of an event can be reconstructed from alternate planes as shown on the right hand side of the diagram..

10 planes of steel and each steel plane is 4.03 inches in width. The steel planes are inter weaved with PVC planes in between. The width of all planes in muon catcher (steel + PVC modules) is ∼3x3 meters. The total number of active PVC planes in near detector is 214.

In the far detector each plane consists of 12 modules and 32 planes make a block. There are total 28 block and 896 planes in far detector. 65% of the Far detector is active material. Fig.2.12 shows the 3-D schematic diagram of the NOνA detectors which are made of alternate planes. The 2-D views of an event are also shown.

Fig.2.13 and Fig.2.14 show the fully constructed near and the far NOνA detectors.

FIGURE 2.13: Near detector at Fermilab.

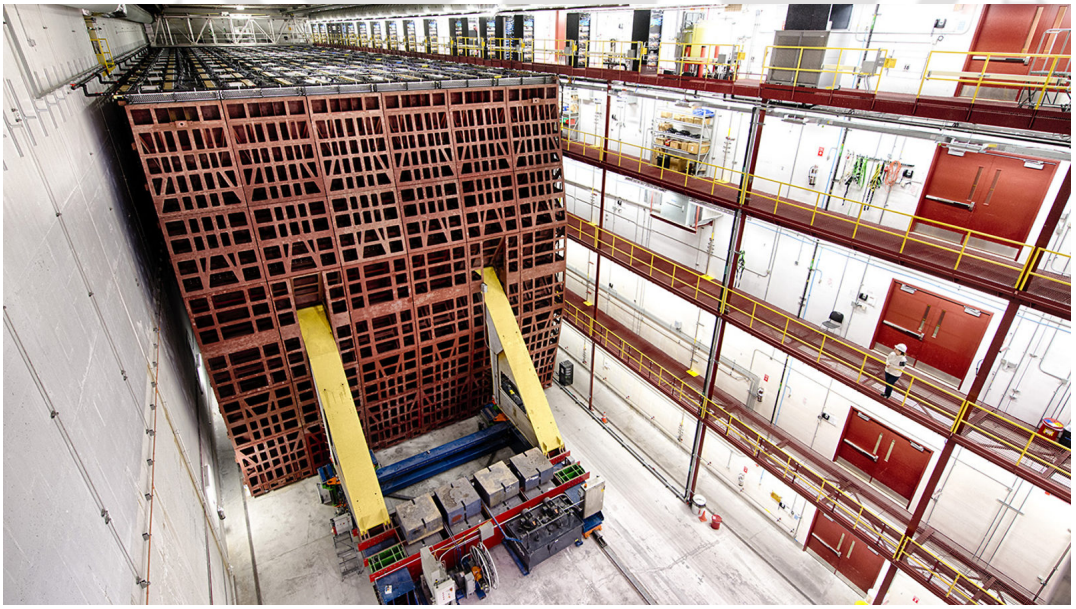

FIGURE 2.14: Far detector at Ash river.

2.3 Data Acquisition (DAQ) System

The main task of NOνA DAQ system [73] is to concentrate data from all of the APDs in the detector and store them into a hard disks permanently for offline analysis. In NOνA an event is recorded hits from the detector in a 500 μsec time window. The detector front end board electronics (FEB) operate in a continuous readout mode with no trigger at the FEB level. At an intermediate level of DAQ system, there is a buffer system which can hold (cached) the data upto 30 minutes. The data is hold so that it can be searched for the various triggers conditions. During buffer time, the data waits for the beam spill trigger. Beam window is 30 μsec window centered around 11 μsec beam spill window. The beam spill data rate is 190 GB per year. The DAQ system interfaces with Data Driver Trigger (DDT) where one can design the customized triggers.

The cached data can be analyzed for calibration and monitoring purpose as well. For this purpose, off-beam spill data (which is not synchronous with beam timing window) is used. The off-spill beam data is mostly cosmic activity in the detector at the rate of 200 kHz which corresponds to 40 MHz hit rate. Such data storage rate is 0.5 GB/sec which translates to 18 TB/year data. The corresponding rate at Near detector is quite low due to its location which is 100 meter underground. Table 2.1 summarizes the count and rates in the NOνA detector.

DAQ system of near detector and far detector is same except that in near detector there is a need for fast sampling due to higher event rate. Also the signal from far end of cell is almost 4 times higher. This requires the multiplexing rate to be 2:1in near detector as compared to 8:1 in case of far detector.

Fig.2.15 shows the schematic diagram of the front end board. Front end board houses avalanche photo diode (APD), application-specific integrated circuit (ASIC), analog to digital converter (ADC) and field programmable gate array (FPGA). 32 channels of APDs are interfaced with 32 cell scintillator detector module. Combined together it is called as APD module. APDs are photosensitive avalanche diodes in which a p-n junction operates at reverse biased mode such that on impact of photon in collector region electron hole pairs are produced which initiate, due to high voltage, a avalanche by further ionization process

TABLE 2.1: Counts and rate in NOνA detectors

	Far Detector	Near Detector
Front end boards	10,752	631
Front end boards per plane	12	2.5
APD channels per board	32	32
Total channels	344,064	20,192
Average Noise hit rate per channel	<30 Hz	$< 30Hz$
Total Noise hit rate	< 12 MHz	500 KHz
Bytes per hit (channel ID, TDC, ADC, status)	10	10
Muon rate	200 kHz	5 Hz
Average Hit channels/muon	200(mulitple hits/plane	50
Muon Hit Rate	40 MHz	250Hz
Total hit rate	52 MHz	500k Hz
Average hits rate per channel	100Hz	50 per hour
Digitizing rate	2MHz	8 MHz
Total data rate	0.4 GB/s	5MB/s
Average occupancy	2.5×10^{-4}	3×10^{-5}

while passing through the junction region. The APDs are kept at a volatage of ∼425 V and run at a gain of 100. In oder to suppress the dark noise, which is greatly due to the generation of thermal electron-hole pair, the APDs are kept at very low temperature (-15 C) during the operation. A thermo electric cooling (TEC) device is mounted on the APD module to provide cooling by applying a voltage across the material. The voltage applied creates a temperature difference on the two sides of the device according to Peltier Effect. The excess heat is taken away from TEC device via a water cooling system. Dry air is also ventilated into the APD in order to prevent moisture.

The APD quantum efficiency is quite high (85%) for the range of wavelength (500-550 nm) that WLS fiber transmits. A high quantum efficiency reduces the multiplicity of the electronics readout.

A customized low noise ASIC amplifier is developed at NOνA such that it maximizes the sensitivity of the low signal from the channels. Signals coming from individual APDs are processed, amplified and pulse shaped by ASIC. Then the signal is multiplexed to an ADC in set of 8 channels in case of far detector and in set of 4 channels in case of near detector. The digitization rate at far detector is 2 MHz and is 8 MHz at near detector (see table 2.1). APD gain is 100 and the signal to noise ratio is 10:1. Field Programmable

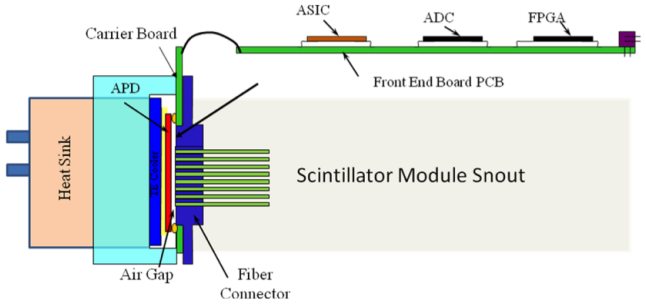

FIGURE 2.15: An APD module in NOνA.

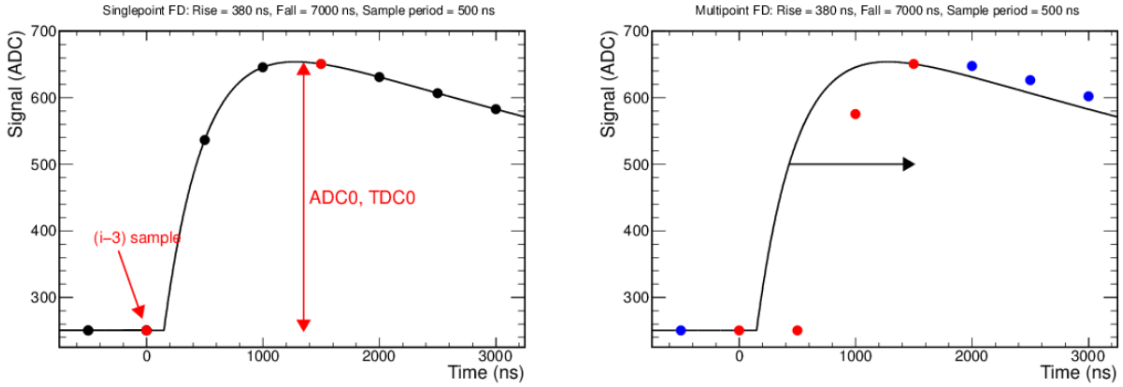

FIGURE 2.16: Multipoint read out ADC [42].

Gate Array (FPGA) on FEB does digital signal processing, I/O functions and monitoring. Signal that is processed at FEB has the time and amplitude information.

FPGA uses the dual correlated sampling (DCS) algorithm to identify the pulse to be a signal as shown in Fig.2.16. Zero suppression of the ADC samples is performed at (FPGA). In DCS algorithm, each ADC sample i is compared to $i-3$. If the difference between those two samples is greater than a fixed threshold, the baseline subtracted ADC value of i sample and it time values are recorded as the hit.

Fig.2.17 represents the various components of the NOνA DAQ system. The flow of data is from left to right. Data in time windows for nanoslices are sent from 64 APDs to one data concentrate module (DCM). DCM is device with an embedded CPU running Linux and custom NOνA DAQ software. At DCM, the data is concentrated to micro slices. The data is sent from DCM to buffer nodes. At buffer nodes data can be held for

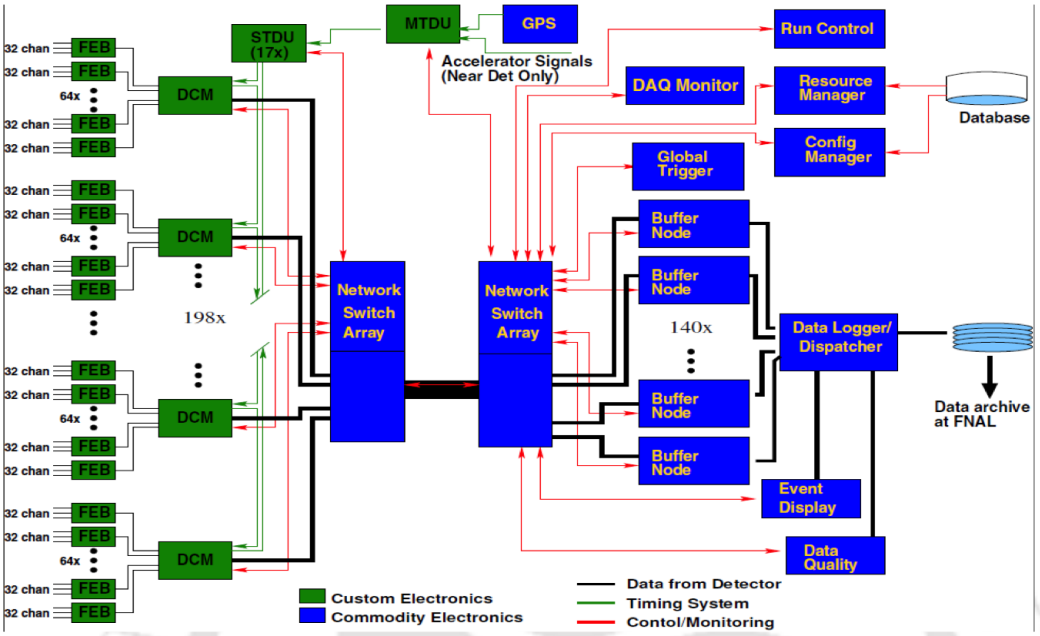

FIGURE 2.17: Schematic diagram of NOνA DAQ system [43].

certain time (20 sec - 30 minutes) awaiting the trigger (beam or DDT). The data which passes triggers is finally sent to data logger for storage using FTS (file transfer system).

NOνA timing system keeps all the DCMs synchronized to GPS clock. The accuracy is less than one 64 MHz tick across the entire system. The timing information is distributed to FEB by DCMs.

There are several DAQ application tasks which are required to be performed such as the control of run start, pause and stop, management of buffer farm, system monitoring etc. Dedicated softwares, are developed for performing all the DAQ application tasks. Run control system is responsible for configuring and controlling the data acquisition system. Run control supports the partitioning of resources . There can be multiple partition such that data collection can be done from different parts of the detector for commissioning and debugging as well. All the resources and configurations (like hardware components and software applications) are provided to partitions by resource manager. Resource manager acts as an interface to configuration database to the run control.

For monitoring purposes, the monitor server send the information concerning with the health of various components of the DAQ system to the graphic interfaces in control room.

Event monitor statistically analyses the data and displays it through the event display in the control room. There are other monitoring tools available at NOνA which are responsible for displaying graphically physical environment of the detectors, like temperature, voltages etc.

Triggered data from buffer nodes is sent to data logger system where the data can be stored for various streams, like reconstruction, data, etc. Logger systems store the data to local disk cache area. Afterwards it is sent to more permanent storage location using file transfer system.

2.3.1 Message Service Analyzer and Error Handling in NOνA Detector

The data acquisition (DAQ) system of NOνA is complex, consisting of more than 400 distributed but closely-interacting components. For the far detector, 12,036 front-end boards are used to continuously read out 385,152 detector channels. The raw data are first concatenated, based on the geometrical location of the detectors, in Data Concentrator Modules (DCMs). They are then pushed downstream to a buffer farm. There are nodes dedicated to management and coordination of the system, e.g., run control, application manager, message server, and DAQ online monitor. Therefore, to maintain a healthy running state of the DAQ system, it is necessary to have a means of continuously monitoring each participating component and for detecting and reacting, with minimum delay, to abnormalities that might put the system or the quality of data in jeopardy. The system which takes care of it is called as Automatic Error Recovery System (AERS)[3].

The AERS is composed of three functional parts:

1. Status Report: Status reports are issued, routed, collected and stored though a distributed message service package MESSAGE FACILITY developed by Scientific Simulation Initiative (SSI) Fermilab.

[3] https://indico.cern.ch/event/304944/contributions/1672234/attachments/578486/796618/CHEP_NOvA_DAQ_MessageAnalyzer.pdf

2. Message Analyzer : Decision making engine performs the correlation analysis based on the status reports. Message analyzer is a rule engine that processes Event-Condition-Action rules (ECA) and sends an email to the shifter informing about the status of the detector .

3. SUPERVISOR: Error handling supervisor reacts to situations and carries out the action. The action part of an ECA rule submits the detected error to the supervisor. Supervisor takes proper action based on error submitted.

2.3.1.1 Message Analyzer

A message analyser is a real time event correlation analysis based on log messages. Log messages issued from various DAQ components are scanned for certain strings to extract the information of the component state which we called as fact extraction. A series of test on incoming messages are done on several aspects like

1. Issuer, severity and category.
2. Regex pattern matching.
3. Occurrence or frequency

Here is an example:

Ex.- During data taking NOvA run control sends out a heartbeat check every second to components and expects a response otherwise emits a status message. A Fact that DCM has a heart attack is asserted as:

- Test whether the sender is RunControl.
- Test whether the severity is Warning or higher.
- Test whether the category is RegularCheck.
- Test whether the body matches regex expression DCM missed heartbeats.

- Test whether the message has been seen over 10 times in the past 60 seconds.

If all the testes are successful, a flag of TRUE is marked under the fact that the "DCM has a heart attack"(output of fact extraction is a BOOLEAN FLAG). Next step is to identify the event using the aforesaid information.

2.3.1.2 Event Identification

Event identification is achieved by a correlation analysis, called as rule, which is based on input from previous steps of extracted facts. Extracted facts are logically correlated to identify events. For example the following correlation -

"an event with dcm_failure is asserted when both dcm_heartattack and dcm_selfcheck_failure flagged during data_taking or hardware_config stage".

Which means that there is a DCM failure while taking data due to the two facts which correspond to not-responsive DCM. There are also scenarios when multiple rules are combined. This is called as Cascading. In cascading process, a rule is triggered for querying its condition on an incoming event and executing its action may in turn trigger new rules for further assessment.

Due to high number of DAQ components, it becomes complicated to write facts and rules for each connecting component, which differ slightly. So to cope with this issue, we have a grouping approach called the Extended Approach. This extends a primitive fact into a collection of related facts by distinguishing sources and targets of a message.

The facts and rules form the knowledge base of the Message Analyzer and can be written in the form of a configuration file. One simple rule is shown in Fig.2.18

Sometimes, more complex situation can not be handled just by above rules. In such cases, user-defined functions which enable the Message Analyzer to handle a great deal of more complex situations using customized logic implemented in generic programming language such as C++ are developed.

```
Rule_simple :
{
  type        : simple
  description : "this is an example of a basic rule"

  # filtering conditions
  severity    : warning
  sources     : ["dcm.*", "rcWindow"]
  categories  : ["*"]

  # match by regex, or match by 'if contains', not both
  regex       : "corruption\s.*\sdcm-03-\d{2}-\d{2}"
  contains    : "corruption"

  # frequency count can be for each source
  frequency   : { count     : 10
                  timespan  : 60   # in seconds
                  persource : true }
  # actions can be empty
  actions     : { alert : { level:warning message:"%s corrupted" }
                  popup : { message : "…" }
                  RunControl : { … } }
}
```

FIGURE 2.18: A simple rule in message analyser.

In summary the Message Analyzer is a light-weight correlation analysis tool with great flexibility and extensibility. It is being used for monitoring run state and data quality in the NOvA DAQ system. The Message Analyzer, including the configuration language for defining conditions and rules, were developed at Fermilab by Qiming Lu. Please see [74] for more details.

2.4 Simulations

As described in [75], NOνA simulation chain comprised of multiple steps. Division of simulation in several steps reduces the complexity and also allows to possibility of re-tune the previous step in case of any error without worrying about reproducing the whole simulation again, which is a very time consuming process in terms of computing resources. The simulation comprises broadly of following processes:

- Beam simulation.
- Neutrino interactions and Cosmic rays simulations.

- Detector simulation.
- Parametrized front end simulation.

The first simulation is NuMI beam or the flux simulation. it begins with simulation of hadron production in the target to produce pions and kaons. Then simulation of magnetic horns to focus produced pions and kaons in the beam line and subsequently their decay to neutrinos in the decay pipe. The simulation is done with FLUGG which consists of FLUKA [76, 77] for generation particles and GEANT4 [78, 79] to simulate the geometry of environment of NuMI beam line. The geometry includes target hall, the horns and the beam decay pipe as shown in Fig. 2.3. As an output, the neutrinos and the parent informations are stored in the flux files.

Using the information previous step, neutrino interactions in the detectors are simulated using the GENIE generator. Alongside with the GENIE particle generation step, cosmic particles are also generated separately using the Cosmic RaY (CRY) generator [80]. GENIE uses the nuclear models like Bodek and Ritchie version of the Relativistic Fermi Gas (RFG) which are modified to include interaction involving nucleon-nucleon interaction (termed as meson-exchange current (MEC)). GENIE uses the sophisticated interaction specific cross section models to allow for quasi elastic (QE), deep inelastic scattering (DIS) and resonant (RES). It takes into account any intranuclear scattering which is followed by initial interaction and is termed as final state interactions (FSI). FSI simulation is done using a GENIE subpackage called INTRANUKE. Once primary particle are generated using GENIE or CRY, the particles are propagated in the detector, so that they can further interact and decay to secondaries and deposit energy in the active medium, using GEANT4. The detector geometry is simulated using GEANT4. Final state hadrons are modelled using physics list QGSP_BERT_HP . QGSP specifies quark gluon string model and high energy hadrons ($>$ 10GEV) are modelled using with it and hadrons ($<$ 10GeV) are modelled using Bertini cascade model. The low energy neutrons are modelled using the high precision (HP) neutron model. The generated particles along with its informations are kept in form of what is called as MC truth. Rock muons which are produced in the surrounding rock of the near detector are simulated separately than

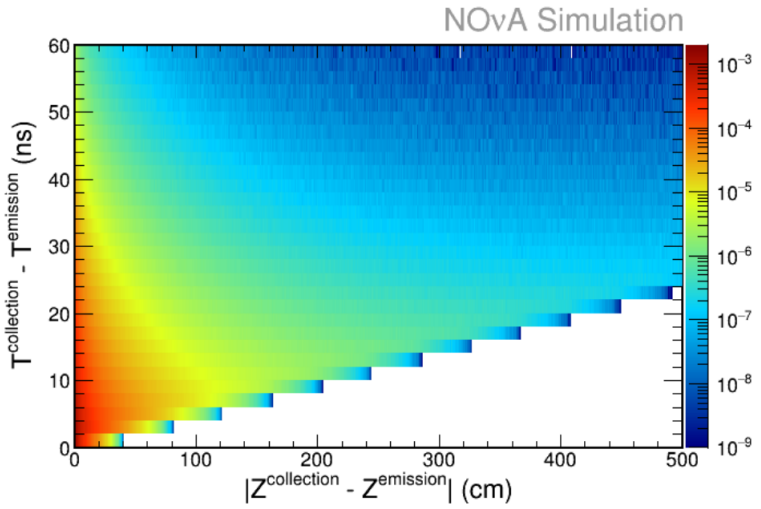

FIGURE 2.19: The collection rate of scintillation photons by a wavelength shifting fiber loop as estimated by a custom ray tracing simulation.

the GENIE stage and only the rock muons which are having the interaction in the detector are stored. The generated rock muon events are overlayed with GENIE events, after the GEANT4 stage.

After the aforesaid steps, the next step is to convert the deposited energy into the light and its transportation to the electronics through wavelength shifting fibers. Though GEANT4 has the capability to simulate it, however it is a time consuming process. Considering all the cells and readout in the NOνA are identical, the NOνA has generated template to parameterize photon transport. The templates once generated for a cell can be used for other cells as well. The collection of scintillation photon by fiber and its transport through it and the APD response of the captured light are parameterized. A ray tracing algorithm is developed which uses the the scintillation spectrum of the liquid scintillator, reflectivity of the PVC cell walls, and the absorption spectrum of the fibers and result is shown in Fig. 2.19

Next, the photon are to be transported to the electronics. Owing to the the fact that fiber is looped, half of the photons are transported to each ends. The photons are attenuated as per the attenuation curve which is derived from quality control tests on the detector.

After the photon transport simulation, electronics readout process is simulated. It begins with number of PEs (photo-electrons) collected by APDs. The APD is modelled to have 85% quantum efficiency. Excess noise factor is applied before readout simulation simulates all the three components of FPGA - ASIC for pulse shaping, ADC to digitize the analog signal and FPGA which performs the real-time suppression. The CR-RC circuit in ASIC is modelled as given by equation Eq. 2.4,

$$f(t) = N_{pe}\frac{F}{F-R}\left(e^{-(t-t_0)/F} - e^{-(t-t_0)/R}\right) \tag{2.4}$$

where, t_0 is the time of collection of photo-electrons by APD, F and R are the fall time and rise time of CR-RC circuit respectively. N_{pe} is the number of photo-electrons collected. The pulsed shape is converted on a pre-determined PE to ADC conversion factor which is derived from charge injection studies. A baseline is decided from pedestal data. The FPGA then searches for peaks, above threshold, using a dual correlated sampling trace as:

$$dcs_i = ADC_i - ADC_{i-3} \tag{2.5}$$

where ADC_i is the i_th output from the ADC. A hit is registered in the cell if the difference as given by Eq. 2.5 is higher than threshold.

In the later chapters you will see, how the above simulations are verified by data driven methods by using cosmic induced EM shower.

2.5 Calibration

Estimation of energy deposited in the cell is done by calibrating the hits using reference hits from cosmic muons which stops in the detector as the muon deposited energy in a cell is well defined by the Bethe-Bloch equation. Due to the length of NOνA cell, the photons get attenuated as they travel to one end of the cell to electronic readout. First step of calibration includes the correction to photons due to attenuation. It is done on

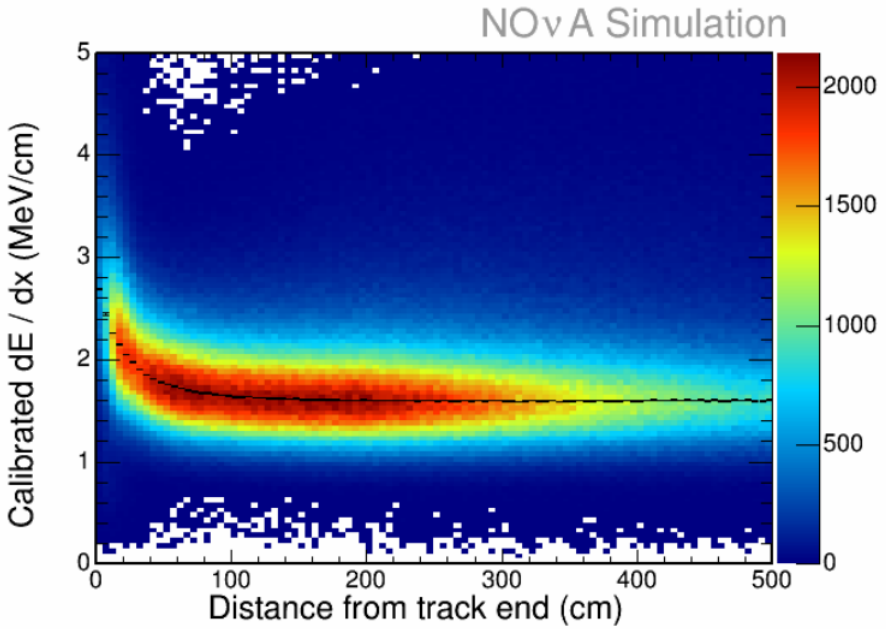

FIGURE 2.20: 2-D histogram depicting deposited energy per path length (dE/dX) [44]

each cell and the reference hits are from cosmic muon which satisfy certain criteria called as tri-cell criteria to reduce the uncertainty on the path length estimation in the cells. The Tri-cell hits are those hits which have two adjacent hits on muon trajectory. As a result of attenuation calibration, ADC value is converted to corrected photo electrons (PECorr). PECorr is independent of the position of the hit in the cell.

Next step is to use the PECorr to convert it to energy in physics unit, GeV. This is called as absolute calibration. For this cosmic muon hits with well understood dE/dx information (Bethe Bloch equation) are taken as reference as shown in Fig.2.20.

An event, in the form of hits, needs to be reconstructed so that the useful physics information can be extracted. Since the thesis work is related to ν_e-appearance analysis, a detailed description of ν_e-CC reconstruction method is given in the next chapter.

2.6 Chapter Summary

This chapter presents an overview of the NOνA beam line and the detectors. The chapter also includes a summary of the NOνA DAQ system including various monitoring tools to ensure good quality of data for physics analysis. NOνA experiment is long baseline neutrino oscillations experiments. It has been designed and optimized to detect ν_e-CC interactions. The cells size and low z material gives NOνA a capability to differentiate π^0 and electron initiated showers and thereby a good selection efficiency over backgrounds. Use of near detector helps in canceling most of the systematic uncertainties. Cosmic data has been utilized for calibrating the detector.

Chapter 3

Reconstruction

To extract useful physics information from the hits deposited by the particles in the detector, we need to associate particle informations to the hits. The process through which kinematics and dynamic information is attached to an event is called as reconstruction. In order to achieve NOνA's main Physics goal, we need to identify various interactions and particles such as - $\nu_e - CC,\ \nu_\mu - CC,\ \pi^0, \mu$ etc. These particles and processes manifest themselves as, EM showers, tracks or prongs in the detector as shown in Fig.3.1.

For $\nu_\mu \rightarrow \nu_e$ oscillation, $\nu_e - CC$ exhibits itself as electromagnetic (EM) shower in the detector. The backgrounds to this process mostly come from neutral currents and energetic π^0 that instantly decays to two photons. The reconstruction for this channel in NOνA constitutes of various steps as shown in Fig.3.2 [81] which are described in details in the following sections.

3.1 Slicer

The current section is a based on the detailed work as found in reference [82]. In NOνA, an event record is a collection of activities that has occurred in the detector within a readout window of 500 μsec. The hits can be originated from noise or signals. In far detector, the readout window contains 50 - 70 cosmic ray particles, whereas in near the

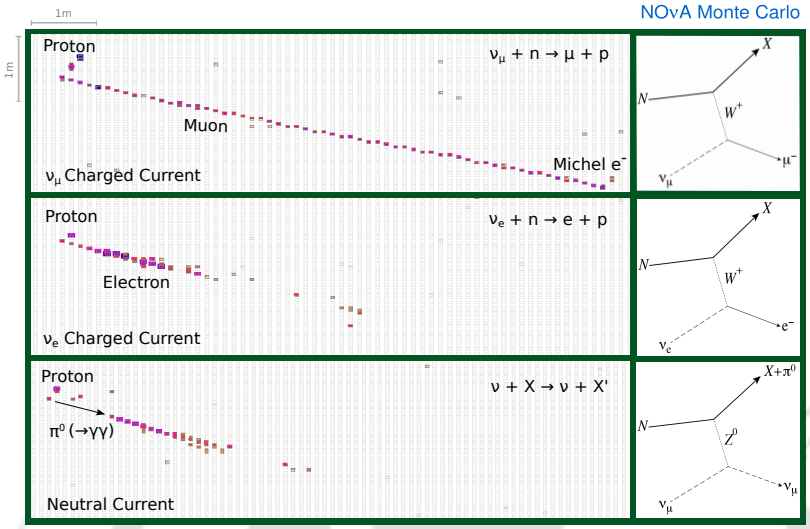

FIGURE 3.1: Various ν interactions' events topologies in NOνA detector.

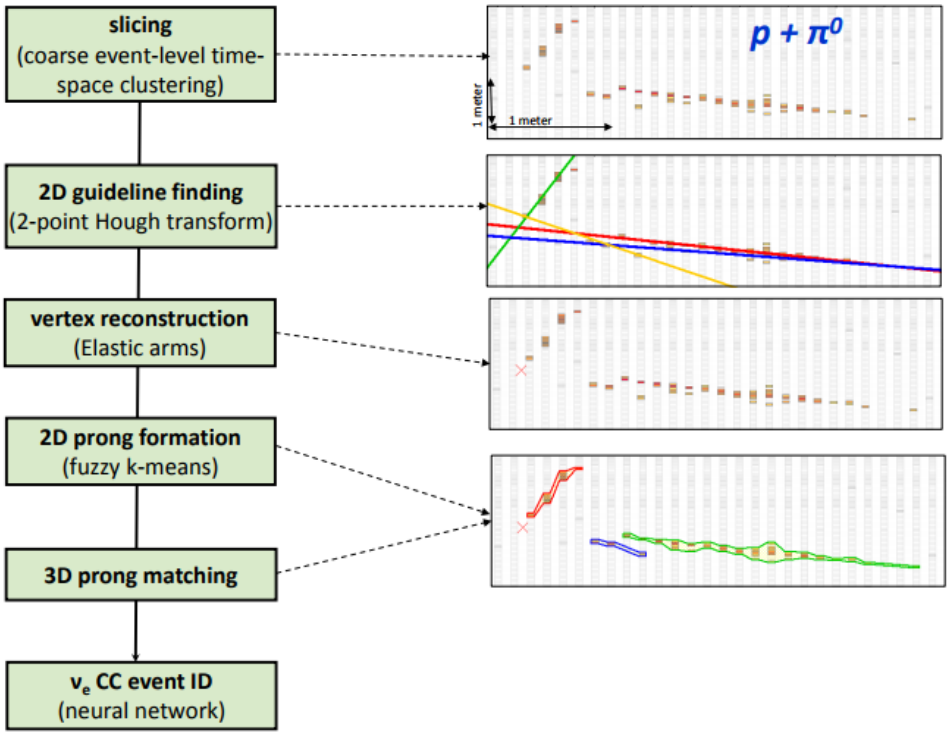

FIGURE 3.2: Reconstruction chain for reconstructing a ν_e-CC event in NOνA far detector.

detector, there are approximately 4 neutrino interactions per 10 μsec beam spill. The first step of reconstruction is slicing, which means to separate the noise and signal hits into different clusters (collection of hits) such that there should be one event (a neutrino or cosmic interaction) per cluster.

This is achieved in NOνA by the implementation of a density-based clustering algorithm (DBSCAN) algorithm. Through this algorithm hits can be characterized as clusters on the basis of density of points. These points are part of some parameter space. The algorithm loops over all points (which correspond to hits) and determine on the basis of density threshold whether the point is a core point or the border point. A core point should have a number of neighboring points ($MinPts$) which are more than certain number inside a distance (ϵ) which is determined by the score function given by Eq.3.1.

$$\epsilon = \left(\frac{\Delta T - \Delta\vec{r}/c}{T_{res}}\right)^2 + \left(\frac{\Delta Z}{D_{pen}}\right)^2 + \left(\frac{\Delta XY}{D_{pen}}\right)^2 + \left(\frac{PE_{pen}}{PE}\right)^5 \tag{3.1}$$

where $\Delta X, \Delta Y, \Delta Z$ are the distance between two hits coordinates, ΔT is the time difference between two hits, T_{res} is the timing resolution of the two hits added in quadrature and D_{pen} is a distance penalty, PE is the sum in quadrature of the number of photoelectrons in the two hits. The last term is raised to power 5 to suppress the low energy hits from the noise as the noise spectrum falls off as $PE^{-2.5}$. The score function determines whether the two hits are neighbor or not. The first term penalizes the hits if they are far separated in time, the second and third term penalize the hits which are far separated in space, $MinPts$, ϵ, D_{pen}, PE_{pen} are free parameters and has been tuned.

Performance is defined in terms of efficiency and purity. Efficiency is defined as the measure of fraction from an interaction that are contained in a slice. Purity is defined as the fraction of hits that come form the leading interaction in the slice. The slicer algorithm free parameters are tuned to maximize the slices with high efficiency and purity. For over 80% of slices in ND have efficiency over 90% and over 95% slices have purity over 90%.

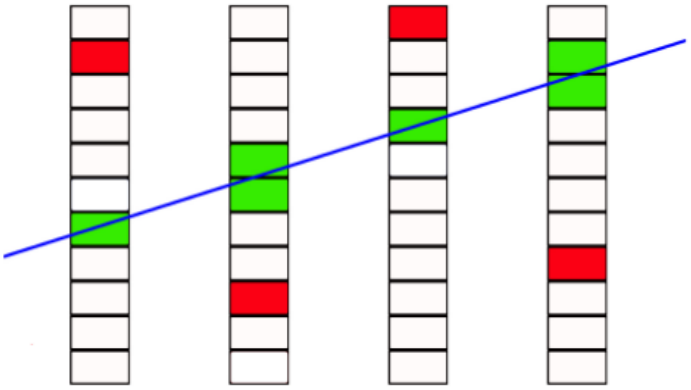

FIGURE 3.3: Tracking approach.

3.2 Tracking

The trajectories of particles are fitted as tracks. Two kinds of tracking algorithms are used. The tracking is useful for the particle which does not produce large electromagnetic or hadronic showers.

3.2.1 Cosmic Tracking

As the name suggest, cosmic track algorithm produces the tracks fitted from the trajectories of cosmic particles (mostly muons) coming from atmosphere in the detector. The muon trajectory is characterized as a straight line as most of the hits along the particle trajectory are confined to one or two cells per plane. The tracks are made out of the long trajectories by fitting a line in either view (X or Y) in the detector. The tracking is performed on slices (as described in the previous section). The hits which are far from the line are removed. By doing this, most of the noise hits are removed as shown in Fig.3.3. The tracks which are formed so are 2-D tracks. The two tracks found in both the views are combined to form 3-D tracks. For more details please see [83].

3.2.2 Kalman Tracking

Kalman filter algorithm [84] works on "predict and probe" method. In this method, the best estimate of track position is determined using the measured hits and noise in the process. The process noise includes measurement error and system error. Measurement error arises due to the finite width and depth of the cell and system error comes from the multiple scattering of the particle on its path. Kalman filter algorithm has three broad steps in NOVA: track finding, track fitting and view matching.

In the track finding process, input hits which are 4 cells apart are used as the seeds to the algorithm. The pair of hits (downstream in the beam direction) in the slice are used to estimate the position and slope of the track. Using the predict and probe method, the cells which are close to the predicted location are added to the track (the hits should be $8\chi^2$ units away from the track position) and track is expanded upstream of the beam direction. This process is updated after every hit is added and goes on till no hit is left to be added to the track. After this the process is repeated in the reverse direction to pick up the left over hits. In neutrino interaction, tracks are emanating in downstream direction (beam direction). Therefore, the longest track are picked up first by expanding the track from downstream direction to upstream. All the hit positions are fit with the piecewise linear segments to allow deviation originating from multiple scattering.

After track is found in either view (X-Z or Y-Z), the two tracks are matched to make a 3-D track. Two tracks are matched in either view by matching the start of the track and end of the track coordinates. The tracks are put in a pair in descending order of proximity of their start and end points in space.

3.2.3 Hough Transform

Hough transform algorithm is utilized in NOvA to find out the patterns or global features like straight lines in a event. The algorithm is applied separately to each view. The hits from the event are taken into pair and a line is drawn with the points. The line is characterized by polar coordinates (ρ, θ), where ρ is the perpendicular distance from

the origin to the line and θ is the angle between ρ and the x axis. The line can thus be represented by a point in parameter space (Hough Space). The hough space is filled with all the point, representing lines, and each point is filled in the hough space with a Gaussian smear vote according to

$$vote = e^{-\frac{(\rho-\rho_0)^2}{2\sigma_\rho^2}} e^{-\frac{(\theta-\theta_0)^2}{2\sigma_\theta^2}} \tag{3.2}$$

where $\sigma_\rho = 3/\sqrt{12}$ and $\sigma_\theta = 3/d\sqrt{6}$ and d is the distance between the two hits in the detector. The values of parameters are found after tuning the algorithm on simulation. After filling the hough space by the the votes as above, the major lines (event features) will occur as peaks. The peak is smoothened by average hit in the map using a Gaussian smoothing weight. A threshold above which a peak will be considered, is determined for each hough space using the average bin height. The peaks which exist above the threshold are taken as lines. However this way it produces many lines. To circumvent this NOνA uses an alternate approach. In the alternate approach, the line corresponding to highest peak is identified. After this the highest hough peak hits are removed from the event and hough transform is again applied to remaining hits and a hough space is generated again. This way smaller lines are identified which were not represented in the first iteration. The process is repeated till no hough peak above the threshold is found or a preset number of lines is reached. This is called as "Multi-Hough" approach [85].

The performance of the Multi Hough is determined by measuring the perpendicular distance between the true interaction vertex and the primary hough line. Using simulation studies it is found that primary hough lines pass within an average distance of 6.9 cm for Neutral Currents, 4.1 cm for $\nu_\mu - CC$ and 2.7 cm for $\nu_e - CC$ cm from the interaction vertex and for secondary hough line it is 9.9 cm for Neutral Currents, 8.2 cm for $\nu_\mu - CC$ and 8.8 cm for $\nu_\mathrm{e} - CC$.

3.2.4 Elastic Arms Vertex

Elastic Arms algorithm [86] is used to find out the global event vertex using the input from the Multi-Hough transform algorithm. Elastic arms is based on algorithm known as method of deformable templates [87]. In NOvA the typical template for a neutrino interaction is the vertex with secondary particle originating from it. The secondary particle usually come out as tracks or shower and usually have angular separation. Each of these secondary particles can be approximated by an "arm" from the vertex.

The method includes the input of Hough lines as the maximum number of "arms" in either XZ or XY view and generation of several vertexes from the intersection of the input line. A vertex is chosen out of list of vertexes by evaluating and optimizing an energy cost function as

$$E = \sum_{i=1}^{N}\sum_{a=1}^{M} V_{ia}M_{ia} + \lambda\sum_{i=1}^{N}\left(\sum_{a=1}^{M} V_{ia} - 1\right)^2 + \frac{2}{\lambda_v}\sum_{a=1}^{M} D_a \tag{3.3}$$

where V_{ia} is the measure how strongly a hit is associated with the arm a, M_{ia} is the distance between cell hit i and arm a, D_a is a measure of the distance between the first hit of arm a and vertex, and λs are the tunable parameters to control the strength of the terms with $\lambda_v = 7/9X_0$, the photon radiation length (30 cm in NOνA). V_{ia} is measured as

$$V_{ia} = \frac{e^{-\beta M_{ia}}}{e^{-\beta\lambda} + \sum_{b=1}^{M} e^{-\beta M_{ib}}} \tag{3.4}$$

where, $e^{-\beta\lambda}$ parameterizes the likelihood of hit being a noise hit and β is a parameter that represents the range over which a hit can have association with the arm a. The first term of Eq.3.3 is the measure of of goodness of fit between the hits and the arms and second term is the penalty for the hits which are not associated with the arms. The third term is a customized NOνA term, which is the penalty for the hits which have a gap between the vertex and first hit. This term is quite important in reducing the $\pi^0 \to \gamma\gamma$ backgrounds.

All the vertex seeds and the arms are evaluated using Eg.3.3 and minimization is done in ROOT using MINUITS[1]. The fit procedure is initialized at low value of β so as to avoid local minima of the energy function and it is gradually increased to reach on the final vertex point.

The algorithm gives the vertex resolution as 10.9 cm, 11.6 cm and 28.8 cm for $\nu_e - CC$, $\nu_\mu - CC$, and NC events, respectively.

3.2.5 Fuzzy K Vertex

Once the global event vertex is located, the next step is to find the number of prongs (possibly the number of particles) coming out of the interaction around the vertex. This is achieved by using the possibilistic fuzzy-k means algorithm [88][39]. Possibilistic term mean that a hit is not supposed to have membership in any of the prong. It can have no membership at all among all the prongs. These hits are considered as noise. The fuzzy term means that a hit is allowed to have multiple membership among different prongs.

The method utilizes the fact that the hits of an event as seen from its vertex can be converted to single dimensional angular space in which each hit is represented by an angle θ_j ranging from $-\pi$ to π, with Z axis (beam axis) as $\theta = 0$. The different density point of hits, on the θ-axis will then be an indication of number of possible prongs. Each hit is alloted an angular uncertainty, as a function of distance from the vertex, arising from the scattering along the path. The scattering angular uncertainty is modelled from simulation by generating 1 and 2 GeV muons and electrons scattering along the path.

Seeds for prongs are set at the local maxima in density of n hits in the angular space by using the density matrix w,

$$w_k = \sum_i^n e^{-\left(\frac{\theta_k - \theta_i}{\sigma_i}\right)^2} \tag{3.5}$$

[1] https://root.cern.ch/doc/master/classTMinuit.html

where $\theta_k = -\pi + \frac{k*\pi}{180}$ with k is varied between 0 and 360 in steps of 1. The membership of hit to a prong is determined as

$$\mu_{ij} = e^{-\frac{md_{ij}\sqrt{a}}{\beta}} \tag{3.6}$$

where $d_{ij} = \left(\frac{\theta_j - \theta_i}{\sigma_j}\right)^2$ and a as the number of prongs and m is a tunable parameter which represents the fuzziness in the cluster and is set to 2. β is a measure of spread of hits around the prong center and its value is set to 4. The prong centers are then updated to θ_i' according to

$$\theta' = \theta_i + \frac{\sum_{j=1}^{n} A_{ij}(\theta_j - \theta_i)}{\sum_{j=1}^{n} A_{ij}} \tag{3.7}$$

where $A_{ij} = \frac{U_{ij}^m}{\sigma_j^2}$ and the process is repeated until $\Delta\theta = |\theta_i' - \theta_i|$ is less than some predetermined tolerance for all the prong centers. Prong angles are updated and additional prongs are found until all the hits in the event have at least 1% membership in a prong. The prongs with significant overlaps are merged as one. The prong thus produced are in 2-D view (one form each X-Z and Y-Z view). To form 3-D prongs from each view, all the prongs are matched with each other by comparing the energy profile in each view. A Kuiper metric is minimized to find the best match as

$$K = D^+ + D^- \tag{3.8}$$

with D^+ and D^- are the distances between the profiles respectively.

The fuzzy k vertex algorithm performance can be judged by the completeness of hits in a prong. The process produces the prongs of completeness of 95% for $\nu_e - CCQE$, 86% for $\nu_e - CC\ non - QE$, 98% for $\nu_\mu - CCQE$ and 92% for $\nu_\mu\ non - CCQE$ events.

3.2.6 ANN Event Classification

To identify the ν_e-CC signal, NOνA has developed likelihood based particle identification algorithm, called as LID [42][39][89]. This method takes the input of the object produced by the previously mentioned methods in the reconstruction chain. Each particle deposits energy in a characteristic way in the cell along its path in the detector. LID takes the information of energy deposited per path length (dE/dX) to classify different types of particle that produce showers or tracks in the NOνA detectors.

As previously mentioned, the ν_e-CC appearance manifest itself as an EM shower. LID takes the input from Fuzzy K prong and recluster the fuzzy prongs so as to confirm the shape of the expected shower. Since Fuzzy K algorithm allows the overlap of prongs, LID performs the deconvolution so as to avoid the double counting of cell hits energy.

After reclustring and deconvolution, the highest shower is picked and dE/dx is calculated in two perpendicular direction: the longitudinal direction which is parallel to the direction of shower and transverse direction which is perpendicular to the direction of shower. In longitudinal direction, dE/dx is measured for each plane a particle has deposited energy in. In transverse direction, dE/dx is measured radially as a function of cell widths, from the center of the core and is a measure of how concentrated is the the core shower, as shown in Fig.3.4 and Fig.3.5.

The dE/dx information of particle is compared with the normalized templates histograms of the expected dE/dx in planes in longitudinal direction and transverse direction for different particles $(e, \gamma, \pi^0, \pi^{\pm}, \mu, p, n)$. The probabilities are drawn by comparing the energy deposition in i_{th} plane from the start of shower in template histograms. The logarithms of these probabilities are summed over all the planes and a likelihood is constructed for the particle hypothesis

$$LL(e^-) = \sum_{i}^{N} log(P_i(e^-))/N \tag{3.9}$$

where, $LL(e^-)$ is the likelihood for an particle to be an electron with N being the total number of planes for longitudinal likelihood or N being the total number of transverse

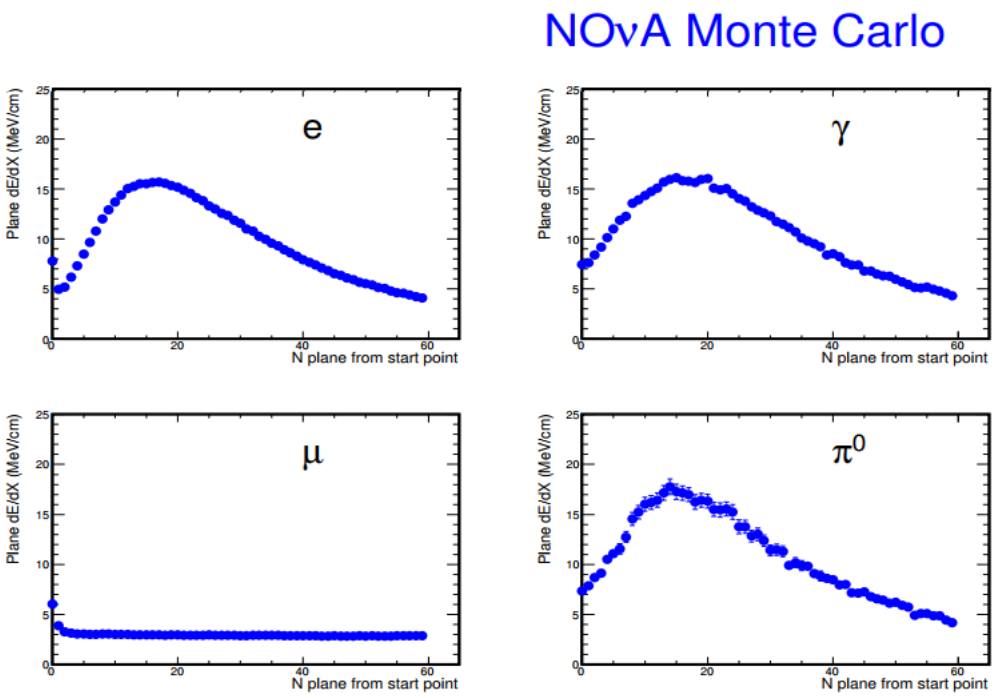

FIGURE 3.4: Longitudinal dE/dx deposition as a function of plane for various particles.

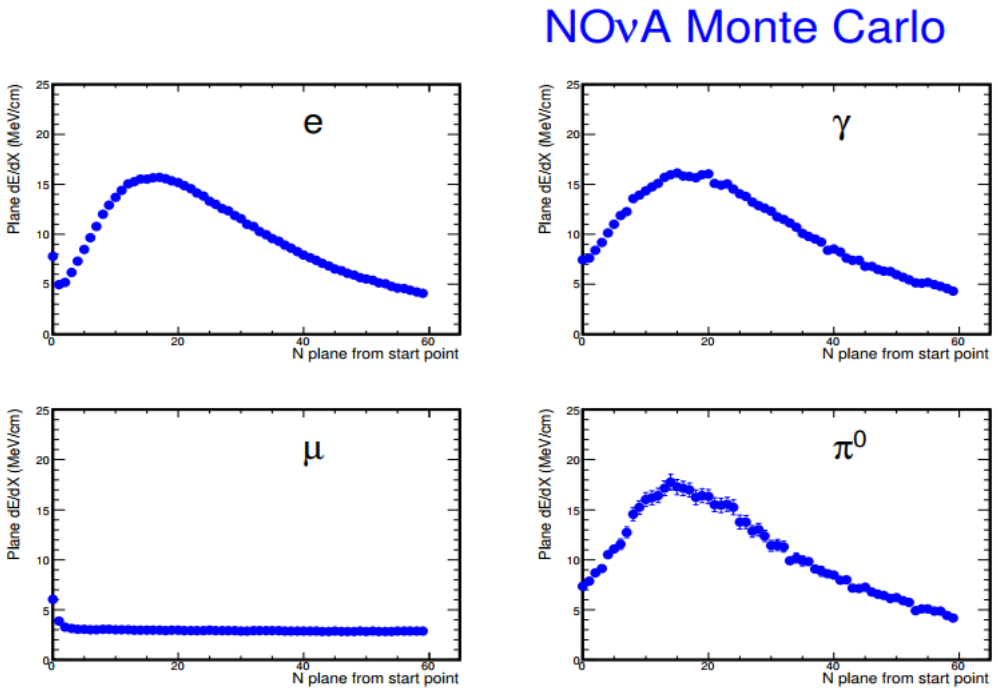

FIGURE 3.5: Transverse dE/dx deposition as a function of plane for various particles.

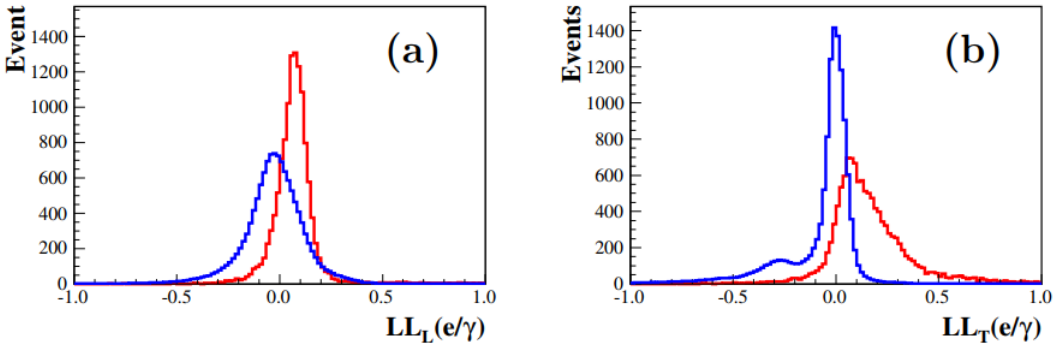

FIGURE 3.6: Likelihood differences between e and γ

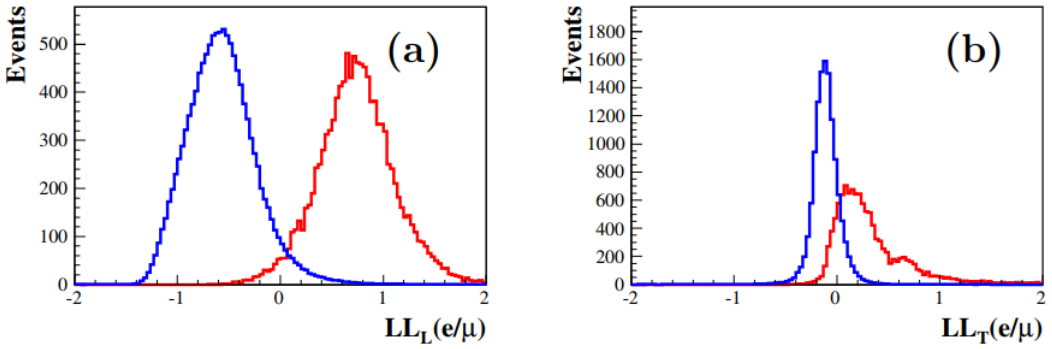

FIGURE 3.7: Likelihood differences between e and μ

cells for the transverse likelihood. The discriminating power can be seen by making more useful quantity out of likelihood- difference likelihoods. Various difference likelihood are shown in the Fig.3.6, Fig.3.7 and Fig.3.8.

An artificial neural network (ANN) is setup and is implemented in ROOT's TMVA package. The input variables are the above likelihood differences along with few more topological variables to identify $\nu_e - CC$ signal from the background. The other variables included are mass of π^0, the fraction of energy in most energetic shower, the energy deposited near the vertex and gap between vertex and the starting point of the shower. The performance of LID in separating the signal from background is shower in the Fig.3.9

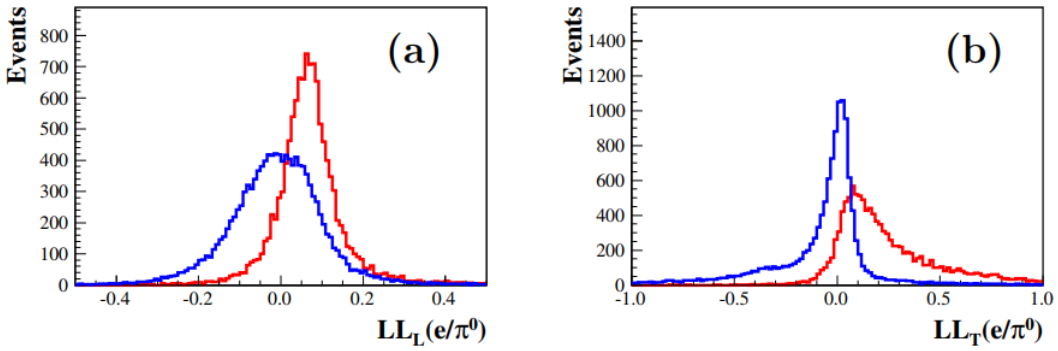

FIGURE 3.8: Likelihood differences between e and π^0

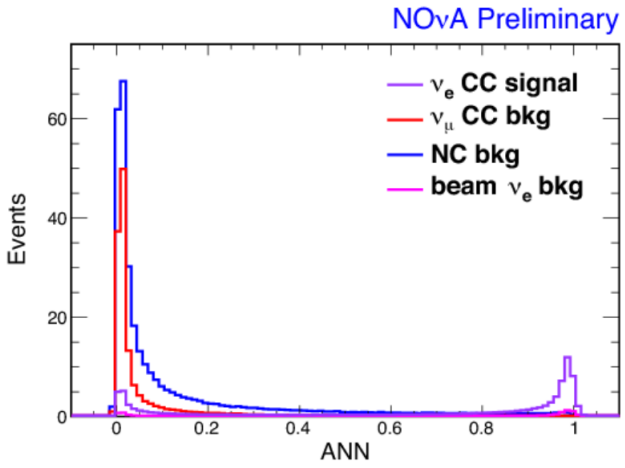

FIGURE 3.9: LID output

NOνA LID achieves the efficiency of 47.57 % in selecting $\nu_e - CC$ shower with purity 84.1%. The cosmic background rejection suppression is 95.8 million to 1.

Another particle classification is Convolution Visual Network (CVN) [90] which is based on deep convolution network from the field of computer vision. The main advantage of CVN is that it is independent of reconstruction chain and thereby free from any biases and errors, coming from reconstruction chain, which lead to misidentification of particles. In NOνA event display in X-Z and X-Y dimension acts as a picture with energy deposited by the particle in the cells as pixels. The two event displays can be fed as pictures with calibrated hits and is trained using the Caffe framework [91] with approximately

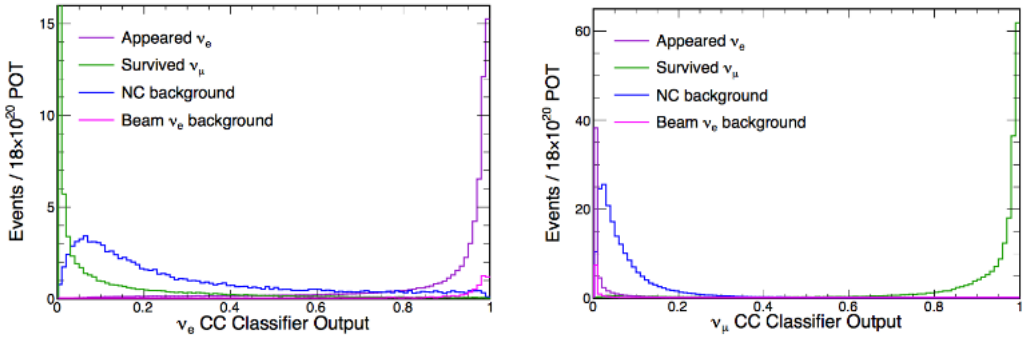

FIGURE 3.10: CVN output

4.7 million simulated neutrino interactions and cosmic ray data in far detector. The convolution layers are created with incorporation of features extracted from the pixels and output in the form of sets of variables and weights are fed through a multi-layer perception.

The output is a classification of events under different particle hypothesis, $\nu_\mu - CC$, $\nu_e - CC$, $\nu_\tau - CC$ and NC etc. as shown in Fig.3.10, where events are associated with score ranging from 0 to 1, where 0 is most likely to be background and 1 is most likely to be a signal event.

Another type of particle-ID has been developed for NOνA, called a Library Enabling Matching (LEM) [45] that classifies the events by matching the events against the simulated events, from vast library of simulated ν_e and background interactions, and characterize the particle with the most suitable matches from the library. The properties of the 1000 most matched library events are fed to decision tree to obtain a PID. Fig.3.11 shows a trial event and a best library match found by matching the energy. Fig.3.12 shows the LEM output for signal (ν_e) and backgrounds.

3.3 Chapter Summary

The signal for ν_e appearance in the NOνA detector is an electron initiated EM shower. To identify the shower, NOνA has developed the event reconstruction tools and particle-ID

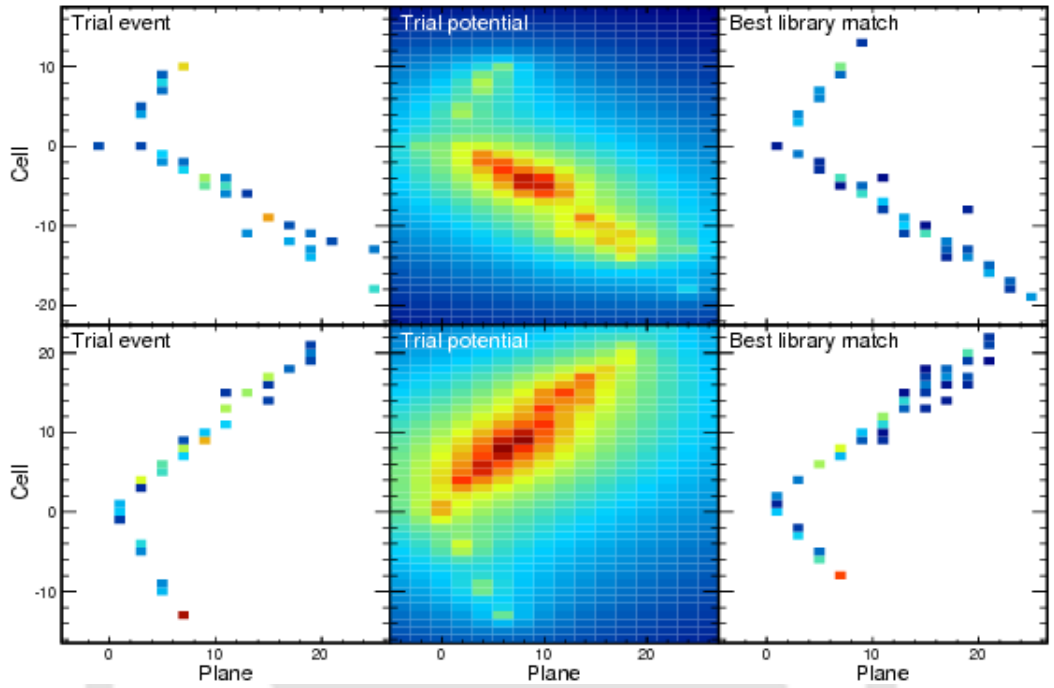

FIGURE 3.11: Example of LEM matching. On the left is a trial ν_e CC event, on the right the best match found. The central panels shows the potential U in which the library events are placed in order to calculate the match energy. The upper panels show one view, and the lower panels show the other. Picture taken from [45]

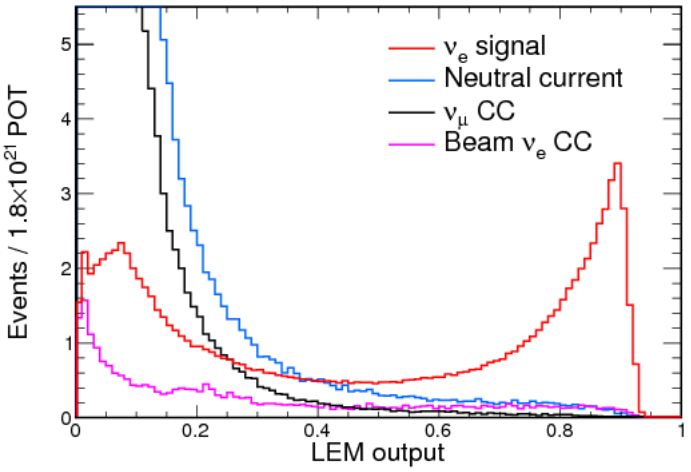

FIGURE 3.12: The distribution of the LEM output variable for ν_e CC signal events compared to the background components: neutral current (blue), ν_μCC (black) and intrinsic beam ν_e CC (magenta) . In order to make the details in the signal-like region visible, the y-axis truncates much of the background peak. 95% of neutral current events and 98% of ν_μ charged current events have LEM<0.15. The distributions are scaled to a nominal 3-year NuMI exposure of 1.8×10^{20} protons-on-target. Plot from [45]

algorithms. The reconstruction of particle in NOνA starts from calibrating the hits using the cosmic muons. After this, event is reconstructed through slicer, tracking, multi hough transform, elastic arms and Fuzzy K vertex algorithms, to get the global and kinematics based features of the event. Then an event is processed through particle identification algorithms so that an event can be classified under a particular particle hypothesis. NOνA has three PIDs- LEM, LID and CVN. LEM and LID have been used in first analysis and CVN has been used in the subsequent analysis as the event classifier. One of the objective of the thesis is to validate these particle ID algorithm using a data driven method.

Chapter 4

Bremsstrahlung EM Shower

In NOνA, the signal for the ν_e-CC interaction is the appearance of an electron signal in the far detector. On production, the electron in the far detector produces an EM shower in the detector. This requires correct EM shower modelling and reconstruction, and good detector calibration of attenuation and alignment. It is important to use independent EM shower sample based on data to test ν_e algorithms and calibration.

Cosmic ray muons are abundant in NOνA detectors. They induce EM showers by three different means: energetic muons undergoing bremsstrahlung radiation (Brem), muons decaying into electrons in flight (DiF), and muons stopping in the detectors and decaying into Michel electrons. Michel electrons have long been used in calibration. However, energy of Michel electrons ($\sim$ 53 MeV) is small compared to energy region of beam ν_e events ($\sim$ 2 GeV). Brem and DiF, on the other hand, provide abundant EM shower samples at few-GeV energy region. Once isolated, they provide large number of EM shower as independent test samples.

Brem and DiF both yield EM showers, but Brem gives photons and DIF gives electrons. While photon-induced EM showers are useful to comparison with electron showers from ν_e-CC, due to the similarities of the showers, DIF has the extra advantage of giving actual electron showers and it will be an ideal sample for comparison with ν_e-CC sample. We will discuss DiF in the next chapter.

This chapter discusses the bremsstrahlung EM shower extraction from cosmic muons in NOνA detectors and its usage as data driven method to validate the following aspects in the experiment:

1. $\nu_e - CC$ signal shower modelling in the form of data/MC comparison.
2. $\nu_e - CC$ particle identification algorithm cross check and validation.
3. $\nu_e - CC$ signal selection efficiency in the detector.

First analysis simulation files are used for the study. Dataset files definition of data and MC are:

- **Data**: "ynitin_prod_artdaq_fd_cosmic_prod3_subset10of12_isgoodrun"
- **MC**: "prod_artdaq_R17-03-09-prod3genie.d_fd_cry_full_v1"

The detector exposure used in files is:

- Data: 2288.9 seconds.
- MC: 2467.7 seconds

4.1 Bremsstrahlung Muons Track Preselection

In NOνA, the cosmic rate at Far detector is 148 kHz and therefore cosmic muons are abundant. In an event, which corresponds to 500 μsec, there are about 30 - 50 muon tracks for each event. Fig.4.1 shows cosmic activity, corresponding to a 5 ms exposure window, in the far detector.

The selection of bremsstrahlung EM shower event starts from selecting the muon tracks. Considering the aim of the study toward ν_e, the preselection cuts are designed such that the brem shower event should be close in mimicking a $\nu_e - CC$ EM shower. Therefore, the muons tracks are selected such that-

[1]Picture taken from http://nusoft.fnal.gov/nova/public/neutrinos.html

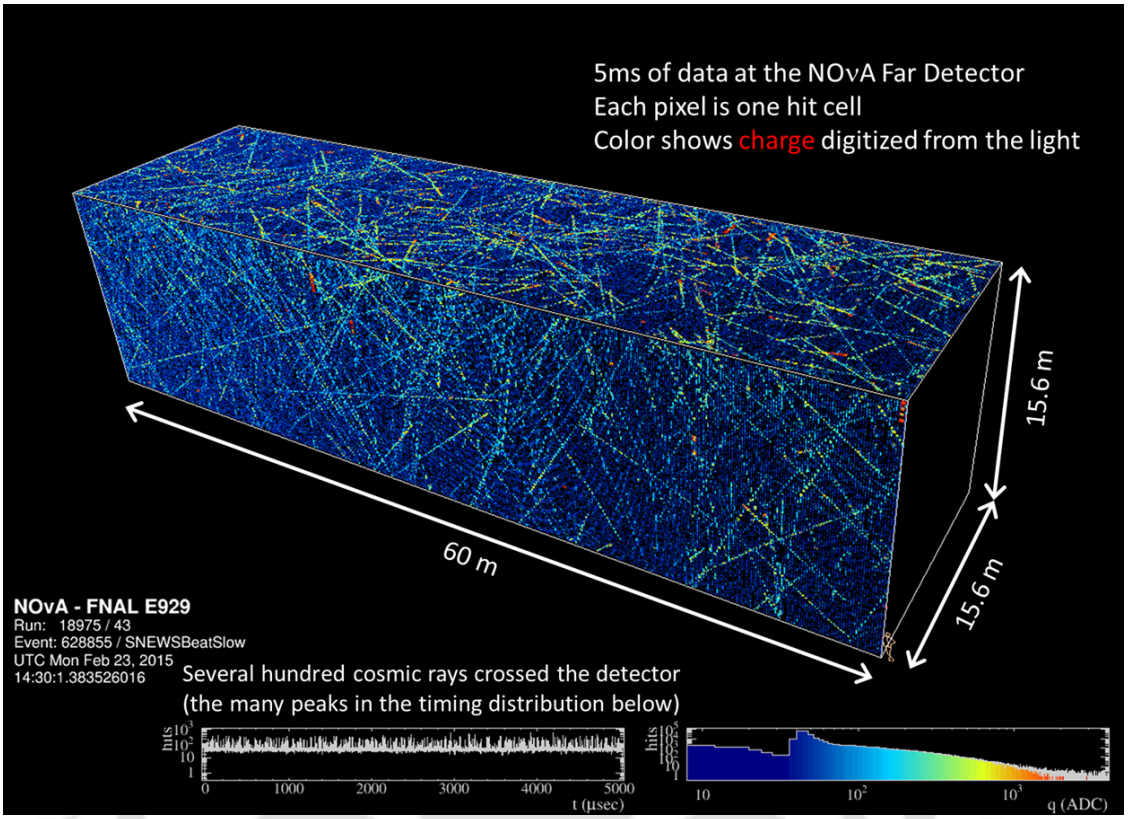

FIGURE 4.1: Cosmic activity in the NOνA far detector[1].

	MC	Data
Containment	2.3×10^7	3.4×10^7
$cos\theta_z$	3.4×10^6 (14.6 %)	4.6×10^6 (13.6 %)
No. of planes	2.8×10^6(12.0 %)	3.7×10^6 (11.0 %)
$E_{shw} > 0.5GeV$	7.6×10^4(0.3 %)	1.0×10^5 (0.3 %)

TABLE 4.1: Pre-selection cuts applied in the analysis.

1. The track should not be too steep compared to the beam direction. The beam direction is the Z- axis of the detector.

2. The track should be long enough to have good chance of generating Brem showers.

3. The Brem shower event energy should be close to energy of ν_e event.

Table 4.1 summarizes the cut flow at the pre-selection level which are similar to the selection criteria applied in the ν_e-CC appearance analysis.

In the detector, the muon is a minimum ionizing particle[2] (MIP) and the energy corresponding to a MIP in NOνA detector is 1.57 MeV/cm.

Considering the direction of muon track in the detector, we redefine the MIP as,

$$MIP = \frac{0.00157 \times \ (2 \times cell\ half\ width)}{track\ Z\ direction\ cosine}\ GeV/cm. \tag{4.1}$$

When a muon undergoes bremsstrahlung process, the amount of energy deposition along the way in the detector increases. Therefore, in the detector, the planes with EM shower hits will be having more energy than the planes with muon hits. As in the Fig.4.2, one can see that energy deposition in the detector planes increases for EM shower region. We have utilized this characteristic of bremsstrahlung radiation in extracting the EM shower which is described in the next section.

4.2 EM Shower Finding in NOνA Detector

An EM shower finding algorithm has been developed to find out the EM shower on a reconstructed muon track. The algorithm is used to identify the muons with possible EM shower using dE/dx information. As shown in Fig.4.2, we can see that a muon deposits energy as a minimum ionization particle (MIP) in the detector cells and in case of muon undergoing bremsstrahlung process, the additional EM shower hits can deposit much more energy in a small detector region with a overlapping muon energy. It is then possible to take the energy deposition of reconstructed cosmic tracks and look for the region where the energy is significantly greater than a MIP. More specifically, if we find 5 consecutive planes with energy greater than 2 MIP, we call it the shower start point; If we find 5 consecutive planes with energy in the range of 0.5 MIP to 1.5 MIP, we call it the shower end point. If we find both the shower start and end point, then the shower is identified as Brem. If we find only the shower start point and reach the end of track without an end point, the shower is identified as an electron from DiF. Showers with energy deposition

[2] http://pdg.lbl.gov/2005/reviews/passagerpp.pdf

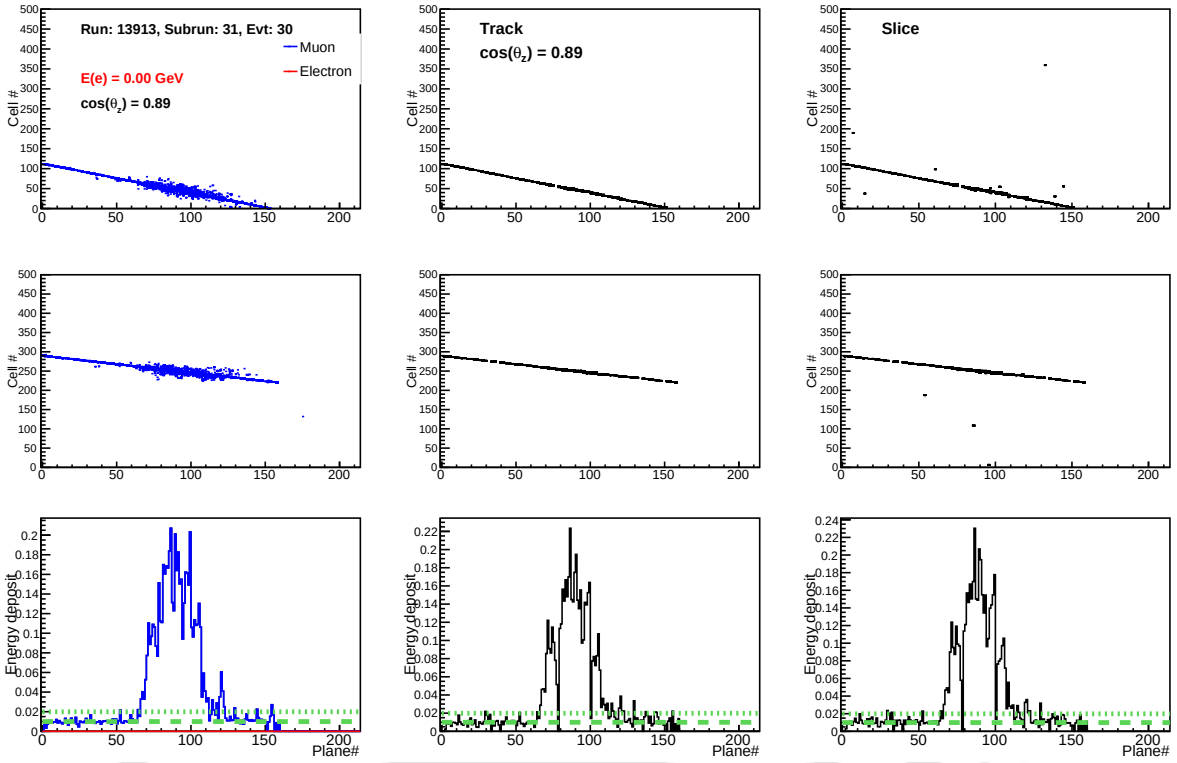

FIGURE 4.2: A brem event in FD before selection. The first column represent the simulated muon track event. The blue dots represent FLS hits (simulated hits) of muons. The middle column represents the reconstructed track and third column represent the cluster of the muon event. Green dash lines show the 1 MIP and 2 MIP mark in a plane.

passing the energy cuts are saved as raw digit objects. The Fig.4.3 shows a NOνA event display in which a cosmic moun is undergoing bremsstrahlung process.

Using the energy deposition characteristics, the shower region is identified. Fig.4.4 shows the shower region located and extracted by application of aforesaid shower finding steps and muon removal algorithm (describe in sections ahead) in an event.

Once the bremsstrahlung shower region is identified, the next step is to extract the EM shower hits and to remove the muon hits from the event. To achieve this, a new muon algorithm is developed which removes the muon hits from the track as described in the next section.

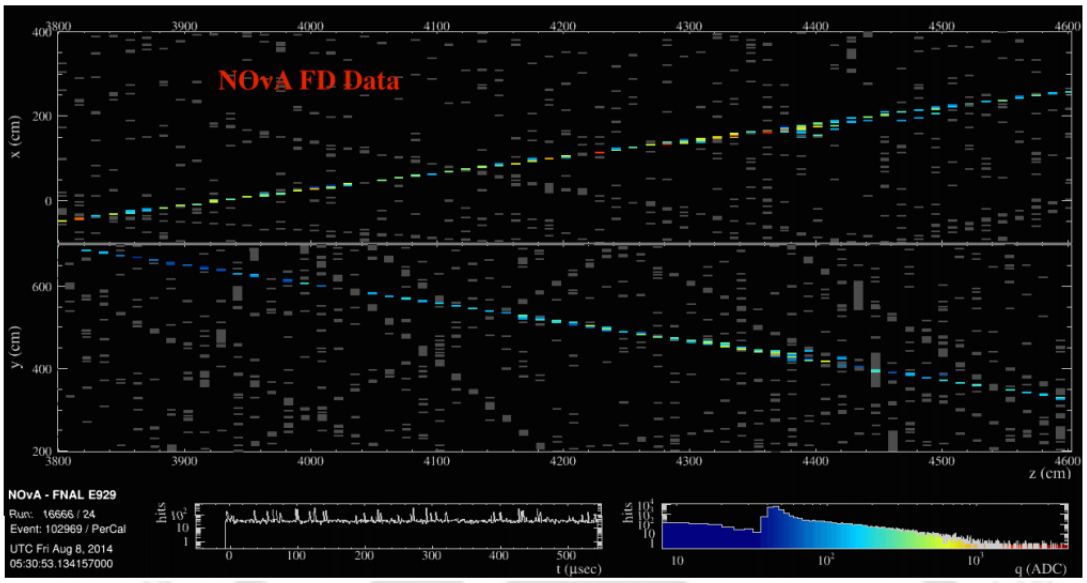

FIGURE 4.3: Cosmic muon with bremsstrahlung EM shower in NOνA far detector. The two view X-Z (up) and Y-Z (down) of detector are shown. The colored hits belong to the muon track with colors correspond to ADC value. Grey hits belong of other particles or noise. As is visible that around at 4400 cm on Z axis the energy deposition (number of colored hits increased) increased due to EM shower energy deposition.

4.3 Muon Removal Algorithm

A new muon remove algorithm for EM showers has been developed which is based on the Muon-Removal algorithm for charged current events [46]. The algorithm first looks at the slice where a cosmic EM shower is found. In the case of decay in flight (DiF) EM shower (discussed in next chapter), with a shower region defined by the shower-finding algorithm, the muon removal algorithm removes all hits outside that region and what left will be pure electron hits. In the case of bremsstrahlung showers one additional problem is that we have a muon track inside the EM shower region as well. Therefore the muon-removal algorithm should remove hits that belong to muon track corresponding to the energy of a MIP in the shower region. The cell hits which are consistent with a MIP on the muon track are removed by scaling down the ADC in the cells by weight $1 - MIP\ fraction$, where $MIP\ fraction$ is the muon energy fraction in the cell. If the cell hit on the track is found to have less energy consistent with a MIP upper range , then the next closest cell on the same plane is taken and is scaled down by weight that would be just enough to make the total hit removal in that plane consistent with a MIP. Fig.4.7 shows the fraction

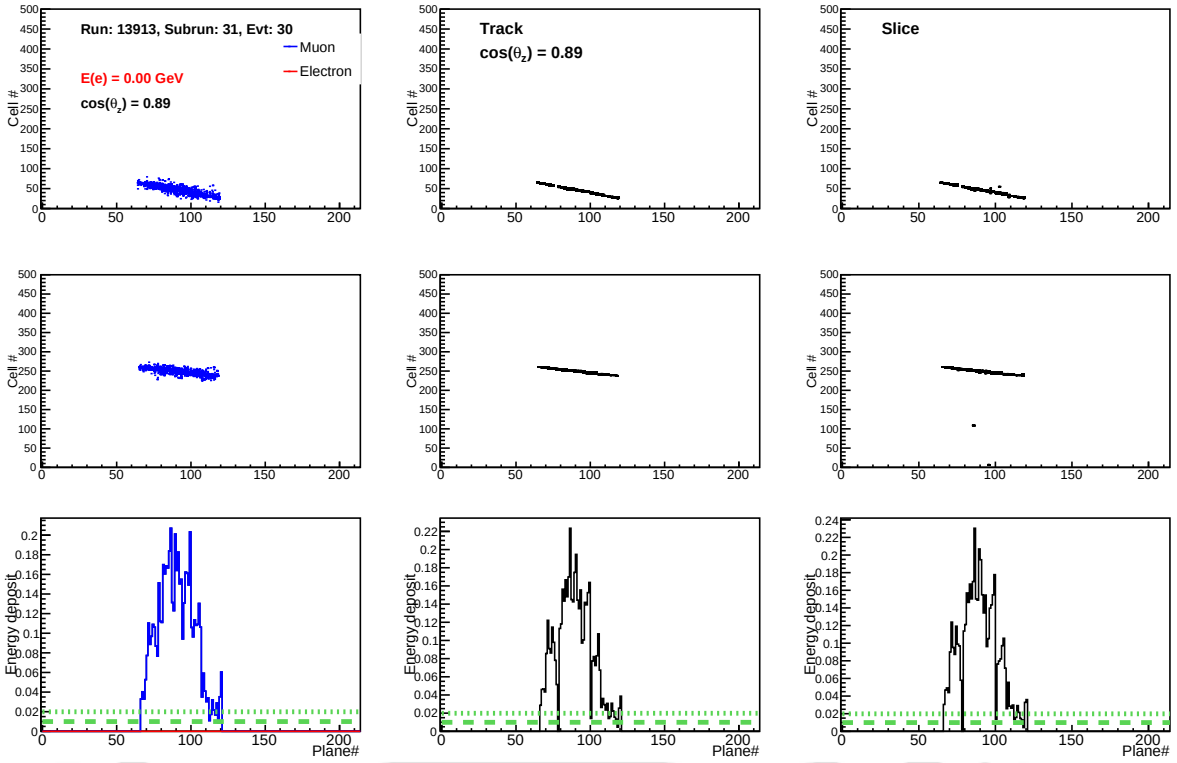

FIGURE 4.4: Extracted Brem hits in blue from muon track in Fig.4.2 after shower finding and muon removal is applied. Track (middle) and slice (right) hits are also shown. Green dash lines show the 1 MIP and 2 MIP mark in a plane.

of muon hits energy ($\sigma =\sim 7\%$) which are left behind after the muon removal algorithm is applied. Fig.4.8 shows the same event as in Fig.4.3 after muon hits are removed.

In Fig. [4.5, 4.6] are shown the vertex distributions Y vs X, Y vs Z, X vs Z, and X, Y, Z of Brem candidates in the detector.

4.4 Data/MC Comparison

4.4.1 EM Shower Reconstruction

The brem shower extracted hits are put through the standard reconstruction for ν_e analysis as described in the previous chapter. After shower reconstruction, the data and MC comparison variables are plotted, as shown in following sections.

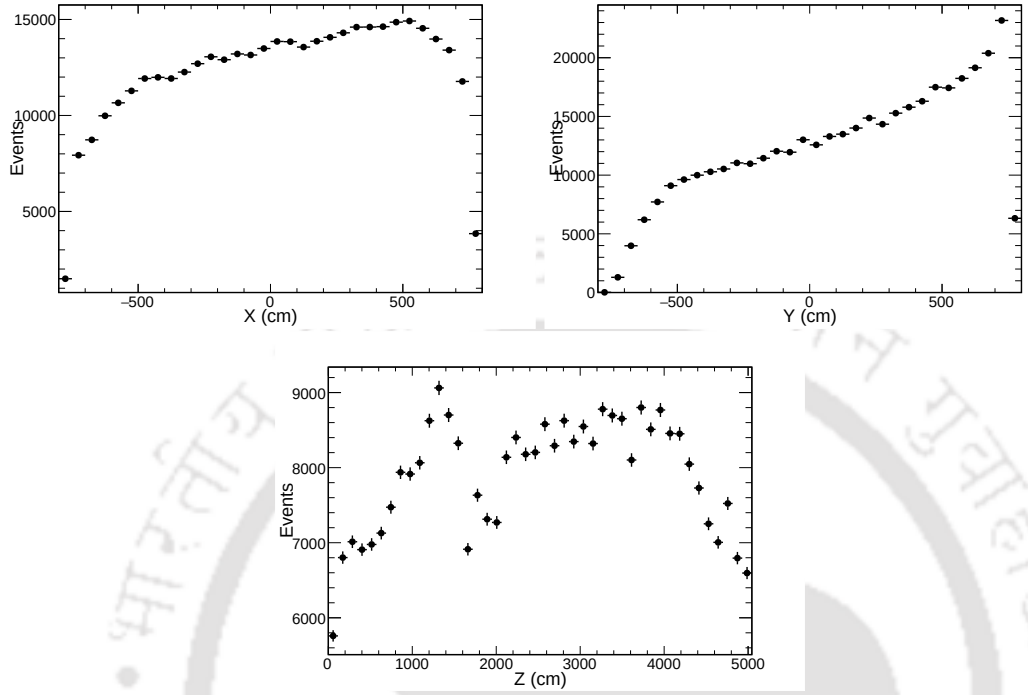

FIGURE 4.5: X, Y and Z distribution of Brem candidates in the far detector.

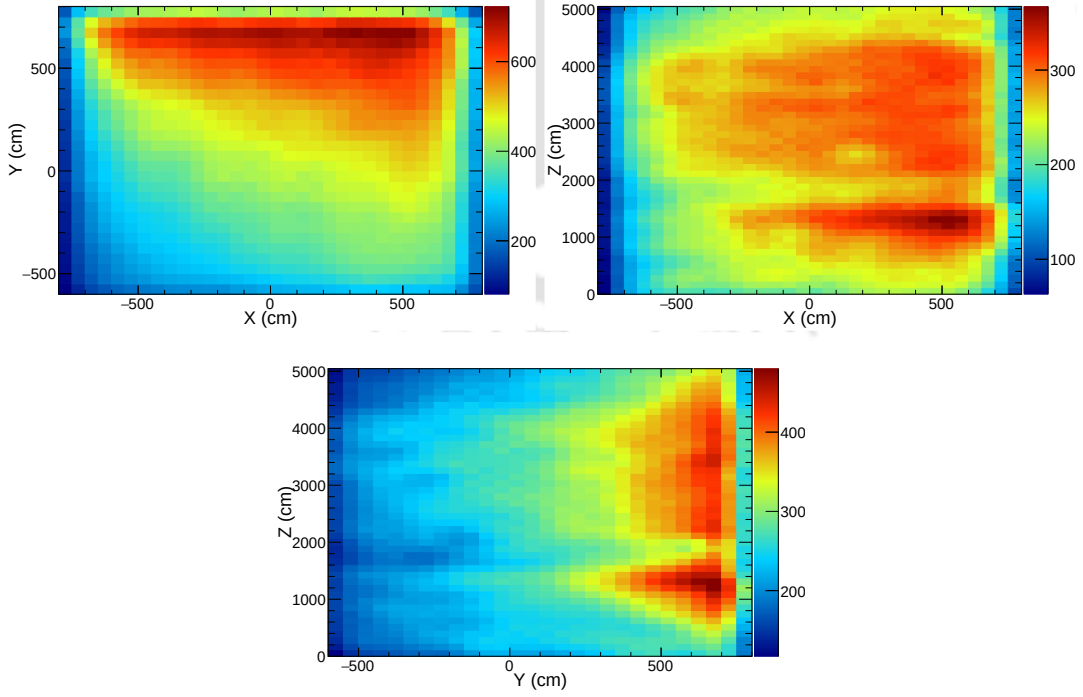

FIGURE 4.6: Y vs X, Y vs Z and X vs Z distribution of Brem candidates in the far detector.

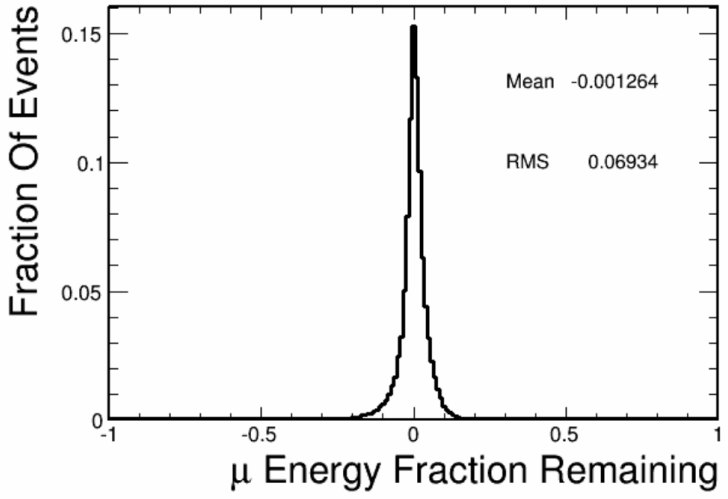

FIGURE 4.7: Fraction of muon hits left behind after muon removal. From [46].

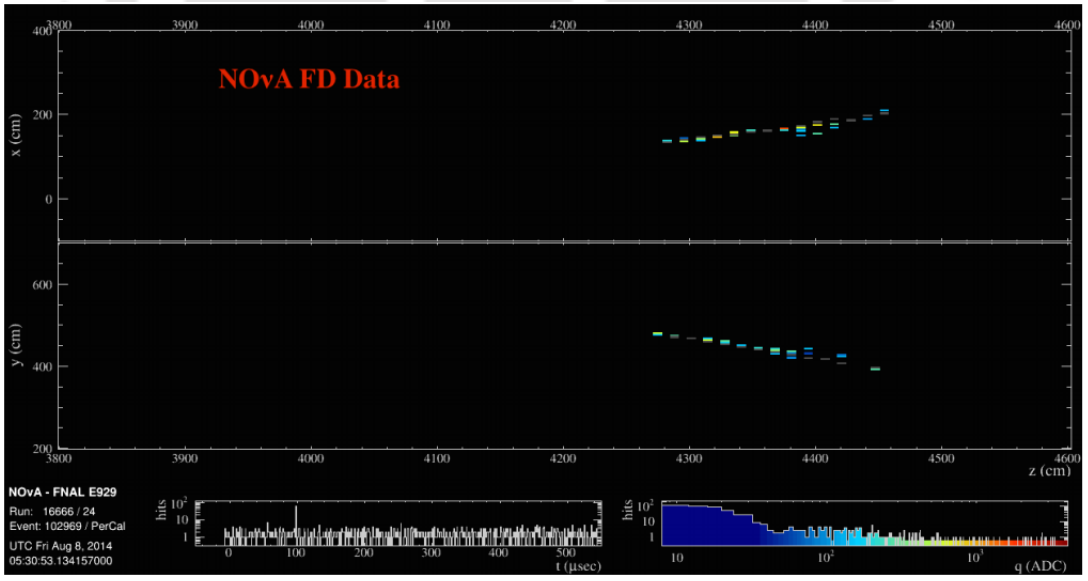

FIGURE 4.8: Extracted bremsstrahlung EM shower hits of muon of Fig.4.3 after shower finding and muon removal algorithms are applied.

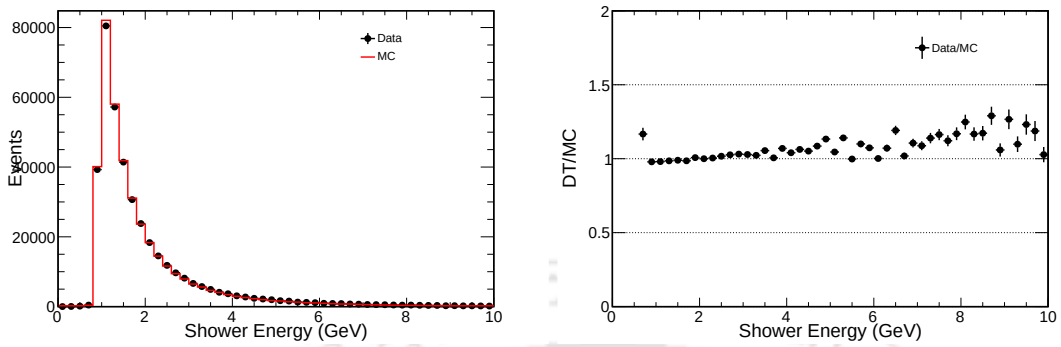

FIGURE 4.9: Energy distribution of Brem (left) and ratio of Data/MC as a function of shower energy (right).

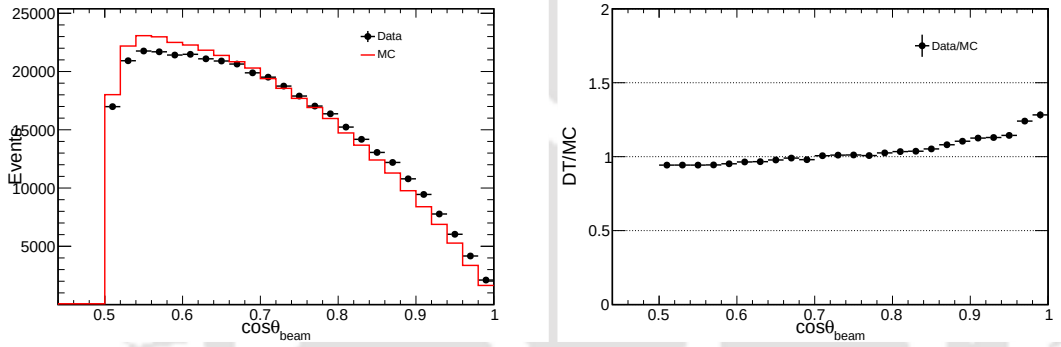

FIGURE 4.10: Angular distribution of Brem (left) and ratio of Data/MC as a function of shower angle with respect to beam direction (right).

4.4.2 Data and MC Comparison

This section contains the data and MC comparison plots for the bremsstrahlung EM shower. Fig.4.9 shows the reconstructed energy distribution of Brem EM shower in data and MC. It is clearly seen that data and MC are in very good agreement in the energy range (1 - 3 GeV) relevant to ν_e signal energy. However, a difference was observed in the data and MC for the angular distribution as shown in Fig.4.10. Fig.[4.11 - 4.13] shows the data and MC distribution and comparison for various shower variables, for which a very good agreement between data and MC is observed.

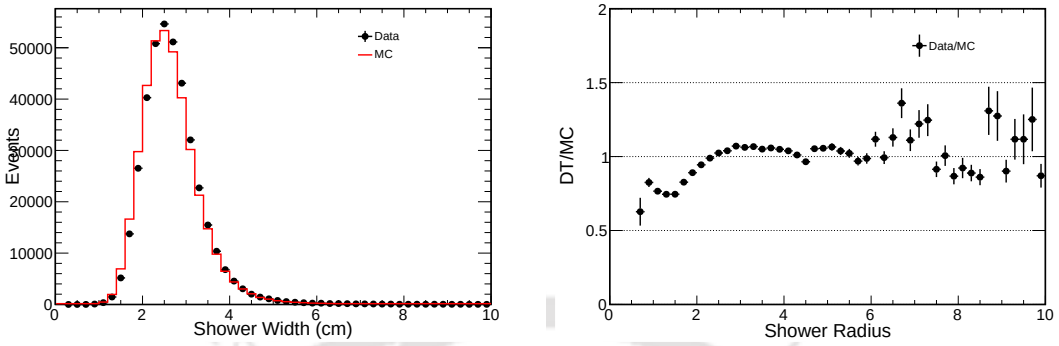

FIGURE 4.11: Bremsstrahlung shower radius distribution (left) and its data/MC comparison (right). The shower radius is defined as the average distance from shower cells to shower core weighted by cell energy.

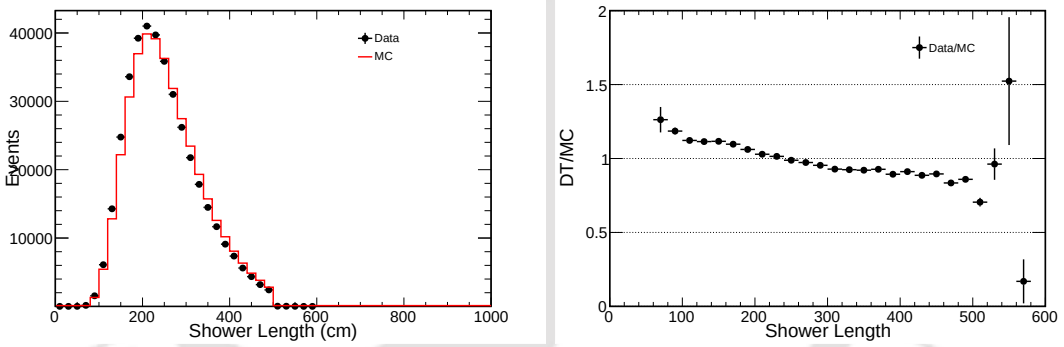

FIGURE 4.12: Length distribution of Brem (left) and and its data/MC comparison (right).

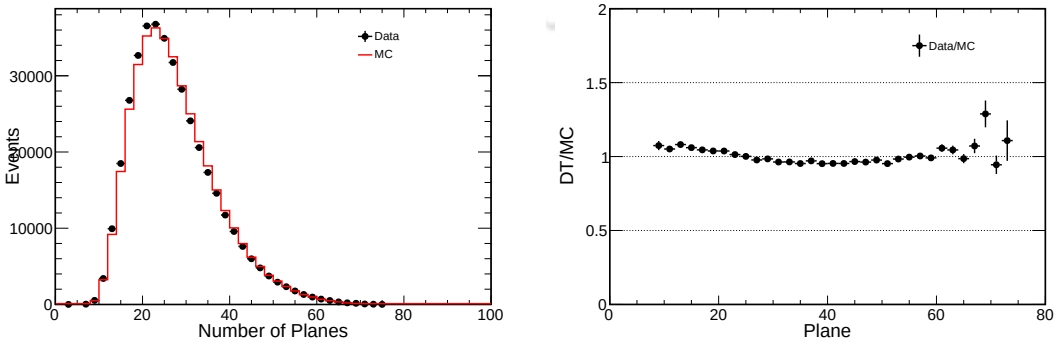

FIGURE 4.13: Planes distribution of Brem (left) and and its data/MC comparison (right).

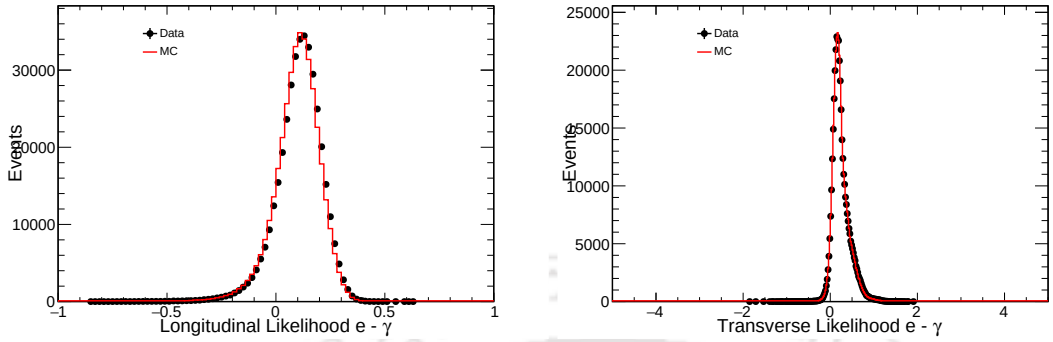

FIGURE 4.14: Longitudinal (left) and transverse (right) likelihood differences between electron and γ particle hypothesis using Brem sample. Brem shower is clearly identified as electron type than γ shower in transverse and in longitudinal direction.

4.4.3 Likelihood Differences for Brem EM Shower

The likelihood based particle identification (LID) is an artificial neural based algorithm in which various likelihoods are fed as an input to select ν_e-CC signal. In this section, likelihood differences are shown for various particle hypothesis. Likelihoods are constructed in two directions: longitudinal likelihood which is along the direction of shower and transverse likelihood which is along the transverse direction of the shower. After that, likelihood differences are plotted. For an event, if likelihood difference (say between two particles a and b) is toward the positive axis, then the event is more likely to be b type or vice versa.

Fig.4.14 shows the plot of likelihood difference between e and γ and we can see that the peak in the plot is toward the positive side, which indicates that Brem events are more likely to be electron type than γ type. This is the first indication of similarity between Brem EM shower and ν_e-CC electron EM shower.

Fig.4.15 shows the plot of likelihood difference between e and μ and we can see that the peak in the plot is clearly toward the positive side, which indicates that Brem events are more likely to be electron type than μ type. This indicates that the Brem sample, extracted from μ, has clear distinction from its source, i.e., muon.

Similarly Fig.4.16 shows the plot of likelihood difference between e and π^0 and we can see that the peak in the plot is toward the positive side in transverse likelihood than longitudinal likelihood. This is obvious due to the fact that $\pi^0 \rightarrow \gamma\gamma$ and it has two

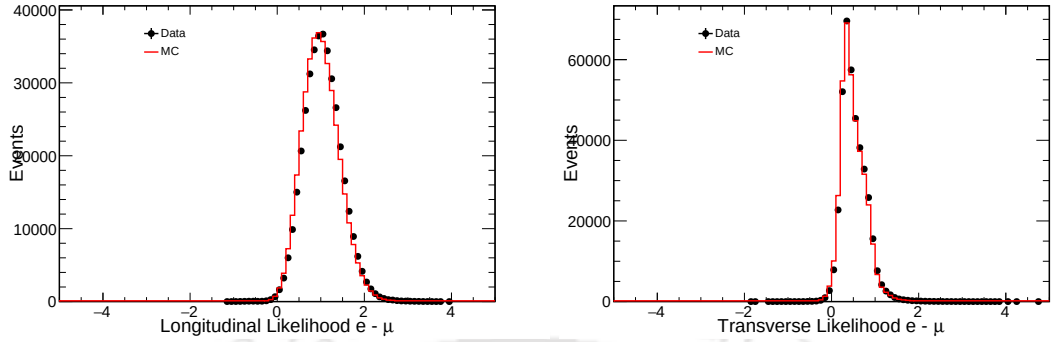

FIGURE 4.15: Longitudinal (left) and transverse (right) likelihood differences between electron and μ particle hypothesis. Brem shower is clearly identified as electron type than μ in transverse and in longitudinal direction .

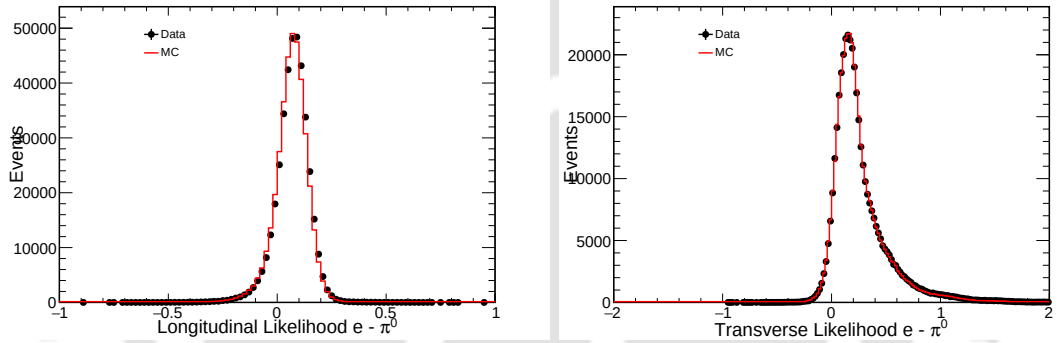

FIGURE 4.16: Longitudinal (left) and transverse (right) likelihood difference between electron and π^0 particle hypothesis. Brem shower is more identified as electron type than π^0 shower in transverse direction. However in longitudinal direction, the separation is less.

shower prongs and because of two shower prongs, the shower is wider (ν_e-CC electron has single shower and so is Brem) in transverse direction and thereby the distinction is more in comparison with single shower produced in ν_e-CC interaction.

4.4.4 Particle-ID Classification for Brem EM Shower

As mentioned earlier, the particle-ID LID, takes the input of aforesaid likelihood difference to classify a particle. Fig.4.17 shows the likelihood based identifier (LID) output using Brem sample. As can be seen very clearly that it peaks toward the high signal region (around LID value 1) which implies that our sample is e-type.

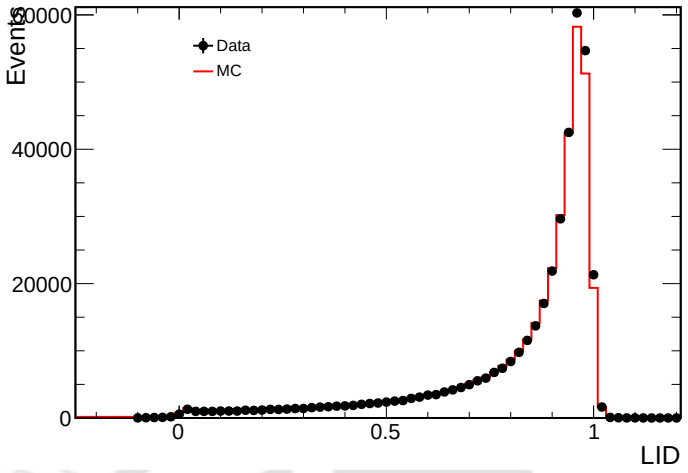

FIGURE 4.17: LID output. Most of the brem showers are identified as ν_e showers.

We also process our sample through another particle-ID, LEM and CVN and Fig.4.18 and Fig.4.19 show the output of the LEM and CVN respectively. Since these particle IDs are made for beam events, unlike LID, LEM and CVN use the harder cuts on directionality and energy of ν_e events. Due to this reason, the Brem sample events are classified more as background type by LEM and CVN.

To address the above issue, we have developed a reweighing method to make the Brem shower sample equivalent to ν_e shower by constructing matrix, from a bin by bin comparison of brem and ν_e-CC energy and angle distribution (reweighing discussed in sections ahead). We will also see how after applying reweighing method, LEM starts to classify more Brem events (Fig.4.18) as e- type.

4.5 Reweighing Method

As seen in the previous section, LEM classifies most of the brem events as non-e type. This section includes a reweighing (RW) method which basically quantifies the difference between the brem and ν_e sample which is further used to make the brem sample more similar to the ν_e-CC sample.

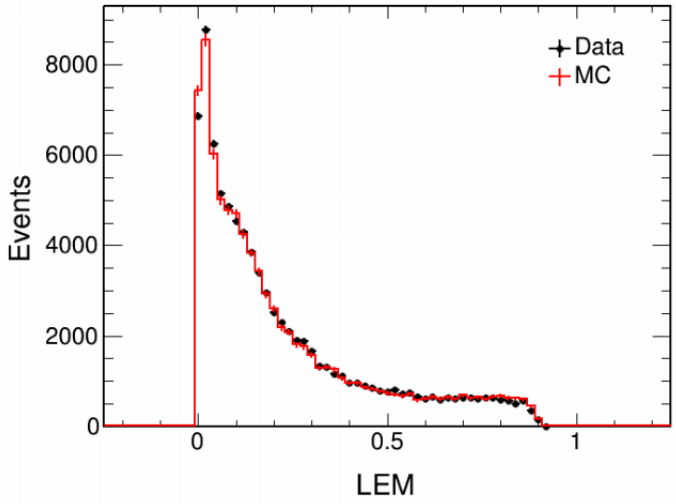

FIGURE 4.18: LEM output for brem events. Most of the events are classified as background type.

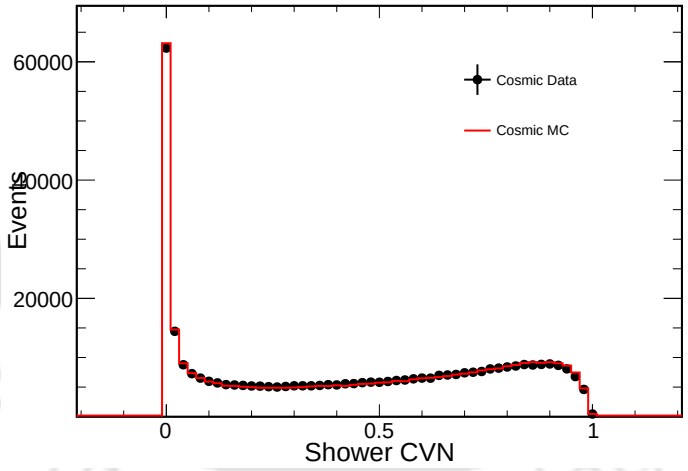

FIGURE 4.19: CVN output for brem events. Most of the events are classified as background type.

4.5.1 Brem EM Shower vs Beam ν_e EM Shower

A ν_e re-weigh method has been developed to make the cosmic brem more like a shower from ν_e-CC interaction. The difference between cosmic induced EM showers and beam ν_e-CC induced EM showers mostly come from difference between two shower energy and angle distributions. As shown in Fig.4.20[3], the EM shower from ν_e-CC shower peaks at ∼2

[3]The color coding in the plots: red MC, black Data, blue ν_e-CC

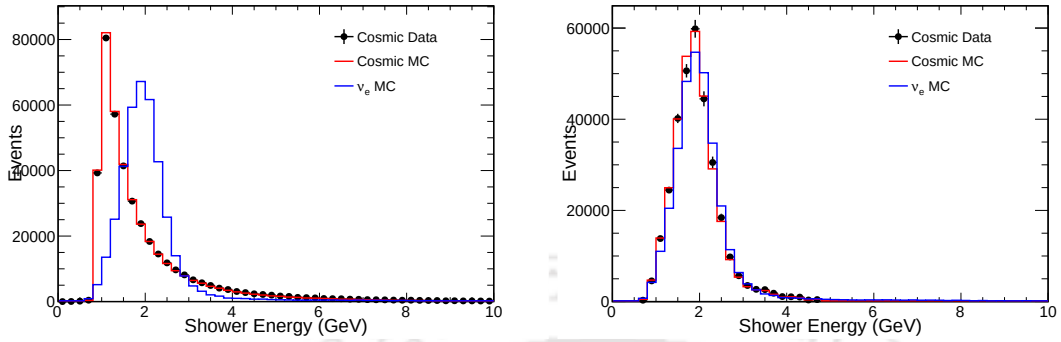

FIGURE 4.20: Energy of Brem EM shower vs ν_e-CC EM shower before (left) and after (right) reweighing method is applied.

GeV, while the cosmic EM shower energy distribution peaks ∼1.5 GeV. This difference is first corrected by the 0.5 GeV cut applied at shower selection, but the difference in energy spectrum shape still exists. The angles of beam EM showers are around the beam direction ($\sim 0^\circ$ with respect to Z axis)[4], while cosmic EM showers originate mostly form particles coming in perpendicular direction in the detector.

We use a 2-D matrix to re-weigh the energy and angle ($\cos Z$) distribution of cosmic EM showers to ν_e events. Swap[5] MC is used to get the ν_e spectrum by selecting electron neutrino according to truth, and applying LID cut at 0.7. After reweighing, both LID and LEM show more events peaking at signal region with significantly higher efficiencies. It is also worth noticing that LEM has better data MC agreement after re-weight, indicating that the previous discrepancy may come from cosmic angle difference in data and MC. In the following plots the Brem sample is reweighed in energy and angle w.r.t ν_e-CC energy and angle. From Fig.[4.20-4.26] it is quite clear that, after reweighing the Brem sample mimics the ν_e sample better.

We can see that in Fig.4.22, the shower radius of both the sample are almost same. Even after applying reweighing method the radius distribution does not change much. This gives the confidence that reweighing method is working as it is supposed to. Also we can make one more inference that Brem and ν_e EM showers have same radius despite of

[4]Remember, Z axis is along beam axis which is horizontal
[5]In swap MC, ν_μ events are swapped by ν_e events so as to get high ν_e statistics.

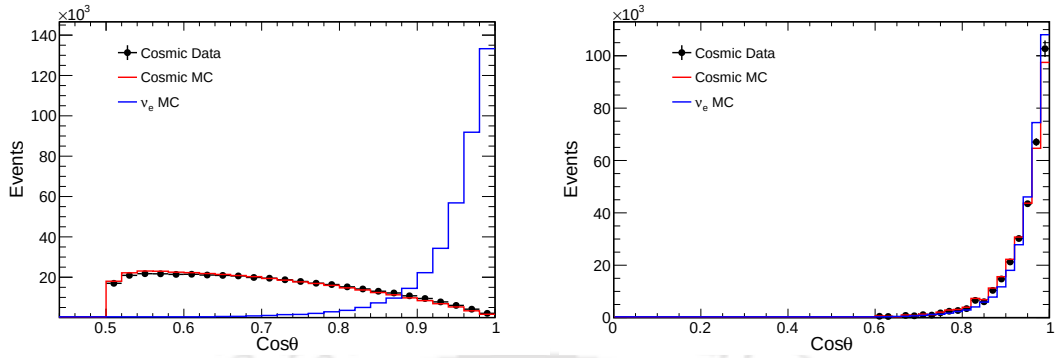

FIGURE 4.21: Angular distribution of Brem EM shower vs ν_e-CC EM shower before (left) and after (right) reweighing method is applied..

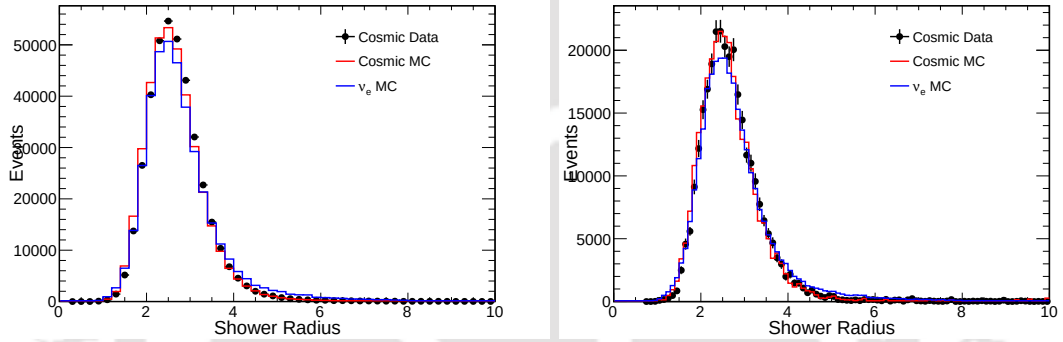

FIGURE 4.22: Radius of Brem EM shower vs ν_e-CC EM shower before (left) and after (right) reweighing method is applied. The shower radius is defined as the average distance from shower cells to shower core weighted by cell energy.

difference in energy as the difference in energy should show up in length of showers and we expect that the length of the two showers will be different.

As can be seen in Fig.4.23, that length of two shower samples are indeed different due to energy difference.

We can also see in Fig.4.23, there is almost no effect of reweighing in the shower length because shower length is independent of the angle. It is calculated by combining the 2-D hits in two views of the detector to form the 3-D event. This way the absolute length becomes independent of the angles with respect to Z axis in either view, as a long track may have shorter projection in one view but it will be larger in the other view or vice versa.

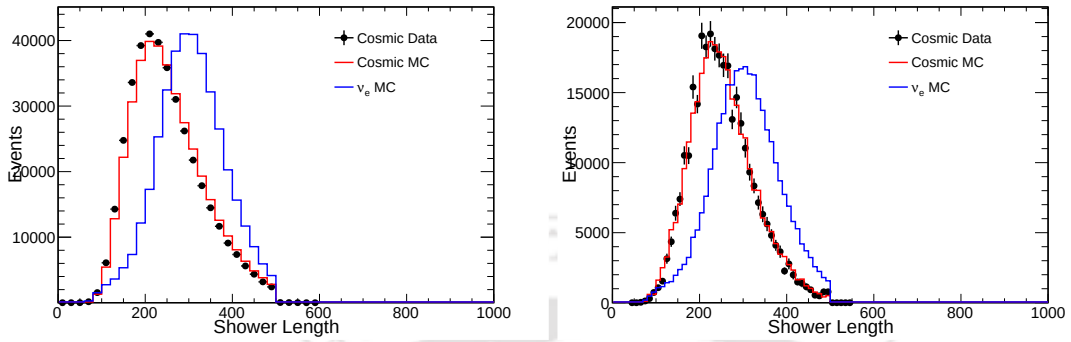

FIGURE 4.23: Length of Brem EM shower vs ν_e-CC EM shower before (left) and after (right) reweighing method is applied.

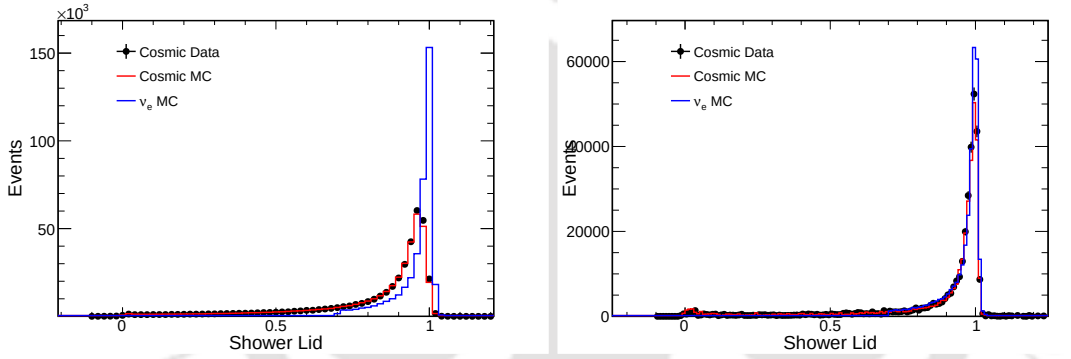

FIGURE 4.24: LID output of Brem EM shower vs ν_e-CC EM shower before (left) and after (right) reweighing method is applied.

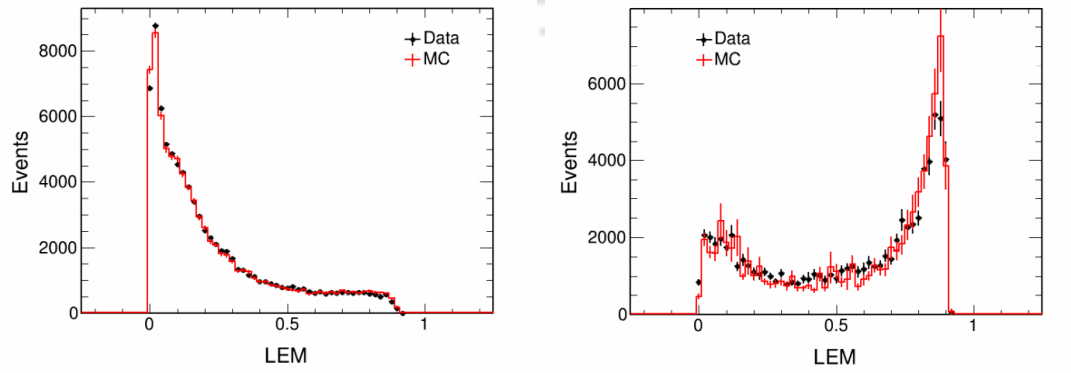

FIGURE 4.25: LEM output of brem EM shower vs ν_e-CC EM shower before (left) and after (right) reweighing method is applied.

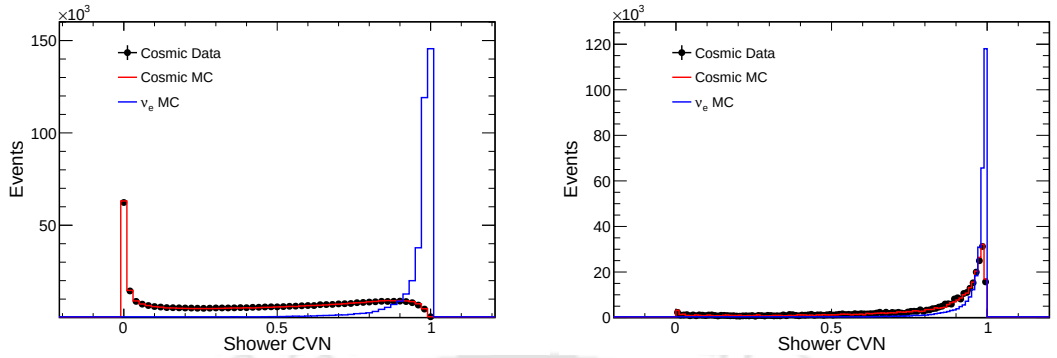

FIGURE 4.26: CVN output of brem EM shower vs ν_e-CC EM shower before (left) and after (right) reweighing method is applied.

4.6 Signal Selection Efficiency in the Far Detector

Brem EM shower has been used to determine the systematics associated with the performance of the detector. Efficiencies are calculated as number of showers passing PID cuts (high signal PID region) divided by all showers selected as function of signal position in X, Y and Z to check and possible calibration effects. This way any possible effect which is not picked up by simulation in such big detectors will be spotted out. Since all of the detector regions are same so ideally the signal selection efficiency in the detector (except edges, where reconstruction if not perfect) should be a flat distribution in X, Y and Z directions.

Overall, in data and MC, we see a relatively flat efficiency distribution across the detector, with good data and MC agreement. The extent to which the efficiencies from data and MC do not agree motivates a systematic error on the predicted electron neutrino signal efficiency for the electron neutrino appearance analysis [92]. In the plots shown in Fig.[4.27-4.50], we have plotted the Brem shower selection efficiency for the following values of particle-ID for high $\nu - e$-CC signal regions:

- LID and CVN >0.7 with and without reweighing.
- LID and CVN >0.9 with and without reweighing.

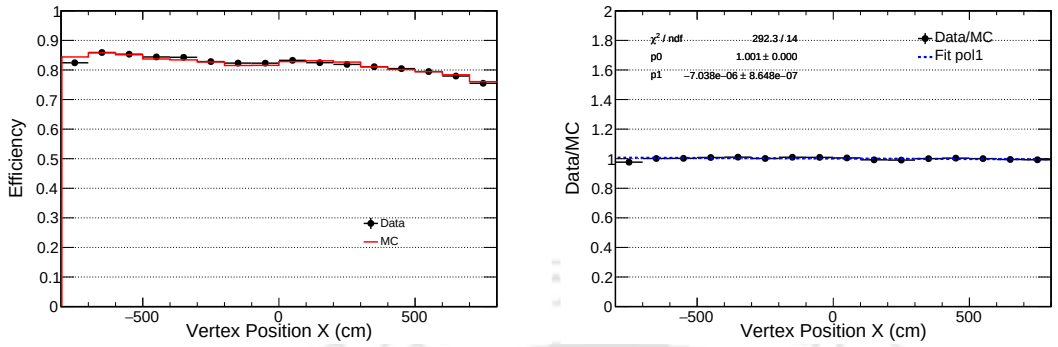

FIGURE 4.27: Signal selection efficiency in X direction (left) and data/MC comparison (right) for LID>0.7.

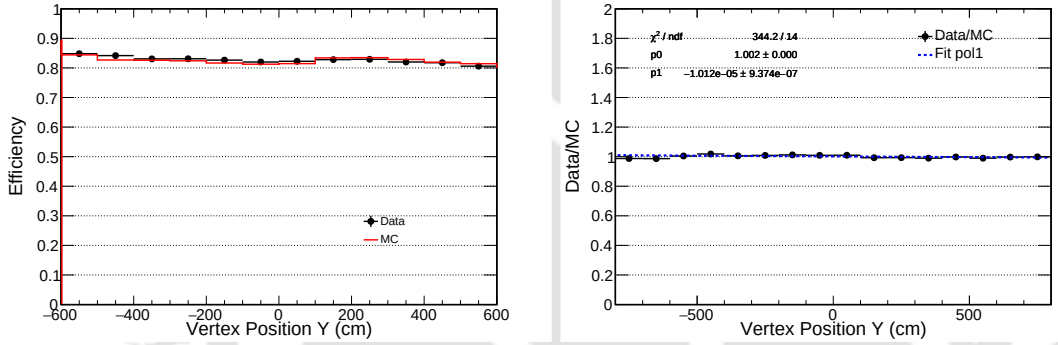

FIGURE 4.28: Signal selection efficiency in Y direction (left) and data/MC comparison (right) for LID>0.7.

4.6.1 Signal Selection Efficiency for LID

4.6.1.1 LID >0.7

In this subsection selection efficiencies are plotted as a function of X, Y and Z axis for Brem shower having LID value > 0.7.

In Fig.4.27 and Fig.4.28 and Fig.4.29, the selection efficiency in X and Y direction, we can see that distribution is overall flat and any slope or distortion (other than ideally flat) in detector is well captured by simulations. The data and simulation in this case, overall agree well within 1%.

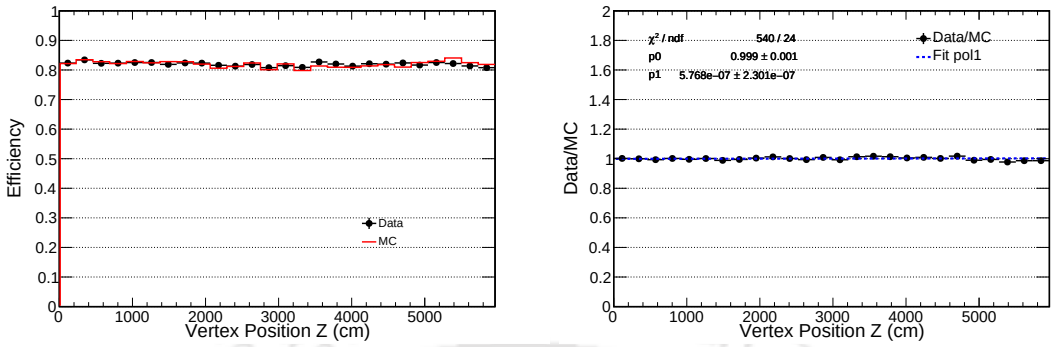

FIGURE 4.29: Signal selection efficiency in Z direction (left) and data/MC comparison (right) for LID>0.7.

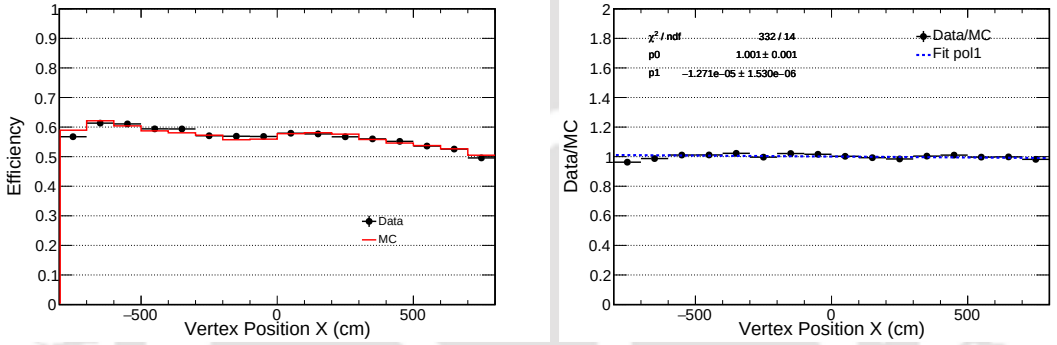

FIGURE 4.30: Signal selection efficiency in X direction (left) and data/MC comparison (right) for LID>0.9.

4.6.1.2 LID >0.9

In this subsection, selection efficiencies are plotted as a function of X, Y and Z directions in the detector, for brem shower having PID value > 0.9.

In Fig.4.30 and Fig.4.31 and Fig.4.32 shown the selection efficiency in X, Y, and Z direction, we can see that distribution is overall flat and any slope or distortion (other than ideally flat) in detector is well captured by simulations. The data and simulation in this case, overall agree well within 1%

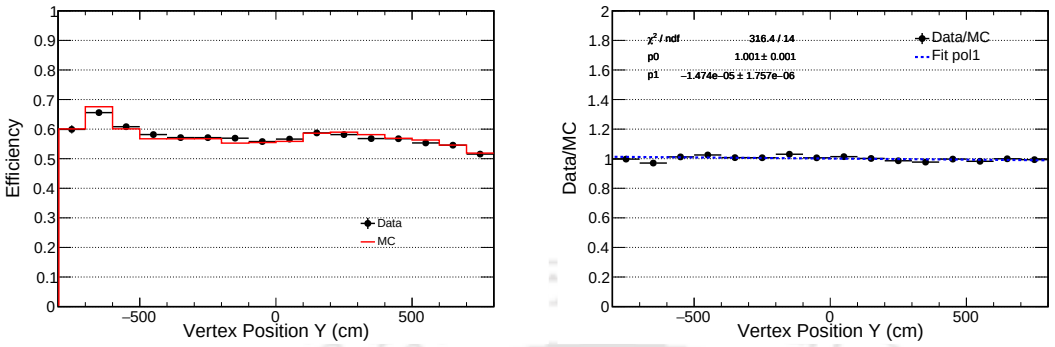

FIGURE 4.31: Signal selection efficiency in Y direction (left) and data/MC comparison (right) LID>0.9.

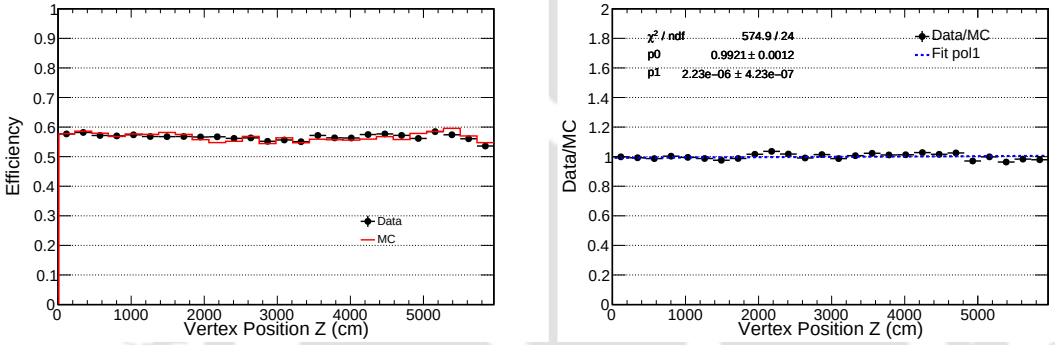

FIGURE 4.32: Signal selection efficiency in Z direction (left) and data/MC comparison (right) LID>0.9.

4.6.1.3 Reweighed Sample LID >0.7

In this subsection, selection efficiencies are plotted as a function of X, Y and Z direction in the detector for Brem shower after reweighing and having LID value > 0.7.

In Fig.4.33 and Fig.4.34 and Fig.4.35, the selection efficiency in X, Y and Z directions, we can see that distribution is overall flat and any slope or distortion (other than ideally flat) in detector is well captured by simulations. The data and simulation in this case, overall agree well within 4%.

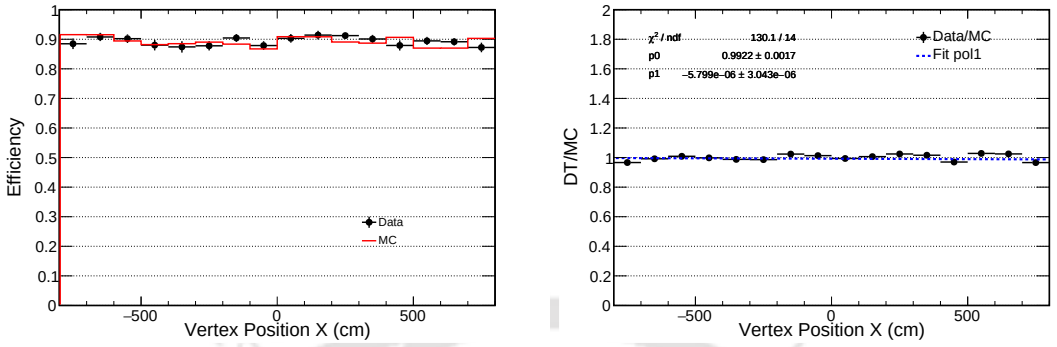

FIGURE 4.33: Signal selection efficiency in X direction (left) and data/MC comparison (right) of reweighed sample LID>0.7.

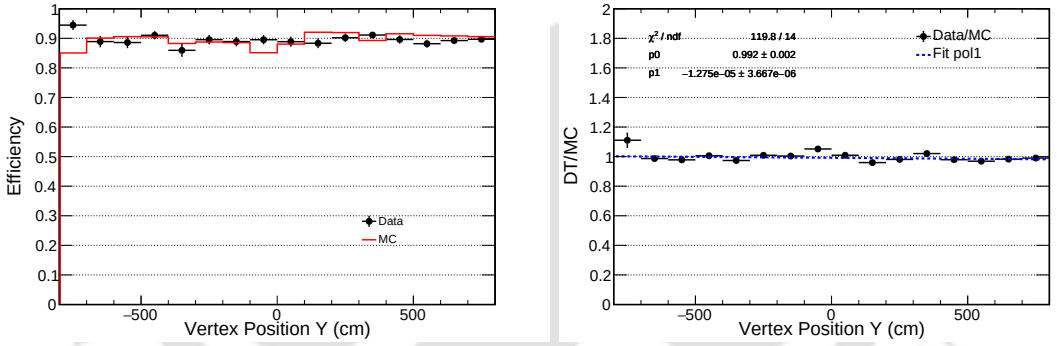

FIGURE 4.34: Signal selection efficiency in Y direction (left) and data/MC comparison (right) of reweighed sample LID>0.7.

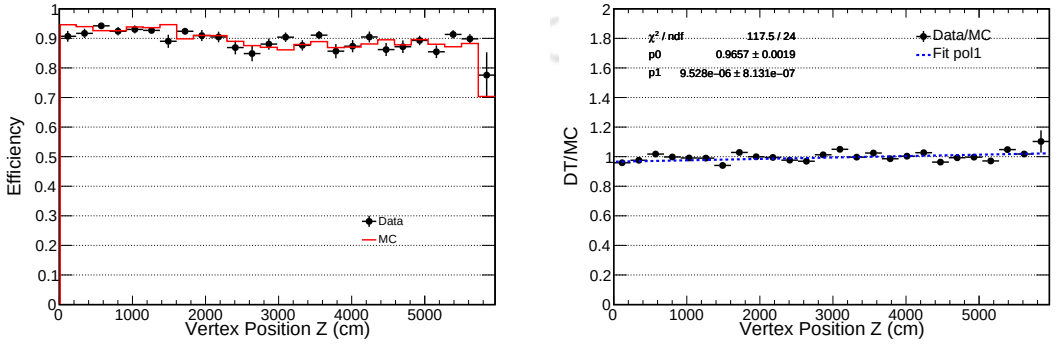

FIGURE 4.35: Signal selection efficiency in Z direction (left) and data/MC comparison (right) of reweighed sample LID>0.7.

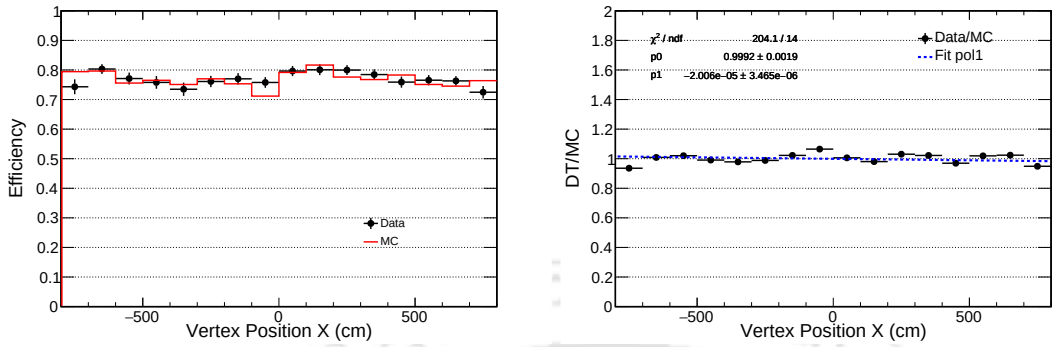

FIGURE 4.36: Signal efficiency in X direction (left) and data/MC comparison (right) of reweighed sample LID>0.9.

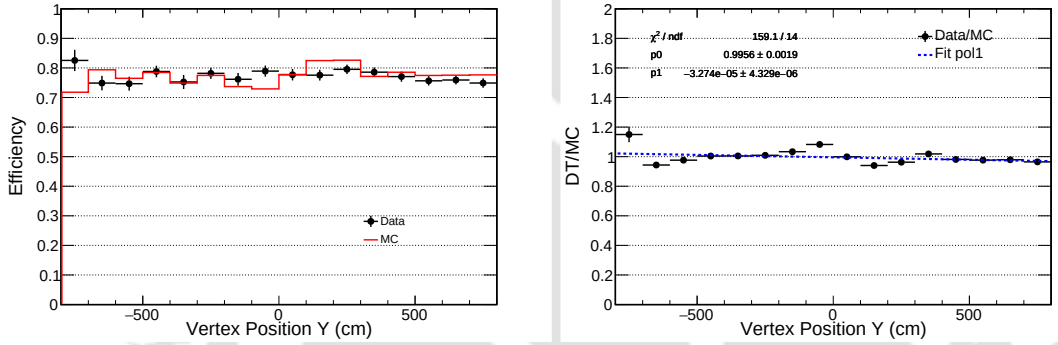

FIGURE 4.37: Signal efficiency in Y direction (left) and data/MC comparison (right) of reweighed sample LID>0.9.

4.6.1.4 Reweighed Sample LID >0.9

In this subsection, selection efficiencies are plotted as a function of X, Y and Z direction in the detector for Brem shower after reweighing having LID value > 0.9.

In Fig.4.36 and Fig.4.37 and Fig.4.38, the selection efficiency in X, Y and Z directions, we can see that distribution is overall flat and any slope or distortion (other than ideally flat) in detector is well captured by simulations. The data and simulation in this case, overall agree well within 2%.

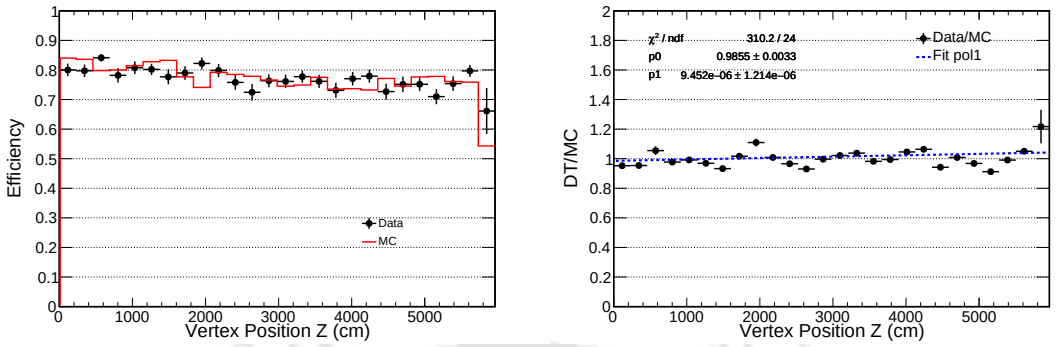

FIGURE 4.38: Signal efficiency in Z direction (left) and data/MC comparison (right) of reweighed sample LID>0.9.

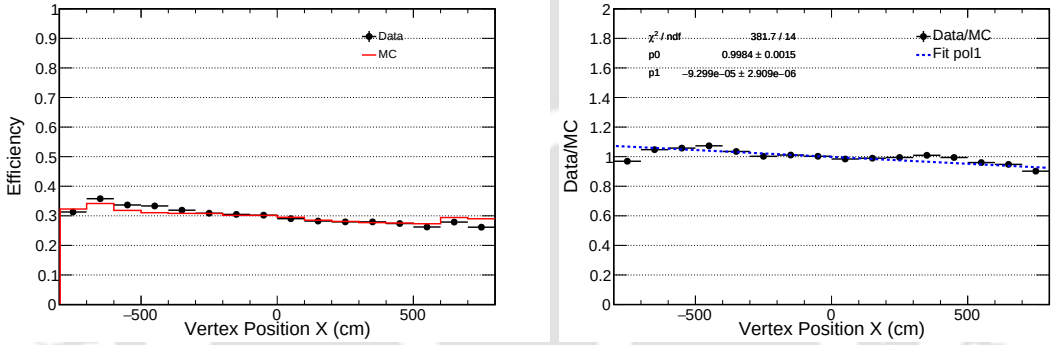

FIGURE 4.39: Signal selection efficiency in X direction (left) and data/MC comparison (right) for CVN >0.7.

4.6.2 Signal Selection Efficiency for CVN

4.6.2.1 CVN >0.7

In this subsection, selection efficiencies are plotted as a function of X, Y and Z direction in the detector for Brem shower having CVN value > 0.7.

As can be seen from Fig.4.40 and Fig.4.41, the signal selection efficiency in Y and Z direction for CVN is quite flat. In Fig.4.39 the selection efficiency in X direction has some slope but that is well modelled in simulations. Overall the data and MC agrees well within 4%.

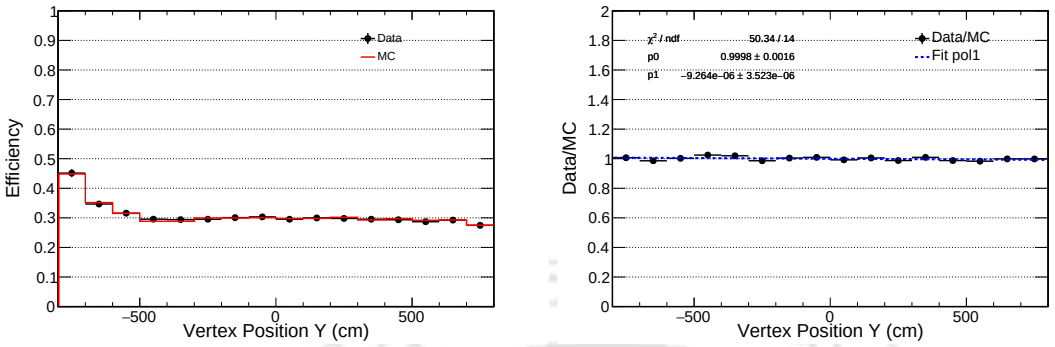

FIGURE 4.40: Signal selection efficiency in Y direction (left) and data/MC comparison (right) for CVN >0.7.

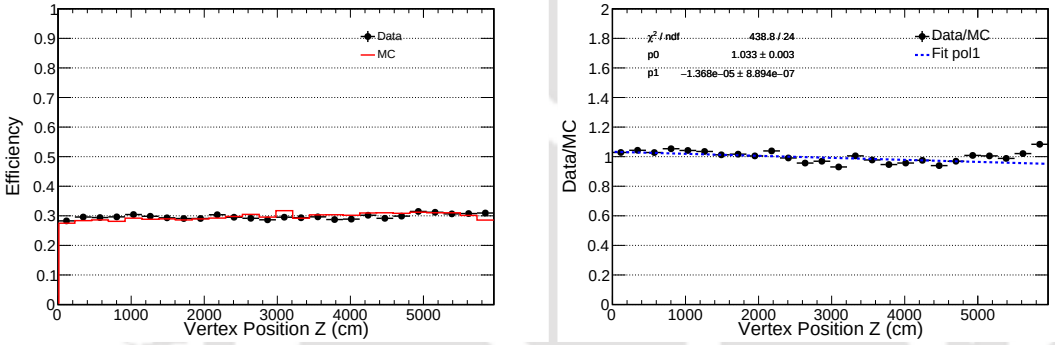

FIGURE 4.41: Signal selection efficiency in Z direction (left) and data/MC comparison (right) for CVN >0.7.

4.6.2.2 CVN >0.9

In this section selection efficiencies are plotted as a function of X, Y and Z direction in the detector for brem shower having CVN value > 0.9.

As can be seen from Fig.4.43 and Fig.4.44, the signal selection efficiency in Y and Z direction for CVN is quite flat. In Fig.4.42 the selection efficiency in X direction has some slope but that is well modelled in simulations. Overall the data and MC agrees well within 3%.

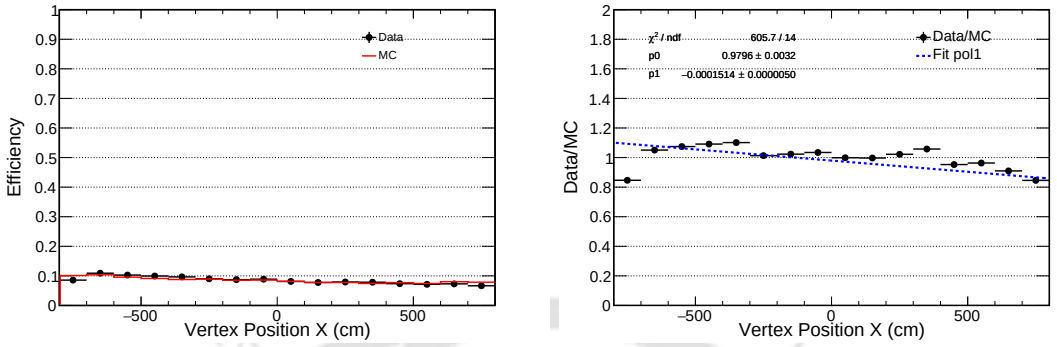

FIGURE 4.42: Signal selection efficiency in X direction (left) and data/MC comparison (right) for CVN >0.9.

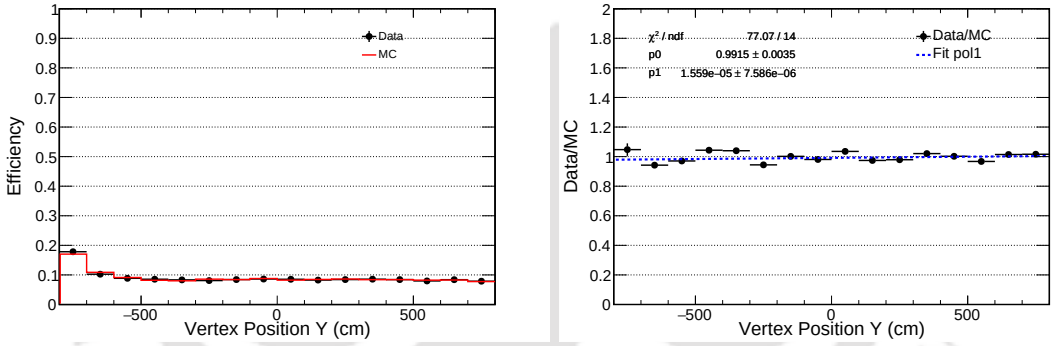

FIGURE 4.43: Signal selection efficiency in Y direction (left) and data/MC comparison (right) for CVN >0.9.

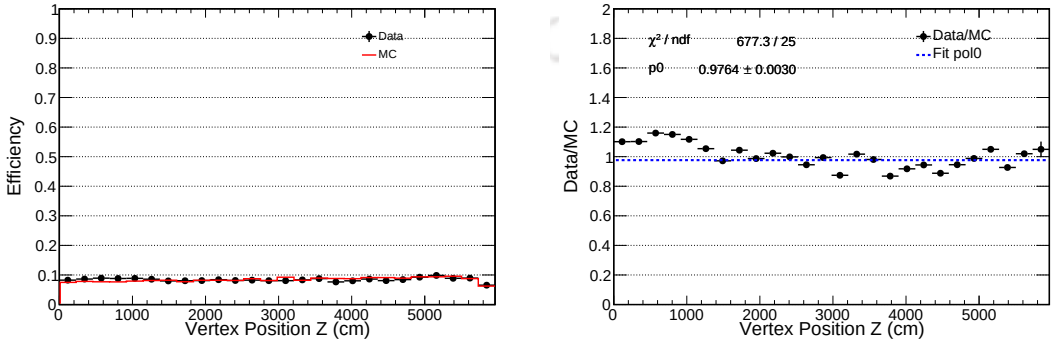

FIGURE 4.44: Signal selection efficiency in Z direction (left) and data/MC comparison (right) for CVN >0.9.

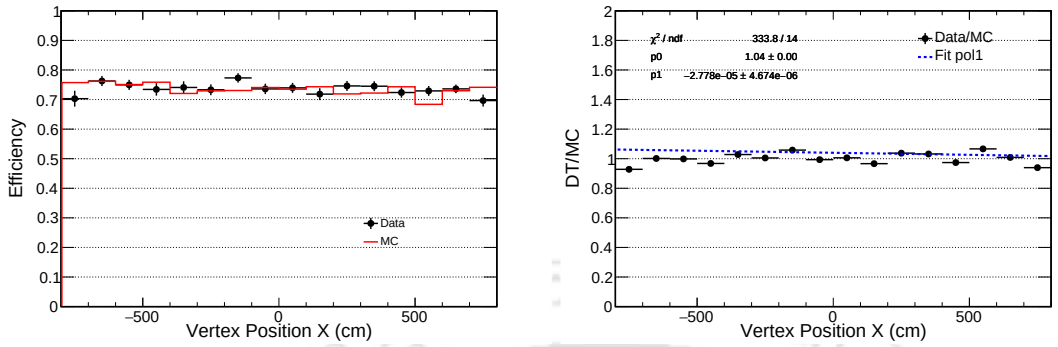

FIGURE 4.45: Signal selection efficiency in X direction (left) and data/MC comparison (right) of reweighed sample for CVN >0.7.

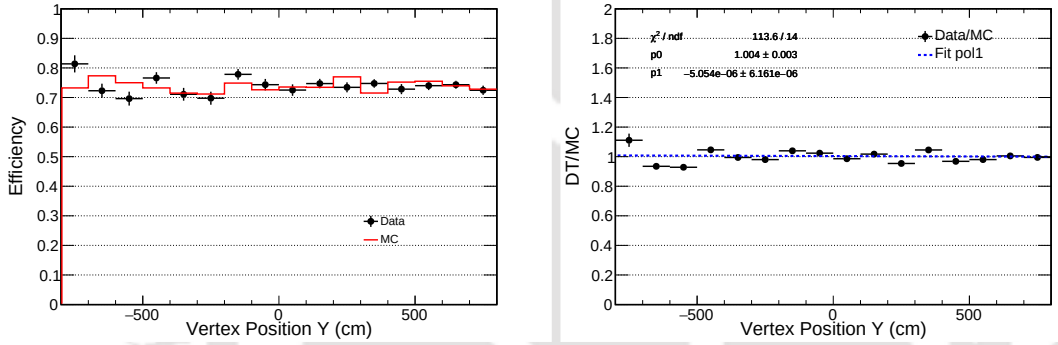

FIGURE 4.46: Signal selection efficiency in Y direction (left) and data/MC comparison (right) of reweighed sample for CVN >0.7.

4.6.2.3 Reweighed Sample CVN >0.7

In this subsection selection efficiencies are plotted as a function of X, Y and Z direction in the detector for Brem shower after reweighing having CVN value > 0.7.

As can be seen from Fig.4.45, Fig.4.46 and Fig.4.47, the signal selection efficiency in X, Y and Z direction for CVN is quite flat. Overall the data and MC agrees well within 4%.

4.6.2.4 Reweighed Sample CVN >0.9

In this subsection selection efficiencies are plotted as a function of X, Y and Z direction in the detector for Brem shower after reweighing having CVN value > 0.9.

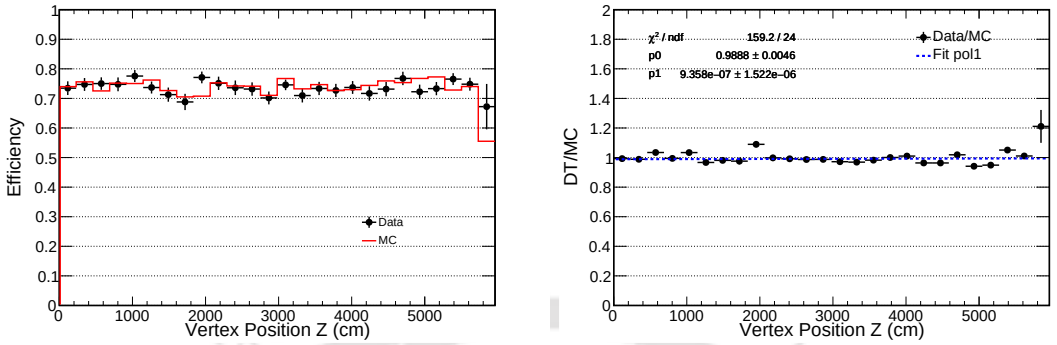

FIGURE 4.47: Signal selection efficiency in Z direction (left) and data/MC comparison (right) of reweighed sample for CVN >0.7.

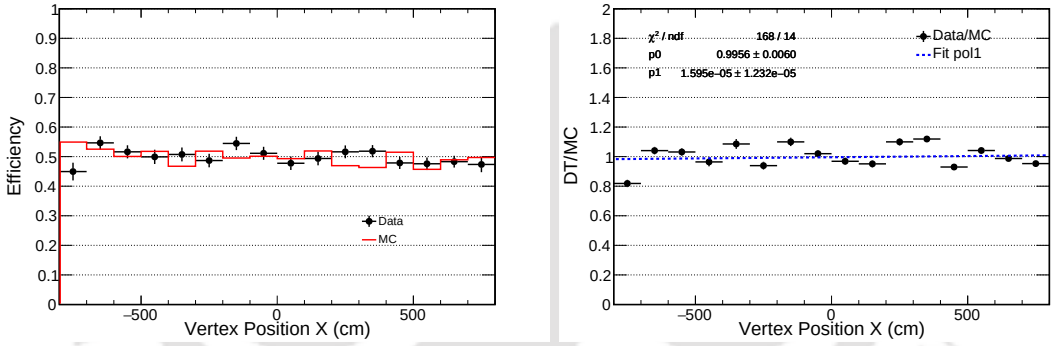

FIGURE 4.48: Signal selection efficiency in X direction (left) and data/MC comparison (right) of reweighed sample for CVN >0.9.

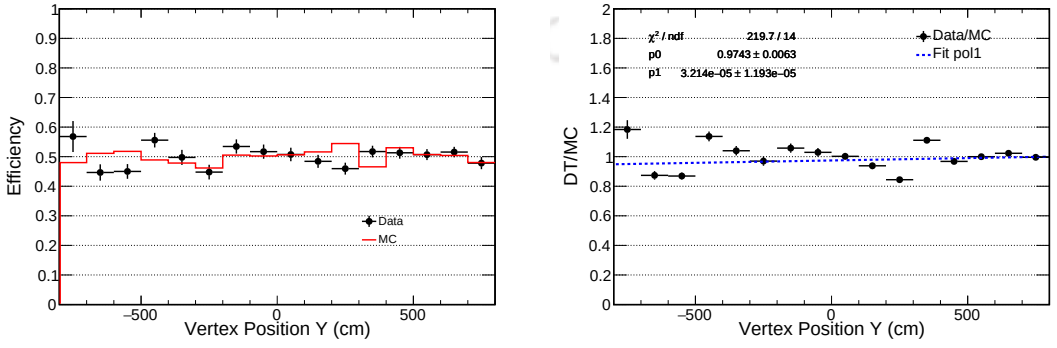

FIGURE 4.49: Signal selection efficiency in Y direction (left) and data/MC comparison (right) of reweighed sample for CVN >0.9.

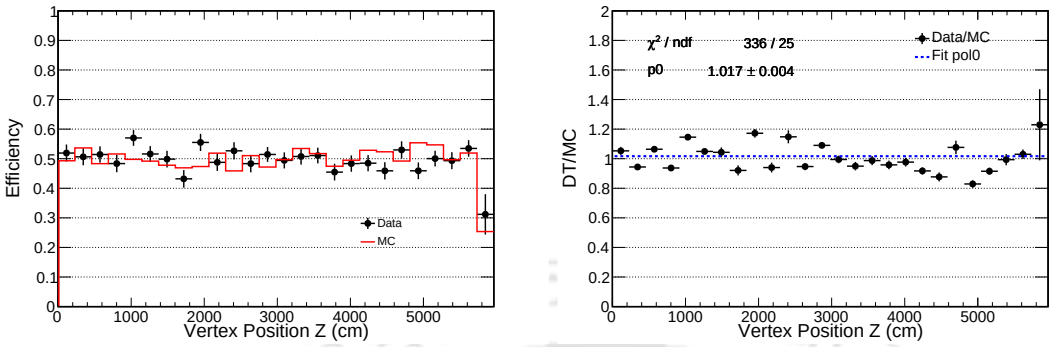

FIGURE 4.50: Signal selection efficiency in Z direction (left) and data/MC comparison (right) of reweighed sample for CVN >0.9.

Data/MC agreement	LID >0.7	LID RW >0.7	LID >0.9	LID RW >0.9
X	0.10 ± 0.00%	0.78 ± 0.17%	0.10 ± 0.10%	0.08 ± 0.19%
Y	0.20 ± 0.00%	0.80 ± 0.20%	0.10 ± 0.10%	0.44 ± 0.19%
Z	0.10 ± 0.10%	3.43 ± 0.19%	0.79 ± 0.12%	1.45 ± 0.33%

TABLE 4.2: Summary of the systematic uncertainty on the performance of the far detector based on the Brem EM shower sample using LID

Data/MC agreement	CVN >0.7	CVN RW >0.7	CVN >0.9	CVN RW >0.9
X	0.16 ± 0.15%	4.00 ± 0.00%	2.04 ± 0.32%	0.44 ± 0.60%
Y	0.02 ± 0.16%	0.40 ± 0.30%	0.85 ± 0.35%	2.57 ± 0.63%
Z	3.33 ± 0.30%	1.12 ± 0.46%	2.36 ± 0.30%	1.70 ± 0.40%

TABLE 4.3: Summary of the systematic uncertainty on the performance of the far detector based on the Brem EM shower sample using CVN.

As can be seen from Fig.4.48, Fig.4.49 and Fig.4.50., the signal selection efficiency in X, Y and Z direction for CVN is quite flat. Overall the data and MC agrees well within 3%.

Table 4.2 and Table 4.3 summarises the data and MC agreement for signal selection efficiency studies for LID and CVN respectively. The disagreement will be assigned as a systematic on signal selection efficiency.

As can be seen from the Table 4.3 and 4.2, the overall data and MC agreement stays within 4% for all cases.

	MC	Data
Number of tracks	433139(100.0%)	17541400 (100.0%)
20 < Number of cells < 500	336030 (77.0%)	12830400 (73.0%)
Vertex position X and Y <180 cm	23591 (54.0%)	6712560 (38.0%)
cos $\theta > 0.9$	70434 (16.0%)	1978690 (11.0%)
Track length >1300 cm	25352 (6.0%)	1453770 (8.0%)
Shower selected	5370 (1.0%)	133538 (0.8%)

TABLE 4.4: Pre-selection cuts applied in the analysis.

4.7 Bremsstrahlung EM Shower from Rock Muons in Near Detector

So far, we have discussed about Brem shower in far detector (FD). This section includes the Brem EM shower extracted in near detector (ND) using rock muons. In ND, we do not have as many energetic cosmic ray muons as in FD, but have an abundant rock muon sample. Neutrinos interact with rocks along the neutrino beam line and generate high energy muons before reaching the near detector. The rock muon events selection are described in reference [93].

Rock Muon Selection:

- 20 < Number of cells < 500
- Vertex position X and Y <180 cm

Choose the longest track in slice.

- cos $\theta > 0.9$
- Track length >1300 cm

Table 4.4 shows the cut flow of the pre selection cuts and Fig.4.51 shows a rock muon event elected in the near detector.

These muons may go through the detector, leave long isolated track, and bremsstrahlung EM showers during their journey. A MR algorithm for ND rock muon samples is developed. We then perform a similar shower reconstruction and PID study as in FD. Both

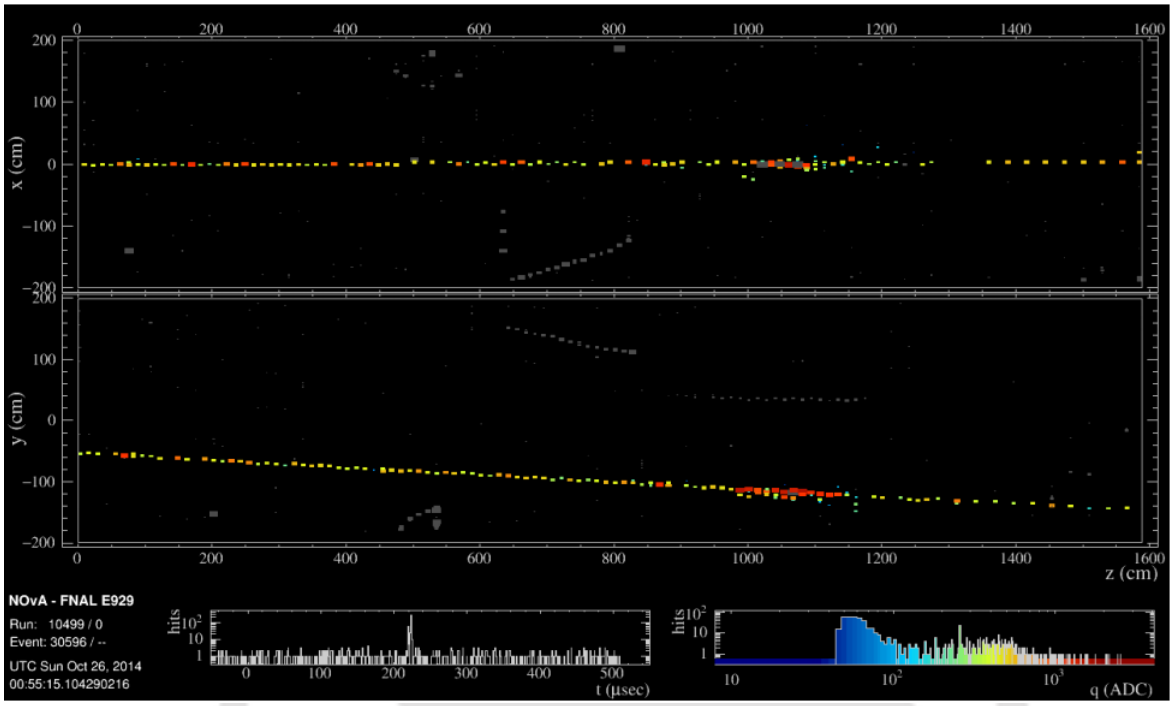

FIGURE 4.51: A rock muon in NOνA near detector.

data and MC showers are then reconstructed. Fig.[4.53-4.61] show the comparisons of reconstructed shower variables between data and MC. All MC plots are area normalized. The plots show good data and MC agreement, which demonstrate that the EM showers are well modelled by ND MC. Fig.4.51 a rock muon events which undergoes Brem shower. Fig.4.52 shows the Brem shower extracted, from the aforesaid rock muon, using MR algorithm.

Fig.4.53 shows the energy distribution of the rock muon brem EM shower. The data and MC agreement is very good. Fig.4.54 is the angular distribution of Brem in ND. We can clearly see that all of the Brem events in ND, unlike Brem in far detector, are along the beam direction as we have selected the rock muon with $\cos\theta > 0.9$. Fig.4.55, Fig.4.56 and Fig.4.57 show the plane traversed, radius and length of rock muon Brem EM shower in ND.

The longitudinal and transverse likelihood difference plots for various particle hypothesis are shown in Fig.4.58, Fig.4.59 and Fig.4.60. As can be seen from the plots the peak

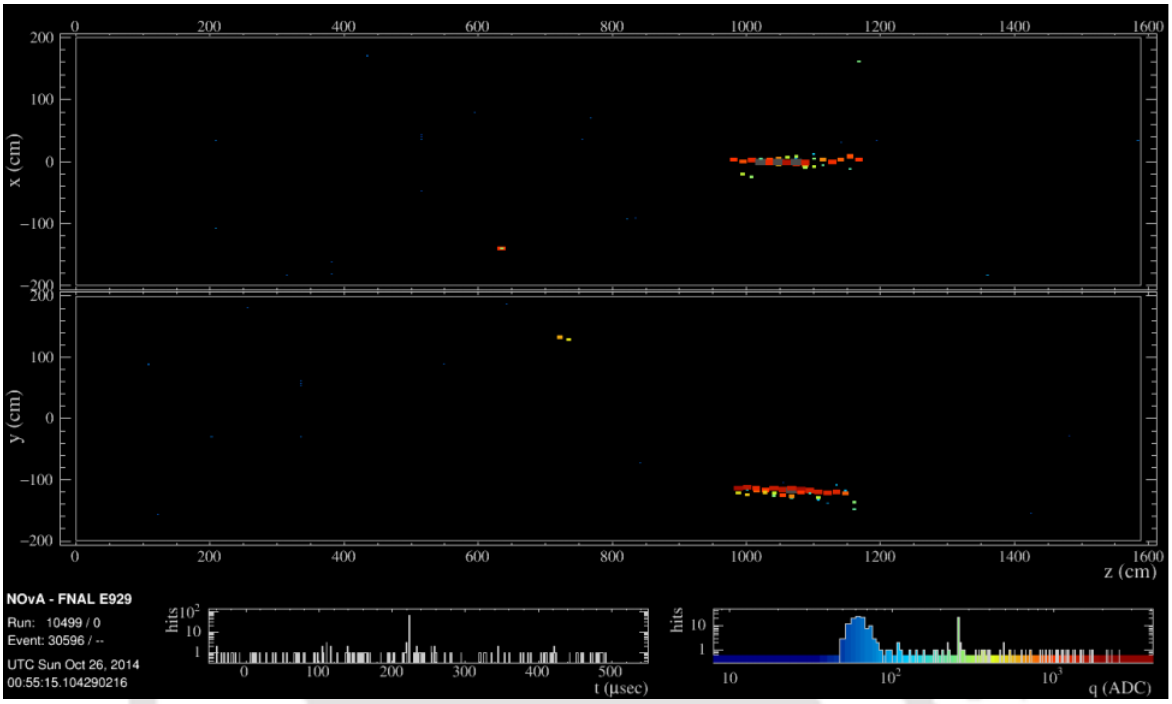

FIGURE 4.52: A rock muon EM shower extracted in NOνA near detector.

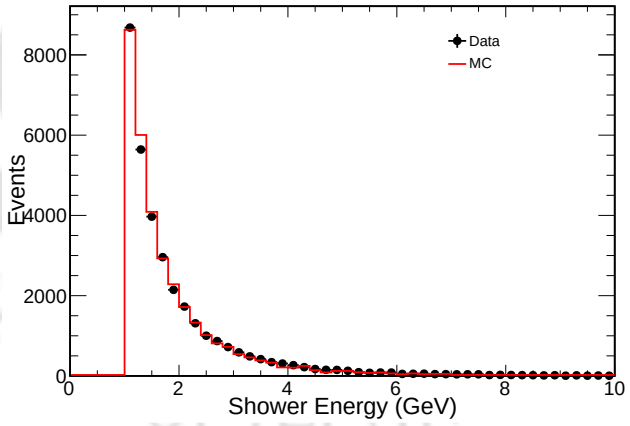

FIGURE 4.53: Data and MC comparison of energy for Brem EM showers generated by rock muons in near detector.

of all the plots is toward positive side. This implies that most of the rock muon brem EM showers are likely to be *e*-type.

We also see the LID output of the rock muon brem EM shower and from Fig.4.61, it is clear that most of the rock muon Brem EM showers are classified as ν_e-CC type to high degree.

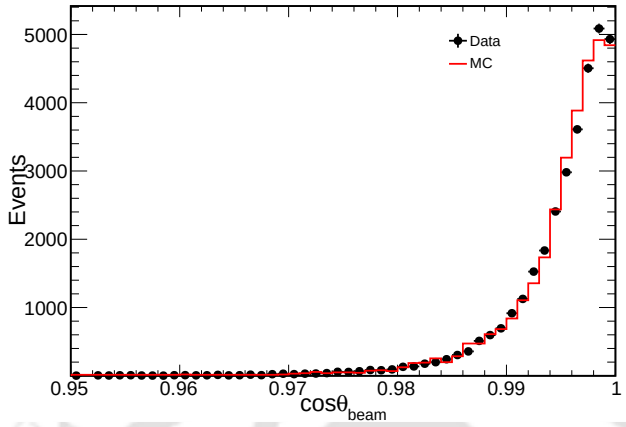

FIGURE 4.54: Data and MC comparison of $cos\theta$ for Brem EM showers generated by rock muons in near detector.

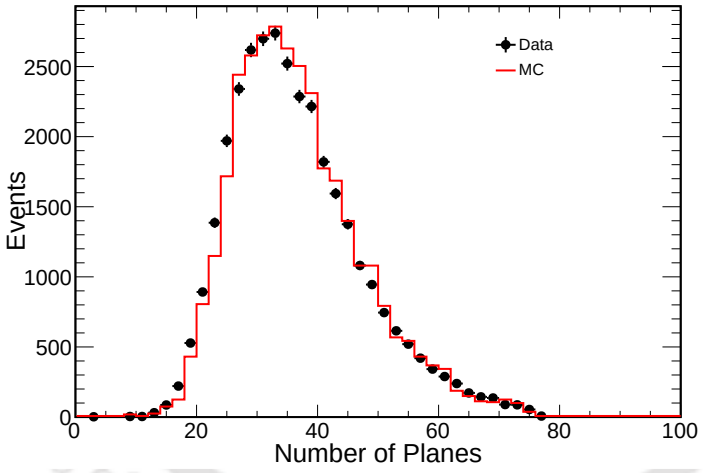

FIGURE 4.55: Data and MC comparison of planes for Brem EM showers generated by rock muons in near detector.

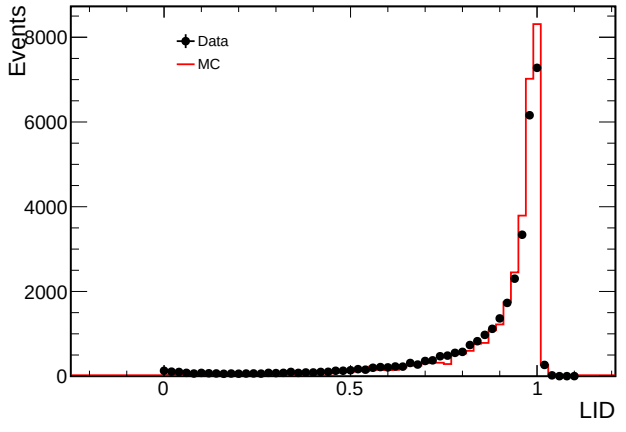

FIGURE 4.61: Data and MC comparison of LID output for Brem EM showers generated by rock muons in near detector. Most of the events are classified as ν_e type.

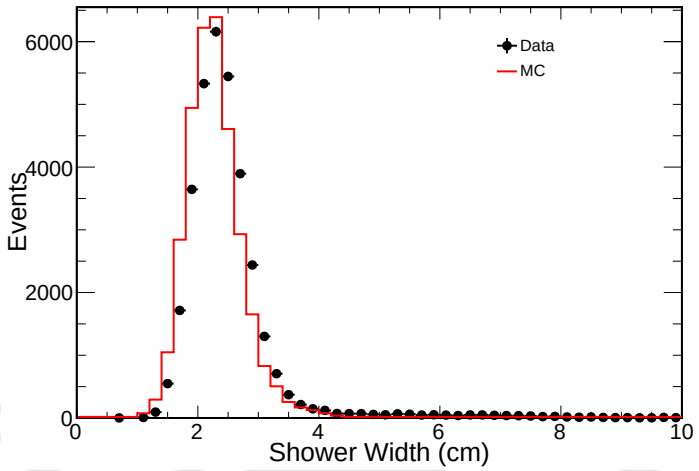

FIGURE 4.56: Data and MC comparison of radius for Brem EM showers generated by rock muons in near detector. The shower radius is defined as the average distance from shower cells to shower core weighted by cell energy.

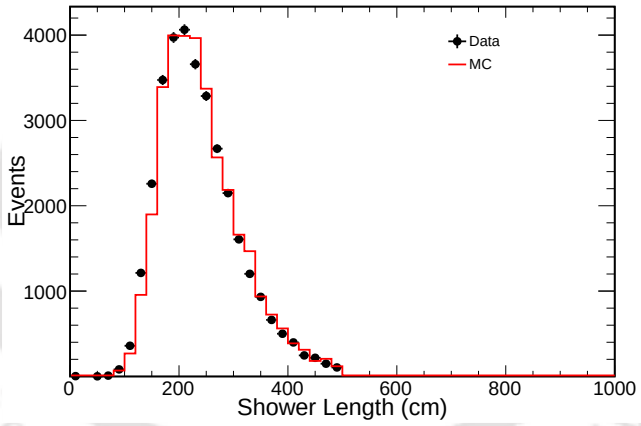

FIGURE 4.57: Data and MC comparison of length for Brem EM showers generated by rock muons in near detector.

ND rock muon brem shower can be used for many studies like data/MC agreement or cross section studies in ND. It has been used by several previous and current going cross section studies at NOνA such as ND physics studies such as ν_e-CC cross-section, ν on e scattering and neutral current coherent π^0 cross section studies [94][71].

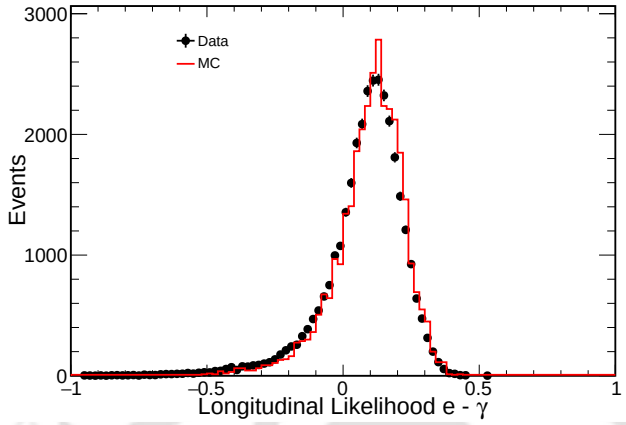

FIGURE 4.58: Data and MC comparison of likelihood difference under hypothesis of e and γ for Brem EM showers generated by rock muons in near detector.

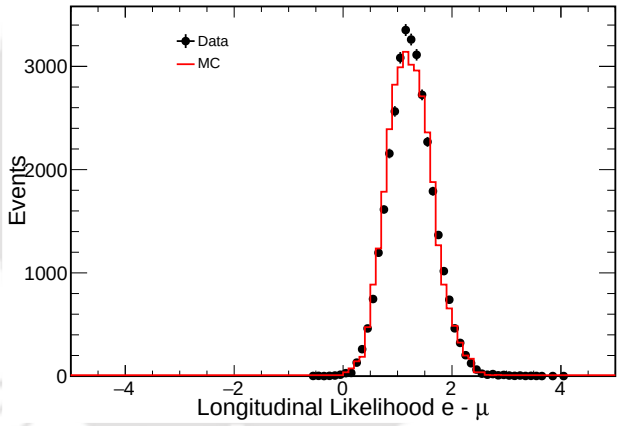

FIGURE 4.59: Data and MC comparison of likelihood difference under hypothesis of e and μ for Brem EM showers generated by rock muons in near detector.

4.8 Chapter Summary

NOνA looks for electron initiated electromagnetic shower in the $\nu_\mu \rightarrow \nu_e$ mode. The ν_e-appearance analysis relies on how correctly the shower simulations are done, how efficient are the algorithms which select and identify the $\nu_e - CC$ signal in the detector. Bremsstrahlung EM shower sample is extracted using Muon Removal algorithm and it is used to validate the signal shower modelling, PID and to study signal selection efficiency. Rock muon bremsstrahlung shower sample is also extracted and is used to study the same for the ND. Based on this study, we conclude that MC is in good agreement with data.

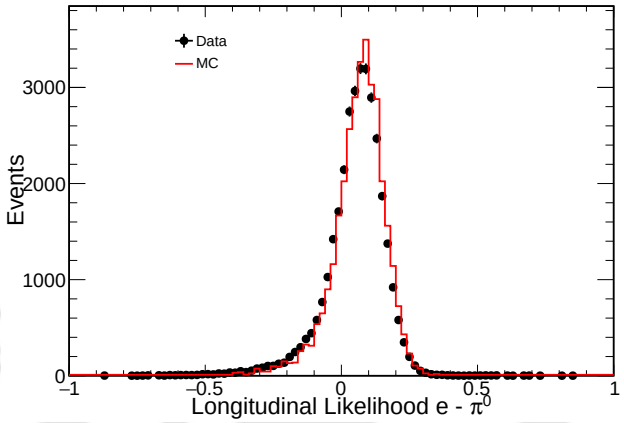

FIGURE 4.60: Data and MC comparison likelihood difference under hypothesis of e and π^0 output for Brem EM showers generated by rock muons in near detector.

Bremsstrahlung sample has been found to be a very useful sample that after removing the muons from the shower regions and can mimic the ν_e-CC signal. The data and simulations agree very well within few percent level. PID's have identified most of the Brem as signal, which enhances the confidence in the PID to select signal EM shower. Overall, bremsstrahlung has emerged as a new method of benchmarking the ν_e-CC signal.

In the next chapter, we discuss about the decay in flight (DiF) muon induced EM shower which is a much cleaner shower sample. The new shower exactly mimics the ν_e-CC initiated EM shower as discussed in the next chapter.

Chapter 5

Decay in Flight EM Shower

As discussed in the previous chapter, the bremsstrahlung EM shower sample has been used to benchmark several tools in the ν_e-CC analysis. It is also pointed out how an EM shower generated by muon decay in flight (DiF), which is the subject of the current chapter, exactly mimics the EM shower generated by the ν_e-CC interaction in the detector. This is because, Brem is a γ induced EM shower and DiF is an electron induced EM shower, exactly similar to ν_e-CC electron EM shower, and that makes the DiF an ideal sample in the context.

The chapter begins with a brief discussion on the motivation of the study, we will discuss in details the selection and the extraction process of EM shower generated by DiF muon. Subsequently, we will present the results on the study of benchmarking several aspects of ν_e-CC analysis using the DiF EM shower sample.

5.1 Motivation

In $\nu_\mu \rightarrow \nu_e$ appearance mode, ν_e interacts in the far detector through a quasi elastic charged current interaction and produces an electromagnetic (EM) shower. The topologies of the interaction in far detector is shown in the Fig.5.1. Apart from signal EM shower, there are also cosmogenic EM showers in the detector. The cosmic rate at far detector is

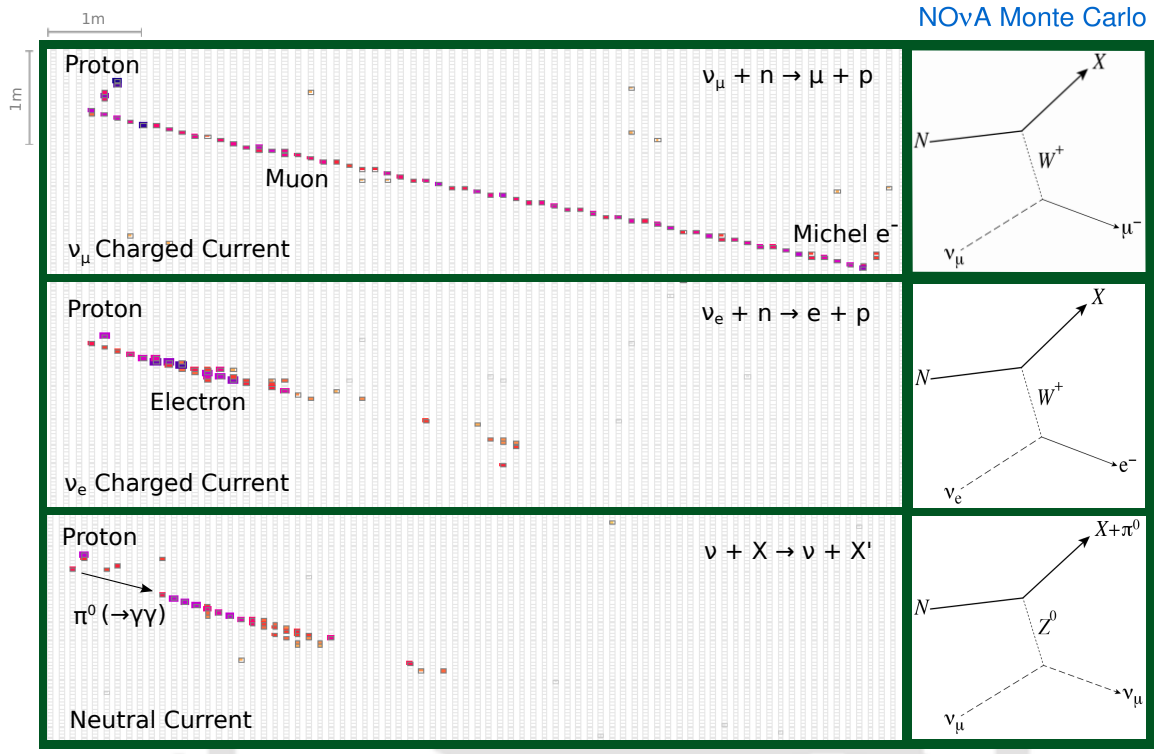

FIGURE 5.1: Event topologies at NOνA far detector. A ν_μ-CC interaction exhibits itself as a long track. ν_e-CC and NC π^0 interactions result in an electromagnetic showers in the detector.

148 kHz, due to its location on surface. Muons can induce an electro-magnetic shower via soft scattering in the form of Bremsstrahlung radiation and it can also produce EM shower initiated by an electron which is produces by a muon decaying in flight before it stops. The total EM shower rate in FD is 72 kHz and DiF rate in FD is 70 Hz [29][28]. The produced electron can initiate an EM shower. EM shower from muon decay in flight can serve as a similar source of these EM shower. A cosmic muon undergoing the DiF process in FD is shown in Fig.5.2. Once the DiF shower is extracted, it can be used as a data-driven method for verifying or improving shower modelling and PID performance in selecting EM shower at NOνA. Previously studied Brem sample has proved to be very useful tool in verifying and testing PID's and shower modelling at NOνA. DiF, unlike Brem, with no muon or very little contamination, can serve as an ideal EM shower sample.

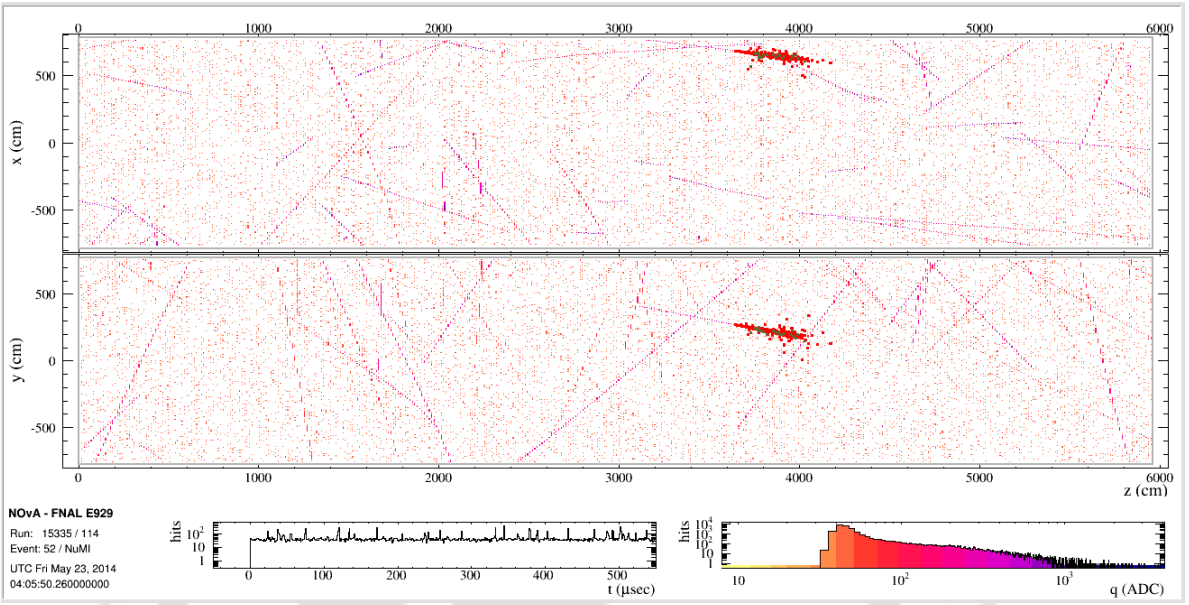

FIGURE 5.2: A decay in flight muon event in the far detector. The thick red hits show the electron initiated EM shower. The above image is an 500 μsec window of a cosmic event in the far detector in X-Y and X-Z view. The bottom left plot shows the timing of the hits and the bottom right plot shows the charge deposited information of the event.

5.2 DiF EM Shower Selection in FD

A DiF event is selected by identifying the amount of energy it deposits per plane while traversing the detector. Since muon is a minimum ionizing particle (MIP) , so in the advent of decay in flight, shower energy deposition along the muon track increases. As a consequence, per plane energy deposition increases. We look for the energy deposition in a series of consecutive planes in the beginning and ending of the shower. If certain criteria are met (as described later), we select the event as a DiF event. However, Brem at the end of a muon track is a biggest source of background for DiF. Brems' selection criteria of energy deposition is very similar to that of DiF. To differentiate DiF from Brem EM shower, additional selection criteria are applied such as distance of selected shower end from the reconstructed track end and number of missing planes in the shower region. Using this, we are able to reject backgrounds to more than 99.9% as shown in Fig.5.7.

The following datasets are used for this study:

- **Data**: "ynitin_prod_artdaq_fd_cosmic_prod3_subset10of12_isgoodrun"

- **MC**: "prod_artdaq_R17-03-09-prod3genie.d_fd_cry_full_v1"

The detector exposure used in files is:

- Data: 7016.3 seconds.
- MC: 2646.6 seconds

All files are reconstructed and processed for standard reconstruction in NOνA software S16-12-07 release. Main PID algorithm for first analysis (FA) is chosen, for validation which is a likelihood based on artificial neural network (LID). Selection of a DiF constitute of three steps. First step is the selection of a muon candidate. In the second step the muon is scanned over its planes for energy deposition and DiF is located. In the last step, if DiF is located in the second step, DiF shower is put in a separate reconstruction objects called as slicer. The steps are described in more detail in the following sections.

5.2.1 Muon Tracks Pre-selection

EM Shower extraction process begins by selecting a candidate muon track in a slice. A good track is selected first which passes the following cuts:

- Should have timing between 25000 ns and 475000 ns.
- Should have number of reconstructed 3-D fuzzy prongs <2.
- Should have track end contained within 20 cm of x,y and z edge of detector and start should not be contained.
- Track end should not be any closer than 50 cm of bad region of the detector.
- Number of planes the muon track traverses is > 16 which is optimized using signal significance ($s/\sqrt{s+b}$) as shown in Fig.5.3.
- The muon track is also required to be in the horizontal direction (as close to beam direction) by requiring $\cos\theta > 0.5$, where θ is the angle of the muon track with respect to the beam axis (Z-axis).

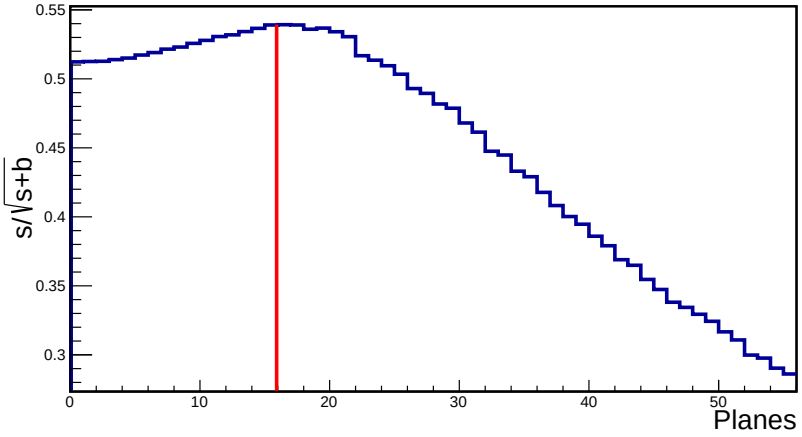

FIGURE 5.3: Signal significance ($s/\sqrt{s+b}$) for number of planes a muon track should traverse. The maximum value is 16 planes.

Table5.1 shows the cut flow table for selection of a muon track for both data and MC. The signal sample is selected based on the MC truth information and at the preselection level, the selection efficiency for signal events is $\sim$ 6.2%.

Preselection cuts	Signal	MC Muon tracks	Data Muon tracks
No. of slices	98319	2.98×10^8(100%)	7.82×10^8 (100%)
No. of fuzzy prong <2	55925	2.29×10^8(76.8%)	6.09×10^8 (77.8%)
Good slice cut[1]	46128	1.83×10^8(61.4%)	4.57×10^8(58.4%)
Contained cosmic track cut	44893	1.66×10^8 (55.7%)	3.97×10^8 (50.7%)
Angle cut on cosmic track cosθ >0.5	7334	2.56×10^7 (8.6%)	5.80×10^7(7.41%)
Plane cut on cosmic track	7278	2.35×10^7 (7.8%)	5.14×10^7(6.57%)
Avoiding bad regions from the end of the track <50 cm	6058	9.31×10^6 (3.12%)	2.11×10^7(2.7%)

TABLE 5.1: Cut flow table for the selection of muon candidates.

5.2.2 DiF EM Shower Extraction

After the selection of the muon candidate, the muon track is searched for shower region. A shower region is determined by energy deposition per plane (dE/dx) information, downstream of the track. In the event of EM shower, as can be seen from Fig.5.4, energy deposition per unit plane increases in comparison to energy deposited in the same plane

[1] slc→IsNoise() && tracks.size()==1 && tracks[0]→Is3D() && slc→MeanTNS() <25000 —— slc→MeanTNS()>475000

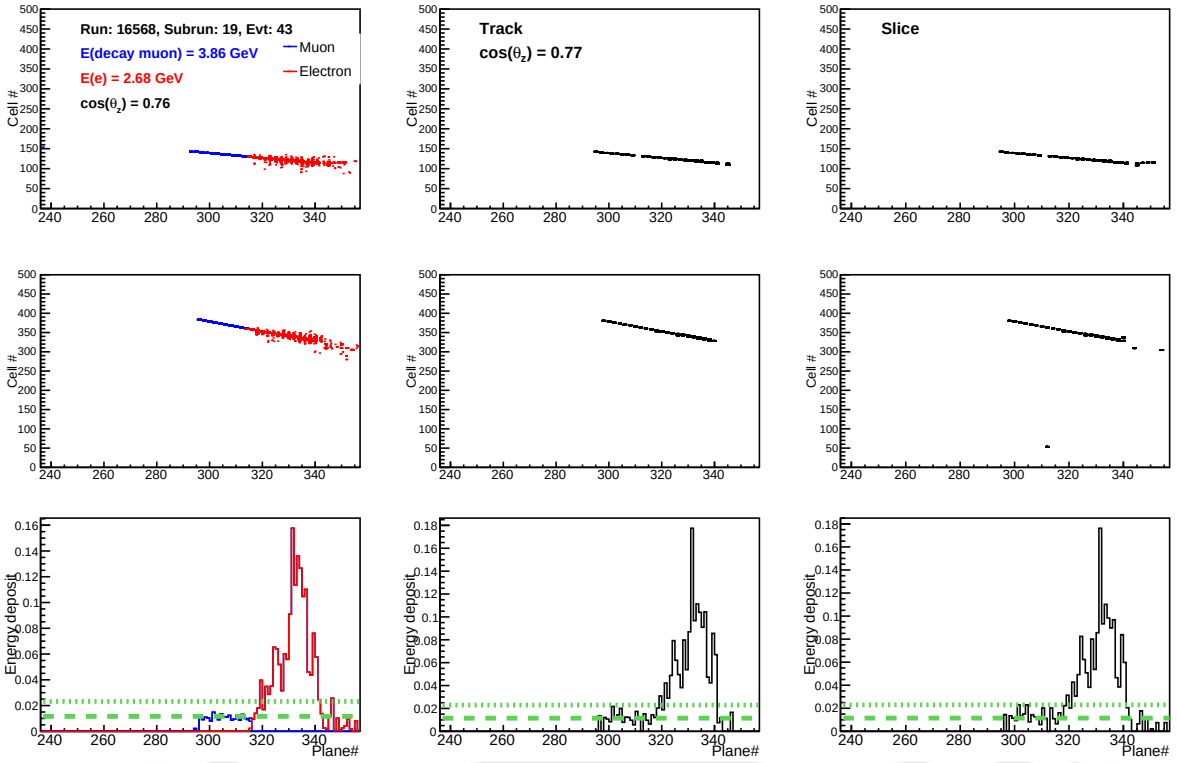

FIGURE 5.4: A DiF event in FD before selection. The blue dots represent fls hits (simulated hits) of muons and red dots represent the hits by an electron. Alongside track (middle) and slice (right) hits are also shown. Green dash lines show the 1 MIP and 2 MIP mark in a plane.

by a muon since muon is a minimum ionizing particle. This characteristic is utilized in finding the muon tracks which are undergoing EM shower processes.

The EM shower finding process starts with a preselected muon track. If on the track, two consecutive planes are found to have energy (energy is averaged over 5 planes) deposition more than 2.5*MIP, it is marked as the starting of the EM shower. Given starting of the EM shower, the end of the EM shower is found if five consecutive planes downstream are having energy deposition between 0.01*MIP and 1.5*MIP. This way shower start and shower end is marked on the muon track. DiF electron remains a MIP particle for few cells before it starts showering. So in order to accommodate the electron hits we scan the planes upstream from the shower start position till $Plane\ average\ energy \leq 0.9 * MIP$ and this is the final shower start position. Hits energy between shower start plane and end plane is summed and we only select shower with energy > 0.5 GeV.

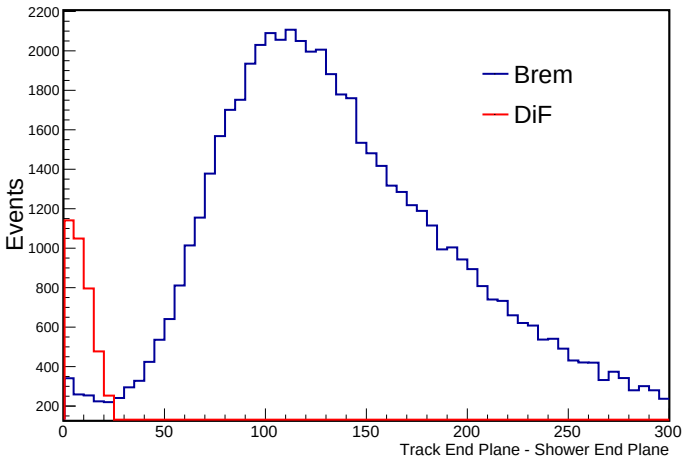

FIGURE 5.5: Track End -Shower End distribution for both DiF and Bremsstrahlung EM shower events.

However, a considerable amount of bremsstrahlung shower are still getting selected by the aforesaid procedure. The bremsstrahlung shower itself is a huge background for the DiF. Such bremsstrahlung process events occurs near the end of the muon track and the events look similar to a decay in flight electron. We developed a variable which drastically reduces the brem background-

$$\begin{aligned} \textit{Track end variable} = \textit{Reconstructed cosmic muon track end plane} \\ - \textit{shower end plane as found by } dE/dx \textit{ process.} \end{aligned} \tag{5.1}$$

The distribution for this variable is shown in Fig.5.5 for both the brem and DiF EM shower.

It is clear that this variable is quite important in removing the brem background events from the DiF signal.

We choose the value of track end variable by maximizing the signal significance, $s/\sqrt{s+b}$, where, s is signal (DiF) and b is background (brem), as shown in Fig.5.6 and the signal significance is found to be maximum for a selection cut at 26 on this variable.

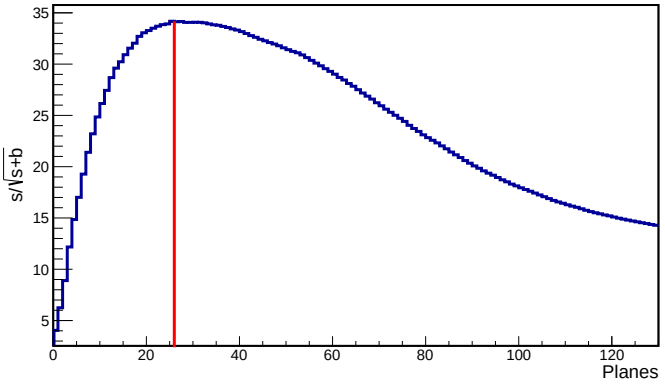

FIGURE 5.6: Signal significance ($s/\sqrt{s+b}$) as a function of end of the track end variable. Maximum significance at 26 planes.

Since muon is a MIP particle so, in case of a brem event there will be a track with MIP energy deposition before and after the shower. So, a cut is designed to remove such kind of backgrounds. The cut, called as MIP at track end is defined as:

$$\begin{aligned} MIP\ at\ track\ end = Scan\ over\ 3\ planes\ from\ the\ reconstructed\ cosmic \\ track\ end\ and\ check\ whether\ it\ is\ consistent\ with \\ 0.5 * MIP\ < average\ plane\ energy\ < 1.5 * MIP. \end{aligned} \tag{5.2}$$

This selection criteria further reduces the background by about by 17%, with about 3% loss in signal efficiency.

Table.5.2 shows the cut flow table for shower extraction. Fig.5.7 shows the reduction of signal and background as a function of different selection criteria applied to extract the EM shower. The shower hits are then stored for further processing such that all the hits in the slice within 80 cm of track are stored into a new cluster as shown in Fig.5.10. A NOνA event display of selected cosmic muon decay in flight is shown in Fig.5.9

Fig.5.8 shows the cut efficiency for each cut as a function of energy and angle of DiF shower.

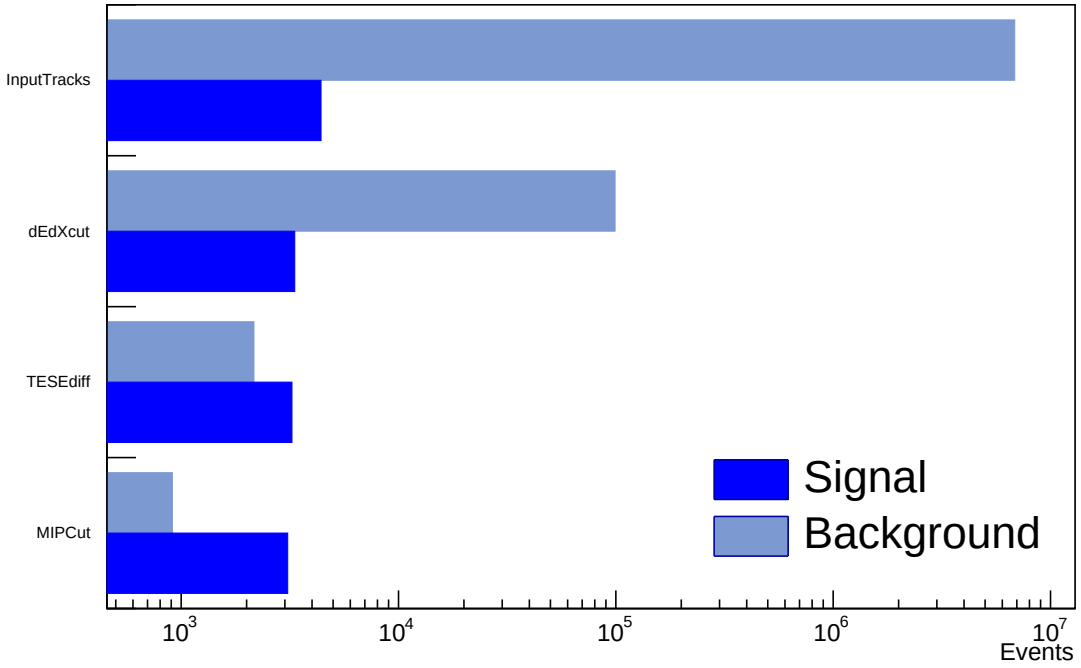

FIGURE 5.7: Bar chart of signal selection and background rejection with consecutive shower extraction cuts. Background is rejected upto 99.9 %.

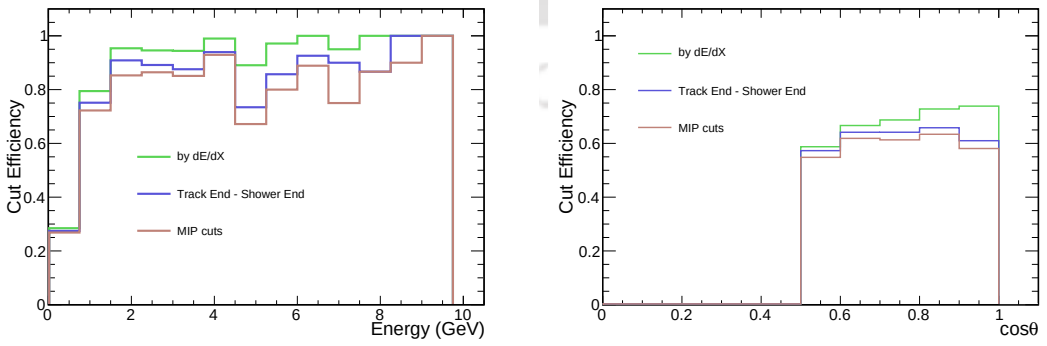

FIGURE 5.8: Efficiency of each selection criteria cut as a function of shower energy and shower angle.

Shower selection	MC truth (sig)	MC (sig+bg)	Data
Input tracks	6058(100%)	9.31×10^{6}(100%)	2.11×10^{7}(100%)
Shower found by DE/dx info	4016(66.3%)	65829(0.71%)	100161(0.47%)
Track end variable <26	3774(62.3%)	5396(0.05%)	10411(0.05%)
MIP at track end	3659(60.4%)	4453(0.04%)	9031(0.04%)

TABLE 5.2: Cut flow table for the selection of the DiF shower after preselection cuts.

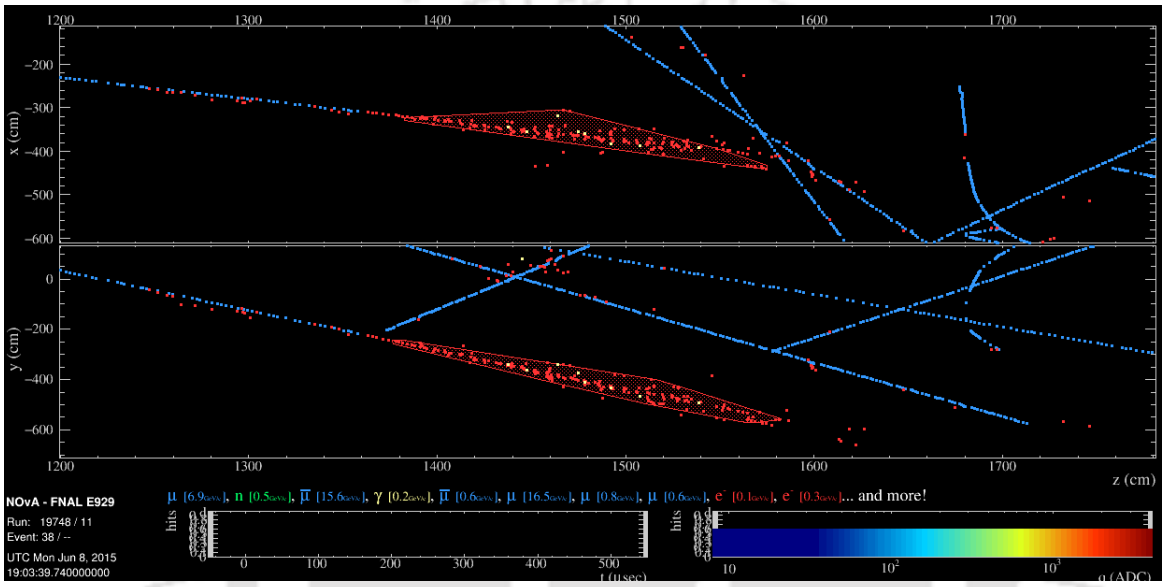

FIGURE 5.9: A selected DiF muon in far detector. The blue dots represent fls hits (simulated hits) of muons and the red dots represent the hits by an electron. The boundary around the red hits represents the shower region which is extracted. The above image is a 500 μsec window of a cosmic event in the far detector in X-Y and X-Z view. The bottom left plot shows the timing information of the hits and the bottom right plot shows the charge deposited information of the event.

Fig.5.11 shows the final selection efficiency of DiF shower selection as a function of energy and angle and Fig.5.12 shows the purity of sample as a function of energy, angle and particle-ID (LID).

In Fig. [5.13, 5.14] are shown the vertex distributions Y vs X, Y vs Z, X vs Z, and X, Y, Z of Brem candidates in the detector.

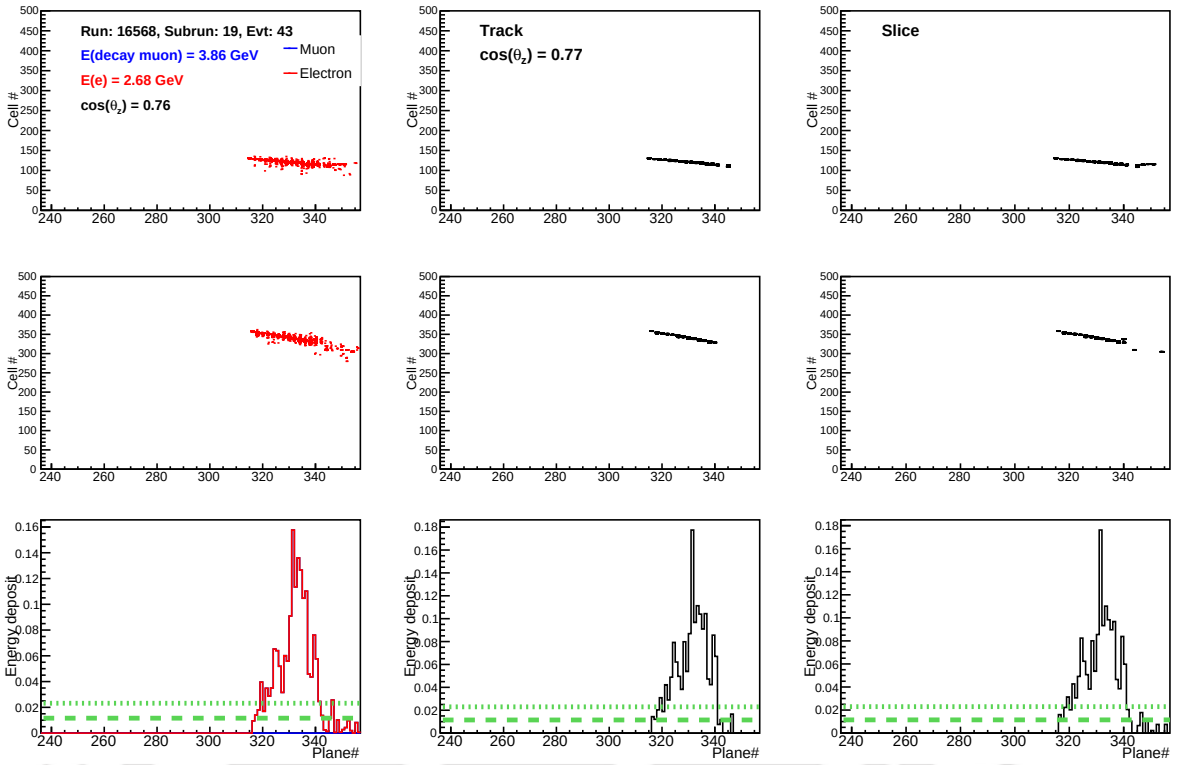

FIGURE 5.10: Selected DiF hits in red after shower finding and muon removal on the muon track of Fig.5.4. Red hits (left) which correspond to electron are selected after the parent muon is removed. Alongside track (middle) and slice (right) hits are also shown. Green dash lines show the 1 MIP and 2 MIP mark in a plane.

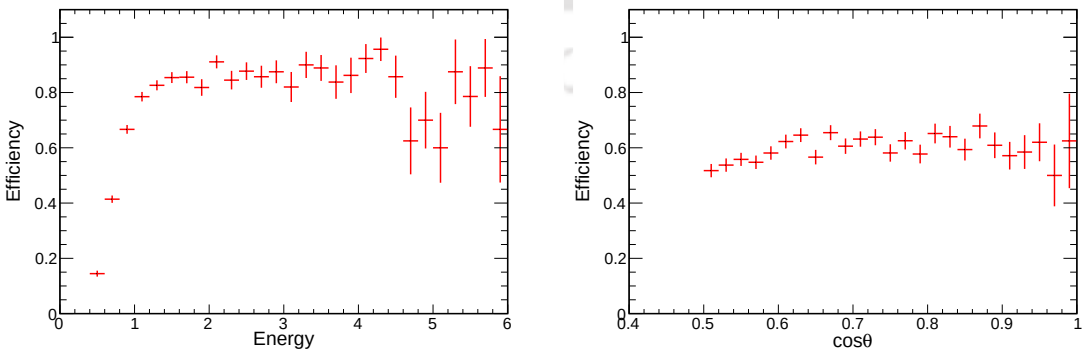

FIGURE 5.11: Efficiency as a function of *energy*, *angle* w.r.t beam of finally selected DiF events.

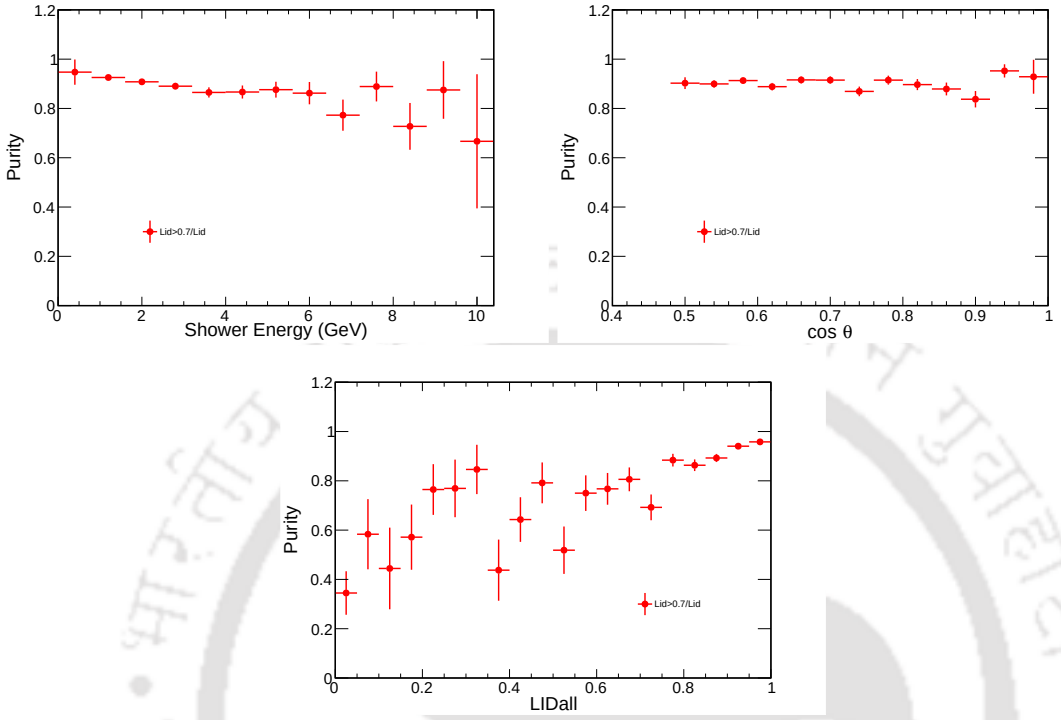

FIGURE 5.12: Purity as a function of *energy*, *angle* w.r.t beam and particle-ID (LID) of finally selected DiF events.

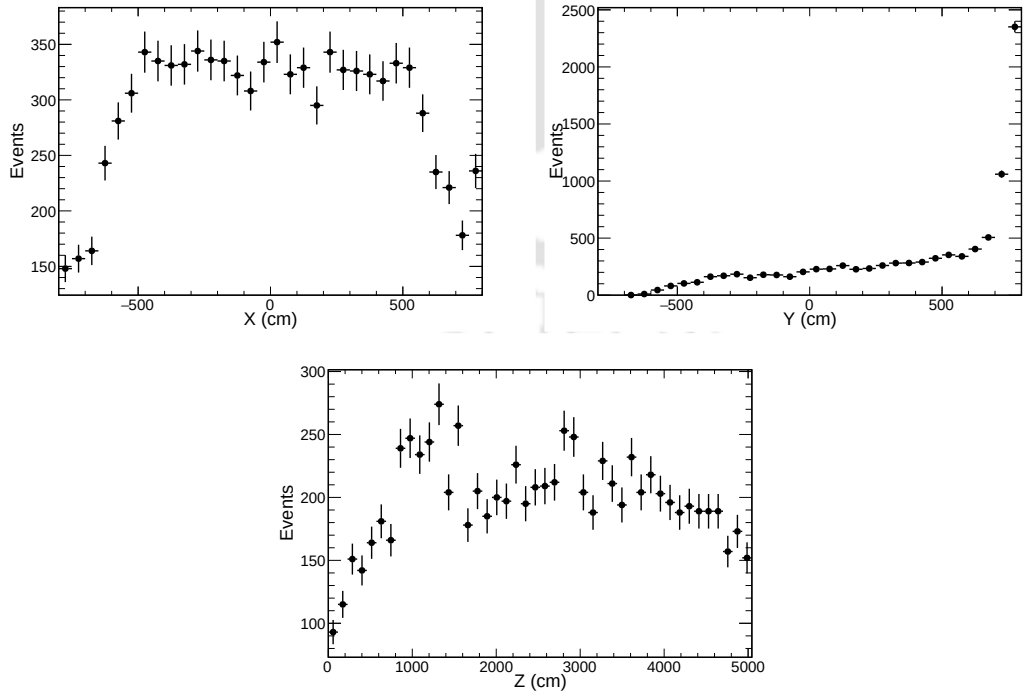

FIGURE 5.13: X, Y and Z distribution of DiF candidates in the far detector.

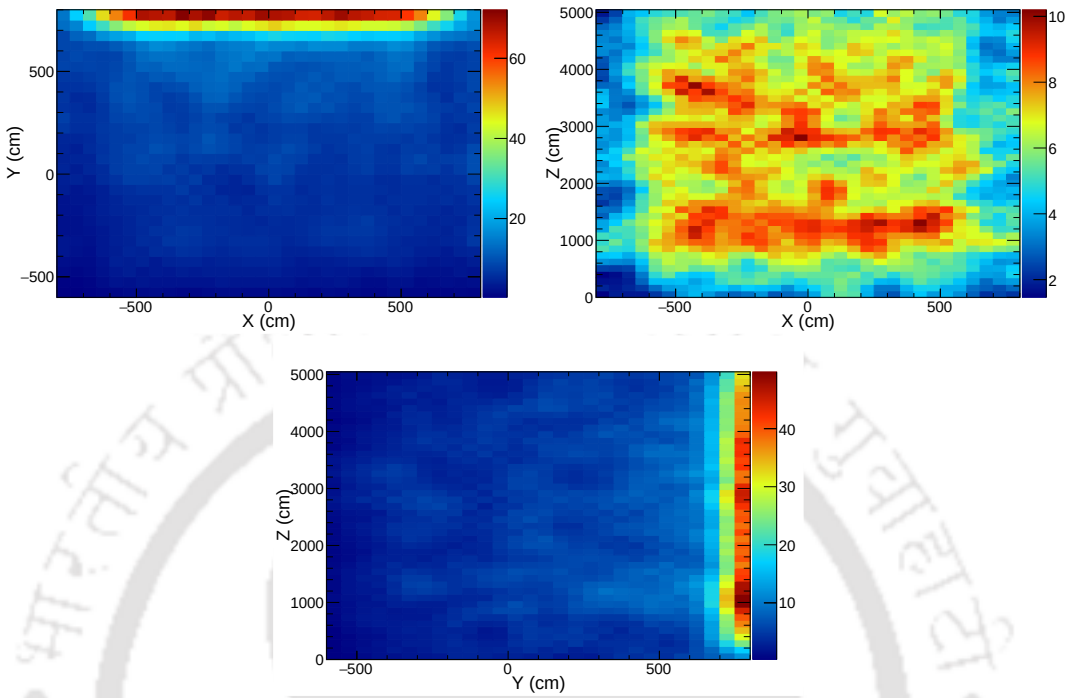

FIGURE 5.14: Y vs X, Y vs Z and X vs Z distribution of DiF candidates in the far detector.

5.2.3 Shower Extracted Cluster Efficiency

In this section, we examine the type of hits stored in the cluster after shower extraction in terms of cluster efficiency and cluster purity. The cluster efficiency reflects the amount of electron hits successfully stored in the cluster. The cluster purity is a measure of non-electron hits which are there in the cluster. A cluster efficiency and purity for selecting DiF hits are plotted in Fig.5.15 which are defined as follows:

- Cluster efficiency: Returns the fraction of all energy in an event from a specific set of Geant4 track IDs that is a collection of weighted hits.
- Cluster purity: Returns the fraction of hits in a collection that come from the specified Geant4 track ids. For DiF case, track IDs refer to electron IDs.

Based on the plots in Fig.5.15, we conclude that the algorithm developed to select the DiF events can perform with an efficiency of about 95% and purity about 97%. In case of

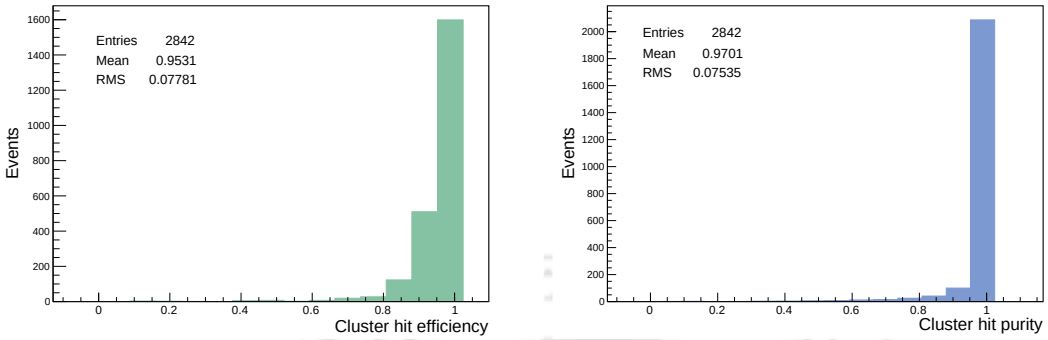

FIGURE 5.15: Left plot shows the cluster efficiency for electron hits and right plot shows the purity of hits in the cluster.

brem shower, these numbers are not available as NOνA cosmic simulations do not store information of bremsstrahlung process, but is expected to be smaller due to the muon contamination.

5.3 DiF Variables and Data/MC comparison

DiF shower sample is reconstructed using the standard reconstruction chain of ν_e-CC interaction in far detector and various distributions are plotted. NOνA has several reconstruction chains which depend on what kind of event one wants to analyze. The standard reconstruction chain for identifying ν_e-CC event constitute of following steps. We process the DiF sample through the same chain as ν_e-CC.

- First raw hits are converted to callibrated hits called as CalHits.
- Space and time correlated hits are put together in a form of C++ containers called as clusters.
- Track reconstruction is done on the clusters.
- HoughTransform algorithm [95] is run on the clusters for any pattern recognition.
- Vertexing is done using ElasticArms [96] algorithm using HoughTransform output.
- FuzzyKVertex algorithm [97] uses the ElasticArms output and HoughTransform output to separate sub clusters in clusters.

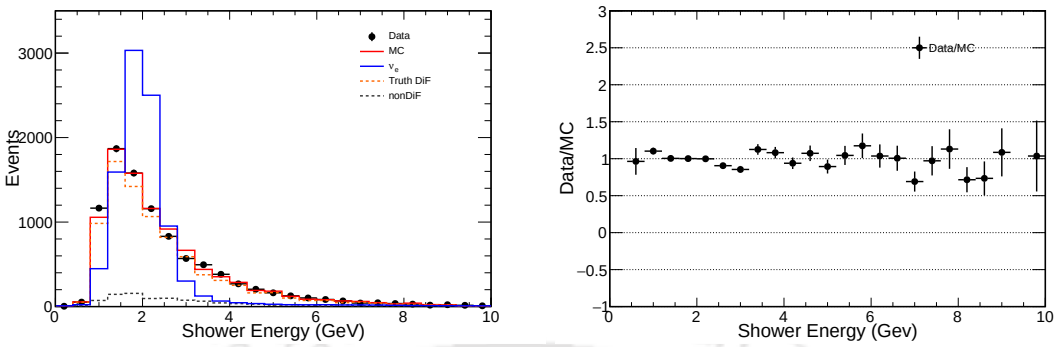

FIGURE 5.16: Energy distribution of DiF (left) and ratio of Data/MC as a function of shower energy (right).

- The selected events are passed through the particle-ID algorithm.

We use the same reconstruction chain as that of ν_e-CC as our main aim is to use DiF shower to check particle identification algorithms, shower modelling and to determine the systematics associated with the selection for ν_e-CC interaction in the detector. Data and MC comparison for energy and angle variables are also shown in the Fig.5.16, Fig.5.17[2]. The agreement is reasonable. The plots includes the ν_e-CC MC (blue), DiF by MC truth (dashed orange) value and non-DiF events (dashed black) . As can be seen in the plots, the ν_e-CC MC is relatively higher in energy than DiF sample (red) . Also as expected, the DiF angular distribution is more perpendicular than ν_e-CC beam sample. This is due to the fact that most of the muon tracks are coming from upward direction in the detector whereas beam direction is horizontal. The right plots show the data and MC ratio of DiF events.

5.3.1 Likelihoods Differences for DiF EM Shower

The particle identification (PID) in NOνA is based on artificial neural network in which various likelihoods are fed as an input to select ν_e-CC signal. In this section, likelihood differences are shown for various particle hypothesis in Fig.[5.18-5.20]. Likelihoods are constructed in two direction: longitudinal likelihood along the direction of shower and

[2]The color coding in the plots: red MC, black Data, blue ν_e-CC, orange MCTruth DiF, gray MCTruth non-DiF

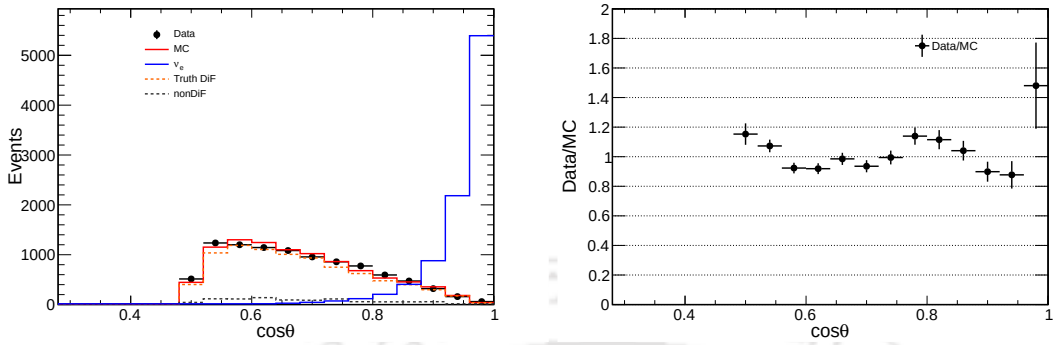

FIGURE 5.17: Angular distribution of DiF (left) and ratio of Data/MC as a function of shower angle with respect to beam direction (right).

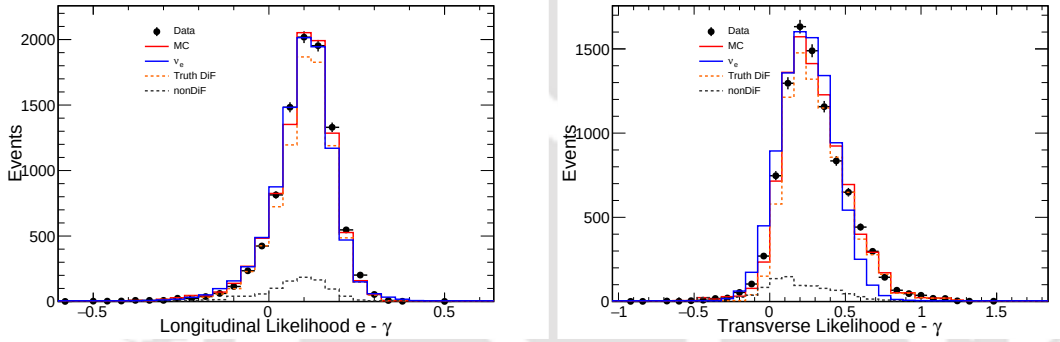

FIGURE 5.18: Longitudinal and transverse likelihood differences between electron and γ particle hypothesis. DiF shower is more identified as electron type than γ shower in transverse direction. However in longitudinal direction, separation is relatively less.

transverse likelihood along the transverse direction of the shower. Transverse likelihoods differences show clear distinction between two particle hypothesis. Longitudinal likelihoods differences however show lesser distinction than transverse direction in case of γ and π^0. This is due to the fact that the characteristics of the shower in the core (longitudinal direction) in γ, e and π^0 seems almost the same.

Fig.5.21 shows the final neural network output for PID (DiF results are shon only for LID, a likelihood based ANN PID). The events near the value 1 are classified as ν_e-CC type and the events which are far from 1 are classified as non-ν_e-CC type. As clearly seen, most of the DiF EM showers are identified as a shower from ν_e-CC interaction.

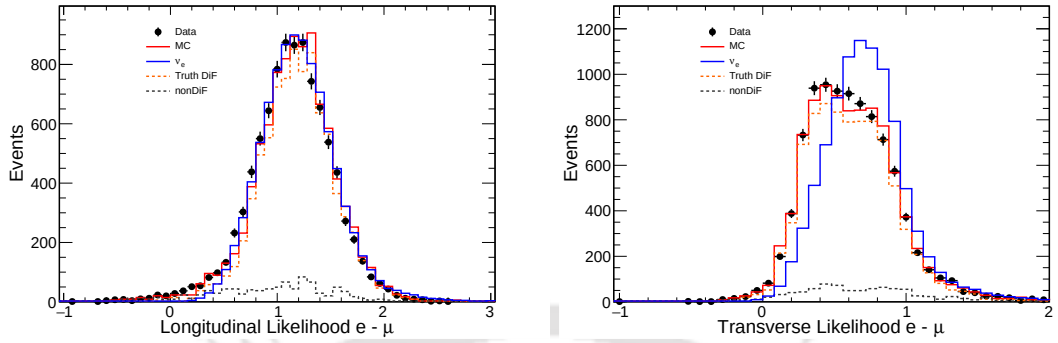

FIGURE 5.19: Longitudinal and transverse likelihood differences between electron and μ particle hypothesis. DiF shower is clearly identified as electron type than μ in transverse and in longitudinal direction.

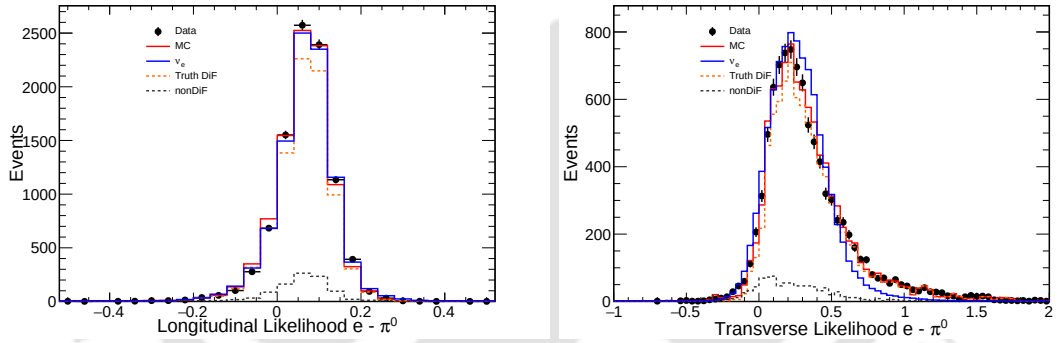

FIGURE 5.20: Longitudinal and transverse likelihood differences between electron and π^0 particle hypothesis. DiF shower is more identified as electron type than π^0 shower in transverse direction. However in longitudinal direction, the separation is relatively less.

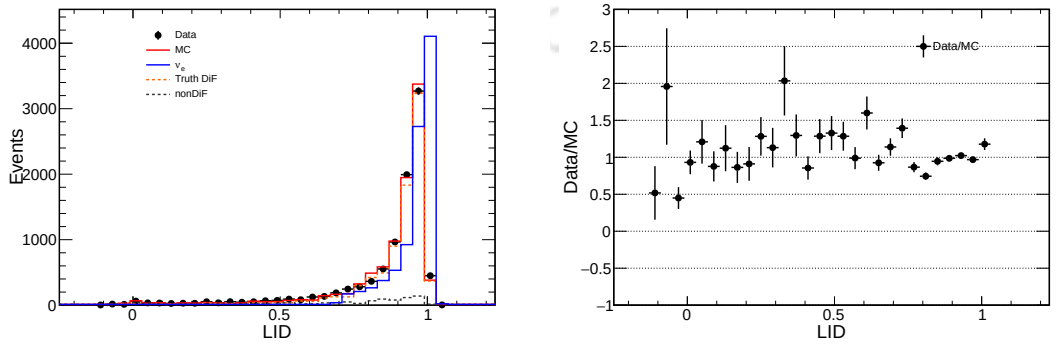

FIGURE 5.21: Artificial neural network based particle-ID, LID output. Most of the DiF showers are identified as ν_e showers.

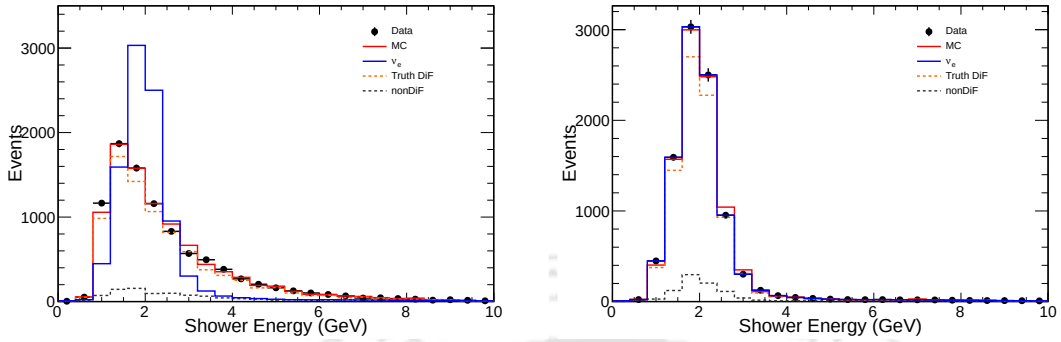

FIGURE 5.22: Energy of DiF vs ν_e-CC before and after reweighing.

5.4 DiF EM Shower vs ν_e-CC EM Shower and Reweighing

As discussed earlier, the goal of this study is to benchmark PID and shower modelling of ν_e-CC shower using DiF sample. However, before we do that, we should see the comparison of the variables from both the shower samples. As already shown in Fig.[5.16-5.17], shower variables have been compared between DiF (both data and MC) and ν_e-CC MC and we can see that the difference between the two types of showers mostly comes from energy and angle distributions. In comparison to ν_e-CC shower, DiF is less energetic as shown in Fig.5.16. The direction of DiF is also different than ν_e-CC shower as shown in Fig.5.17. ν_e is along the beam direction, which is along the horizontal axis of the detector (Z-axis of the detector) and DiF inherits the direction of cosmic muon which is more vertical (Y-axis). To take care of this difference, the variable distributions are reweighed with respect to the ν_e-CC energy spectrum. In the plots shown in Fig.[5.22-5.26], we apply reweighing only in one variable only with respect to ν_e-CC energy, and compare the results. After reweighing the DiF shower, the distributions of two showers have reasonable agreement. However, since we are not reweighing in angle, the cosine distribution remains same. We only apply reweighing in energy variable as the DiF events are statistically limited in the signal region ($\cos\theta > 0.8$) as can be seen in plot of Fig.5.23 and construction of reweighing matrix suffer from limited statistics.

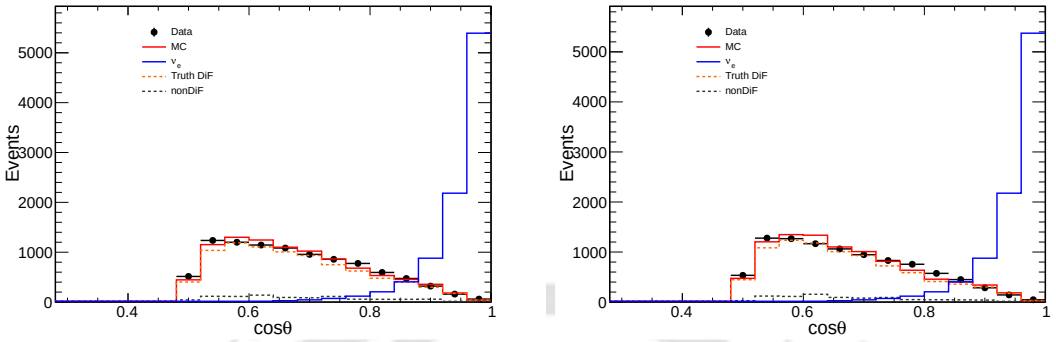

FIGURE 5.23: Direction w.r.t. beam ($\cos\theta$) of DiF vs ν_e-CC before (left) and after (right) reweighing.

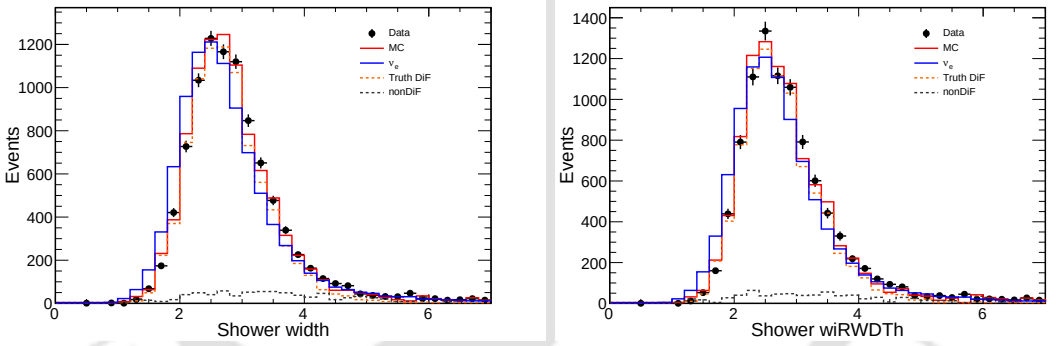

FIGURE 5.24: Shower radius of DiF vs ν_e-CC before (left) and after (right) reweighing.

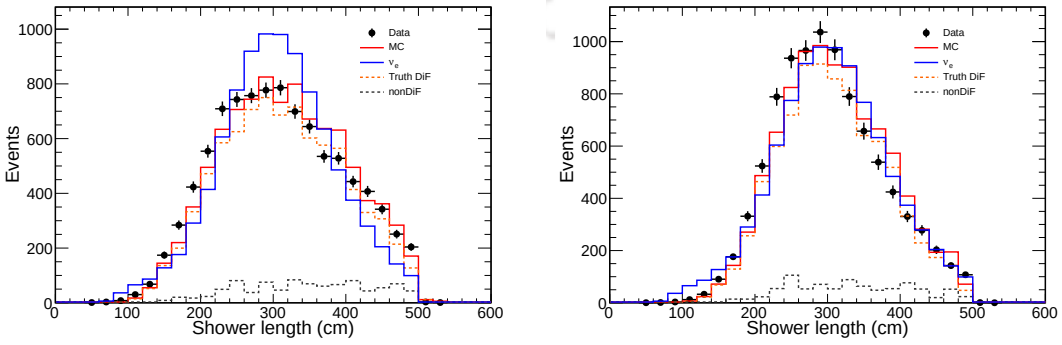

FIGURE 5.25: Shower length of DiF vs ν_e-CC before (left) and after (right) reweighing.

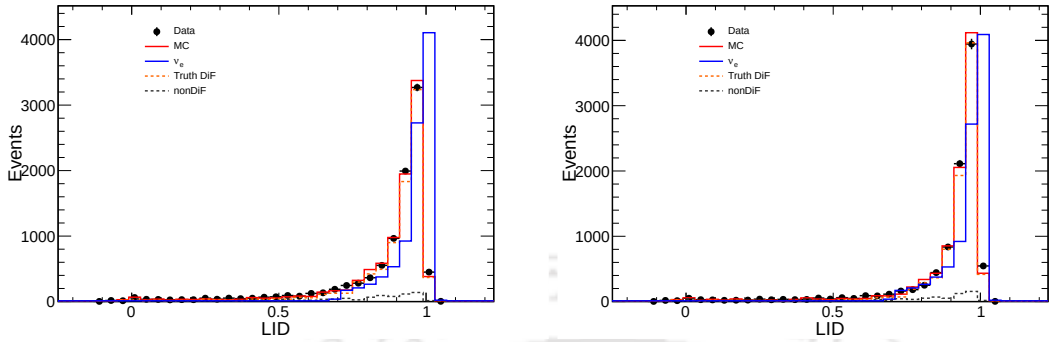

FIGURE 5.26: ANN output of DiF vs ν_e-CC before (left) and after (right) reweighing.

5.5 Signal Selection Efficiency

DiF shower can be used to determine the systematics associated with the performance of the detector. DiF shower selection efficiencies are calculated in X , Y and Z dimensions of the detector. Ideally it should be flat across the detector. Any slope in the detection efficiency will be counted as the systematics associated with the detector performance.

5.5.1 Signal selection Efficiency for LID >0.7

In this subsection, selection efficiencies are plotted for DiF shower having PID value >0.7. Fig.[5.27-5.29] show the detector efficiency as a function of X,Y and Z. The data and MC ratio are also shown in the figures on the right. The results are constrained by the limited statistics but the agreement between data and MC in terms of detector efficiency as a function of X, Y and Z is well within 4%.

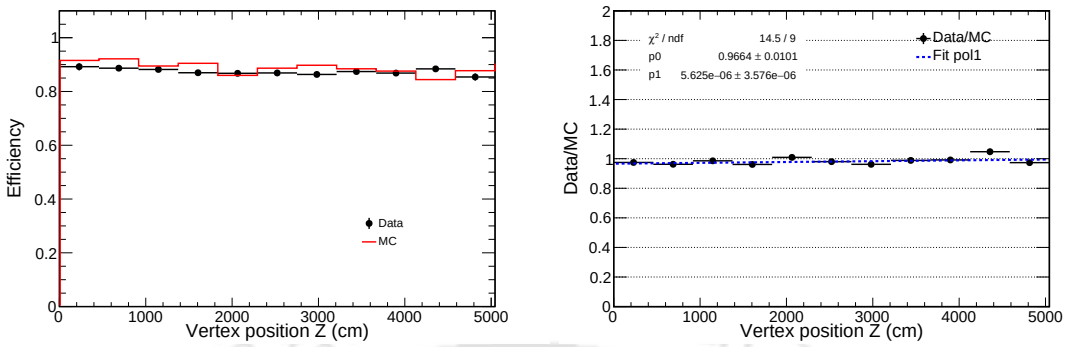

FIGURE 5.29: Signal selection efficiency in Z direction and data/MC comparison for LID >0.7.

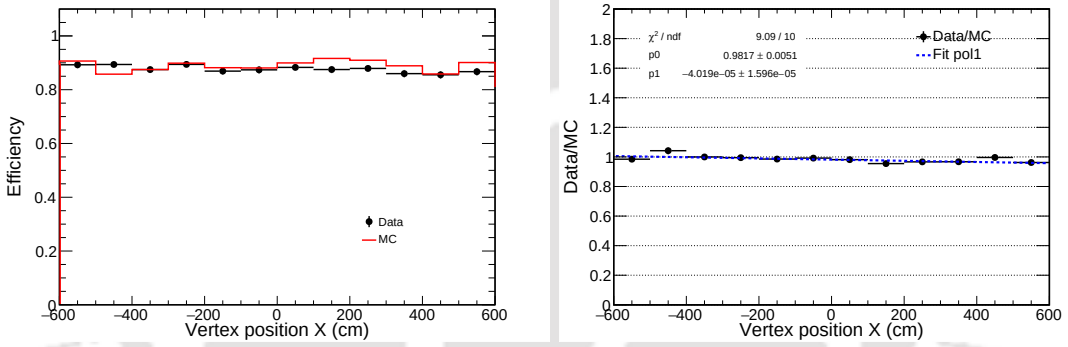

FIGURE 5.27: Signal selection efficiency in X direction and data/MC comparison for LID >0.7.

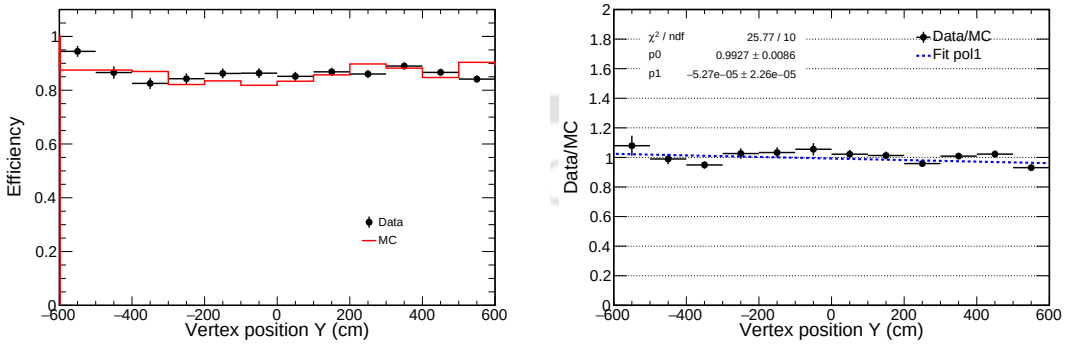

FIGURE 5.28: Signal selection efficiency in Y direction and data/MC comparison for LID >0.7.

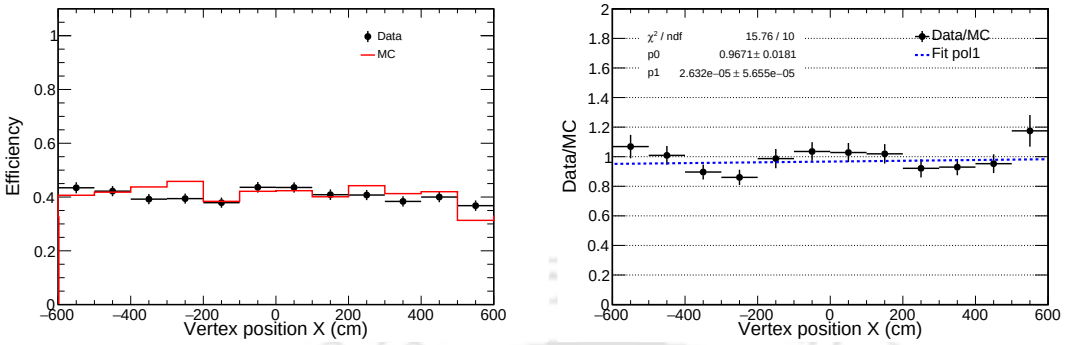

FIGURE 5.30: Signal selection efficiency in X direction and data/MC comparison for LID >0.9.

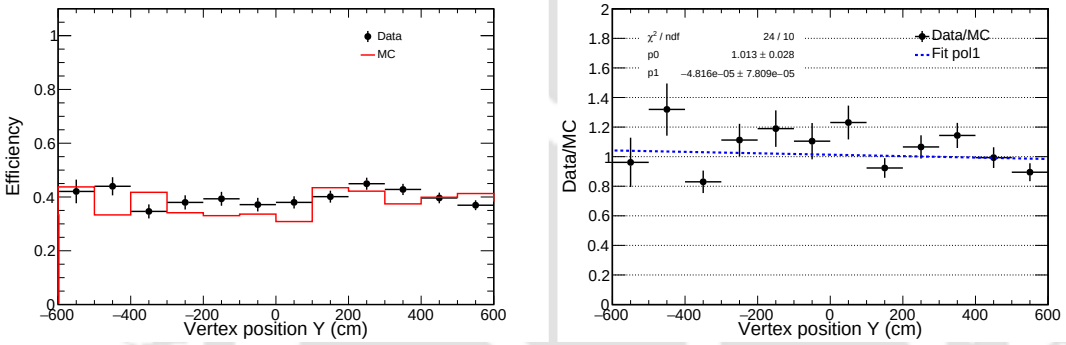

FIGURE 5.31: Signal selection efficiency in Y direction and data/MC comparison for LID >0.9.

5.5.2 Signal Selection Efficiency for LID >0.9

In this subsection, selection efficiencies are plotted for DiF shower having PID value >0.9. Fig.[5.30-5.32] show the detector efficiency as a function of X,Y and Z. The data and MC ratio are also shown in the figures on the right. The results are constrained by the limited statistics but the agreement between data and MC in terms of detector efficiency as a function of X, Y and Z is well within 4%..

5.5.3 Reweighed Signal Selection Efficiency for LID >0.7

In this subsection, selection efficiencies are plotted for reweighed DiF shower having PID value >0.7. Fig.[5.33-5.35] show the detector efficiency as a function of X,Y and Z. The data and MC ratio are also shown in the figures on the right. The results are constrained

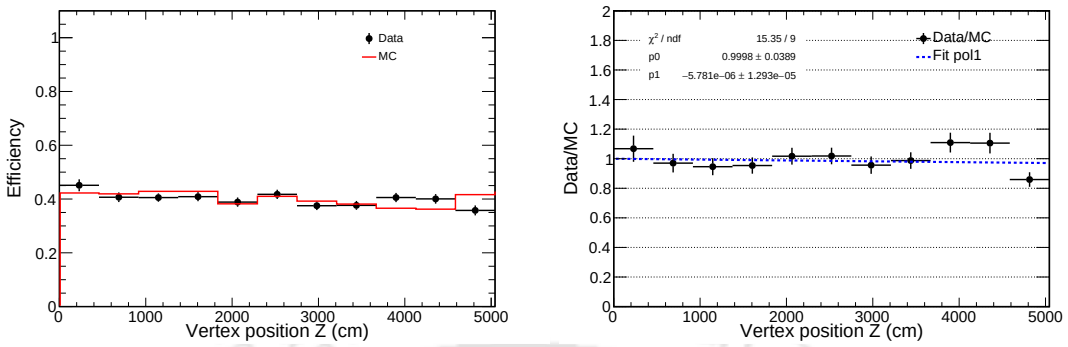

FIGURE 5.32: Signal selection efficiency in Z direction and data/MC comparison for LID >0.9.

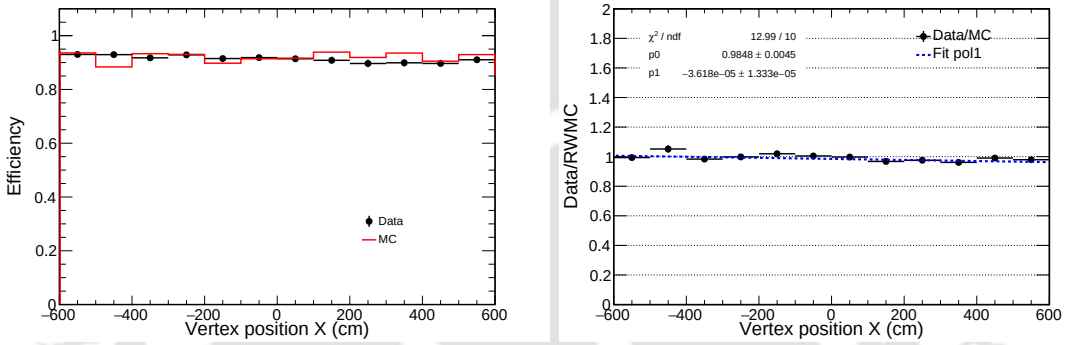

FIGURE 5.33: Signal selection efficiency in X direction and data/MC comparison of reweighed sample for LID >0.7.

by the limited statistics but the agreement between data and MC in terms of detector efficiency as a function of X, Y and Z is well within 2%.

5.5.4 Reweighed Signal Selection Efficiency for LID >0.9

In this subsection, selection efficiencies are plotted for reweighed DiF shower having PID value >0.9. Fig.[5.36-5.38] show the detector efficiency as a function of X,Y and Z. The data and MC ratio are also shown in the figures on the right. The results are constrained by the limited statistics but the agreement between data and MC in terms of detector efficiency as a function of X, Y and Z is within 4%.

Table 5.3 summarises the data and MC agreement for signal selection efficiency studies. The disagreement will be assigned as a systematic on signal selection efficiency.

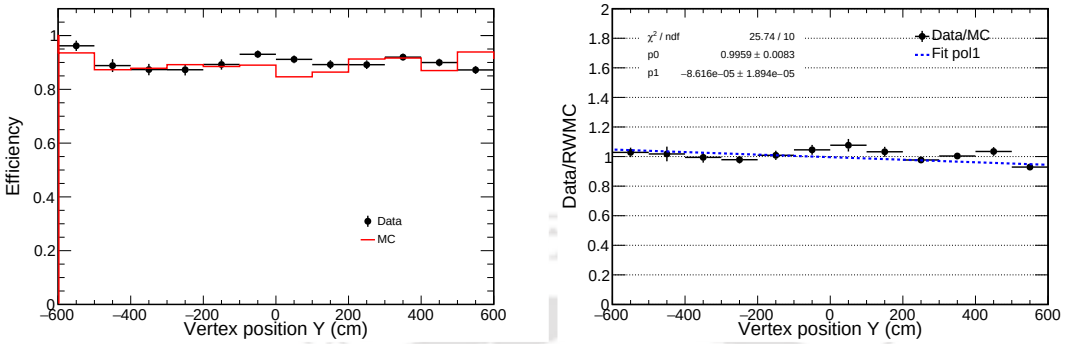

FIGURE 5.34: Signal selection efficiency in Y direction and data/MC comparison of reweighed sample for LID >0.7.

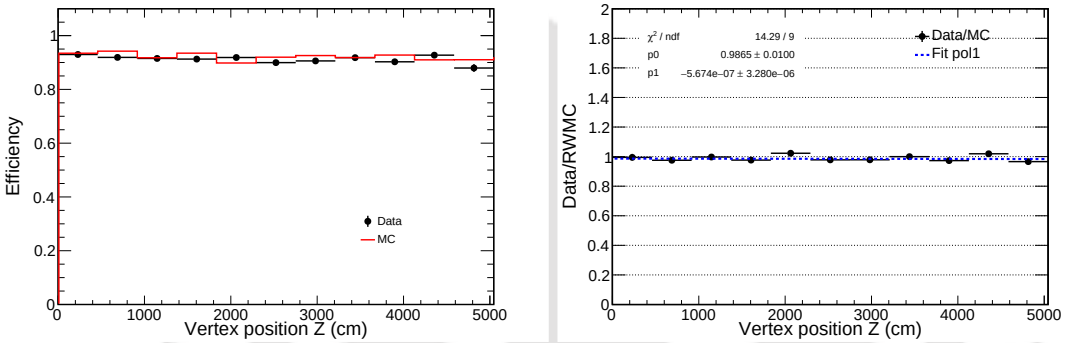

FIGURE 5.35: Signal selection efficiency in Z direction and data/MC comparison of reweighed sample for LID >0.7.

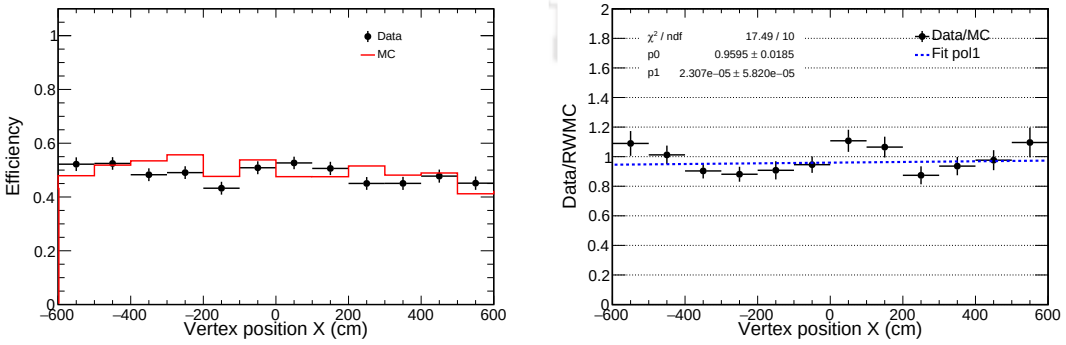

FIGURE 5.36: Signal selection efficiency in X direction and data/MC comparison of reweighed sample for LID >0.9.

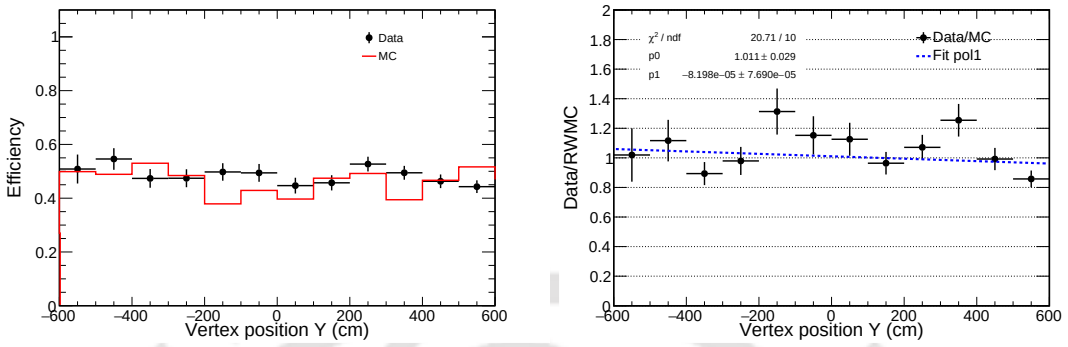

FIGURE 5.37: Signal selection efficiency in Y direction and data/MC comparison of reweighed sample for LID >0.9.

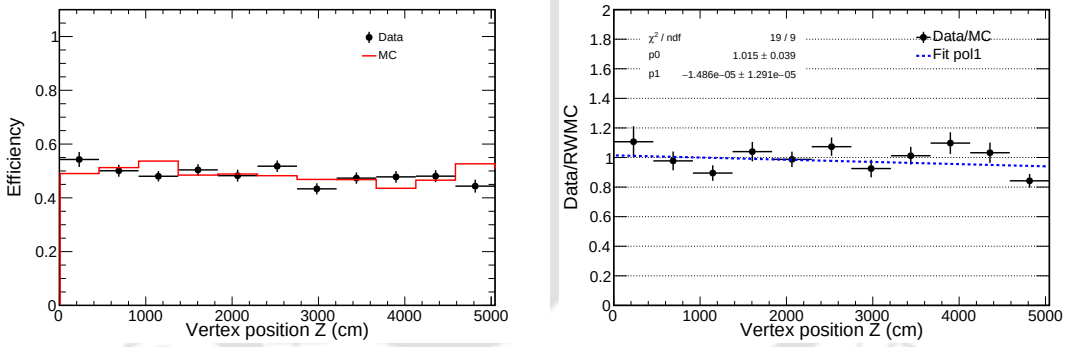

FIGURE 5.38: Signal selection efficiency in Z direction and data/MC comparison of reweighed sample for LID >0.9.

Data/MC agreement	LID >0.7	LID RW >0.7	LID >0.9	LID RW >0.9
X	1.83 ± 0.51%	1.52 ± 0.45%	3.29 ± 1.81%	4.05 ± 1.85%
Y	0.73 ± 0.86%	0.41 ± 0.83%	1.30 ± 2.80%	1.10 ± 2.90%
Z	3.36 ± 1.01%	1.35 ± 1.00%	0.02 ± 3.89%	1.50 ± 3.90%

TABLE 5.3: Summary of the systematic uncertainty on the performance of the far detector based on the DiF EM shower sample.

5.6 Chapter Summary

In this chapter the DiF EM shower is extracted and used as data driven method to verify the shower simulation and particle identification in NOνA. Decay in flight shower is a pure and cleanest EM shower and there is no need of removing muon MIP hits inside the shower region, whereas in the previously studied bremsstrahlung shower, purity depends on the efficiency of the muon removal algorithm. The data/MC disagreement of angular ($\cos\theta$) distribution which was present in brem, is not seen in DiF. We reweigh the DiF sample with respect to ν_e-CC energy. EM shower selection efficiencies are found to be stable and within 4%. However, since DiF is rarer than bremsstrahlung,the statistics of the DiF sample are smaller than the bremsstrahlung events, making this study statistically limited at this time.

Chapter 6

Summary and Conclusions

In this chapter, we perform a comparison study of the two EM shower samples in understanding the aforementioned detector effects and also possibility of performing similar studies for other current and generation future of neutrino experiments.

In the previous two chapters, we have studied the selection of both bremsstrahlung EM shower and muon DiF induced EM shower. The usage of these samples in the study of detector performance and performance of the particle-ID algorithm are also discussed. The Brem EM shower sample is larger by a factor of about 21 compared to the DiF EM shower sample.

Fig.6.1 shows the energy distribution for DiF EM shower and brem EM shower side by side. As we can see in both the cases, the data and MC is in very good agreement (within 2%) or less in the energy range (around 2 GeV), relevant to ν_e-CC appearance oscillation analysis.

Fig.6.2 shows the angle (w.r.t to beam) distribution for DiF EM shower and Brem EM shower. There is a clear slope in data/MC ratio for Brem EM shower, which is not found in the DiF EM shower sample.

Fig.[6.3-6.5] show the likelihood differences under certain particle hypothesis for DiF and Brem EM shower sample. Both the showers show similar behavior in these plots.

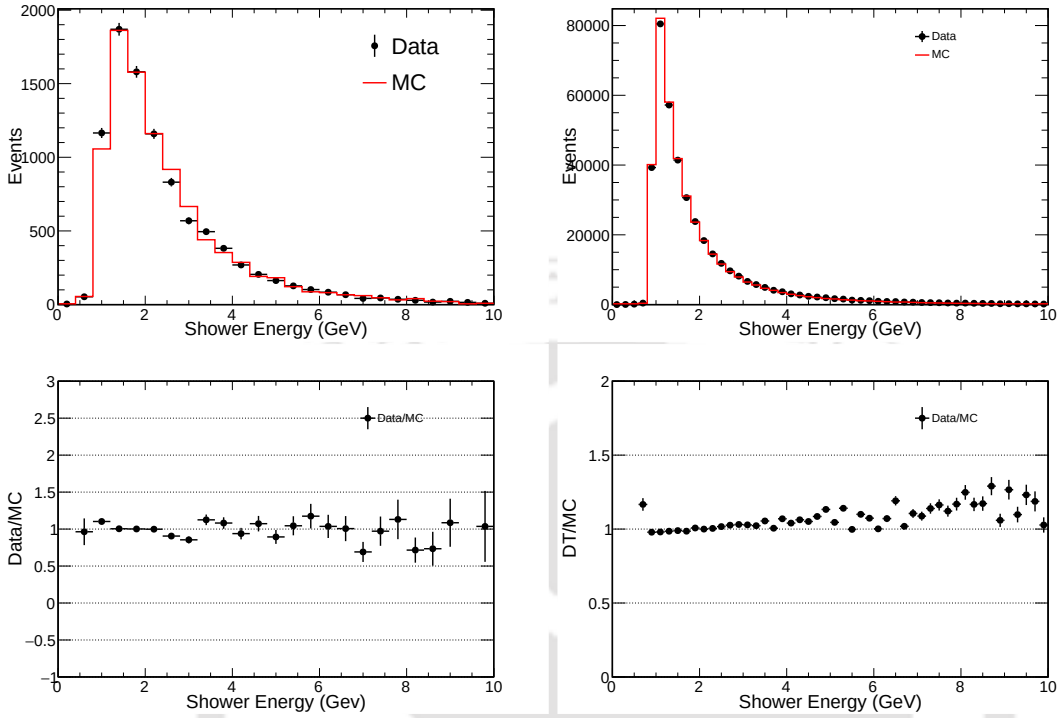

FIGURE 6.1: Left: Top plot shows the DiF data and MC comparison and bottom plot shows the ratio of data/MC. Right: Top plot shows the Brem data and MC comparison and bottom plot shows the ratio of data/MC

Fig.6.6 shows the particle-ID (LID) distribution for both DiF and Brem sample together. Both the showers events are mostly classified as ν_e-CC type. As discussed earlier, DiF is a pure electron shower whereas in case of Brem EM shower we need to remove the muon hits from the shower region which is not 100% efficient. Due to this, we expect that particle-ID, which are developed to identify ν_e-CC shower, should classify more DiF events as ν_e-CC type than that of Brem. To quantify this performance, we define:

$$R = \frac{shower\ events\ above\ LID\ >\ 0.7}{Total\ shower\ events} \tag{6.1}$$

and we calculate the R for both the shower samples to find:

$$\begin{aligned} R_{brem} &= 0.79 \\ R_{DiF} &= 0.87 \end{aligned} \tag{6.2}$$

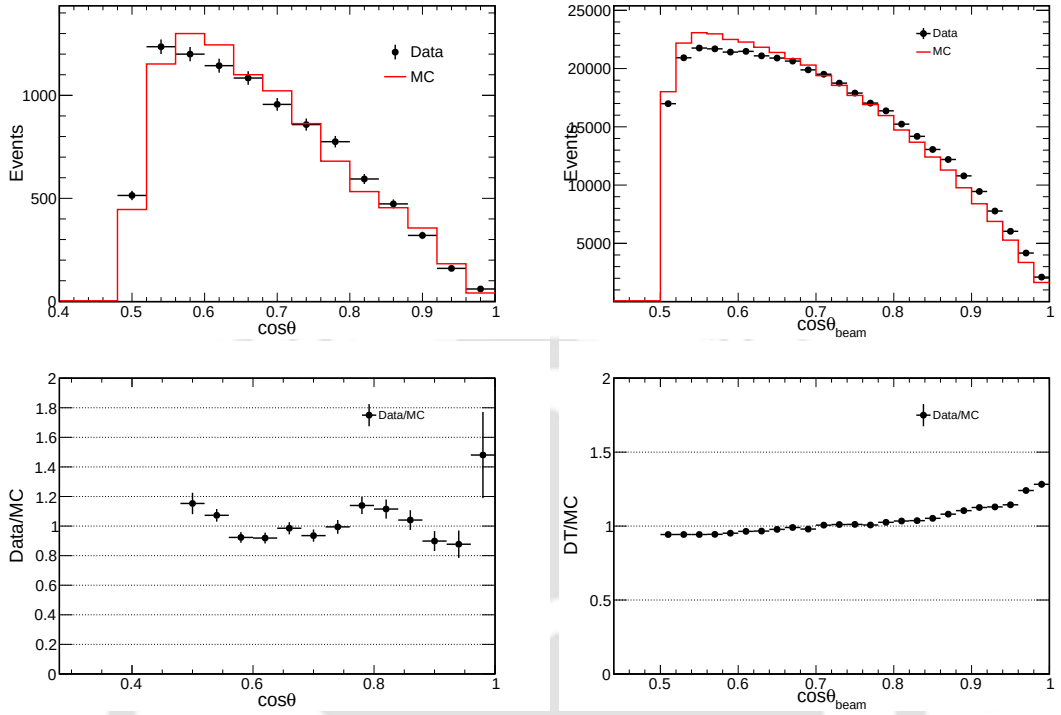

FIGURE 6.2: Left: Top plot shows the DiF data and MC comparison and bottom plot shows the ratio of data/MC. Right: Top plot shows the Brem data and MC comparison and bottom plot shows the ratio of data/MC.

It clearly indicates that performance of the DiF EM shower in imitating ν_e-CC induced EM shower is better than the Brem EM shower by about 10%. This also proves that the particle-ID algorithm developed for ν_e-CC signal are working as expected.

6.1 Comparison with ν_e and Reweighing

In this section, we compare the DiF and Brem EM showers before and after reweighing with respect to ν_e-CC sample. As mentioned earlier, the DiF sample is statistically limited and due to which we only reweigh the DiF sample with respect to ν_e-CC energy. Brem sample is reweighed with respect to both ν_e-CC energy and direction.

Fig.6.7 shows the energy distribution of all samples before and after reweighing. After reweighing, all three distribution are matching perfectly as expected.

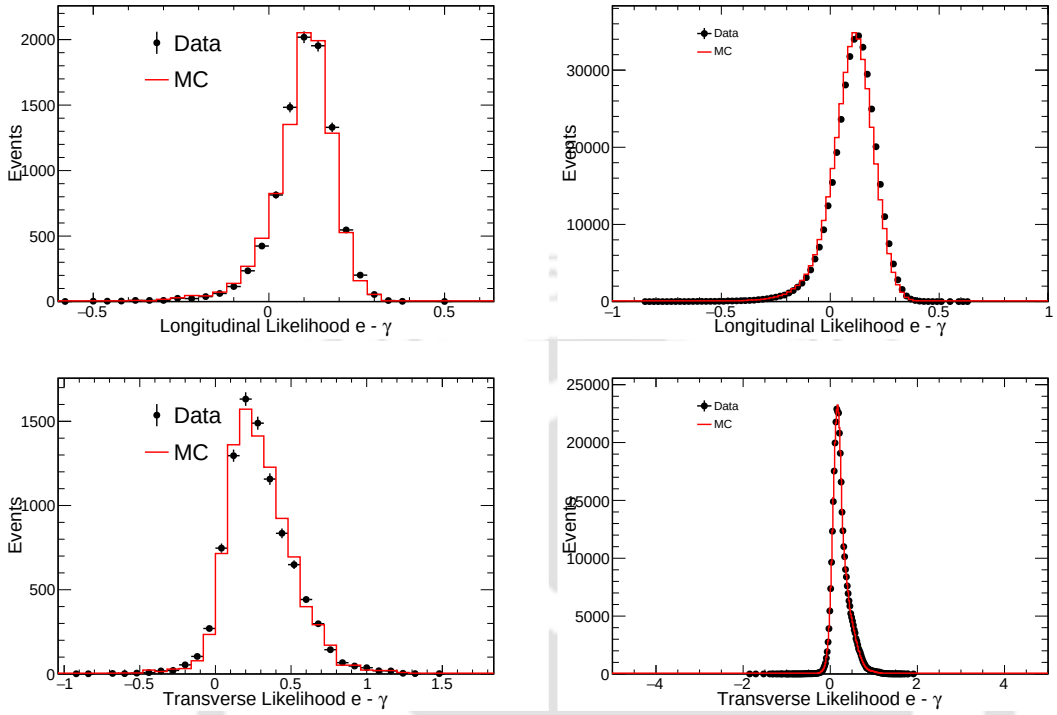

FIGURE 6.3: Left: Top plot shows the DiF longitudinal likelihood difference plot for $e-\gamma$ hypothesis. Bottom plot shows the DiF transverse likelihood difference plot for $e-\gamma$ hypothesis. Right: Top plot shows the Brem longitudinal likelihood difference plot for $e-\gamma$ hypothesis. Bottom plot shows the Brem transverse likelihood difference plot for $e-\gamma$ hypothesis

Fig.6.8 shows the angular distribution of all samples before and after reweighing. After reweighing, only Brem sample matches with ν_e sample. This is obvious as DiF is reweighed only with respect to energy.

Fig.6.9 shows the radius distribution of all samples before and after reweighing and both the showers match well with ν_e sample.

Fig.6.10 shows the length distribution of all samples before and after reweighing. We see in the figure that only DiF shower matches well with ν_e sample in shape. At this point it is important to note that, DiF, as expected, behaves more similarly to ν_e-CC sample than Brem. It is clear from Fig.6.9 and Fig.6.10 that even with one dimensional reweighing, the DiF sample is performing better than the Brem EM shower sample.

Fig.6.11 shows the particle-ID distributions of DiF and Brem along with ν_e-CC, before and after reweighing which clearly demonstrate the performance of the PID algorithm in

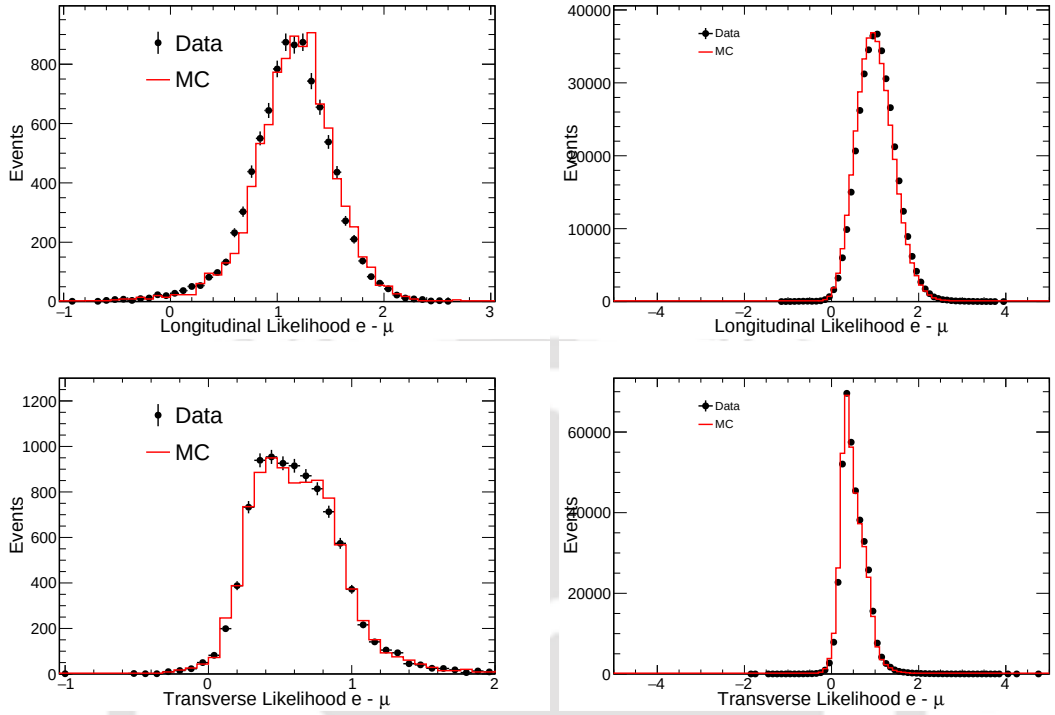

FIGURE 6.4: Left: Top plot shows the DiF longitudinal likelihood difference plot for $e-\mu$ hypothesis. Bottom plot shows the DiF transverse likelihood difference plot for $e-\mu$ hypothesis. Right: Top plot shows the Brem longitudinal likelihood difference plot for $e-\mu$ hypothesis. Bottom plot shows the Brem transverse likelihood difference plot for $e-\mu$ hypothesis

NOνA is selecting ν_e-CC interaction.

6.2 Detector Efficiency as a function of Vertex Position

In this section, we study the detector efficiency as a function of vertex position, X, Y, Z, using both DiF and Brem EM shower samples and as a function of different LID cuts. Fig.[6.12-6.14] show the efficiency of the NOνA far detector as a function of vertex position, X, Y, Z and for LID > 0.7 before reweighing and Fig.[6.18-6.20] show the same plots after reweighing.

Similarly, Fig.[6.15-6.17] show the efficiency of the NOνA far detector as a function of vertex position, X, Y, Z and for LID > 0.9 before reweighing and Fig.[6.21-6.23] show the same plots after reweighing.

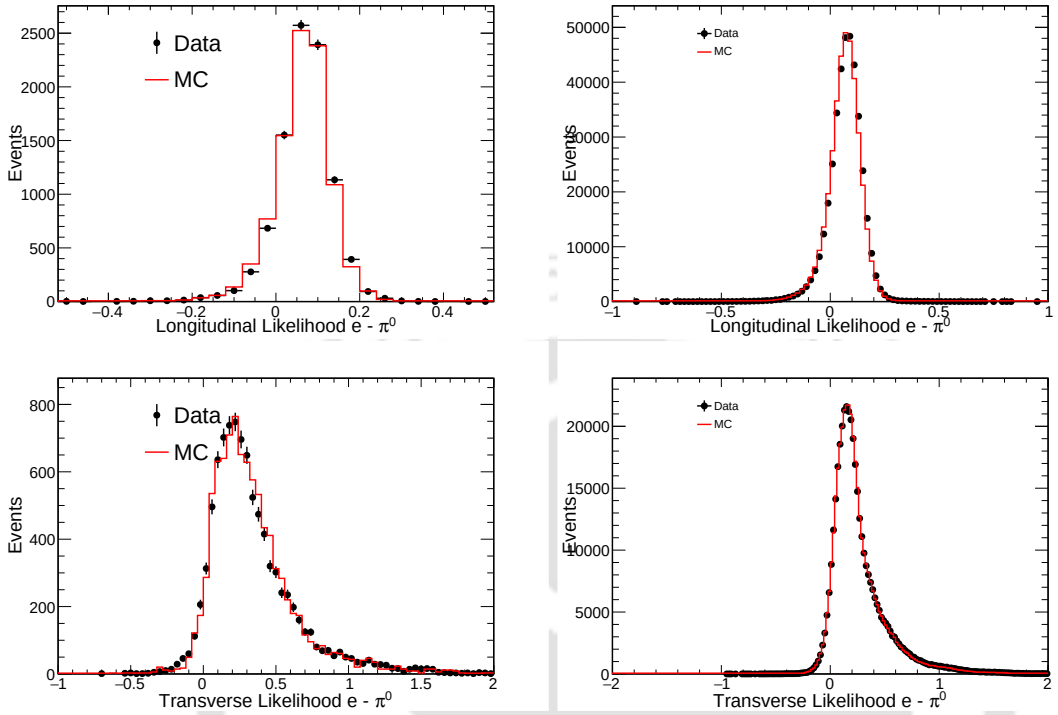

FIGURE 6.5: Left: Top plot shows the DiF longitudinal likelihood difference plot for $e - \pi^0$ hypothesis. Bottom plot shows the DiF transverse likelihood difference plot for $e - \pi^0$ hypothesis. Right: Top plot shows the Brem longitudinal likelihood difference plot for $e - \pi^0$ hypothesis. Bottom plot shows the Brem transverse likelihood difference plot for $e - \pi^0$ hypothesis.

Data/MC agreement	LID >0.7	LID RW >0.7	LID >0.9	LID RW >0.9
X	1.83 ± 0.51%	1.52 ± 0.45%	3.29 ± 1.81%	4.05 ± 1.85%
Y	0.73 ± 0.86%	0.41 ± 0.83%	1.30 ± 2.80%	1.10 ± 2.90%
Z	3.36 ± 1.01%	1.35 ± 1.00%	0.02 ± 3.89%	1.50 ± 3.90%

TABLE 6.1: Summary of the systematic uncertainty on the performance of the far detector based on the DiF EM shower sample.

Table 6.1 summarizes data and MC differences in the study of the detector efficiency as a function of X, Y and Z positions in the detector using the DiF EM shower sample. These differences can be assigned as a systematic uncertainty associated with the detector performance. It is clearly seen that the study in limited by statistics currently. By considering the error on the ratio as well, as conservative estimate of about 4% can be assigned as systematic uncertainty on the detector performance.

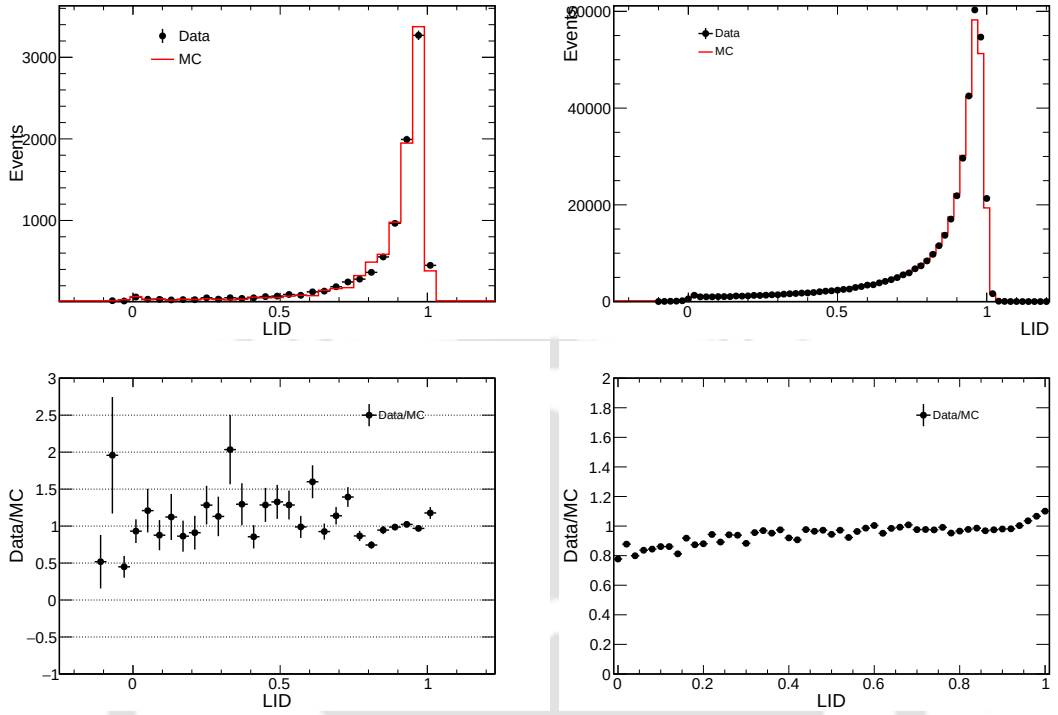

FIGURE 6.6: Left: Top plot shows the DiF data and MC comparison and bottom plot shows the ratio of data/MC. Right: Top plot shows the Brem data and MC comparison and bottom plot shows the ratio of data/MC

Data/MC agreement	LID >0.7	LID RW >0.7	LID >0.9	LID RW >0.9
X	0.10 ± 0.00%	0.78 ± 0 .17%	0.10 ± 0.10%	0.80 ± 0.19%
Y	0.20 ± 0.00%	0.80 ± 0.20%	0.10 ± 0.10%	0.44 ± 0.19%
Z	0.10 ± 0.10%	3.43 ± 0.19%	0.79 ± 0.12%	1.45 ± 0.33%

TABLE 6.2: Summary of the systematic uncertainty on the performance of the far detector based on the brem EM shower sample.

Table 6.2 summarizes data and MC differences in the study of the detector efficiency as a function of X, Y and Z positions in the detector using the DiF EM shower sample. These differences can be assigned as a systematic uncertainty associated with the detector performance. By considering the error on the ratio as well, as conservative estimate of about 4% can be assigned as systematic uncertainty on the detector performance, based on Brem EM shower sample, as well.

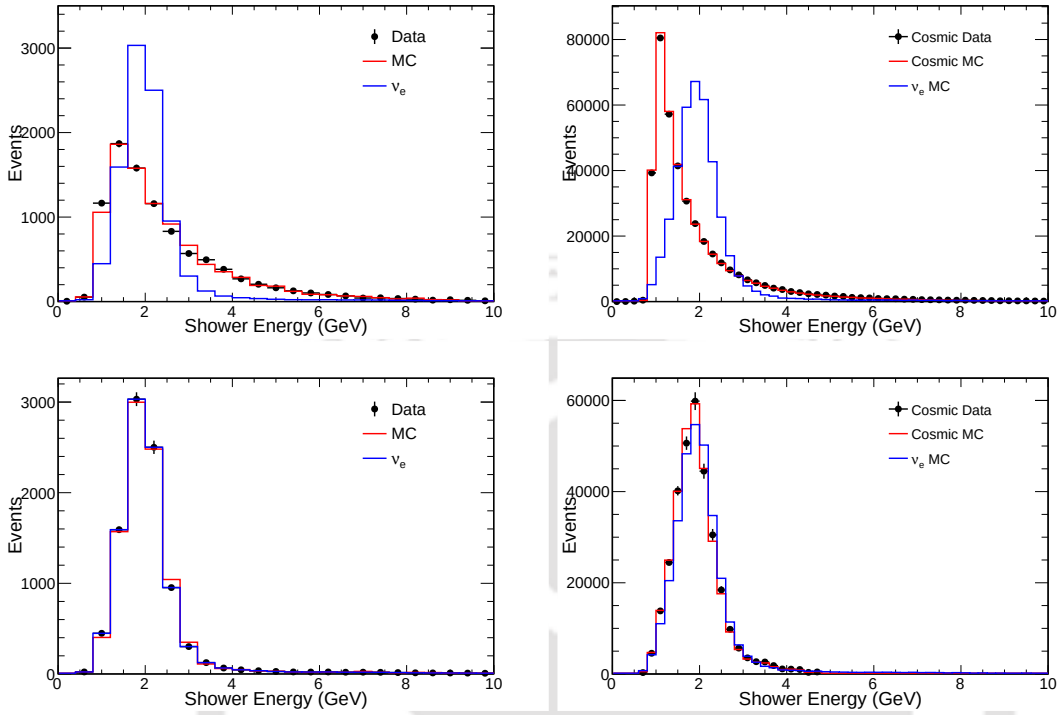

FIGURE 6.7: Left: Top plot shows the DiF data and MC comparison along with ν_e MC and bottom plot shows the DiF reweighted w.r.t energy of ν_e. Right: Top plot shows the Brem data and MC comparison and ν_e and bottom plot shows the Brem after reweighed w.r.t ν_e energy and angle.

6.3 Conclusions

We use Cosmic Muon-Removal algorithm for finding and isolating EM showers from data and MC in the NOνA far detector. The shower digits are put into standard NOνA ν_e-CC reconstruction and PID algorithms. A comparison using first analysis data and MC dataset shows consistent distributions, validating the EM shower modeling and reconstruction. The PID efficiencies as functions of vertex position are calculated and observed to have good data/MC agreement. It shows the calibration effects are well controlled.

An alternate EM shower, the decay in flight EM shower is successfully extracted from cosmic muons which is a purer electron initiated EM shower that also mimics the EM shower initiated by ν_e-CC electron. The shower modeling is verified by using the data and MC samples and is found to be in good agreement. The PIDs are also benchmarked as most of the decay in flight EM showers are found to be selected as ν_e-CC types.

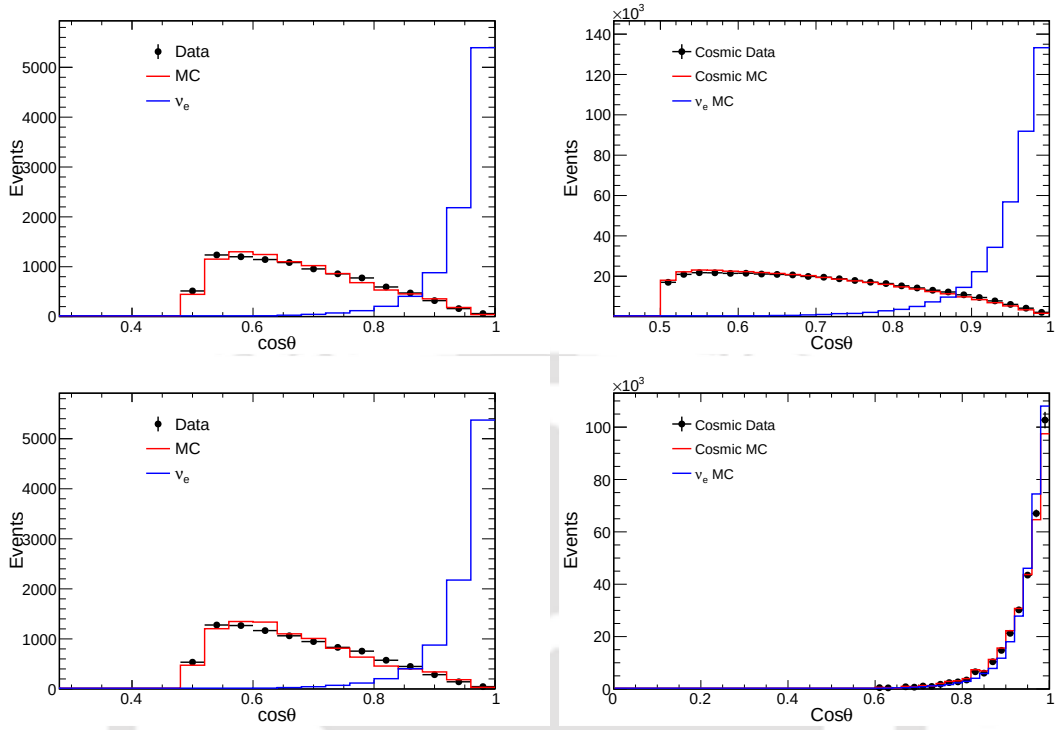

FIGURE 6.8: Left: Top plot shows the DiF data and MC comparison along with ν_e MC and bottom plot shows the DiF reweighed w.r.t energy of ν_e. Right: Top plot shows the Brem data and MC comparison and ν_e and bottom plot shows the Brem after reweighted w.r.t ν_e energy and angle.

Reweighing technique is used to mitigate the difference between cosmic EM shower and ν_e-CC shower as major difference between two showers is because of difference in energy and direction. Decay in flight shower is reweighed with respect to ν_e energy and it is found that after reweighing it behaves more like a ν_e-CC shower. The PID efficiencies are plotted as a function of vertex position and are found to be in good agreement between data and simulations which implies callibration effects to be in well controlled. The muon decay in flight initiated EM shower sample is found to be a better one compared to the Brem EM shower sample. However, the DiF EM shower sample is statistically limited currently.

Since every ν_e-appearance or ν_e-disappearance neutrino oscillation experiment in sub GeV range would have EM shower as its signal, therefore, similar studies immediately find its applicability in such experiments as a data driven method to validate the shower

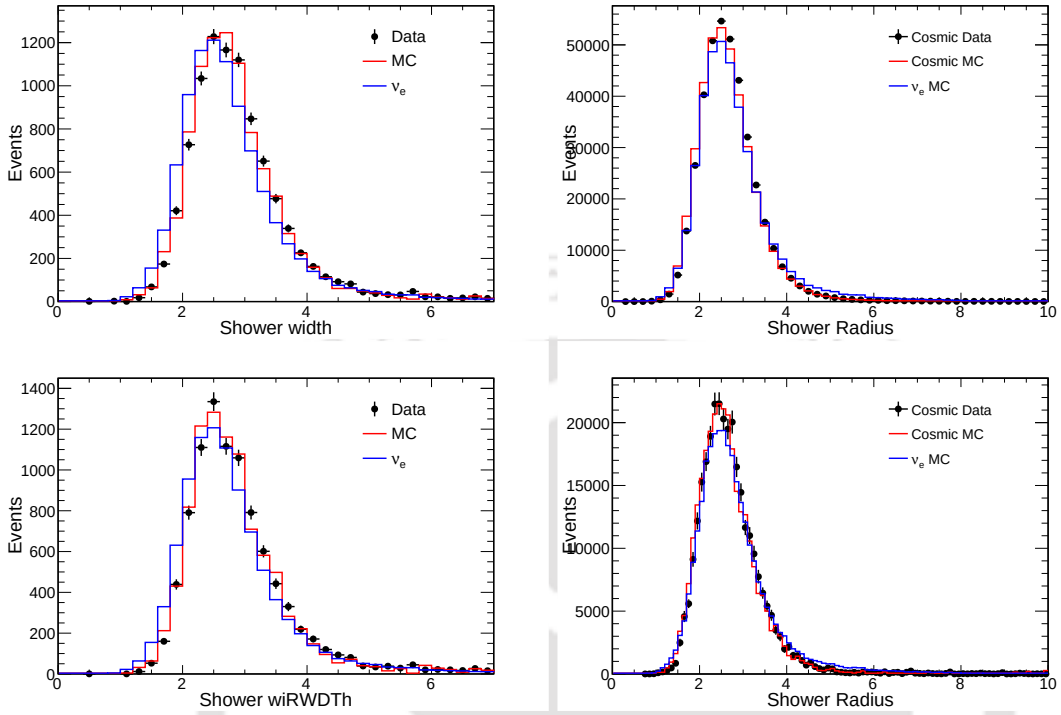

FIGURE 6.9: Left: Top plot shows the DiF data and MC comparison along with ν_e MC and bottom plot shows the DiF reweighed w.r.t energy of ν_e. Right: Top plot shows the Brem data and MC comparison and ν_e and bottom plot shows the Brem after reweighed w.r.t ν_e energy and angle.

modelling, understand the performance of the particle-ID algorithms and detector efficiency.

Using these samples, we also studies the detector performance as a function of EM shower position and data and MC is found to be in good agreement. Based on these studies, we recommend assigning an overall systematic uncertainty of about 4% on the detector performance. The largest systematic effect on the ν_e appearance analysis comes from cross-section models and it is estimated to be 7.7% on ν_e signal and 8.6% on beam background [98] and the corresponding statistical uncertainty is 15% and 22% on ν_e signal and beam background respectively.

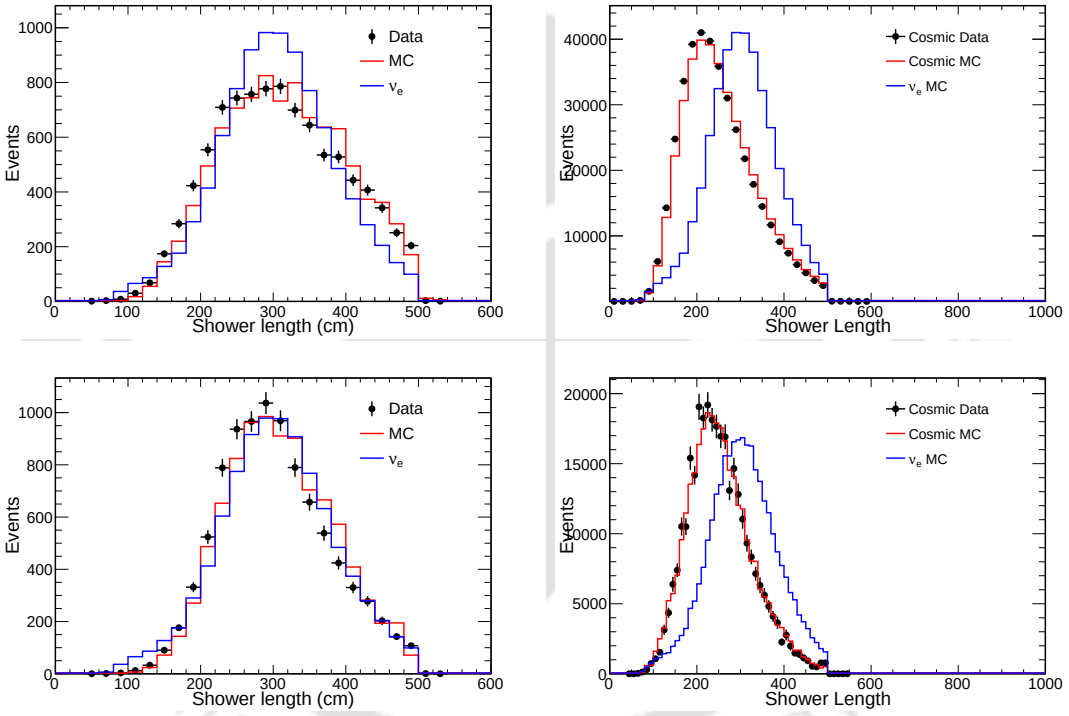

FIGURE 6.10: Left: Top plot shows the DiF data and MC comparison along with nu_e MC and bottom plot shows the DiF reweighed w.r.t energy of ν_e. Right: Top plot shows the Brem data and MC comparison and ν_e and bottom plot shows the Brem after reweighed w.r.t ν_e energy and angle.

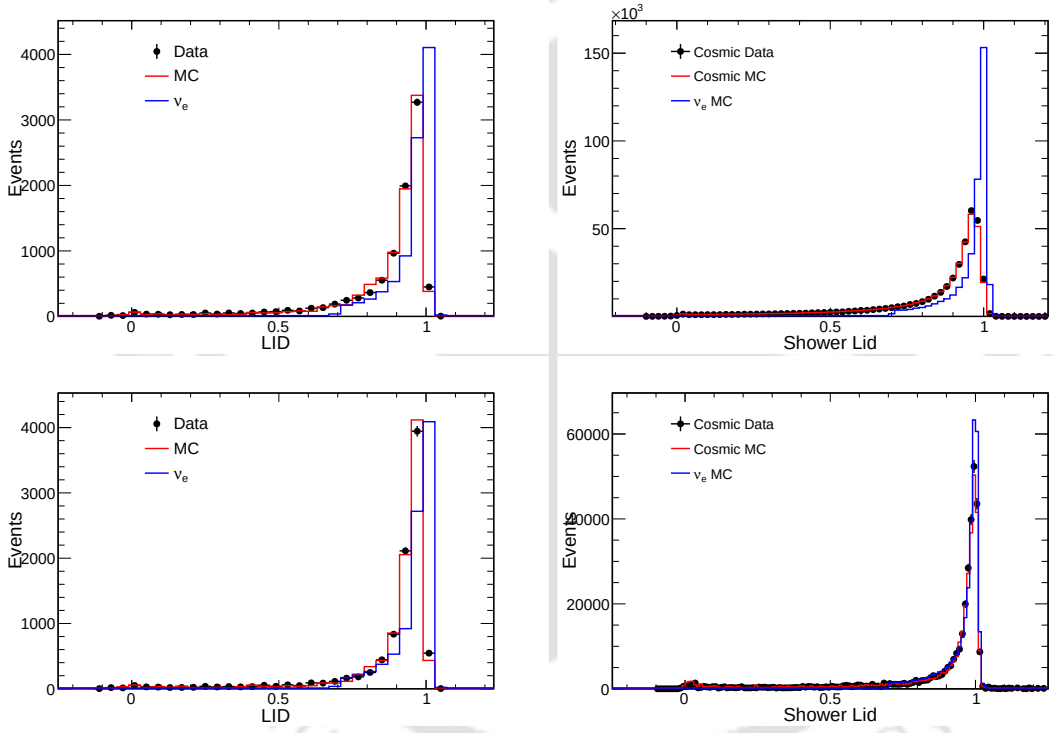

FIGURE 6.11: Left: Top plot shows the DiF data and MC comparison along with ν_e MC and bottom plot shows the DiF reweighed w.r.t energy of ν_e. Right: Top plot shows the Brem data and MC comparison and ν_e and bottom plot shows the Brem after reweighed w.r.t ν_e energy and angle..

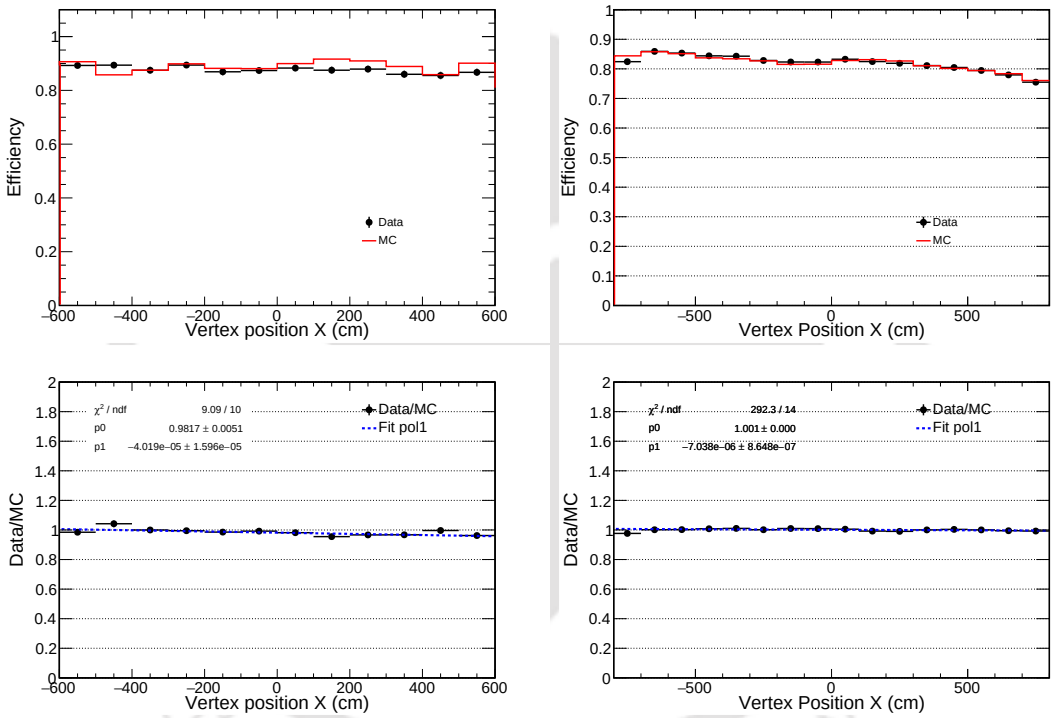

FIGURE 6.12: Left: Top plot shows the DiF data and MC comparison of detector efficiency, for LID >0.7, as a function of vertex position X and bottom plot shows the ratio of data/MC. Right: Top plot shows the Brem data and MC comparison of detector efficiency, for LID >0.7, as a function of vertex position X and bottom plot shows the ratio of data/MC.

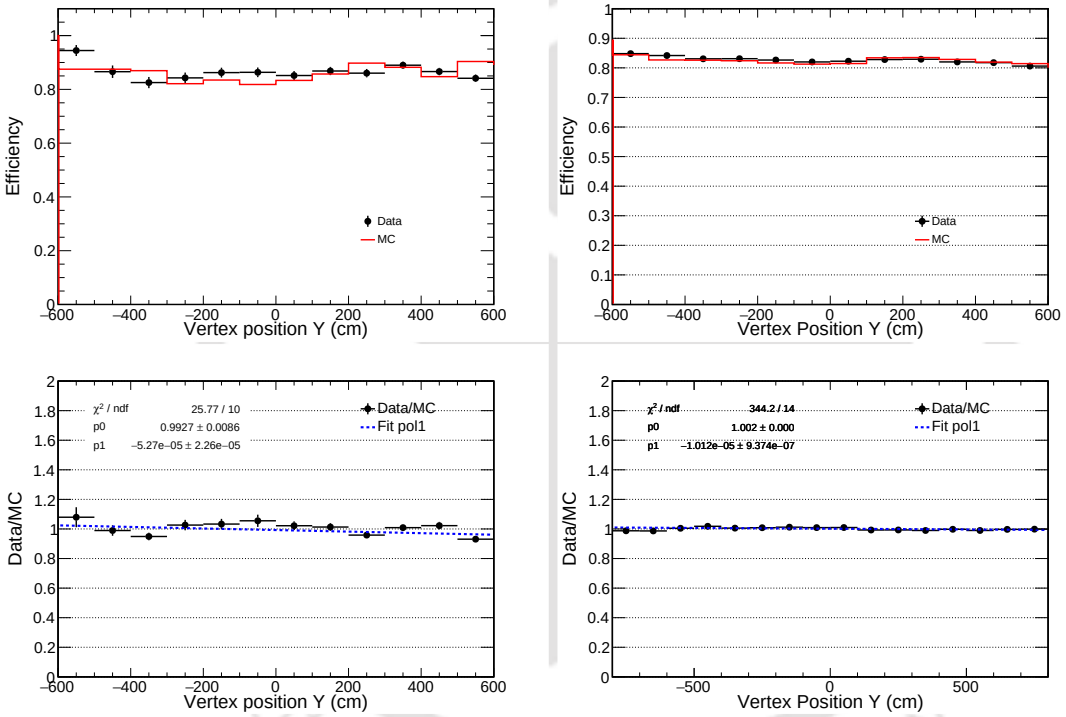

FIGURE 6.13: Left: Top plot shows the DiF data and MC comparison of detector efficiency, for LID >0.7, as a function of vertex position Y and bottom plot shows the ratio of data/MC. Right: Top plot shows the Brem data and MC comparison of detector efficiency, for LID >0.7, as a function of vertex position Y and bottom plot shows the ratio of data/MC.

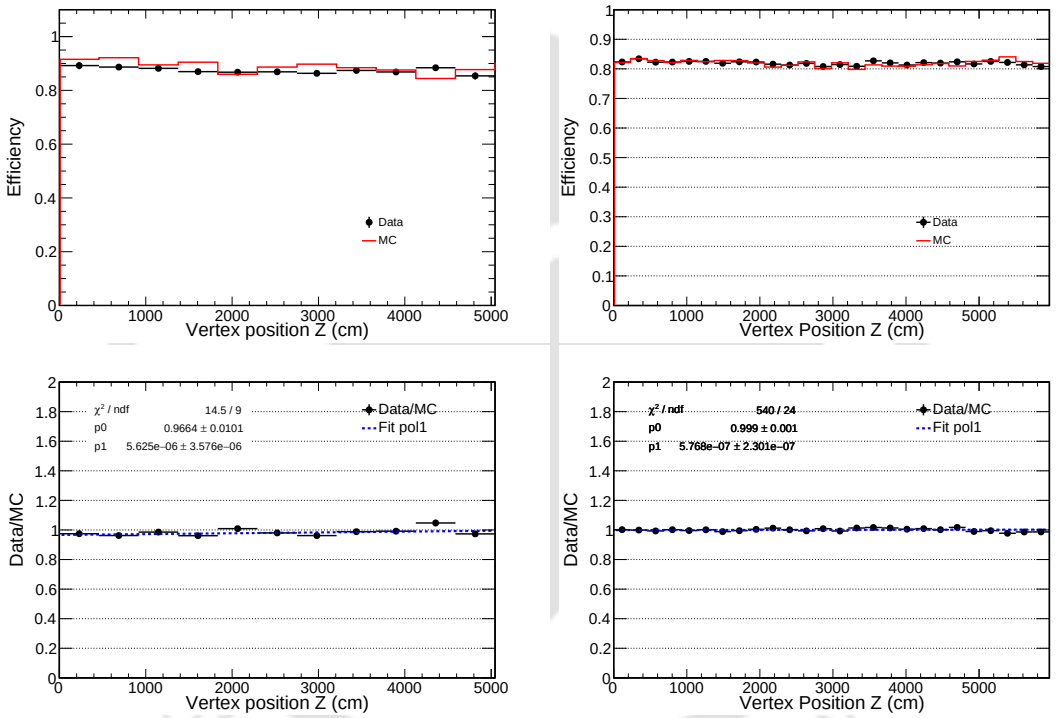

FIGURE 6.14: Left: Top plot shows the DiF data and MC comparison of detector efficiency, for LID >0.7, as a function of vertex position Z and bottom plot shows the ratio of data/MC. Right: Top plot shows the Brem data and MC comparison of detector efficiency, for LID >0.7, as a function of vertex position Z and bottom plot shows the ratio of data/MC.

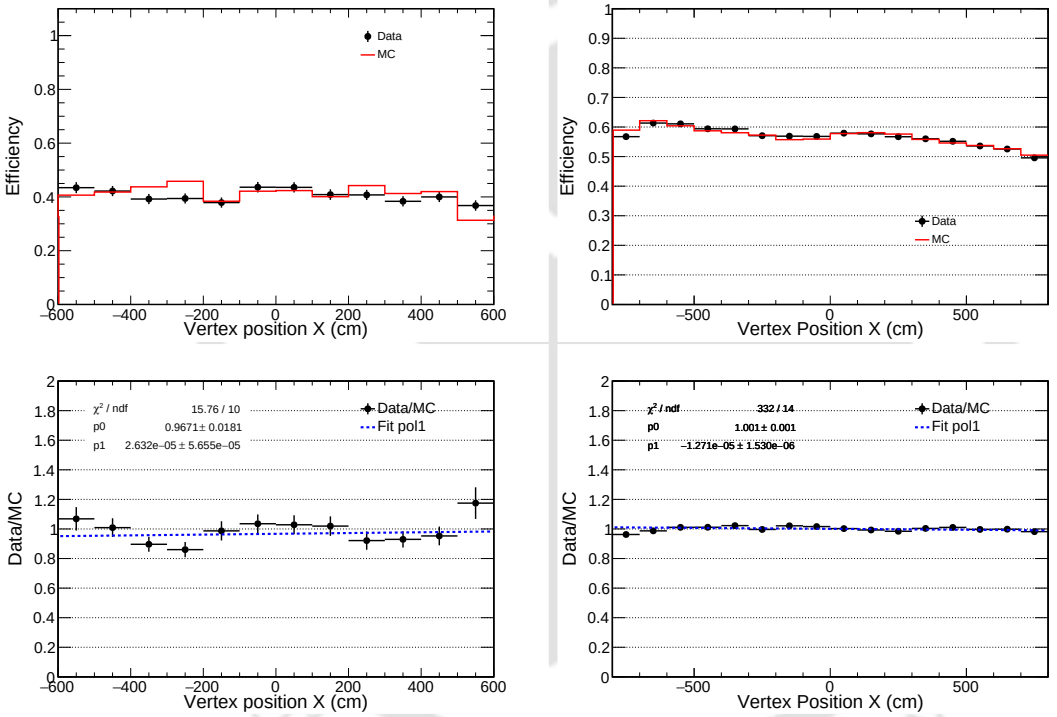

FIGURE 6.15: Left: Top plot shows the DiF data and MC comparison of detector efficiency, for LID >0.9, as a function of vertex position X and bottom plot shows the ratio of data/MC. Right: Top plot shows the Brem data and MC comparison of detector efficiency, for LID >0.9, as a function of vertex position X and bottom plot shows the ratio of data/MC.

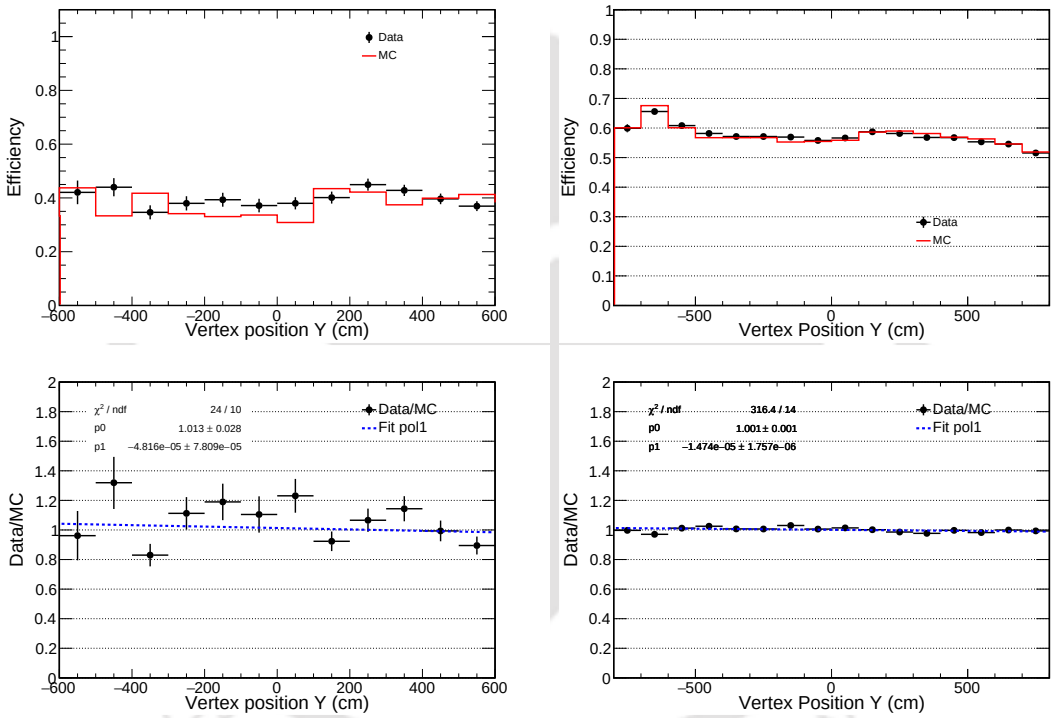

FIGURE 6.16: Left: Top plot shows the DiF data and MC comparison of detector efficiency, for LID >0.9, as a function of vertex position Y and bottom plot shows the ratio of data/MC. Right: Top plot shows the Brem data and MC comparison of detector efficiency, for LID >0.9, as a function of vertex position Y and bottom plot shows the ratio of data/MC.

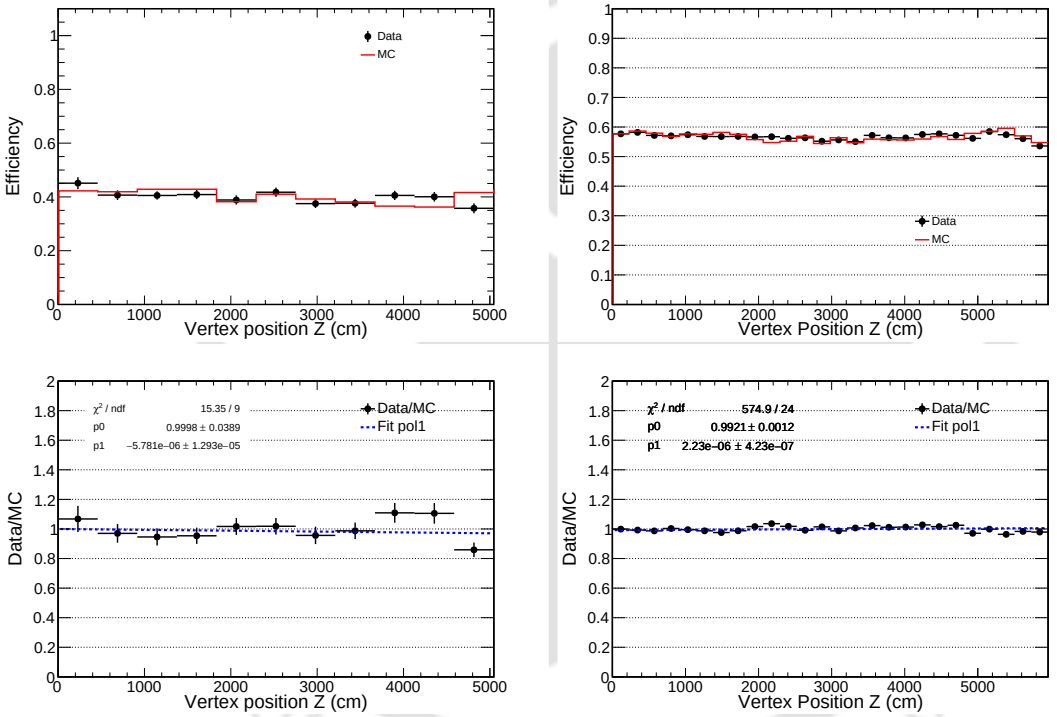

FIGURE 6.17: Left: Top plot shows the DiF data and MC comparison of detector efficiency, for LID >0.9, as a function of vertex position Z and bottom plot shows the ratio of data/MC. Right: Top plot shows the Brem data and MC comparison of detector efficiency, for LID >0.9, as a function of vertex position Z and bottom plot shows the ratio of data/MC.

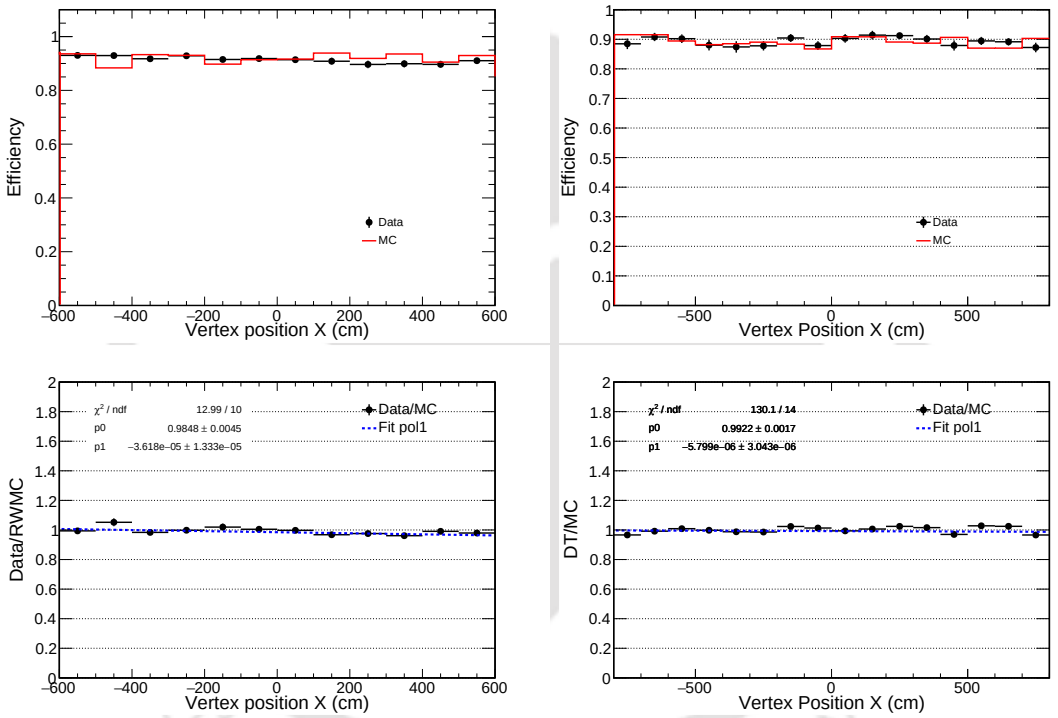

FIGURE 6.18: Left: Top plot shows the reweighed DiF data and MC comparison of detector efficiency, for LID >0.7, as a function of vertex position X and bottom plot shows the ratio of data/MC. Right: Top plot shows the reweighed Brem data and MC comparison of detector efficiency, for LID >0.7, as a function of vertex position X and bottom plot shows the ratio of data/MC.

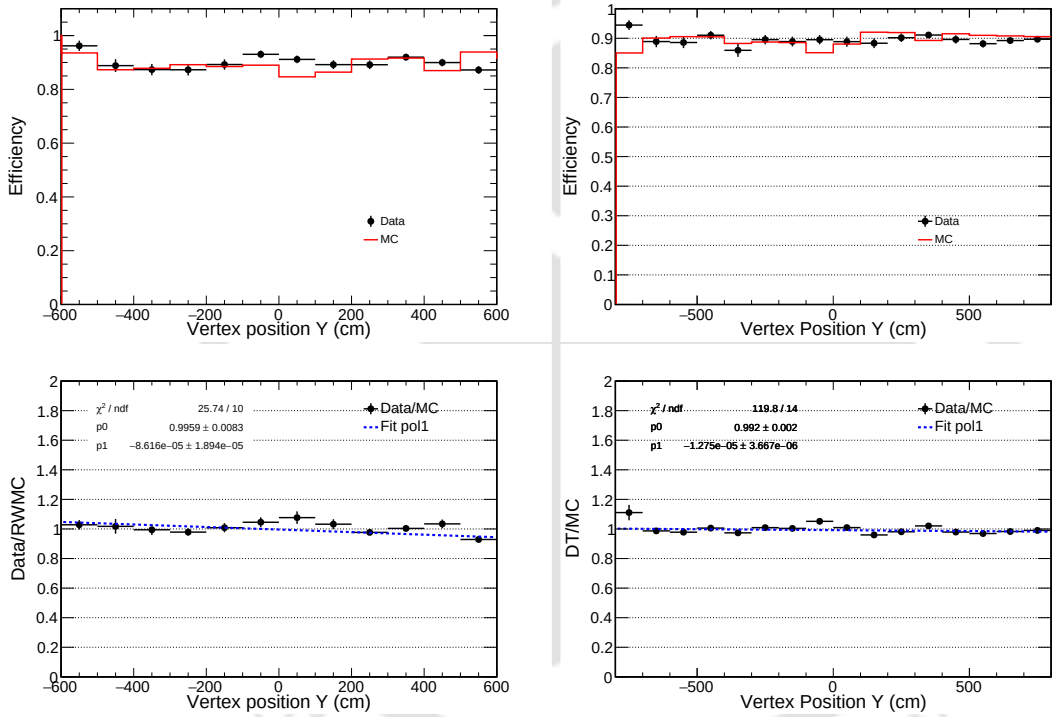

FIGURE 6.19: Left: Top plot shows the reweighed DiF data and MC comparison of detector efficiency, for LID >0.7, as a function of vertex position Y and bottom plot shows the ratio of data/MC. Right: Top plot shows the reweighed Brem data and MC comparison of detector efficiency, for LID >0.7, as a function of vertex position Y and bottom plot shows the ratio of data/MC.

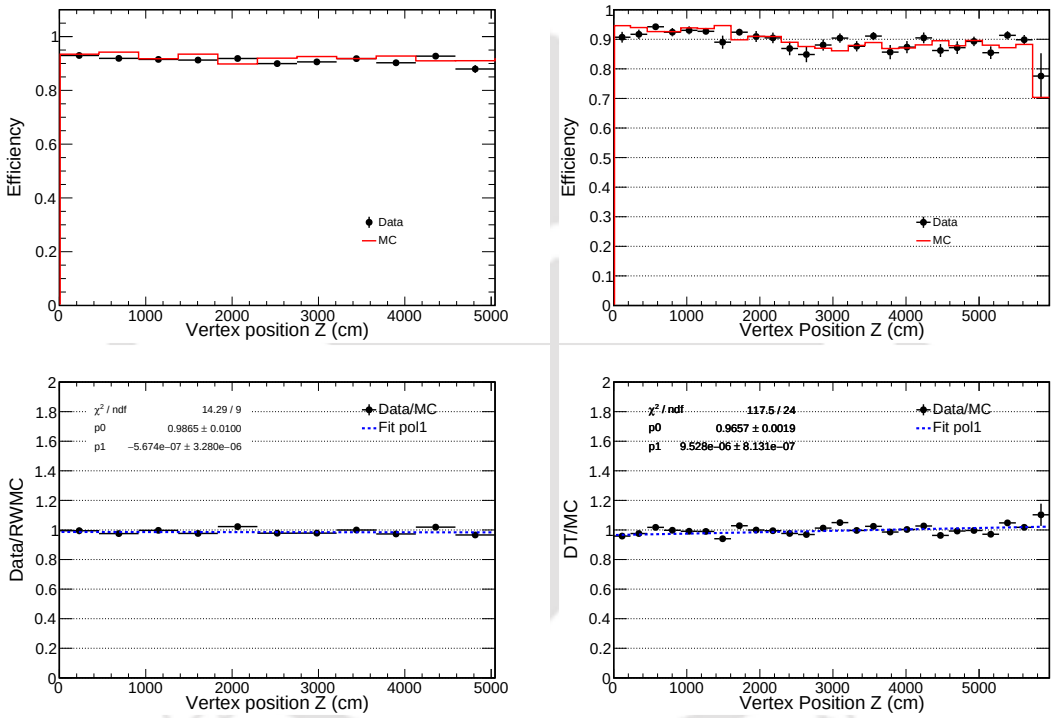

FIGURE 6.20: Left: Top plot shows the reweighed DiF data and MC comparison of detector efficiency, for LID >0.7, as a function of vertex position Z and bottom plot shows the ratio of data/MC. Right: Top plot shows the reweighed Brem data and MC comparison of detector efficiency, for LID >0.7, as a function of vertex position Z and bottom plot shows the ratio of data/MC.

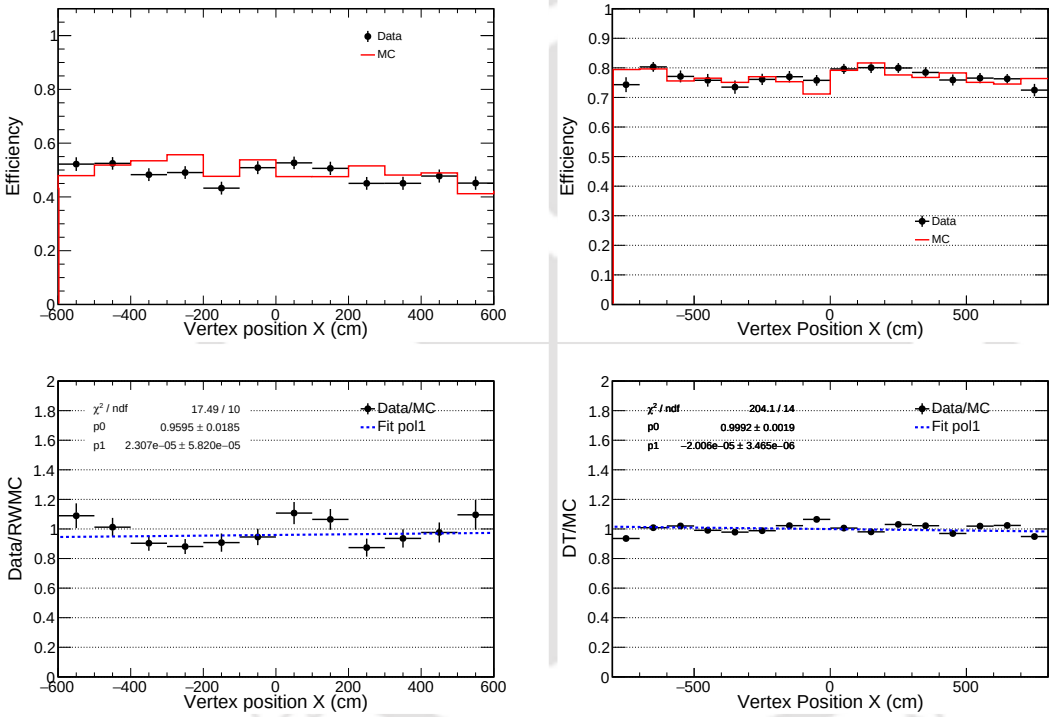

FIGURE 6.21: Left: Top plot shows the reweighed DiF data and MC comparison of detector efficiency, for LID >0.9, as a function of vertex position X and bottom plot shows the ratio of data/MC. Right: Top plot shows the reweighed Brem data and MC comparison of detector efficiency, for LID >0.9, as a function of vertex position X and bottom plot shows the ratio of data/MC.

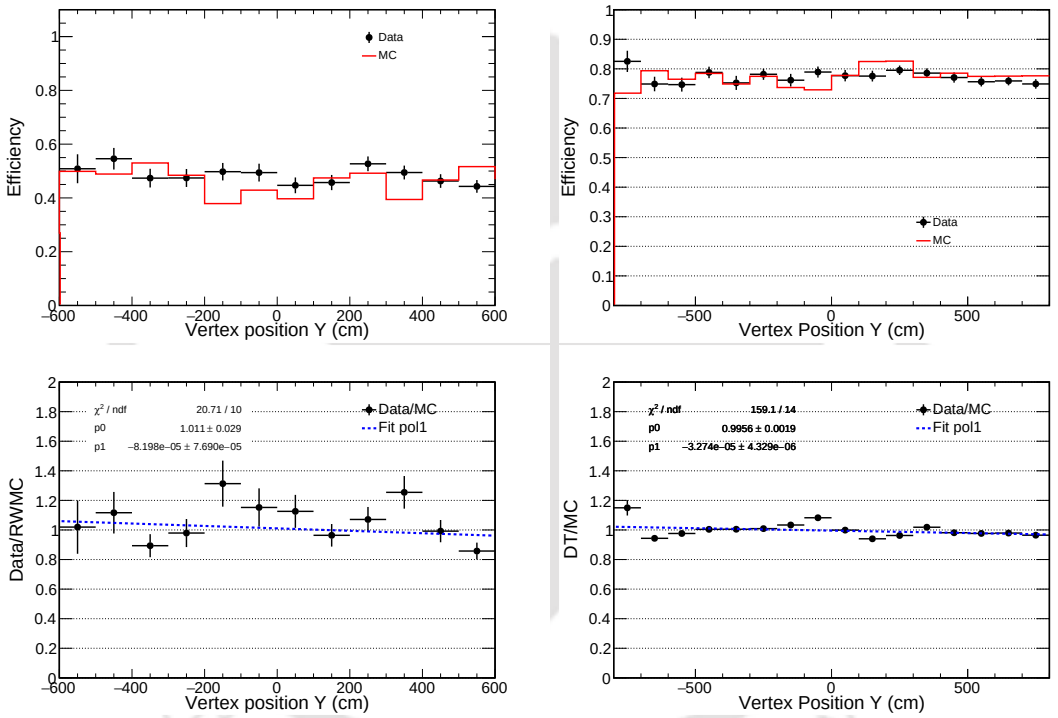

FIGURE 6.22: Left: Top plot shows the reweighed DiF data and MC comparison of detector efficiency, for LID >0.9, as a function of vertex position Y and bottom plot shows the ratio of data/MC. Right: Top plot shows the reweighed Brem data and MC comparison of detector efficiency, for LID >0.9, as a function of vertex position Y and bottom plot shows the ratio of data/MC.

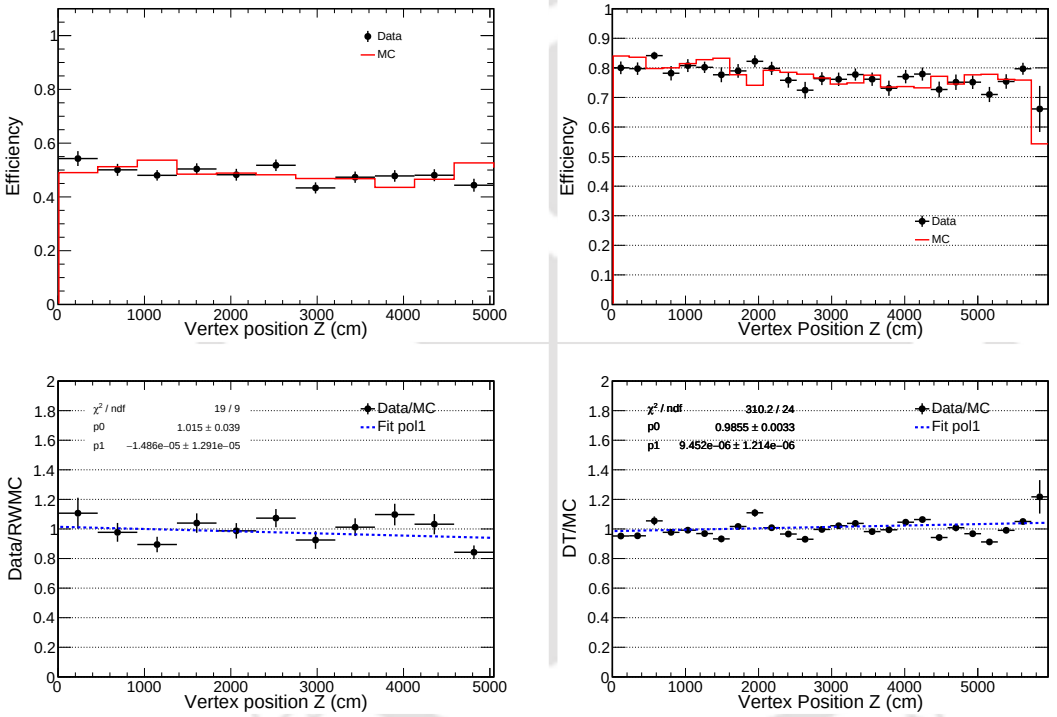

FIGURE 6.23: Left: Top plot shows the reweighed DiF data and MC comparison of detector efficiency, for LID >0.9, as a function of vertex position Z and bottom plot shows the ratio of data/MC. Right: Top plot shows the reweighed Brem data and MC comparison of detector efficiency, for LID >0.9, as a function of vertex position Z and bottom plot shows the ratio of data/MC.

Appendix A

Acronyms and Abbreviations

NOνA: NuMI Off-Axis ν_e Appearance

NuMI: Neutrino from the Main Injector

ν_μ: Muon Neutrino

ν_e: Electron Neutrino

π^0: Pion with zero EM charge

μ: Muon

e: Electron

δ_{CP}: Dirac CP-Violating Phase

$CCQE$: Charged-Current Quasi Elastic

CC: Charged-Current

NC: Neutral-Current

EM: Electromagnetic

Brem: Bremsstrahlung

DiF: Decay in Flight

MR: Muon Removal

ADC : Analogue-to-Digital Converter

ASIC: Application Specific Integrated Circuit

FPGA: Field Programmable Gate Array

APD: Avalanche Photodiode

CAF: Common Analysis Framework

DCM: Data Concentrator Module

DDT: Data-Driven Triggers

TEC: Thermo Electric Cooler

DAQ: Data Acquisition

FEB: Front End Board

FD: Far Detector

ND: Near Detector

LEM: Library Event Matcher

MC: Monte Carlo Simulation

MIP: Minimum Ionizing Particle

PID: Particle Identification

LID: Likelihood Based Identifier

PMNS: Pontecorvo-Maki-Nakagawa-Sakata

POT: Proton on Target

FHC: Forward Horn Current

RHC: Reverse Horn Current

SNO: Sudbury Neutrino Observatory

Super-K: Super Kamiokande

Hyper-K: Hyper Kamiokande

T2K: Tokai to Kamioka

WLS: Wavelength Shifting

JUNO: Jiangmen Underground Neutrino Observatory

RENO: Reactor Experiment for Neutrino Oscillation

Bibliography

[1] *Close, Frank (2010). Neutrinos (softcover ed.). Oxford University Press. ISBN 0-199-69599-7.*

[2] S. Mertens, J. Phys. Conf. Ser. **718**, 022013 (2016).

[3] A. J. Cuesta, V. Niro, and L. Verde, Phys. Dark Univ. **13**, 77 (2016).

[4] J. Lesgourgues and S. Pastor, Adv. High Energy Phys. **2012**, 608515 (2012).

[5] ETH Library's Knowledge Portal, Wolfgang pauli and modern physics.

[6] E. Fermi, Z. Phys. **88**, 161 (1934).

[7] H. Bethe and R. Peierls, Nature **133**, 532 (1934).

[8] B. Pontecorvo, Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol. **1**, 25 (1991).

[9] M. R. Krishnaswamy et al., Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences **323**, 511 (1971).

[10] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, Physical Review **105**, 1413 (1957).

[11] K. Lande et al., AIP Conf. Proc. **243**, 1122 (1992).

[12] J. N. Bahcall, eConf **C030626**, THAT04 (2003).

[13] T. J. Haines et al., Phys. Rev. Lett. **57**, 1986 (1986).

[14] Q. R. Ahmad et al., Phys. Rev. Lett. **87**, 071301 (2001).

[15] Y. Fukuda et al., Phys. Rev. Lett. **81**, 1562 (1998).

[16] Wikipedia NOνA Experiment.

[17] A. Yu. Smirnov, The MSW effect and solar neutrinos, in *Neutrino telescopes. Proceedings, 10th International Workshop, Venice, Italy, March 11-14, 2003. Vol. 1+2*, pages 23–43, 2003.

[18] P. Adamson et al., Phys. Rev. **D93**, 051104 (2016).

[19] K. Abe et al., Phys. Rev. **D83**, 052010 (2011).

[20] B. Aharmim et al., Phys. Rev. **C81**, 055504 (2010).

[21] A. Gando et al., Phys. Rev. **D83**, 052002 (2011).

[22] F. P. An et al., Phys. Rev. Lett. **112**, 061801 (2014).

[23] S.-H. Seo, AIP Conf. Proc. **1666**, 080002 (2015).

[24] J. I. Crespo-Anadn, Nucl. Part. Phys. Proc. **265-266**, 99 (2015).

[25] K. Abe et al., Phys. Rev. Lett. **112**, 061802 (2014).

[26] P. Adamson et al., Phys. Rev. Lett. **112**, 191801 (2014).

[27] K. Nakamura and S. T. Petcov, Neutrino masses, mixing and oscillations.

[28] H. Duyang, NOνA Internal Document no.7742: Cosmic Ray Muon Induced EM Shower in NOνA.

[29] H. Duyang, NOνA Internal Technote no.8973: A Study of Decay-in-Flight Cosmic Muon Induced Electron Shower.

[30] A. Yu. Smirnov, Phys. Scripta **T121**, 57 (2005).

[31] P. Adamson et al., Nucl. Instrum. Meth. **A806**, 279 (2016).

[32] D. D'Angelo, Pramana **79**, 757 (2012).

[33] B. Aharmim et al., Phys. Rev. **C75**, 045502 (2007).

[34] S. Abe et al., Phys. Rev. Lett. **100**, 221803 (2008).

[35] B.-Z. Hu, Recent Results from Daya Bay Reactor Neutrino Experiment, in *Proceedings, 50th Rencontres de Moriond Electroweak Interactions and Unified Theories: La Thuile, Italy, March 14-21, 2015*, pages 229–234, 2015.

[36] T. Kajita, E. Kearns, and M. Shiozawa, Nuclear Physics B **908**, 14 (2016), Neutrino Oscillations: Celebrating the Nobel Prize in Physics 2015.

[37] A. Radovic, Joint Experimental-Theoretical Physics Seminar at Fermilab on Latest Oscillation Results from NOνA, 2018.

[38] I. Esteban, M. C. Gonzalez-Garcia, M. Maltoni, I. Martinez-Soler, and T. Schwetz, Journal of High Energy Physics **2017**, 87 (2017).

[39] E. D. Niner, *Observation of Electron Neutrino Appearance in the NuMI Beam with the NOvA Experiment*, PhD thesis, Indiana U., 2015.

[40] R. M. Zwaska, *Accelerator Systems and Instrumentation for the NuMI Neutrino Beam*, PhD thesis, Texas U., 2005.

[41] D. Ayres et al., The NOvA Technical Design Report, Technical report, 2007.

[42] K. Sachdev, *Muon Neutrino to Electron Neutrino Oscillation in NOνA*, PhD thesis, Minnesota U., 2015.

[43] S. M. Kasahara, Physics Procedia **37**, 1876 (2012), Proceedings of the 2nd International Conference on Technology and Instrumentation in Particle Physics (TIPP 2011).

[44] D. R. Rocco, *Muon Neutrino Disappearance in NOvA with a Deep Convolutional Neural Network Classifier*, PhD thesis, Minnesota U., 2016.

[45] C. Backhouse and R. B. Patterson, Nucl. Instrum. Meth. **A778**, 31 (2015).

[46] K. Sachdev, NOνA Internal Technote no.9729: Muon-Removed Charged Current.

[47] O. Hahn and L. Meitner, Phys. Zeit. **11**, 493497 (1910).

[48] W. Pauli, Phys. Today **31N9**, 27 (1978).

[49] J. Chadwick, Nature **129**, 312 (1932).

[50] S. L. Glashow, Nucl. Phys. **22**, 579 (1961).

[51] S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967).

[52] S. M. Bilenky and J. Hosek, Phys. Rept. **90**, 73 (1982).

[53] R. N. Mohapatra and A. Y. Smirnov, Ann. Rev. Nucl. Part. Sci. **56**, 569 (2006).

[54] P. B. Pal, Am. J. Phys. **79**, 485 (2011).

[55] B. Kayser, Neutrino Oscillation Phenomenology, in *Neutrinos in particle physics, astrophysics and cosmology. Proceedings, 61st Scottish Universities Summer School in Physics, SUSSP61, St. Andrews, UK, August 8-23, 2006*, pages 51–64, 2008.

[56] Z. Maki, M. Nakagawa, and S. Sakata, Progress of Theoretical Physics **28**, 870 (1962).

[57] K. Abe et al., Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **659**, 106 (2011).

[58] R. B. Patterson, (2012), [Nucl. Phys. Proc. Suppl.235-236,151(2013)].

[59] F. P. An et al., Chinese Physics C **37**, 011001 (2013).

[60] Y. Abe et al., Phys. Rev. Lett. **108**, 131801 (2012).

[61] J. K. Ahn et al., Phys. Rev. Lett. **108**, 191802 (2012).

[62] J. N. Bahcall and M. Cribier, The standard solar model, in *Inside the Sun*, edited by G. Berthomieu and M. Cribier, pages 21–41, Dordrecht, 1990, Springer Netherlands.

[63] R. Davis, Phys. Rev. Lett. **12**, 303 (1964).

[64] J. N. Abdurashitov et al., J. Exp. Theor. Phys. **95**, 181 (2002), [Zh. Eksp. Teor. Fiz.122,211(2002)].

[65] F. Kaether, W. Hampel, G. Heusser, J. Kiko, and T. Kirsten, Phys. Lett. **B685**, 47 (2010).

[66] J. N. Bahcall, M. H. Pinsonneault, and S. Basu, Astrophys. J. **555**, 990 (2001).

[67] T. Kajita, Adv. High Energy Phys. **2012**, 504715 (2012).

[68] B. Aharmim et al., Phys. Rev. **C88**, 025501 (2013).

[69] B. H. J. and H. W., Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **159**, 432 (1937).

[70] B. Adhikary, A. Ghosal, and P. Roy, JHEP **10**, 040 (2009).

[71] H. Duyang, Measurement of Neutral Current Coherent π^0 Production In The NOvA Near Detector, in *Meeting of the APS Division of Particles and Fields (DPF 2017) Batavia, Illinois, USA, July 31-August 4, 2017*, 2017.

[72] J. Bian, Measurement of Neutrino-Electron Elastic Scattering at NOvA Near Detector, in *Meeting of the APS Division of Particles and Fields (DPF 2017) Batavia, Illinois, USA, July 31-August 4, 2017*, 2017.

[73] J. Zlek et al., Journal of Physics: Conference Series **513**, 012041 (2014).

[74] Q. Lu, J. B. Kowalkowski, and K. A. Biery, Journal of Physics: Conference Series **331**, 022017 (2011).

[75] A. Aurisano et al., Journal of Physics: Conference Series **664**, 072002 (2015).

[76] T. Bhlen et al., Nuclear Data Sheets **120**, 211 (2014).

[77] A. Ferrari, P. R. Sala, A. Fasso, and J. Ranft, Fluka: A multi-particle transport code (program version 2005), Technical report, 2005.

[78] S. Agostinelli et al., Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **506**, 250 (2003).

[79] J. Allison et al., IEEE Transactions on nuclear science **53**, 270 (2006).

[80] C. Hagmann, D. Lange, and D. Wright, Lawrence Livermore National Laboratory, UCRL-TM-229453, March (2012).

[81] M. Baird et al., Journal of Physics: Conference Series **664**, 072035 (2015).

[82] M. D. Baird, NOνA Internal Technote no.9195: A Side By Side Comparison of Slicer, Cosmic Slicer, and Slicer4D.

[83] B. Rebel, NOνA Internal Technote no.4855: Cosmic Ray Tracker.

[84] N. J. Raddatz, NOνA Internal Technote no. 13545: KalmanTrack Technical Note.

[85] M. D. Baird, NOνA Internal Technote no.8241: Global Vertex Reconstruction beginning with a Modified Hough Transform.

[86] M. Messier, NOνA Internal Technote no.7530: Vertex reconstruction based on elastic arms.

[87] M. Ohlsson, C. Peterson, and A. L. Yuille, Computer Physics Communications **71**, 77 (1992).

[88] R. Krishnapuram and J. M. Keller, IEEE Transactions on Fuzzy Systems **1**, 98 (1993).

[89] J. B. et. al, NOνA Internal Technote no.9923: ν_e identification using shower-shape-based particle likelihoods and artificial neural network.

[90] A. Aurisano et al., JINST **11**, P09001 (2016).

[91] Y. Jia et al., CoRR **abs/1408.5093** (2014).

[92] P. Adamson et al., Phys. Rev. Lett. **116**, 151806 (2016).

[93] H. Duyang, NOνA Internal Document no.13325: Rock Muon Induced EM Showers in ND.

[94] X. Bu, Measurement of Electron Neutrino Charged-Current Inclusive Cross Section in 1-3 GeV energy region with the NOvA Near Detector, in *Proceedings, 10th International Workshop on Neutrino-Nucleus Interactions in the Few GeV Region (NuInt15): Osaka, Japan, November 16-21, 2015*, 2016.

[95] M. D. Baird, NOνA Internal Technote no.8241: Global Vertex Reconstruction beginning with a Modified Hough Transform.

[96] M. Messier, NOνA Internal Technote no.7530: Vertex Reconstruction Based on Elastic Arms.

[97] E. Niner, NOνA Internal Technote no.7648: Vertex Clustering with Possibilistic Fuzzy-K Means Algorithm.

[98] NOνA Collaboration, New constraints on oscillation parameters from ν_e appearance and ν_μ disappearance in the NOνA experiment, arXiv:1806.00096 (2018).